

ÁNGEL HERRÁEZ

Texto ilustrado e interactivo de

Biología molecular e ingeniería genética

— 2.^a edición —

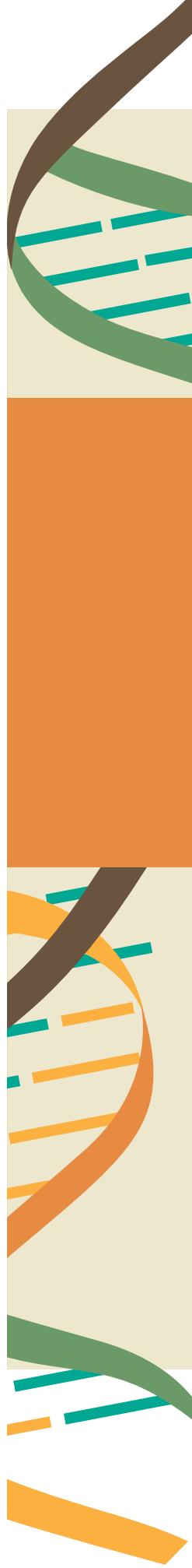
Conceptos, técnicas y aplicaciones
en ciencias de la salud

booksmedicos.org



Sección I

Estructura y función del material genético





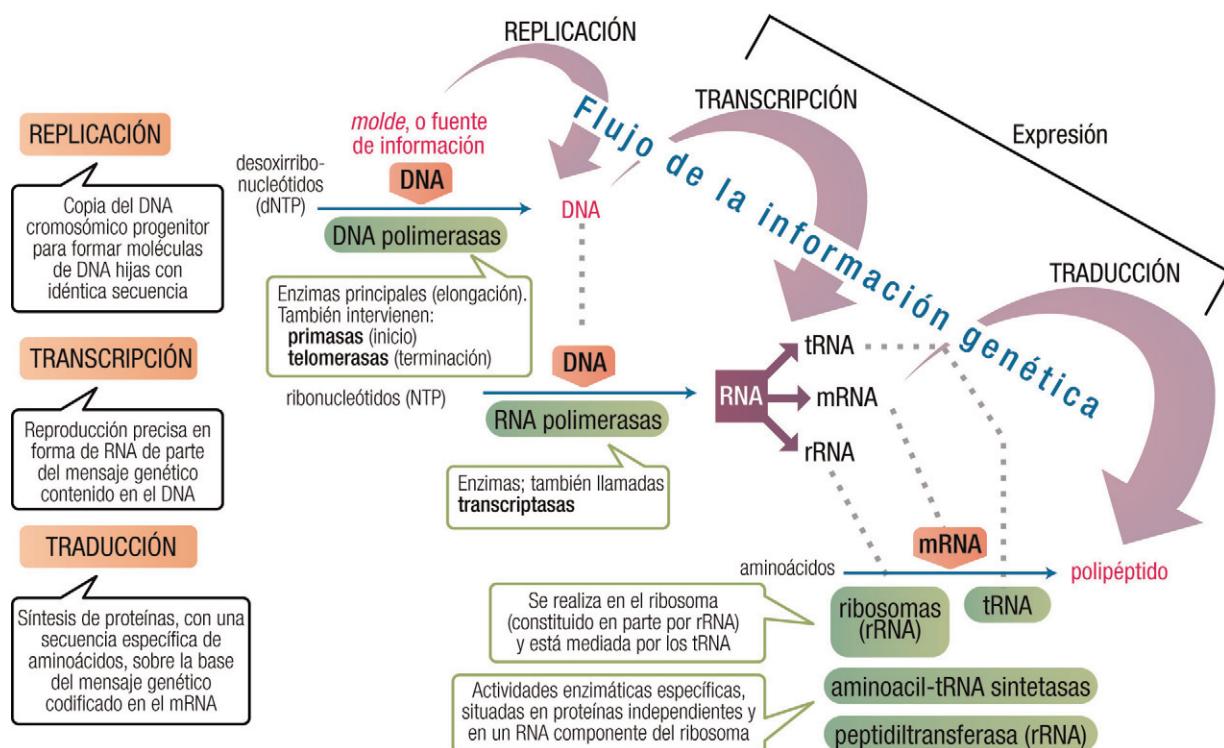
Introducción y aspectos generales

1.1 TRANSMISIÓN DE LA INFORMACIÓN GENÉTICA	3
1.2 EL DNA COMO MATERIAL GENÉTICO	4
1.3 CARACTERÍSTICAS GENERALES DEL GENOMA	5

1.3.1 Genoma de procariotas y eucariotas	5
1.3.2 Magnitud del genoma nuclear de eucariotas	7
1.3.3 Magnitud y características del genoma mitocondrial	10

A modo de introducción general al planteamiento de la biología molecular y la ingeniería genética en este texto, se comenzará describiendo de modo sucinto los tres aspectos que mejor ponen de manifiesto el papel de los ácidos nucleicos en la célula y en especial del DNA como portador de la información genética: por un lado, la descripción global de las rutas de transmisión de la información genética; en segundo lugar, los antecedentes experimentales que demostraron el importante papel de los ácidos nucleicos; finalmente, las características generales del genoma de eucariotas, como punto de partida para el estudio posterior en profundidad de los conceptos teóricos, las tecnologías y las aplicaciones de esta materia en las ciencias de la salud.

1.1 TRANSMISIÓN DE LA INFORMACIÓN GENÉTICA



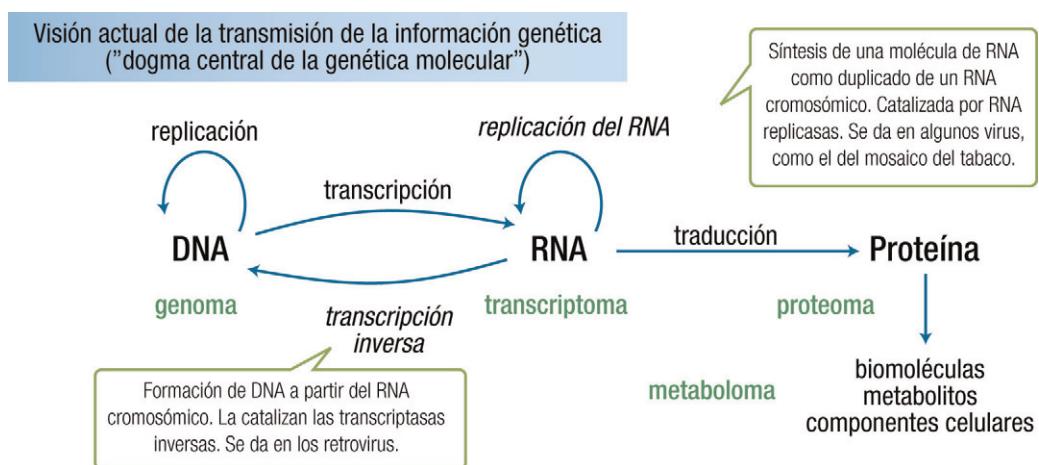
1.1

Francis Crick introdujo el término “dogma central de la genética molecular” para describir, incluso antes de conocerse por completo su relevancia en los seres vivos, el flujo de la información genética y la utilización

de dicha información en las células. En concreto, la transmisión de la información contenida en la secuencia del DNA a las células y organismos descendientes mediante la replicación, y su expresión en biomoléculas funcionales mediante los procesos de transcripción y traducción. Esta descripción comprende, pues, el conjunto de relaciones generales existentes entre el DNA, el RNA y las proteínas.

Debe evitarse la posible interpretación errónea de la representación abreviada de este “dogma”, **DNA → RNA → proteína**, que no indica la transformación de una molécula en otra, tal como se utiliza normalmente en las rutas metabólicas, sino que cada elemento en este esquema aporta la información necesaria para la síntesis del siguiente. En este sentido, es importante el concepto de *plantilla* o *molde* molecular: aunque cada paso está catalizado por una enzima, los sustratos (nucleótidos, aminoácidos) no son seleccionados por dicha enzima, sino que vienen especificados por la intervención del molde, en concreto por su secuencia de nucleótidos.

El progreso en el conocimiento de los procesos de transmisión de la información genética obligó a la ampliación de este esquema básico original para acomodar otros procesos que también ocurren, al menos en algunos organismos particulares, como es el caso de algunos virus:



1.2

Debe resaltarse también el papel esencial de los productos finales (proteínas y algunos RNA) como componentes estructurales de todas las células y como responsables de la catálisis de las reacciones bioquímicas que conducen a la formación de las biomoléculas.

I.2 EL DNA COMO MATERIAL GENÉTICO

El conocimiento de la estructura de los ácidos nucleicos se inicia en el siglo XIX con el aislamiento a partir de núcleos celulares de un compuesto denominado inicialmente **nucleína**, formado por una parte ácida (hoy, DNA) y una parte básica (proteína), así como con el estudio parcial de las propiedades de la nucleína y de su relación con la herencia celular. Sin embargo, la estructura sólo se conoció a mediados del siglo XX, tras demostrarse que el DNA era el componente cromosómico depositario de la información genética. Cabe destacar en este sentido, como antecedentes clave, los experimentos de Frederick Griffith, Oswald T. Avery, Colin MacLeod y Maclyn McCarty, en 1928 y 1944, y de Alfred D. Hershey y Martha Chase, en 1952, que constituyen sendos hitos en la identificación del DNA como material genético.

www

Web 1.1. Experimentos que demostraron la identidad molecular del material genético.

Tras estas aportaciones iniciales identificando el DNA como portador de la información genética, el verdadero inicio de la biología molecular moderna lo constituyó la propuesta de estructura en doble hélice para el DNA, formulada en 1953 por James D. Watson y Francis Crick. A partir de este momento, casi todo el esfuerzo realizado se centra en el estudio del DNA genómico, desde el punto de vista estructural y funcional, en procariotas y eucariotas, con una atención especial a mamíferos en general y humanos en particular. Los avances logrados, tanto teóricos como técnicos y aplicados, han sido enormes. A ello ha contribuido la aparición de la ingeniería genética, como principal área derivada de la biología molecular. Una vez concluido en 2003 el Proyecto Genoma Humano (genómica estructural) se inició el análisis de las funciones de cada gen

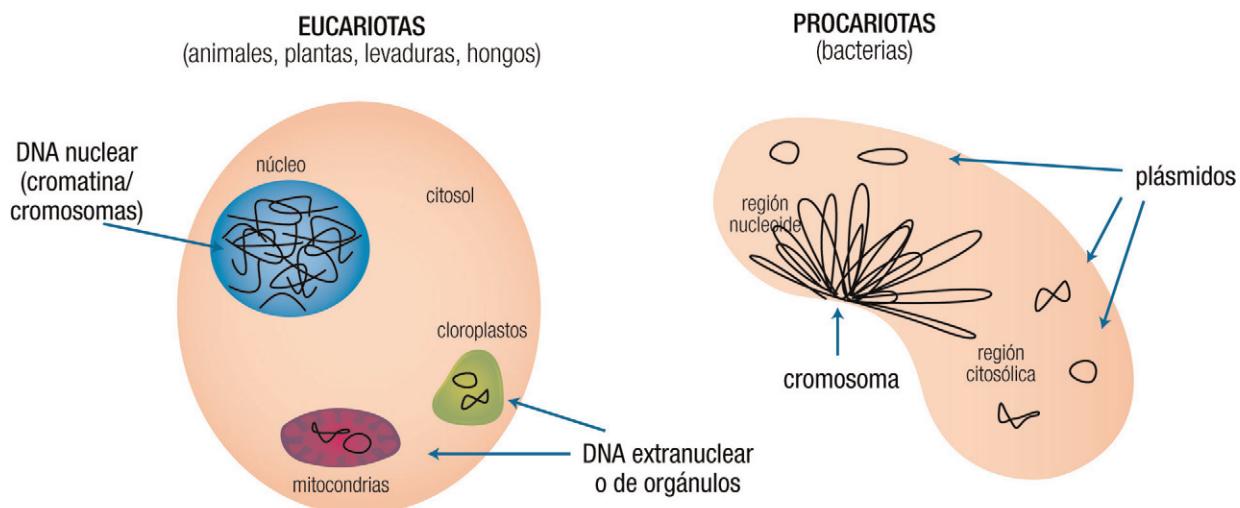
(genómica funcional) y la disección de la estructura y función de las proteínas (proteómica). Como expresión de estos nuevos conocimientos ha surgido una gran variedad de campos de actividad en las ciencias de la salud (medicina, farmacia y veterinaria), tales como anticuerpos monoclonales, métodos inmunoquímicos, biochips para el diagnóstico, clonación molecular y clonación animal, farmacogenes, ingeniería de proteínas, proteínas recombinantes de interés diagnóstico y terapéutico, sondas de hibridación, terapia celular y terapia génica, transferencia de genes, vectores, virus, etc. Nuestro objetivo no es describir los aspectos concretos de estos campos, sino sólo poner de manifiesto las bases moleculares que les sirven de soporte, como punto de partida para un estudio detenido de cada campo en su correspondiente área de conocimiento. Con este fin se presentan en este primer capítulo las características generales del genoma de eucariotas.

I.3 CARACTERÍSTICAS GENERALES DEL GENOMA

I.3.1 Genoma de procariotas y eucariotas

Conceptualmente, el genoma (de la raíz griega *gen*, origen) puede ser considerado como el “patrimonio genético” de un organismo, es decir, el conjunto total de material genético hereditario presente en una célula. Debe incluirse en este concepto tanto el material genético existente en el núcleo como el que contiene mitocondrias y cloroplastos. Por otra parte, es más adecuada esta definición que la que se emplea en ocasiones, “conjunto de genes de un organismo” pues, como se estudiará más adelante, existe gran cantidad de material genético que no se expresa. La mayoría de organismos poseen un genoma constituido por DNA, pero algunos virus tienen un genoma de RNA. La parte codificante del genoma especifica, en forma de genes, la información necesaria para dar lugar a productos génicos, moléculas funcionales que son las proteínas y los RNA. Estos productos son responsables de toda la actividad celular, de sus características morfológicas y de muchas de las claves del comportamiento de un organismo, desde la embriogénesis a su crecimiento, maduración física e intelectual, estado adulto, reproducción, y todas sus funciones metabólicas y actividades fisiológicas.

Con respecto al estudio del genoma, comenzaremos por resaltar las características diferenciales de procariotas y eucariotas:



Nota: Aunque, por claridad, en la figura no se haya reflejado, el tamaño de las células eucarióticas es, en general, bastante mayor comparativamente con las procarióticas. Además, la cantidad de DNA que forma los cromosomas eucarióticos es muy superior a la correspondiente al cromosoma procariótico. Igualmente, los cromosomas de orgánulos son mayores que los plásmidos.

1.3

Recuérdese que los **procariotas** (o **procariontes**) son organismos unicelulares (bacterias y arqueas) de tamaño pequeño (entre 0,2 y 5 µm de diámetro), con membrana plasmática rodeada generalmente por una pared celular rígida y con diversas estructuras superficiales (pili, flagelos, etc.), pero carecen de membrana nuclear, de orgánulos y de cualquier otro tipo de compartimentación interna. El **genoma de procariotas** es, comparativamente, de pequeño tamaño y relativamente sencillo. Se reduce a un único *cromosoma* formado por una molécula circular de DNA, localizada en suspensión en el citosol pero anclada a la membrana

constituyendo la *zona nuclear o nucleoide*, sin membrana que la delimita. Es frecuente encontrar, además, una pequeña cantidad de material genético extra, no esencial, en forma de pequeñas moléculas circulares de DNA, llamadas *plásmidos*. Los plásmidos poseen la propiedad de replicación autónoma, es decir, sin coordinación alguna con la división celular. Desempeñan un papel importante, al codificar proteínas que confieren resistencia a los antibióticos u otros materiales tóxicos, y ofrecen una importante aplicación como herramientas en la tecnología del DNA recombinante, concretamente como vectores para la incorporación de DNA a células.

Las células de **eucariotas** (o **eucariontes**), tanto multicelulares (animales y plantas) como unicelulares y multicelulares simples (protozoos, hongos y algas), son más variadas y complejas. Son de mayor tamaño (entre 10 y 50 μm), poseen membrana plasmática, pueden poseer *pared celular* (plantas), tienen citoesqueleto interno, y muestran una gran compartimentación citoplasmática, con diferentes tipos de orgánulos subcelulares (*mitocondrias* en animales y plantas, *cloroplastos* en plantas, *lisosomas* en animales, etc.) y otras estructuras (*retículo endoplásmico*, *complejo de Golgi*, etc.). El *núcleo* tiene como característica esencial la presencia de una *membrana nuclear* doble. En el genoma de eucariotas debe distinguirse entre el genoma nuclear y el de orgánulos. El **genoma nuclear** aparece bajo dos formas estructuralmente diferentes a lo largo del ciclo celular: la **cromatina** como material filamento disperso por la mayor parte del núcleo en la interfase (entre divisiones), y los **cromosomas**, fácilmente observables como entidades morfológicas o corpúsculos independientes en los momentos previos a la división celular. Ambas formas están constituidas por las mismas moléculas lineales muy largas de DNA bicatenario, estrechamente asociado a proteínas. El **genoma de orgánulos** (cloroplastos y mitocondrias) está formado por cromosomas circulares, al igual que el de procariotas.

Algunas características diferenciales del genoma humano (célula somática, diploide, 2n)

	NUCLEAR	MITOCONDRIAL
Tamaño	$6,4 \text{ Gb} = 6.400 \text{ Mb} = 6.400.000 \text{ kb} = 6,4 \cdot 10^9 \text{ pb}$	$16,5 \text{ kb} \times m$
Tipo de DNA	Lineal, bicatenario	Circular, bicatenario
Número de cromosomas	23 distintos, 2 copias: 46 en total	Uno, con varias copias en cada mitocondria; entre 200 y 2.000 copias por célula
Presencia de nucleosomas	Sí	No
Proteínas asociadas	Alrededor del 50% de la masa, histonas y otras	Poco numerosas o ausentes
DNA codificante	Entre 2 y 5%	~ 93%
DNA repetitivo	Abundante (~ 40%)	Prácticamente nada
Número de genes	~ 25.000	37
Densidad de genes	~ 1 gen cada 128 kb	1 gen cada 0,45 kb
Presencia de intrones en los genes	En la mayoría	Ausentes

Las cifras relativas al número de genes han sufrido fuertes variaciones en los años recientes y aún son dispares dependiendo de la fuente que se consulte. Ello se debe a que no es sencillo definir qué región del DNA genómico es o no un gen, e incluso identificar con certeza sus puntos de comienzo y de fin, los que definen un “marco de lectura abierto” (pág. 314); las técnicas de detección e identificación de genes en la secuencia del genoma continúan evolucionando. Como ejemplos de situaciones que hacen difícil esta precisión pueden mencionarse la existencia de copias de genes que no son funcionales (pseudogenes, pág. 113); el tamaño mucho menor de las regiones codificantes (exones) en la mayoría de genes frente a las no codificantes; la posibilidad de expresión en formas alternativas de algunos genes, dando lugar a diferentes productos (ayuste alternativo, pág. 306); los genes que codifican RNA funcionales son cortos y carecen de marco de lectura, etc. Todo ello complica la definición e identificación inequívocas de un segmento del genoma como un gen. Igualmente varían las cantidades derivadas, como la fracción del genoma constituida por regiones codificantes. Lo importante es, en todo caso, transmitir una idea de las cifras aproximadas, los diferentes tipos de secuencia implicados y las diferencias entre el genoma nuclear de los eucariotas y los genomas de orgánulos y de procariotas.

1.3.2 Magnitud del genoma nuclear de eucariotas

El material genético contenido en el núcleo supone generalmente más del 90% del total de DNA celular. El primer aspecto a destacar es su **magnitud** con respecto al genoma mitocondrial y al de procariotas. Además, a diferencia de éstos, el genoma nuclear está repartido en varios cromosomas, en número diferente según la especie, generalmente muy grandes y con el DNA muy condensado al estar estrechamente asociado con histonas y otras proteínas.

La magnitud del genoma nuclear viene determinada por la cantidad total de DNA en el conjunto de cromosomas de una célula. Ésta se puede expresar de varias formas:

- En número de cromosomas, designado por ***n*** en organismos haploides y ***2n*** en diploides, en los que hay 2 copias de cada cromosoma.
- En masa de DNA (**pg** = picogramos).
- Como **valor C**, notación que expresa la cantidad de DNA total en el genoma de una célula con respecto a la presente en una célula haploide de la misma especie. Como se verá, en células diploides el valor puede ser de **2C** o **4C**, dependiendo del estadio del ciclo celular.
- Lo más habitual es expresarlo en longitud: **pb** = pares de bases en bicatenario, o **nt** = nucleótidos en monocatenario. Para moléculas largas se emplean kilobases (**kb** o **kpb** = 10^3 pb) o megabases (**Mb** o **Mpb** = 10^6 pb).

Comparaciones que resaltan la magnitud del DNA nuclear en eucariotas:

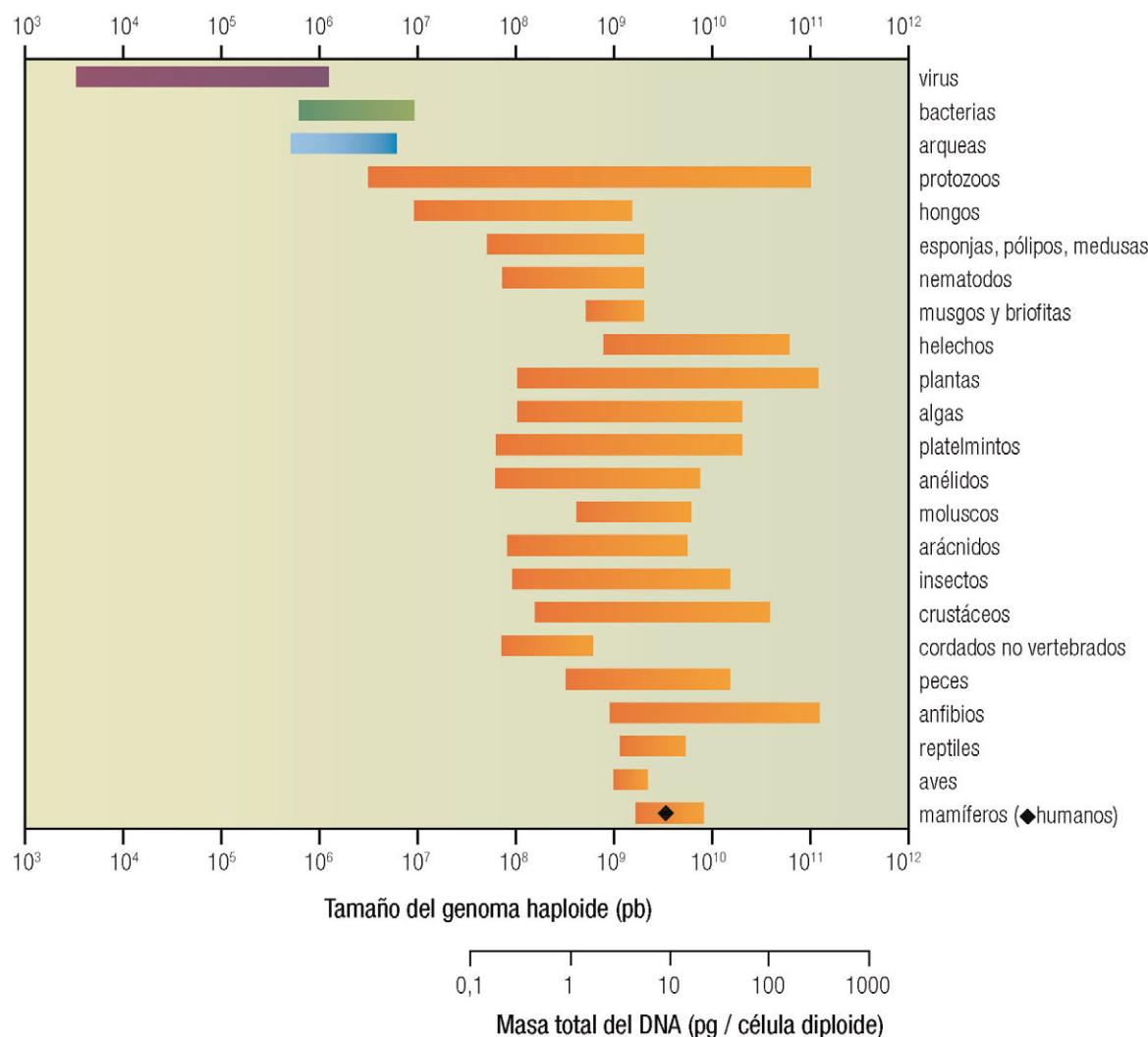
- Por regla general, la cantidad total de DNA en una célula diploide eucariótica es entre 10 y 1.000 veces superior que la de una célula procariótica, aunque la diferencia puede llegar hasta un factor de 200.000, por ejemplo, considerando los casos extremos de algunas plantas y anfibios respecto a la bacteria de menor tamaño, el micoplasma.
- En forma lineal, la longitud total del DNA de una célula humana somática (diploide, 46 cromosomas) sería de unos **2,2 metros**.
- Al referir esta cifra al número de células del cuerpo humano, la longitud total resultante, **110 billones de metros** ($110 \cdot 10^{12}$ m), equivale a casi 3 millones de vueltas a la Tierra, 143.000 viajes de ida y vuelta a la Luna o 370 viajes de ida y vuelta al Sol.
- Para escribir la secuencia completa del genoma de una célula humana diploide se necesitarían 1.280 tomos, que ocuparían 64 m de estantería.
- Para leer en voz alta dicha secuencia se necesitarían 20 años y 3 meses.
- Para almacenar la secuencia completa del genoma de una célula humana diploide en formato electrónico se requieren 1,5 GB (gigabytes).
- Las cifras se hacen casi incomprensibles si se quiere considerar la magnitud total del DNA de la humanidad (actualmente $7 \cdot 10^9$ individuos).
- Datos empleados:
 - N.º de células en el cuerpo: estimado al menos en unos 50 billones = $50 \cdot 10^{12}$ células.
 - Circunferencia de la Tierra: 40.000 km.
 - Distancia promedio entre la Tierra y la Luna: 384.000 km.
 - Distancia promedio entre la Tierra y el Sol: 150 millones de km.
 - 5.000 letras por página y 1.000 páginas por tomo, con un grosor de 5 cm.
 - Velocidad de lectura: 10 letras por segundo.

Procede, por otra parte, analizar la magnitud del genoma de forma comparativa tanto entre individuos de una misma especie como entre especies diferentes:

- a) Todas las células de los individuos de **una misma especie** poseen la misma cantidad total de DNA y número de cromosomas. De esta generalización deben exceptuarse las células germinales, dedicadas a la reproducción del individuo, que poseen la mitad. En todos los casos, estas cifras se refieren a células en un mismo estado de división celular (normalmente la interfase) pues, como se verá posteriormente

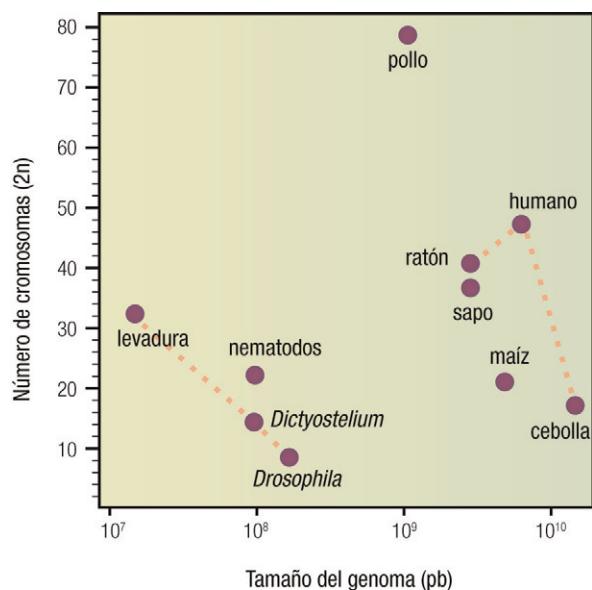
(capítulo 8), la cantidad total de DNA, aunque no el número de cromosomas, varía en una célula a lo largo del ciclo celular.

- b) El contenido total de DNA, expresado en pg o en n.º de pb, y el número de cromosomas varían ampliamente entre especies diferentes. A menudo se ha asumido que el tamaño del genoma pueda estar relacionado con el grado de complejidad del organismo, o su posición en la escala evolutiva, pero hoy en día está claro que esto no es así. Solamente al comparar virus, procariotas y eucariotas se aprecia una diferencia significativa en la magnitud del genoma. Dentro de los eucariotas, el tamaño del genoma varía ampliamente pero de forma independiente de la posición en la escala evolutiva o de lo que pudiera, bajo un punto de vista antropocéntrico, entenderse como *complejidad* de la especie. Así, por ejemplo, la célula humana contiene 700 veces más DNA que la de *E. coli*, pero 300 veces menos que las células de algunas salamandras y plantas angiospermas, e incluso menos que algunos protozoos (eucariotas unicelulares).



1.4

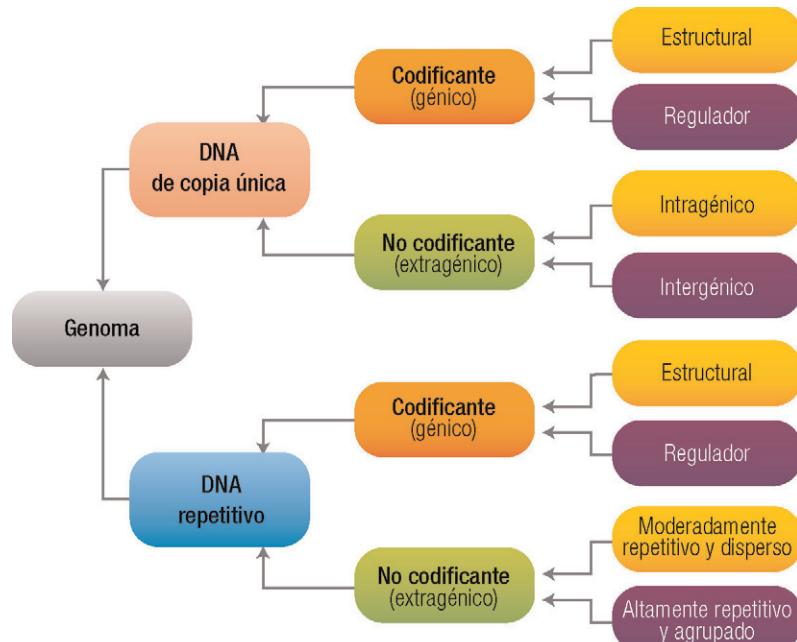
Por otro lado, no existe correlación alguna entre el tamaño del genoma de una especie y el número de cromosomas que lo componen. Por ejemplo, mientras que el genoma humano diploide ($6,4 \cdot 10^9$ pb) está repartido en 46 cromosomas (23 pares, $n = 23$), las células de ratón, con un genoma menor ($3 \cdot 10^9$ pb), poseen menor número de cromosomas (40), mientras que en las células de cebolla un genoma 2 veces mayor que el humano ($1,5 \cdot 10^{10}$ pb) está contenido en sólo 16 cromosomas. De forma similar, la mosca de la fruta, con un genoma ($1,7 \cdot 10^8$ pb) mucho mayor que el de la levadura ($1,6 \cdot 10^7$ pb), posee un número muy inferior de cromosomas (8 frente a 32).



1.5

Esta ausencia de correlación entre cantidad de material genético y complejidad del organismo es una de las pautas para sugerir que el tamaño de los genomas de organismos superiores es muy superior al necesario y, por tanto, que una gran parte es *no codificante*, es decir, no está organizado en genes, nunca traduce su información a un producto génico (RNA o proteína). De todos modos, se debe ser cauto al considerar una categorización de los organismos de acuerdo a su *complejidad*, pues es muy discutible que un ser humano sea más complejo que, digamos, un gusano. En cualquier caso, la hipótesis de una importante fracción no codificante en el genoma se ha confirmado. Por ejemplo, se ha encontrado que los genes tienen una longitud media de alrededor de 1.000 pb. Asumiendo esta cifra, el genoma humano haploide podría contener unos 3,2 millones de genes y, sin embargo, se estima hoy en día que sólo contiene unos 25.000, de modo que la fracción codificante sólo alcanza alrededor del 1% del DNA total.

Aunque la función de los genes está bien establecida para una pequeña fracción del total presente en el genoma humano, la diversidad en la cantidad total de DNA en los genomas puso de manifiesto la necesidad de estudiar con mayor precisión la organización estructural de genes y genomas a nivel molecular (capítulos 9 y 17). Las características más relevantes del genoma nuclear son la presencia de grandes regiones que no codifican producto génico alguno (*DNA no codificante*) y la existencia de secuencias de DNA que se repiten un número más o menos elevado de veces (*DNA repetitivo*). Como se estudiará con detenimiento posteriormente, hay que unir a ello la existencia de una enorme diversidad en la secuencia (polimorfismo, capítulo 24) en distintas regiones del DNA, tanto sencillas como repetitivas, codificantes o no.



1.6

1.3.3 Magnitud y características del genoma mitocondrial

El genoma de orgánulos (mitocondrias en animales, mitocondrias y cloroplastos en plantas) es, posiblemente, un vestigio del cromosoma de bacterias arcaicas que accedieron al citoplasma de eucariotas primitivos, para dar lugar tras la evolución a dichos orgánulos (hipótesis endosimbionte). Se cree que la mitocondria (orgánulo que aporta casi toda la energía que necesitan las células) surgió al acumularse el oxígeno en la atmósfera terrestre. Posiblemente, la mitocondria y el núcleo de la célula eucariótica se formaron en paralelo, al incorporarse por endocitosis células procariotas aeróbicas al interior de la célula eucariótica anaeróbica y fusionarse ambas. Con el tiempo, la mayoría de los genes procarióticos (genes protomitocondriales) se integraron en el genoma nuclear, con lo que el eucariota primitivo anaeróbico ya podía vivir en una atmósfera rica en oxígeno; sólo una pequeña fracción del genoma procariótico primigenio permaneció en la mitocondria.

Las células animales poseen un número muy variable de mitocondrias, dependiendo de la especie y del tejido. Cada mitocondria posee varias copias de un único cromosoma (habitualmente menos de una decena), situadas en la matriz mitocondrial y ancladas a la membrana interna. En consecuencia, el número de copias de DNA mitocondrial (mtDNA) en una célula oscila entre 200 y 2.000 (algunas referencias hablan de hasta 100.000 en casos puntuales). El cromosoma mitocondrial es bícatenario y circular, como el de procariotas, aunque de tamaño muy inferior: 16.569 pb en humanos, con una longitud de 5 µm y una masa molecular de 10 MDa; esto supone una longitud 3.000 veces inferior a la del cromosoma nuclear más pequeño. Teniendo en cuenta el número de copias totales (2 del genoma nuclear frente a múltiples, y muy variables, del mitocondrial), se puede calcular que el DNA mitocondrial total supone, dependiendo del tejido, entre un 0,05 y un 0,5% del DNA total de la célula (quizás hasta el 20% en los casos extremos mencionados). Finalmente, la situación es similar para el cromosoma de cloroplastos, que es igualmente bícatenario y circular, pero de mayor tamaño que el mitocondrial.

A diferencia del DNA nuclear, con abundantes repeticiones y regiones no codificantes, el mtDNA es en su casi totalidad no repetitivo y codificante. En concreto, el cromosoma mitocondrial humano contiene un total de 37 genes que suponen el 93% del DNA. Entre ellos se cuentan 2 genes para RNA ribosómico, 22 para RNA transferente y 13 para proteínas; éstas forman parte de los complejos enzimáticos respiratorios I, III, IV y V, pero suponen sólo el 5% de las proteínas mitocondriales, estando el resto codificadas por DNA nuclear.



Componentes de los ácidos nucleicos

2.1 COMPONENTE ÁCIDO: FOSFATOS	11	2.4.2 Tipos de nucleósidos	18
2.1.1 Difosfatos	12	2.4.2.1 Nucleósidos que forman parte de ácidos nucleicos	18
2.1.2 Trifosfatos	13	2.4.2.2 Otros nucleósidos de interés biológico y clínico	19
2.1.3 Tipos de enlaces fosfato	13		
2.2 COMPONENTE NEUTRO: AZÚCARES	14	2.5 ESTRUCTURA DE NUCLEÓTIDOS	19
2.3 COMPONENTE BÁSICO: BASES NITROGENADAS	14	2.5.1 Características generales	19
2.3.1 Bases nitrogenadas más frecuentes: 2 tipos	15	2.5.2 Nucleótidos sencillos	20
2.3.2 Otras bases de interés biológico y clínico	16	2.5.2.1 Nomenclatura	21
2.3.3 Propiedades fisicoquímicas de las bases purínicas y pirimidínicas	16	2.5.2.2 Estructura de los nucleósidos-monofosfato que forman parte de los ácidos nucleicos	23
2.3.3.1 Existencia de dipolos	16	2.5.3 Propiedades fisicoquímicas de nucleósidos y nucleótidos	23
2.3.3.2 Hidrofobia	16	2.5.3.1 Propiedades iónicas	23
2.3.3.3 Disposición coplanar de los anillos	16	2.5.3.2 Formación de complejos	23
2.3.3.4 Tautomería o isomería dinámica	16	2.5.3.3 Absorción de la luz en el ultravioleta	24
2.3.3.5 Propiedades ácido-base	16	2.5.3.4 Nucleótidos como moléculas de reserva energética	24
2.3.3.6 Absorción de la luz en el ultravioleta	16	2.5.3.5 Hidrólisis	24
2.4 ESTRUCTURA DE NUCLEÓSIDOS	17	2.5.4 Nucleótidos cíclicos y nucleótidos complejos	25
2.4.1 Características generales	17		

Los ácidos nucleicos (DNA y RNA) están constituidos por la unión de numerosos nucleótidos, de 4 tipos. Como introducción a su estudio, se consideran algunos aspectos de interés para cada uno de los componentes estructurales de dichas unidades, los nucleótidos:

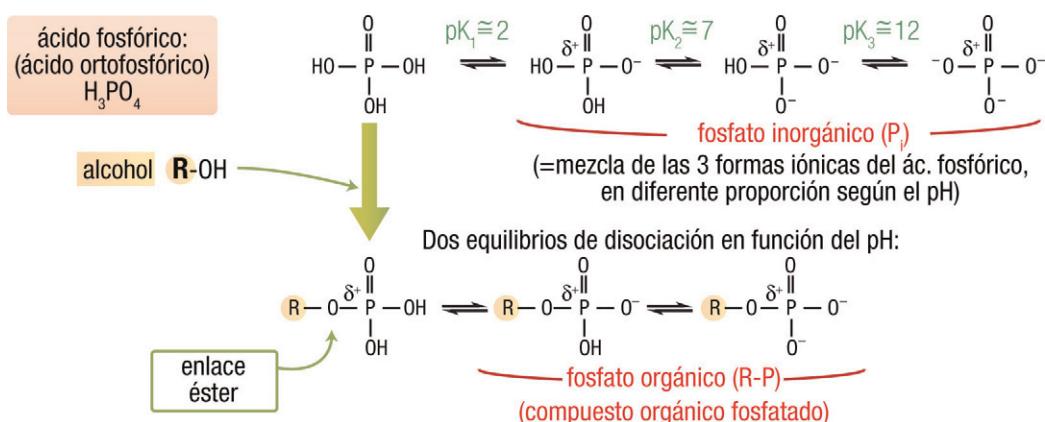


2.1

2.1 COMPONENTE ÁCIDO: FOSFATOS

El fosfato inorgánico (P_i), mezcla de las especies químicas resultantes de la disociación de la molécula de ácido fosfórico, aparece como anión a pH ácido, como dianión a pH fisiológico y como trianión a pH alcalino. En la célula, el fosfato se encuentra tanto en forma libre (P_i) como formando parte de numerosos compuestos orgánicos ($R-P$), resultado de un enlace éster entre el fosfórico y un grupo hidroxilo (alcohol); entre ellos se encuentran los nucleótidos.

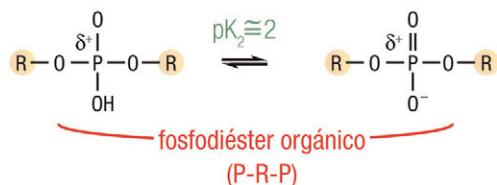
Tres equilibrios de disociación en función del pH:



2.2

Cuando un nucleótido se encuentra formando parte de un ácido nucleico, su grupo fosfato está esterificado al mismo tiempo con dos grupos alcohol (formando un enlace fosfodiéster, pág. 28), por lo que sólo existe un equilibrio de ionización. El pK de tal ionización es similar al pK_1 del ácido fosfórico, razón por la que se denominan ácidos nucleicos:

Un solo equilibrio de disociación en función del pH:

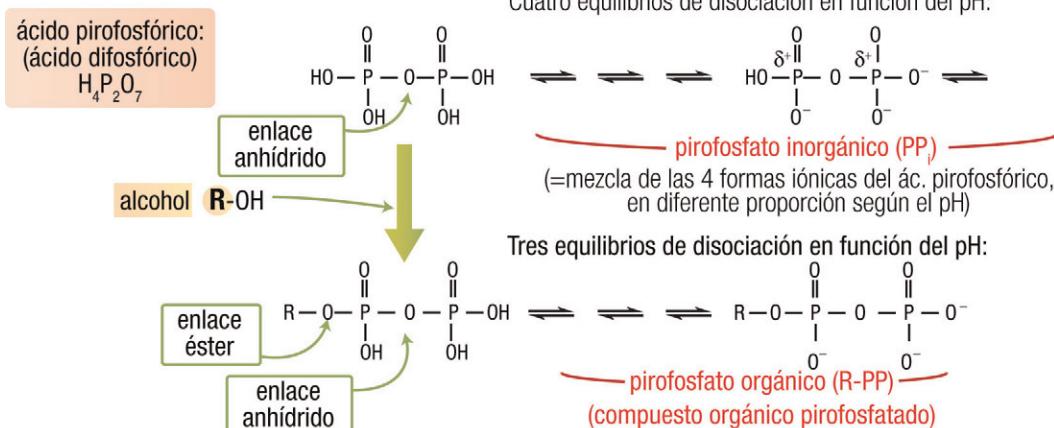


2.3

2.1.1 Difosfatos

El difosfato o pirofosfato (PP_i) se presenta a pH fisiológico como trianión (una de las 4 formas iónicas del ácido pirofosfórico). Es estable en disolución a pH neutro o alcalino; sin embargo, dentro de la célula es inestable por acción de las pirofosfatases, que lo hidrolizan formando 2 P_i . Aparece combinado fundamentalmente en nucleósidos-difosfato (por ejemplo, ADP).

Cuatro equilibrios de disociación en función del pH:

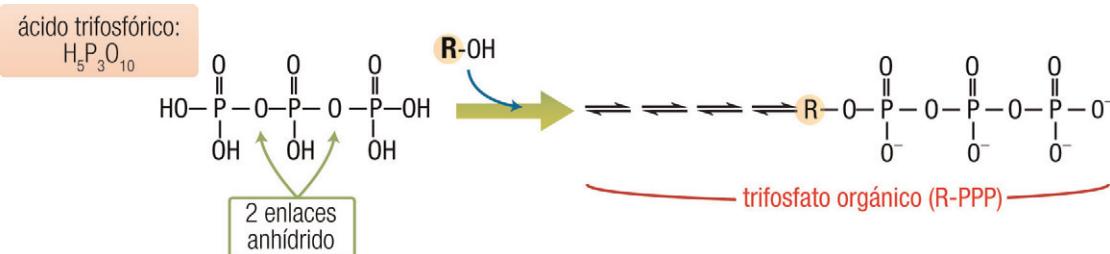


2.4

2.1.2 Trifosfatos

El trifosfato (PPP_i) se presenta a pH fisiológico como tetraanión. No aparece en forma libre en las células, sino formando parte, esencialmente, de nucleósidos-trifosfato (por ejemplo, ATP).

2.5

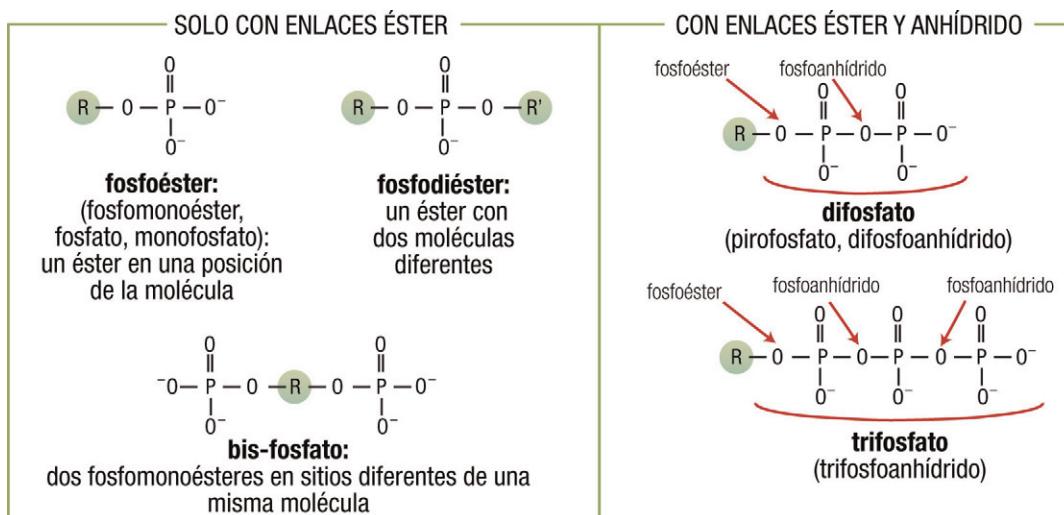


La elevada densidad de carga negativa presente en estos compuestos es responsable de que los nucleósidos-trifosfato, los nucleósidos-difosfato y el PP_i , e incluso el P_i , se encuentren en la célula en su mayor parte formando complejos con iones Mg^{2+} (págs. 23-24).

2.1.3 Tipos de enlaces fosfato

Pueden distinguirse cinco tipos de enlaces fosfato que van a formar parte de los nucleótidos. Obsérvese la estructura y nombre de todos ellos, así como la diferencia entre enlaces éster y anhídrido en los nucleósidos difosfatados y trifosfatados.

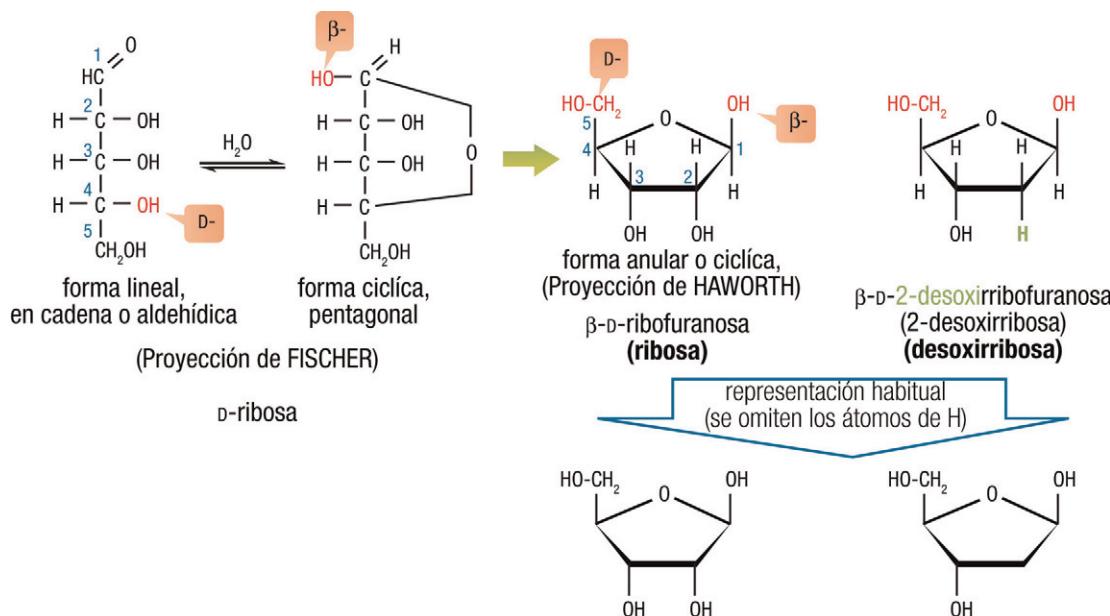
2.6



Más adelante se irán encontrando ejemplos de biomoléculas con estos enlaces, tales como el AMP (fosfoéster, pág. 20), los dinucleótidos, oligonucleótidos y ácidos nucleicos (fosfodiéster), la vitamina B₁₂ (fosfodiéster, pág. 19), el ADP (difosfato, pág. 20), el ATP (trifosfato, pág. 20), o las coenzimas NAD, FAD y CoA (diésteres pero también difosfatos, pág. 25).

2.2 COMPONENTE NEUTRO: AZÚCARES

El azúcar (en su denominación histórica, una *osa*) siempre es una pentosa (monosacárido con 5 átomos de carbono), bien la D-ribosa o la D-desoxirribosa. La **ribosa** (nombre derivado de *osa* del *Rockefeller Institute of Biochemistry*, 1908) interviene en nucleósidos y nucleótidos bajo su forma más estable, β-D-ribofuranosa.



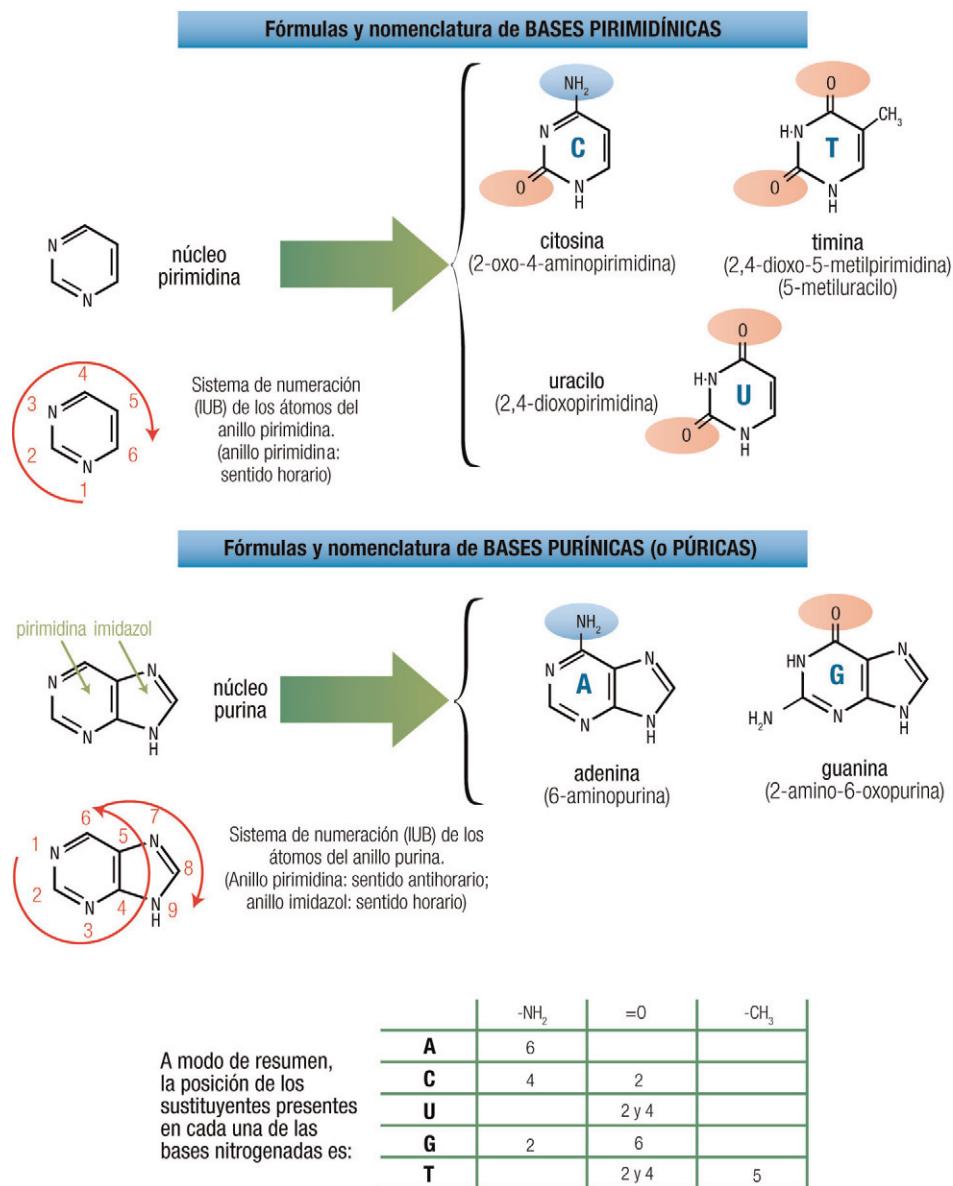
2.7

La fórmula representada como ejemplo puede extenderse a otras muchas moléculas de interés bioquímico, como D-glucosa, D-galactosa, 3-desoxi-D-ribosa, etc., aunque éstas no forman parte de los ácidos nucleicos, por lo que no se exponen aquí.

2.3 COMPONENTE BÁSICO: BASES NITROGENADAS

Son moléculas heterocíclicas, con varios átomos de nitrógeno, formadas bien por un anillo único (pirimidina) o bien por dos anillos condensados (purina). Obsérvense en los ejemplos de bases que intervienen en los ácidos nucleicos la numeración de los átomos del heterociclo y la posición de los sustituyentes externos.

2.3.1 Bases nitrogenadas más frecuentes: 2 tipos



2.8

2.3.2 Otras bases de interés biológico y clínico

Existen otras bases purínicas y pirimidínicas, en general poco frecuentes, que bien se sintetizan a partir de o bien pueden considerarse como derivados estructurales de las bases principales A, G, C, T y U o de la hipoxantina (H), una base purínica emparentada metabólicamente. Algunas existen de forma natural, mientras que otras se han obtenido de modo sintético. Algunas forman parte de la estructura de ácidos ribonucleicos, o presentan una función en el DNA (metilación de las citosinas como señal de regulación epigenética, pág. 278). Otras, en particular sintéticas, muestran utilidad como agentes antibióticos, antiviricos o antitumorales.



Web 2.1. Bases nitrogenadas infrecuentes y derivados de bases.



Web 2.2. Agentes antivirales: análogos de bases, nucleósidos y nucleótidos.

2.3.3 Propiedades fisicoquímicas de las bases purínicas y pirimidínicas

2.3.3.1 Existencia de dipolos

Todas las bases poseen átomos electronegativos de oxígeno (excepto la adenina), en posición exocíclica o extranuclear, y de nitrógeno, tanto exocíclicos como en el anillo, nucleares. Como consecuencia, son abundantes los enlaces polares, lo que les permite interaccionar entre sí mediante puentes de hidrógeno, que poseen una importancia capital en la estructura secundaria de los ácidos nucleicos, en especial del DNA (capítulo 4).

2.3.3.2 Hidrofobia

La naturaleza aromática de los anillos hace que las bases tengan un marcado carácter apolar, sean hidrófobas y poco solubles en agua al pH celular, cercano a la neutralidad. A pH ácido o alcalino adquieren carga y se hacen más solubles en agua. Como ya se estudiará, el efecto hidrófobo (es decir, el aislamiento de las bases evitando el contacto con el agua) y las **interacciones de apilamiento**, que tienen lugar entre bases dispuestas paralelamente (a modo de monedas apiladas), son esenciales para estabilizar la estructura tridimensional de los ácidos nucleicos.

2.3.3.3 Disposición coplanar de los anillos

La observación de modelos tridimensionales permite apreciar la coplanariedad de los anillos y de buena parte de sus sustituyentes, asociada con el carácter aromático.



Web 2.3. Estructura tridimensional de los anillos.

2.3.3.4 Tautomería o isomería dinámica

Los grupos funcionales ceto y amino de las bases están sometidos a equilibrios tautoméricos que afectan, entre otras cosas, a su capacidad de ionización y de formación de enlaces de hidrógeno. A pesar de ello, en las bases libres las formas tautoméricas lactama (ceto) y amina son las predominantes.



Web 2.4. Tautomería de las bases nitrogenadas.

2.3.3.5 Propiedades ácido-base

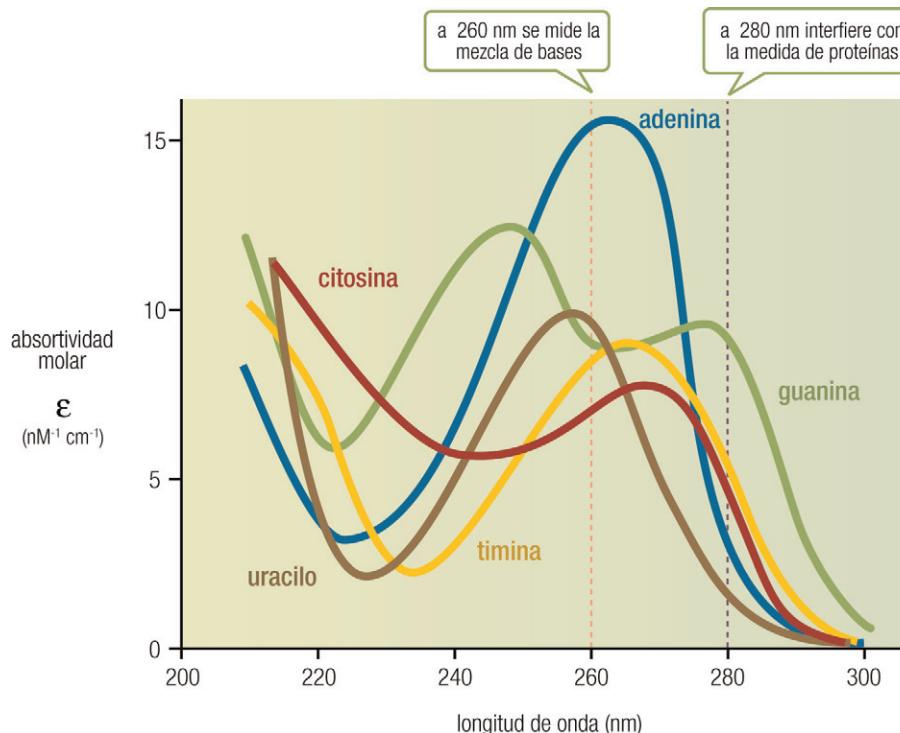
Todas las bases nitrogenadas son bases débiles, pues sus átomos de N, tanto nucleares como extranucleares, se pueden protonar, con valores de pK_b comprendidos entre 9 y 10. Sin embargo, este carácter básico, propio de los anillos purina y pirimidina, se ve disminuido por la presencia de algunos sustituyentes. En concreto, es ligeramente ácido el grupo -OH de las formas lactima tautoméricas en las bases nitrogenadas con grupos lactama (ceto).



Web 2.5. Propiedades ácido-base de las bases nitrogenadas.

2.3.3.6 Absorción de la luz en el ultravioleta

Ésta es una propiedad característica de las bases purínicas y pirimidínicas, debida a su carácter **aromático**. En los espectros de absorción, a pH = 7, se observa una fuerte absorción de la luz en el ultravioleta (UV), con un máximo cerca de 260 nm para todas ellas.



2.9

El mantenimiento de esta propiedad en nucleósidos, nucleótidos (pág. 24) y ácidos nucleicos, de los que las bases forman parte, permite su utilización analítica en estos compuestos; por ejemplo, para la cuantificación de ácidos nucleicos (pág. 32). Debe adelantarse, sin embargo, que la magnitud de la absorción se ve reducida cuando las bases forman parte de ácidos nucleicos, en especial si son bicanterios (págs. 32 y 163).



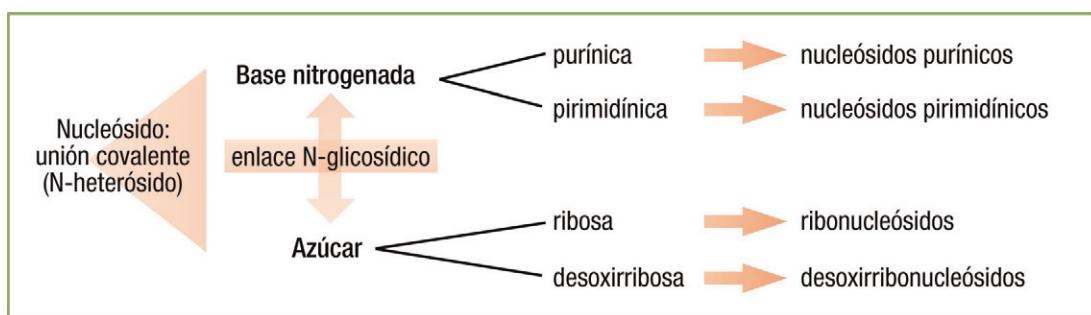
Web 2.6. Absorción de la luz por las bases nitrogenadas.

2.4 ESTRUCTURA DE NUCLEÓSIDOS

Corresponde al esquema general: nucleósido = base nitrogenada + azúcar.

También pueden considerarse como el producto de la hidrólisis de los nucleótidos, con liberación de fosfato inorgánico. Equivale, por tanto, a un nucleótido carente de fosfatos.

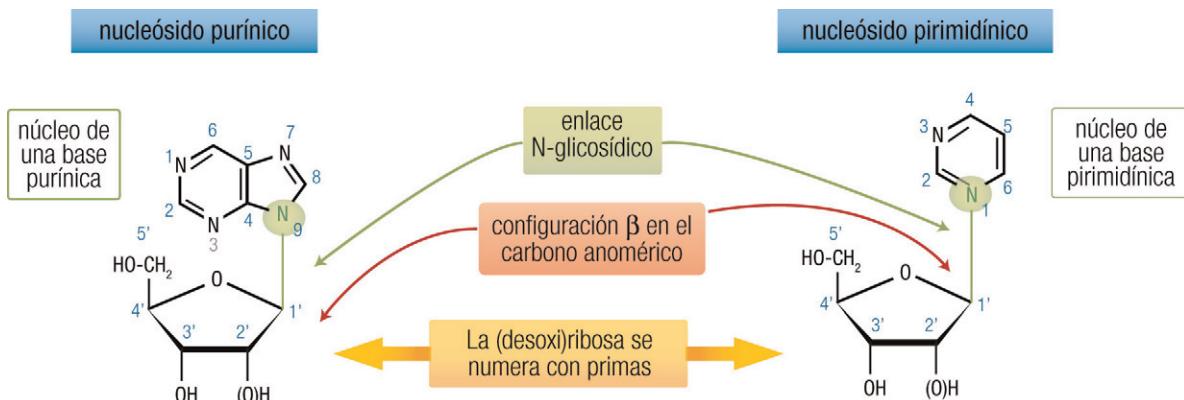
2.10



2.4.1 Características generales

La unión covalente (enlace N-glicosídico) entre la base y el azúcar (ribosa o desoxirribosa) se establece entre el C-1' (carbono anomérico) del azúcar y el N-1 del anillo pirimidínico o el N-9 del anillo purínico, con pérdida de una molécula de agua.

2.11



2.4.2 Tipos de nucleósidos

Debe distinguirse entre nucleósidos constituyentes de ácidos nucleicos y otros nucleósidos de interés biológico y clínico. Su descripción siempre se hace en función de la naturaleza química de la base y del azúcar, como se verá a continuación.

2.4.2.1 Nucleósidos que forman parte de ácidos nucleicos

Tipo	Nucleósidos purínicos		Nucleósidos pirimidínicos	
	ribonucleósidos (se encuentran en el RNA) y desoxirribonucleósidos (se encuentran en el DNA)	ribonucleósido (en RNA) y desoxirribonucleósido (en DNA)	desoxirribonucleósido (en DNA)	ribonucleósido (en RNA)
Estructura				
sufijo -osina			sufijo -idina	
Nombre	adenosina 9-(1'-β-D-ribofuranosil) adenina	guanosina 9-(1'-β-D-ribofuranosil) guanina	citidina 1-(1'-β-D-ribofuranosil) citosina	ribositimina (o ribotimidina) (no existe en RNA, excepcionalmente en tRNA)
	2'-desoxiadenosina 9-(2'-desoxi-1'-β-D-ribofuranosil)adenina	2'-desoxiguanosina 9-(2'-desoxi-1'-β-D-ribofuranosil)guanina	2'-desoxicitidina 1-(2'-desoxi-1'-β-D-ribofuranosil)citosina	uridina 1-(1'-β-D-ribofuranosil) uracilo 2'-desoxiuridina (no existe en DNA)

La distinción entre ribonucleósidos y desoxirribonucleósidos estriba, obviamente, en la presencia de ribosa o desoxirribosa en la molécula.

2.4.2.2 Otros nucleósidos de interés biológico y clínico

Además de los componentes de nucleótidos y ácidos nucleicos, existen en la naturaleza otros nucleósidos en forma libre o combinada. Algunos de ellos son significativos por su localización o acción bioquímica, e incluso por sus aplicaciones clínicas. Cabe citar los siguientes:

- Nucleósidos componentes de los RNA, como inosina, pseudouridina, dihidouridina, etc.
- Nucleósidos como formas activadas de los precursores biosintéticos: UDP-glucosa, CMP-siálico, etc.
- Componentes de coenzimas: S-adenosilmetionina, vitamina B₁₂, etc.
- Antibióticos, como la puromicina.
- Análogos de nucleósidos naturales con aplicación antiviral o anticancerosa. Destaca una amplia variedad de fármacos inhibidores de la transcriptasa inversa, que se emplean para combatir los retrovirus de la inmunodeficiencia humana (VIH-1 y VIH-2), causantes del sida.



Web 2.7. Estructuras de otros nucleósidos.

2.5 ESTRUCTURA DE NUCLEÓTIDOS

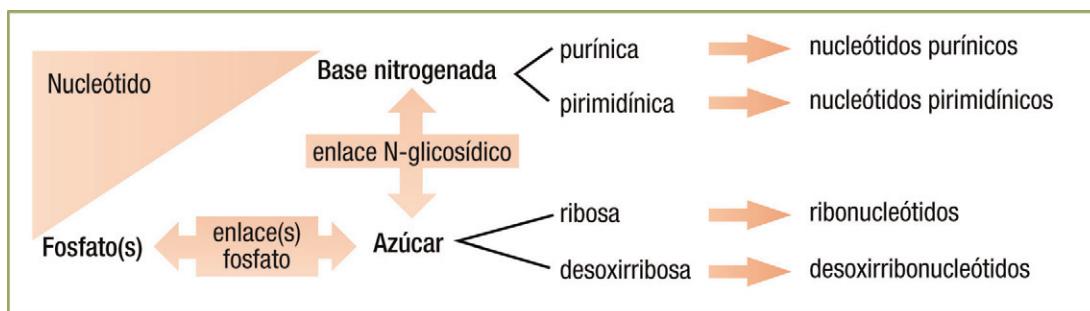
Corresponde al esquema general:

$$\text{nucleótido} = \text{nucleósido} + \text{fosfato(s)} = \text{base nitrogenada} + \text{azúcar} + \text{fosfato(s)}$$

También se puede considerar como:

$$\text{nucleótido} = \text{azúcar-fosfato} + \text{base nitrogenada}$$

2.13



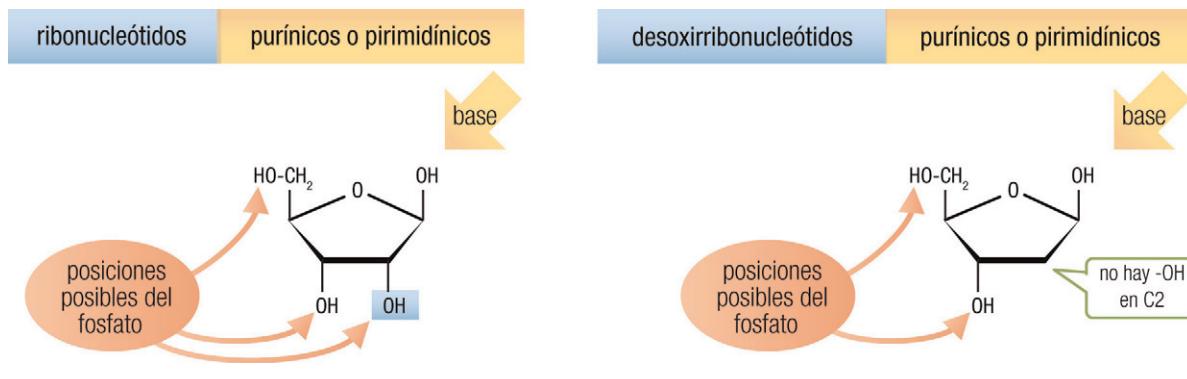
2.5.1 Características generales

La unión, de carácter covalente (fosfomonoéster), se establece entre uno o varios grupos fosfato y sendos grupos OH del azúcar del nucleósido.

La importancia de los nucleótidos es excepcional, principalmente porque son los monómeros constituyentes de los ácidos nucleicos y, por ello, tanto sus precursores biosintéticos como los productos de su hidrólisis. Al mismo tiempo, porque aparecen en forma libre en distintas localizaciones subcelulares, formando parte de moléculas de gran interés bioquímico (especialmente coenzimas).

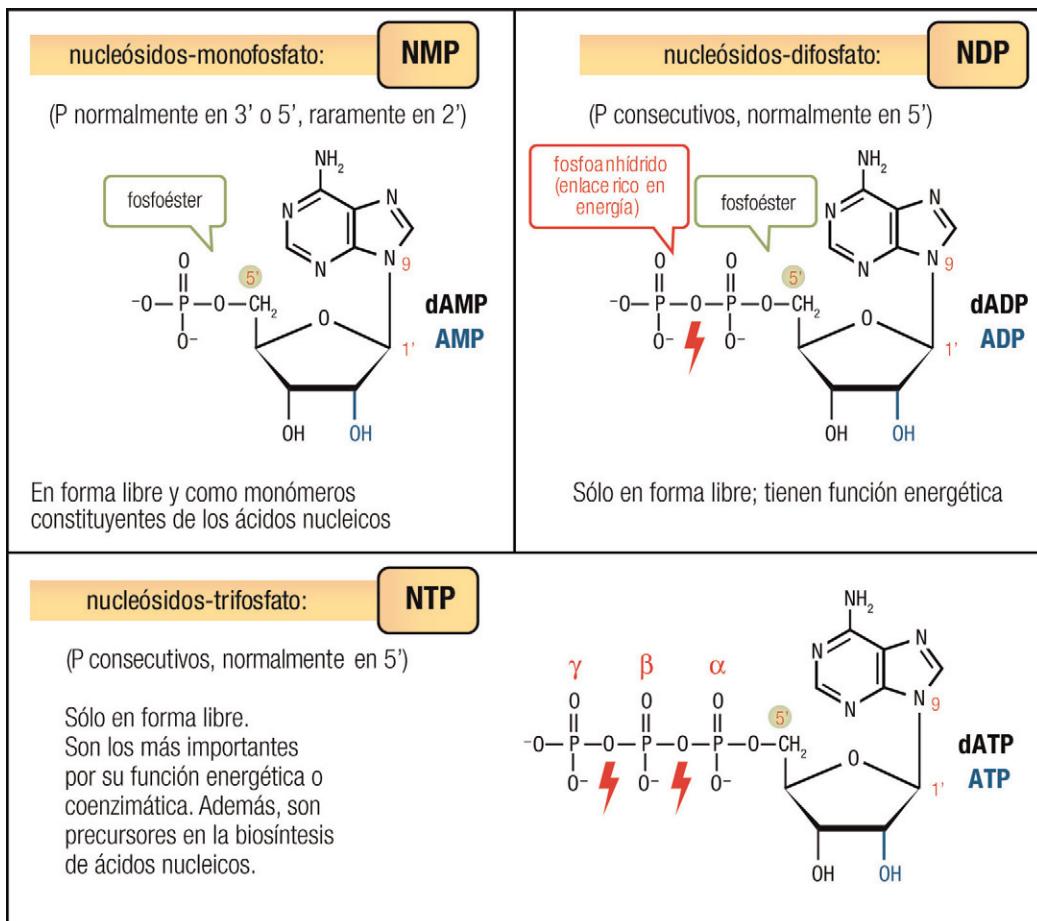
Para estudiar los nucleótidos se emplean distintos criterios de clasificación, atendiendo a:

- Características estructurales, criterio utilizado normalmente:
 - El número y posición en el azúcar de los grupos fosfato: nucleótidos-*monofosfato*, -*bifosfato* y -*trifosfato*.
 - El tipo de azúcar: ribonucleótidos y 2'-desoxirribonucleótidos.
 - El tipo de base nitrogenada: nucleótidos purínicos y pirimidínicos.
 - Su carácter lineal o cíclico (determinado por los enlaces del fosfato).
- Otras características, a las que sólo se alude ocasionalmente: procedencia o estado en la célula, función, etc.



2.5.2 Nucleótidos sencillos

Se consideran nucleótidos sencillos los que presentan fosfato(s) en una sola posición del azúcar; esto incluye tanto los que tienen un único fosfato (*monofosfatos*) como los que tienen dos o tres fosfatos que están unidos entre sí (*difosfatos* y *trifosfatos*). Se muestran como ejemplos los correspondientes a la adenina:



Los nucleótidos que forman parte de los ácidos nucleicos son siempre NMP con un fosfato en posición 5'. Su estructura y nomenclatura se estudian de forma detallada a continuación. Los nucleótidos con 2 o 3 fosfatos

consecutivos unidos en 5' (*difosfatos o NDP y trifosfatos o NTP*, respectivamente), dada su importante función bioquímica, son objeto de estudio en cursos de bioquímica general. Finalmente, otros nucleótidos, con fosfatos en múltiples posiciones del azúcar, se estudian posteriormente como nucleótidos cílicos o complejos (pág. 25).

2.5.2.1 Nomenclatura

Los nucleótidos se **nombran** según dos criterios:

- Como **ésteres fosfato del nucleósido**. Por ejemplo: adenosina-monofosfato, guanosina-difosfato, uridina-trifosfato, etc. (abreviados, respectivamente, como AMP, GDP, UTP).
- Como **ácidos**, sustituyendo la terminación del nombre de la base por el sufijo *-ílico*. Este criterio sólo se utiliza normalmente para los nucleósidos-monofosfato. Por ejemplo: ácido adenílico, ácido guanílico, ácido uridílico, etc. (respectivamente, para AMP, GMP y UMP). En el caso de nucleósidos-difosfato y -trifosfato, se emplearían los nombres diadenílico, triadenílico, diguanílico, etc. (para ADP, ATP y GDP). Si se quiere tener en cuenta que los restos fosfato están ionizados a pH fisiológico, se utiliza el sufijo *-ilato*: adenilato, guanilato, uridilato, etc.

Correspondencia entre los nombres de bases, nucleósidos y nucleótidos:

abreviatura común:	A	G	U	T	C
base:	adenina	guanina	uracilo	timina	citosina
nucleósido:	adenosina	guanosina	uridina	timidina	citidina
nucleótido: (nucleósido- monofosfato, NMP)	adenilato, ác. adenílico o AMP	guanilato, ác. guanílico o GMP	uridilato, ác. uridílico o UMP	timidilato, ác. timidílico o TMP	citidilato, ác. citidílico o CMP



Web 2.8. Desarrollo en profundidad de la nomenclatura y de las abreviaturas normalizadas.

Bajo ambos criterios, para indicar de forma más detallada la estructura, se especifica el número y la posición de los fosfatos. Por ejemplo:

adenosina-5'-monofosfato	ácido 5'-adenílico	5'-AMP
adenosina-3'-monofosfato	ácido 3'-adenílico	3'-AMP
adenosina-5'-difosfato	ácido 5'-diadenílico	5'-ADP
adenosina-5'-trifosfato	ácido 5'-triadenílico	5'-ATP

Cuando no se especifica el número, implícitamente se asume que el fosfato ocupa la posición 5'. En este sentido, los dos últimos ejemplos se nombran simplemente como ADP y ATP.

Estos sistemas de nomenclatura se aplican tanto a desoxirribonucleótidos como a ribonucleótidos. Lógicamente, para los primeros debe anteponerse el prefijo *desoxi-* o su forma abreviada, la letra *d*. Por ejemplo: desoxicitidina monofosfato, dCMP, desoxicitidilato, 5'-dCMP.

Una forma aún más abreviada de representar los nucleósidos-monofosfato es mediante la inicial de su base, anteponiendo la letra “p” si el fosfato se une en 5' o posponiendo la “p” si la unión es en 3'. Por ejemplo: pA, pG, pU, Ap, Gp, Up. Este criterio es de interés para la descripción de la hidrólisis química o enzimática de los ácidos nucleicos (págs. 31 y 210).

Las tablas siguientes resumen todos los nucleótidos sencillos (NMP, NDP y NTP) existentes en la célula, incluyendo sus distintas denominaciones.

Nomenclatura de nucleótidos sencillos purínicos

Bases:	Adenina		Guanina	
	Ribonucleótidos	Desoxirribonucleótidos	Ribonucleótidos	Desoxirribonucleótidos
monofosfato	2'-AMP 2'-adenilato adenosina-2'-fosfato 2'-adenosina-monofosfato	— (*)	2'-GMP 2'-guanilato guanosina-2'-fosfato 2'-guanosina-monofosfato	— (*)
	3'-AMP (Ap) 3'-adenilato adenosina-3'-fosfato 3'-adenosina-monofosfato	3'-dAMP desoxi-etc.	3'-GMP (Gp) 3'-guanilato guanosina-3'-fosfato 3'-guanosina-monofosfato	3'-dGMP desoxi-etc.
	5'-AMP (pA) 5'-adenilato adenosina-5'-fosfato 5'-adenosina-monofosfato	5'-dAMP desoxi-etc.	5'-GMP (pG) 5'-guanilato guanosina-5'-fosfato 5'-guanosina-monofosfato	5'-dGMP desoxi-etc.
difosfato	5'-ADP (ppA) adenosina-5'-difosfato 5'-adenosina-difosfato	5'-dADP desoxi-etc.	5'-GDP (ppG) guanosina-5'-difosfato 5'-guanosina-difosfato	5'-dGDP desoxi-etc.
trifosfato	5'-ATP (pppA) adenosina-5'-trifosfato 5'-adenosina-trifosfato	5'-dTATP desoxi-etc.	5'-GTP (pppG) guanosina-5'-trifosfato 5'-guanosina-trifosfato	5'-dTGTP desoxi-etc.

* La desoxirribosa no tiene -OH en C2.

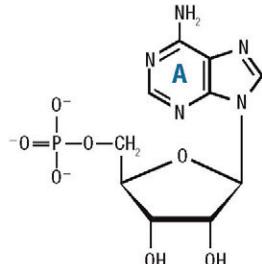
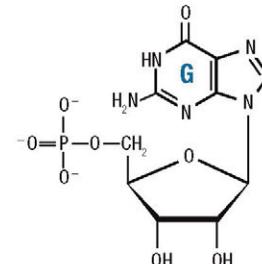
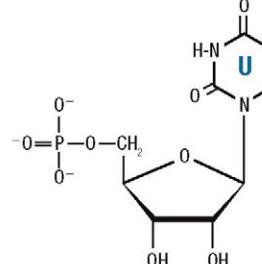
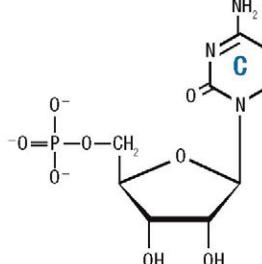
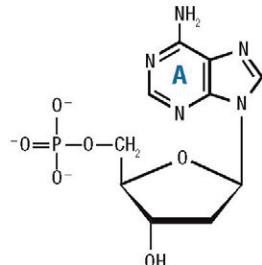
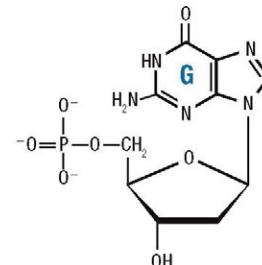
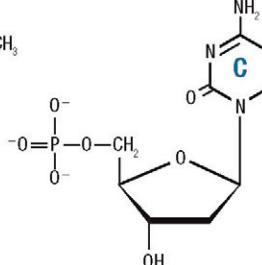
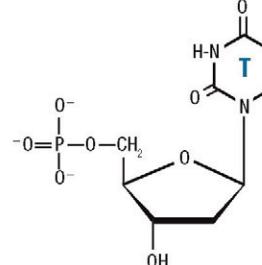
Nomenclatura de nucleótidos sencillos pirimidínicos

Bases:	Uracilo		Timina	
	Ribonucleótidos	Desoxirribonucleótidos	Ribonucleótidos	Desoxirribonucleótidos
monofosfato	2'-UMP 2'-uridilato uridina-2'-fosfato 2'-uridina-monofosfato	— (*)	2'-CMP 2'-citidilato citidina-2'-fosfato 2'-citidina-monofosfato	— (*)
	3'-UMP (Up) 3'-uridilato uridina-3'-fosfato 3'-uridina-monofosfato	3'-dTMP (Tp) 3'-desoxitimidilato desoxitimidina-3'-fosfato 3'-desoxitimidina-monofosfato	3'-CMP (Cp) 3'-citidilato citidina-3'-fosfato 3'-citidina-monofosfato	3'-dCMP desoxi-etc.
	5'-UMP (pU) 5'-uridilato uridina-5'-fosfato 5'-uridina-monofosfato	5'-dTTP (pT) 5'-desoxitimidilato desoxitimidina-5'-fosfato 5'-desoxitimidina-monofosfato	5'-CMP (pC) 5'-citidilato citidina-5'-fosfato 5'-citidina-monofosfato	5'-dCMP desoxi-etc.
difosfato	5'- UDP (ppU) uridina-5'-difosfato 5'-uridina-difosfato	5'- dTDP (ppT) desoxitimidina-5'-difosfato 5'-desoxitimidina-difosfato	5'- CDP (ppC) citidina-5'-difosfato 5'-citidina-difosfato	5'-dCDP desoxi-etc.
trifosfato	5'- UTP (pppU) uridina-5'-trifosfato 5'-uridina-trifosfato	5'- dTTP (pppT) desoxitimidina-5'-trifosfato 5'-desoxitimidina-trifosfato	5'- CTP (pppC) citidina-5'-trifosfato 5'-citidina-trifosfato	5'-dCTP desoxi-etc.

Aunque la nomenclatura que se muestra es la sistematizada, es habitual —y está reconocido por los organismos responsables de normalizar la nomenclatura— omitir el prefijo “desoxi-” y utilizar simplemente “timidina” y “timidilato” para los nucleósidos y nucleótidos de timina con desoxirribosa. Esto se debe a la presencia sistemática de T en el DNA y su ausencia en los RNA (salvo excepciones puntuales).

* La desoxirribosa no tiene -OH en C2.

2.5.2.2 Estructura de los nucleósidos-monofosfato que forman parte de los ácidos nucleicos

Nucleósidos-monofosfato componentes del RNA (= ribonucleótidos)			
Purínicos		Pirimidínicos	
			
adenilato, AMP; 5'-AMP; A, adenosina-5'- fosfato	guanilato, GMP; 5'-GMP; G, guanosina-5'-fosfato	uridilato, UMP; 5'-UMP; U, uridina-5'-fosfato	citidilato, CMP; 5'-UMP; C, cytidina-5'-fosfato
Nucleósidos-monofosfato componentes del RNA (= desoxirribonucleótidos)			
Purínicos		Pirimidínicos	
			
desoxiadénilato, dAMP; 5'-dAMP; dA, desoxiadenosina-5'- fosfato	desoxiguanylato, dGMP; 5'-dGMP; dG, desoxiguanosina-5'-fosfato	desoxitimidilato, dTTP; 5'-dTTP; dT, desoxitimidina-5'-fosfato	desoxicitidilato, dCMP; 5'-dCMP; dC, desoxicitidina-5'-fosfato

2.16

2.5.3 Propiedades fisicoquímicas de nucleósidos y nucleótidos

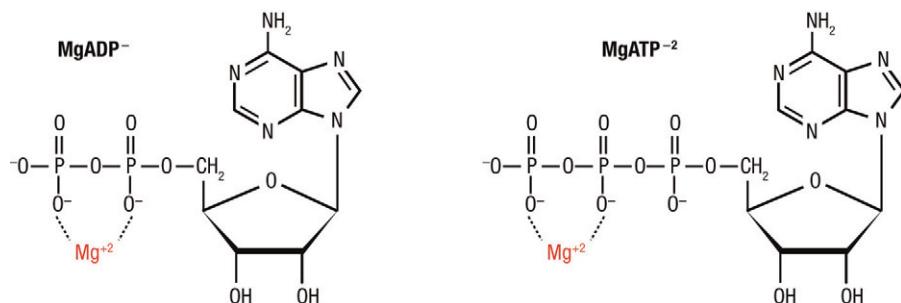
2.5.3.1 Propiedades iónicas

Como consecuencia de la disociación –en disolución acuosa– de los grupos P, PP y PPP, todos los nucleótidos (NMP, NDP y NTP) muestran un marcado carácter ácido, con 2, 3 y 4 valores de pK_a , respectivamente. En general, al ser sus valores de pK_a inferiores a 7, se encuentran cargados negativamente (existen como aniones) a pH fisiológico. Gracias a este carácter aniónico, los nucleótidos forman sales con metales, son cristalizables y son más solubles en agua que los nucleósidos o sus bases nitrogenadas.

2.5.3.2 Formación de complejos

Los grupos PP de los NDP y PPP de los NTP muestran gran afinidad (NTP 10 veces más que NDP) por cationes divalentes, como Mg^{2+} , Mn^{2+} y Ca^{2+} . Al estar éstos presentes en la célula en elevada concentración, los NDP

y NTP no existen como aniones libres, sino formando complejos quelato 1:1, principalmente con Mg^{2+} . Bajo estas “formas activas” ($MgADP^{-}$ y $MgATP^{-2}$) ambos nucleótidos participan en las reacciones enzimáticas de transferencia de fosfato.



2.17



Web 2.9. Estructura del complejo ATP:Mg.

2.5.3.3 Absorción de la luz en el ultravioleta

Los nucleósidos y los nucleótidos muestran el característico máximo de absorción de la luz en el UV, próximo a 260 nm, debido a la base nitrogenada correspondiente (págs. 16-17).



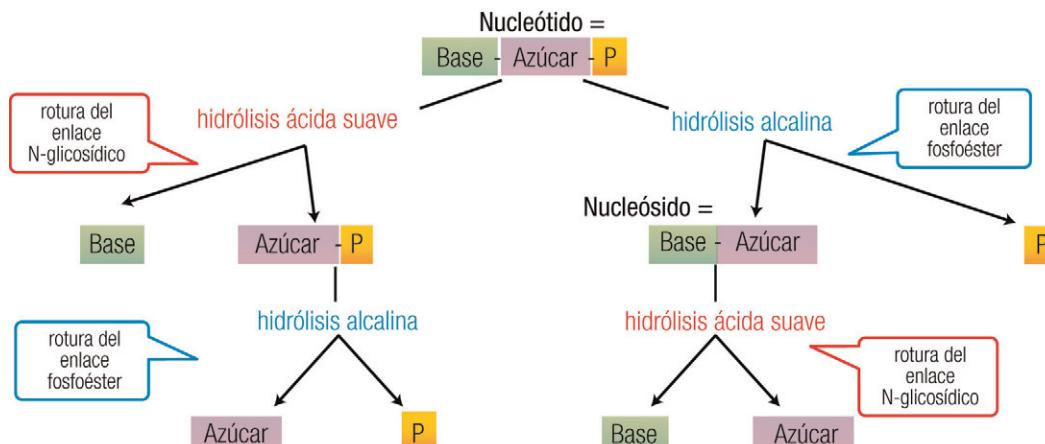
Web 2.10. Absorción de la luz por nucleósidos y nucleótidos.

2.5.3.4 Nucleótidos como moléculas de reserva energética

Muchos nucleótidos (fundamentalmente, los NDP y NTP) actúan en todo tipo de células y en multitud de procesos metabólicos como compuestos ricos en energía. Entre ellos, los nucleósidos-trifosfato intervienen como precursores de la síntesis de DNA y RNA, con liberación de pirofosfato.

2.5.3.5 Hidrólisis

Los nucleósidos y nucleótidos pueden ser hidrolizados en sus componentes, en medio ácido o alcalino. Esta hidrólisis química posee cierto interés en el estudio de los ácidos nucleicos. Para más detalles, véase pág. 31.



2.18

2.5.4 Nucleótidos cílicos y nucleótidos complejos

Con intención de un desarrollo completo del capítulo y debido a su relación estructural, cabe mencionar por último la existencia de otros nucleótidos que no forman parte de los ácidos nucleicos, pero que desempeñan funciones, a menudo esenciales, en los seres vivos.

En primer lugar, existen nucleótidos cílicos, que son nucleósidos-monofosfato en los que un mismo fosfato se une a dos posiciones del azúcar (dos -OH diferentes), generalmente 3' y 5'. Tienen un papel relevante en la transmisión de señales desde el entorno (por ejemplo, hormonas o primeros mensajeros) hacia el interior de la célula. Destacan en esta función de “segundos mensajeros” el AMP cíclico (cAMP) y el GMP cíclico (cGMP).

Bajo la expresión genérica “nucleótidos complejos” recogemos los que difieren estructuralmente de los nucleótidos sencillos en uno o varios de los siguientes aspectos:

- Varios grupos fosfato esterificados en posiciones múltiples del azúcar (bis- y tris-fosfatos, CoA, NADP⁺).
- Presencia de dos nucleósidos, unidos entre sí por un “puente” de grupos fosfato (FAD, NAD⁺, Ap_nA); éstos entran dentro de la categoría de dinucleótidos.
- Presencia de otras bases nitrogenadas (FMN, FAD, NAD⁺, NADP⁺).
- Presencia de azúcares distintos de la ribofuranosa (FMN, FAD).
- Grupos funcionales adicionales (CoA).



Web 2.11. Estructura de nucleótidos cílicos y nucleótidos complejos.



Estructura primaria de ácidos nucleicos

3.1 ASPECTOS GENERALES	27	3.5 PROPIEDADES FISICOQUÍMICAS DE LOS ÁCIDOS NUCLEICOS	30
3.1.1 Niveles estructurales	28	3.5.1 Propiedades en disolución	30
3.2 ESTRUCTURA PRIMARIA	28	3.5.2 Reactividad	31
3.3 DOS TIPOS DE ÁCIDOS NUCLEICOS SEGÚN SU COMPOSICIÓN	28	3.5.3 Hidrólisis química de ácidos nucleicos	31
3.4 FORMAS DE REPRESENTACIÓN LINEAL	29	3.5.3.1 Hidrólisis ácida	31
3.4.1 Fórmula completa	29	3.5.3.2 Hidrólisis alcalina	31
3.4.2 Representaciones esquemáticas	29	3.5.4 Absorción en el ultravioleta	32
3.4.3 Representaciones abreviadas	30		

3.1 ASPECTOS GENERALES

Tradicionalmente se considera que los ácidos nucleicos, como componentes del material genético, forman parte de los **nucleoproteidos**, el grupo más importante de heteroproteidos o “proteínas conjugadas”, entre los que también se incluyen glicoproteidos y lipoproteidos y, en ocasiones, también fosfoproteidos y cromo- o metaloproteidos. Efectivamente, los nucleoproteidos resultan de la interacción, mediante enlaces electrostáticos fuertes, de los residuos básicos (catiónicos) de ciertas proteínas (histonas, protaminas) y los grupos fosfato (aniónicos) del ácido nucleico. El estudio detenido de estas asociaciones nucleoproteicas se hará una vez conocida la estructura de los ácidos nucleicos a todos sus niveles.

Por su **composición química elemental** (C, H, O, N y P), derivada de la presencia exclusiva de nucleótidos en su estructura, los ácidos nucleicos se diferencian de las proteínas, esencialmente, en su mayor contenido de fósforo (10%) y en la ausencia total de azufre.

Existen dos tipos de ácidos nucleicos, denominados **ribonucleico** o RNA y **desoxirribonucleico** o DNA, que difieren en los nucleótidos componentes. Antes de describir su estructura, recordemos brevemente que ambos ácidos nucleicos se encuentran en todo tipo de células, tanto procariotas como eucariotas; como excepción, los virus contienen sólo DNA o sólo RNA. En todos los casos, sus diversas funciones en la transmisión y expresión de la información genética vienen determinadas por su localización subcelular y su estructura de orden superior (conformación tridimensional).

	Localización			Función
	Eucariotas	Procariotas	Virus	
DNA	en núcleo celular (varios cromosomas), en matriz mitocondrial y en estroma de cloroplastos (varias copias del cromosoma)	en la zona nucleoide del citosol (un cromosoma y varios plásmidos)	dentro de la cápsida (sólo en algunos tipos de virus)	Material genético: depositario y transmisor de la información genética, organizada en genes que codifican productos génicos (proteínas o RNA)
RNA	en núcleo celular (temporalmente), en citosol, en matriz mitocondrial y en estroma de cloroplastos	en citosol	dentro de la cápsida (en otros tipos de virus)	Interviene en la transmisión de la información desde el DNA hasta los productos génicos. Interviene en el control de la expresión génica. Constituye el genoma de algunos virus

3.1.1 Niveles estructurales

Desde un punto de vista estructural, para el estudio de los ácidos nucleicos se pueden considerar **varios niveles**, similares, en cierto modo, a los descritos tradicionalmente para proteínas:

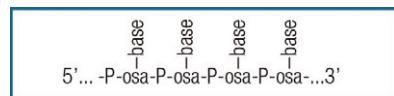
- **Estructura primaria**, descrita por la unión covalente de numerosos nucleósidos mediante enlaces fosfodiéster, dando lugar a un polímero lineal. El orden de los nucleósidos (o sus bases) en la cadena define la **secuencia** del ácido nucleico.
- **Estructura secundaria**, con la que se inicia el análisis espacial de la molécula, correspondiente a interacciones no covalentes que definen la conformación local, es decir, la disposición relativa de nucleótidos que están próximos en la secuencia. Para el DNA este nivel estructural viene definido por la asociación, mediante enlaces de hidrógeno, de dos cadenas polinucleotídicas a través de las bases nitrogenadas que sobresalen del esqueleto azúcar-fosfato. En el RNA sólo se presenta en determinadas regiones de la molécula.
- **Estructuras de orden superior**: Se engloban bajo este término genérico todas aquellas estructuras tridimensionales que surgen a partir de los niveles primario y secundario. En el caso del DNA pueden incluirse aquí las estructuras resultantes del superenrollamiento y de la asociación con proteínas básicas para formar la cromatina. Sin embargo, y a diferencia de lo que ocurre en proteínas, en general la estructura adoptada a este nivel no viene determinada por las de los niveles inferiores. En el caso de algunos RNA (especialmente los tRNA y rRNA) sí se observa un plegamiento tridimensional bien definido, equiparable a la estructura terciaria de las proteínas. Aun en este caso, no se observa una estructura cuaternaria, ya que dicho plegamiento sólo afecta a una molécula y no a varias asociadas entre sí, como sí ocurre en las proteínas.

3.2 ESTRUCTURA PRIMARIA

La estructura primaria, común para DNA y RNA, consiste en la existencia de cadenas, generalmente largas, de carácter macromolecular, formadas por la unión de nucleósidos mediante **enlaces covalentes 3'-5'-fosfodiéster**. Por tanto, la unidad monomérica es el nucleótido, o nucleósido-monofosfato (generalmente contemplado como 5'-NMP).

Los ácidos nucleicos son, por tanto, polímeros de nucleótidos o **polinucleótidos**, con un *esqueleto* lineal formado por unidades alternas de fosfato y pentosa (parte constante de la secuencia), del cual sobresalen lateralmente las bases (parte variable, que define la secuencia), unidas al azúcar respectivo por enlaces *N*-glicosídicos.

Por convenio, la secuencia siempre se indica en **dirección 5' → 3'**, es decir, con el extremo 5' terminal (5'-P) a la izquierda y el extremo 3' terminal (3'-OH) a la derecha.



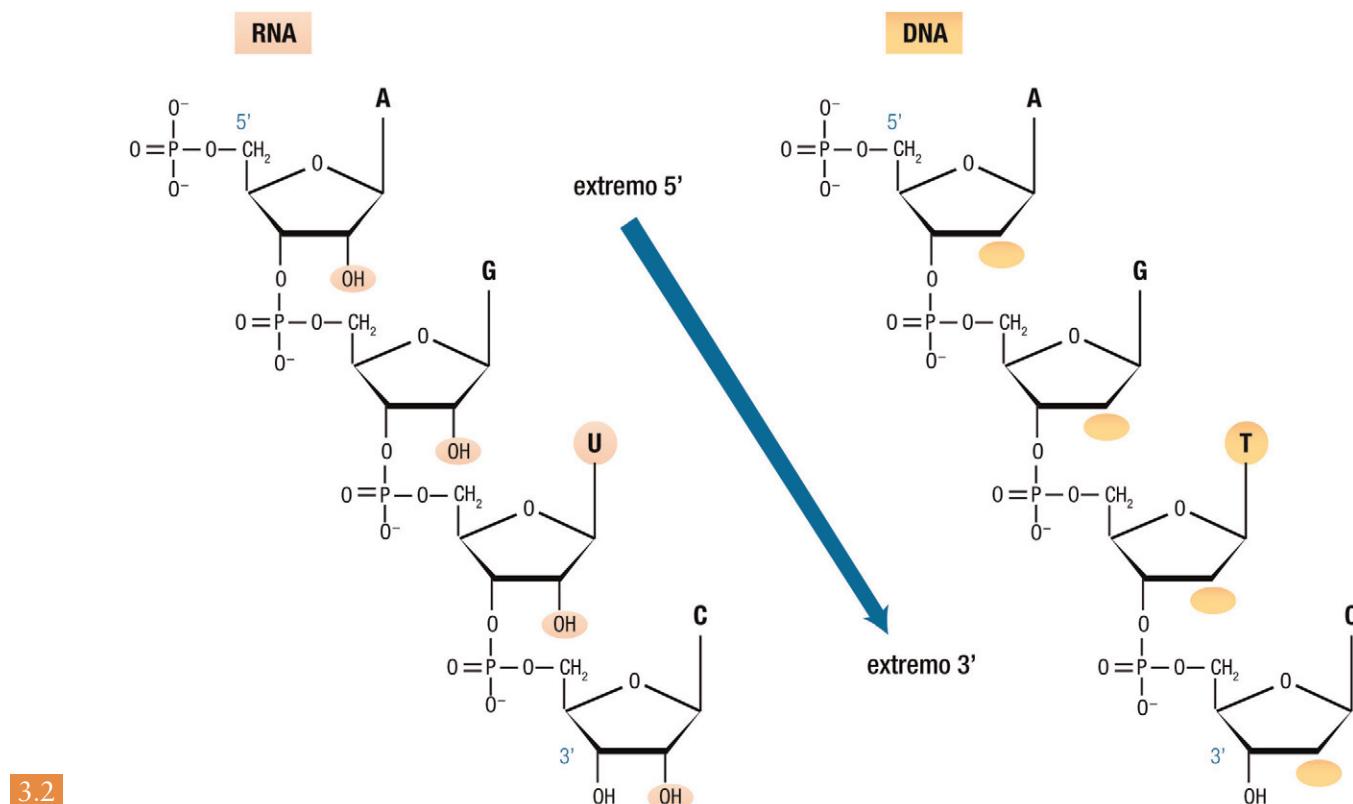
3.1

Dentro de este nivel estructural se estudiarán a continuación las diferencias en composición química de los dos tipos de ácidos nucleicos, las distintas formas de representación y sus propiedades fisicoquímicas más representativas.

3.3 DOS TIPOS DE ÁCIDOS NUCLEICOS SEGÚN SU COMPOSICIÓN

Tipo de ácido nucleico	Ácido ribonucleico (RNA)	Ácido desoxirribonucleico (DNA)
Monómeros	Ribonucleótidos	Desoxirribonucleótidos
Componentes del monómero	Fosfato	enlazando los monómeros (enlace fosfodiéster)
	Azúcar	Ribosa
	Purinas	Adenina, A (6-aminopurina) Guanina, G (2-amino-6-oxopurina)
	Pirimidinas	Citosina, C (2-oxo-4-aminopirimidina) Uracilo, U (2,4-dioxopirimidina)
		Timina, T (5-metiluracilo)

Según la notación de fórmula completa (exceptuando el detalle de las bases), la representación de ambos ácidos nucleicos es la siguiente:



3.4 FORMAS DE REPRESENTACIÓN LINEAL

Pueden emplearse hasta 6 niveles, de simplificación creciente. En general, se hace uso de la letra inicial de las bases, en mayúsculas, para representar a éstas, a sus nucleósidos o a sus nucleótidos. El ácido nucleico propuesto como ejemplo es un RNA (azúcar: ribosa y bases: A, G, U y C), pero las representaciones son igualmente válidas para un DNA (azúcar: desoxirribosa y bases: A, G, T y C).

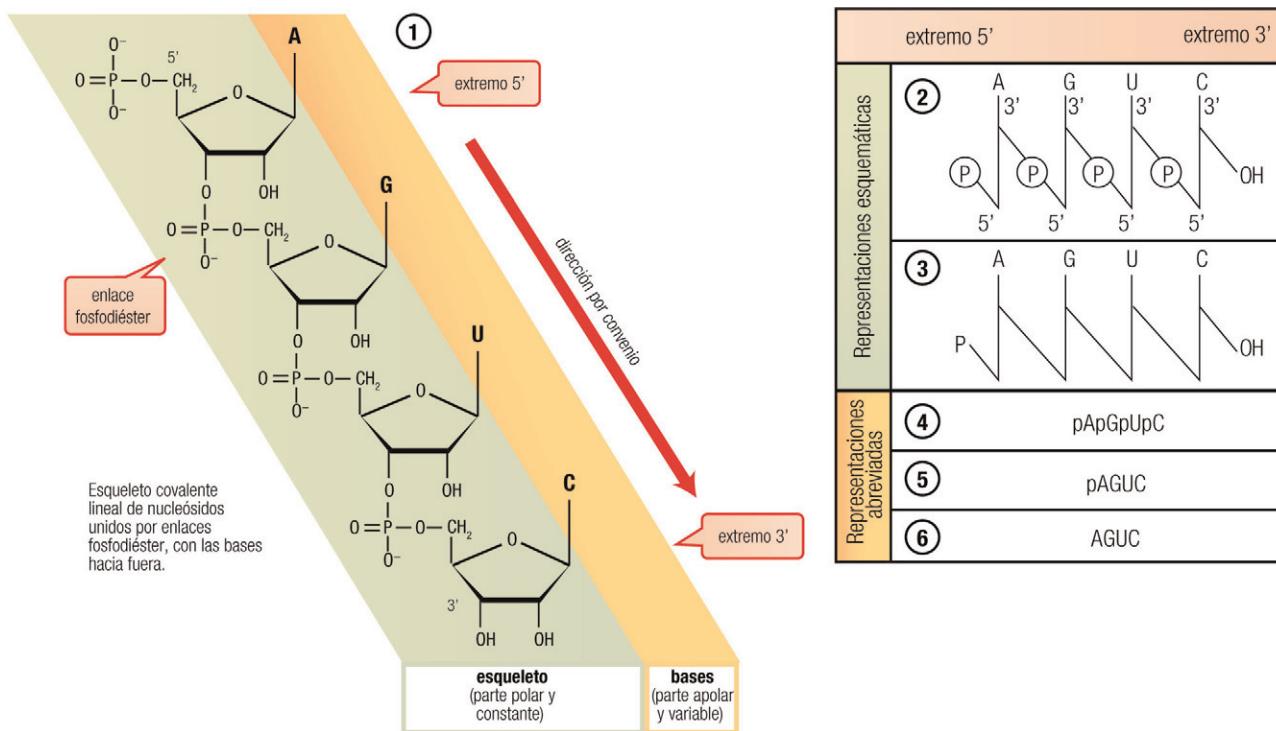
3.4.1 Fórmula completa

Por razones de espacio y fundamentalmente didácticas, es aconsejable comenzar con una representación vertical del esqueleto del ácido nucleico (notación 1 en la figura), que permite apreciar con detalle los enlaces fosfodiéster entre nucleósidos consecutivos (uniones azúcar-fosfato) y la situación transversal al esqueleto de los enlaces N-glicosídicos (uniones azúcar-base). Para mantener la orientación 5' → 3', se sitúa el extremo 5'-P en la parte superior y el extremo 3'-OH en la inferior.

3.4.2 Representaciones esquemáticas

Como peculiaridad de esta representación, el anillo furanósico del azúcar se simplifica mediante un solo trazo vertical, cuyo extremo superior corresponde al C-1' y el inferior al C-5' (notación 2 en la figura).

En una forma aún más simple se omiten los fosfatos intermedios y la numeración de los azúcares (notación 3).



3.4.3 Representaciones abreviadas

La simplificación se puede aumentar indicando cada nucleósido mediante la letra inicial de su base nitrogenada (lo único variable en la secuencia). El polinucleótido queda así reducido a una sucesión de letras mayúsculas intercaladas con letras *p* minúsculas (notación 4). El **extremo 5'-P** (a la izquierda) viene representado por una letra *p* seguida por la letra del primer nucleósido. El **extremo 3'-OH**, a la derecha, viene dado por la letra del último nucleósido. Los nucleósidos **intermedios** llevan la *p* a izquierda y derecha, correspondientes a las posiciones 5' y 3', respectivamente.

En la descripción cotidiana de secuencias de ácidos nucleicos se suprime las letras *p* intermedias, con lo que la representación se reduce a una secuencia de letras precedida por una letra *p* para el fosfato en el extremo 5' (notación 5), o se asume la presencia del fosfato en 5' y se omite incluso esa primera *p* (notación 6).

3.5 PROPIEDADES FISICOQUÍMICAS DE LOS ÁCIDOS NUCLEICOS

Los ácidos nucleicos presentan propiedades importantes relacionadas con sus funciones de almacenamiento (estructura), transmisión (replicación) y expresión (transcripción y traducción) de la información genética. Se comentan aquí brevemente algunas de ellas, relacionadas con la estructura, y el resto se irán estudiando a medida que surja su aplicación al propio estudio de estas moléculas (por ejemplo, la renaturalización de un DNA desnaturalizado se analizará como fundamento de la hibridación, capítulo 12).

3.5.1 Propiedades en disolución

Las cadenas polinucleotídicas de DNA y RNA son **hidrófilas**, debido a la posibilidad de formación de enlaces de hidrógeno con el agua por parte de los numerosos grupos fosfato y -OH del azúcar a lo largo del esqueleto.

Aunque las bases son esencialmente **hidrófobas**, como se discutirá pronto su contribución a la polaridad de la molécula se reduce debido a su ubicación aislada del medio acuoso.

Por otro lado, a pH fisiológico los grupos fosfato (con un pK_a próximo a 2) se ionizan casi por completo, haciendo que las moléculas de DNA y RNA se comporten como **ácidos** y tengan numerosas cargas negativas. Debido a este carácter **polianiónico**, los ácidos nucleicos se encuentran casi siempre neutralizados por interacción iónica con las cargas positivas de otras moléculas. Concretamente, el DNA del genoma nuclear eucariótico se encuentra unido a **histonas**, proteínas ricas en arginina y lisina, cargadas positivamente a pH fisiológico, dando lugar a nucleoproteídos; la importancia de esta asociación se considerará posteriormente (pág. 84). En los espermatozoides, el papel de las histonas lo desempeñan pequeñas proteínas ricas en arginina, denominadas **protaminas** y también con fuerte carga positiva. En otras ocasiones, el DNA se une a **poliaminas** (principalmente espermina y espermidina, así llamadas por haber sido detectadas originalmente en el semen), distribuidas de forma amplia, pero especialmente en algunos virus y bacterias en estado de rápida proliferación. Así, por ejemplo, en el fago T₄ las poliaminas neutralizan un 40% de la carga negativa total del DNA, estabilizando su conformación. Por último, como resultado experimental de esta propiedad, la cristalización de ácidos nucleicos se facilita si su carga negativa se compensa por la unión de iones metálicos.

Debido a la relativa rigidez de la molécula (en doble hélice) y a su inmensa longitud en relación al diámetro, las disoluciones de DNA son muy viscosas. Esta propiedad tiene interés en relación con el estudio del proceso de desnaturación (pág. 165).

3.5.2 Reactividad

En general, los ácidos nucleicos son químicamente muy estables, especialmente el DNA. Ello se debe a que la desoxirribosa carece de grupos reactivos: el 4'-OH está comprometido en el cierre del anillo, y los 3'-OH y 5'-OH están formando enlaces fosfodiéster. El RNA es algo más reactivo, gracias a los grupos 2'-OH, lo que se refleja en sus propiedades, por ejemplo, su hidrólisis en medio alcalino, como se verá a continuación.

3.5.3 Hidrólisis química de ácidos nucleicos

3.5.3.1 Hidrólisis ácida

Los **ácidos fuertes** escinden tanto los enlaces fosfodiéster entre nucleósidos como los enlaces *N*-glicosídicos de cada nucleótido. Por tanto, sirven para analizar la composición de bases de un ácido nucleico, pero no para determinar su secuencia.

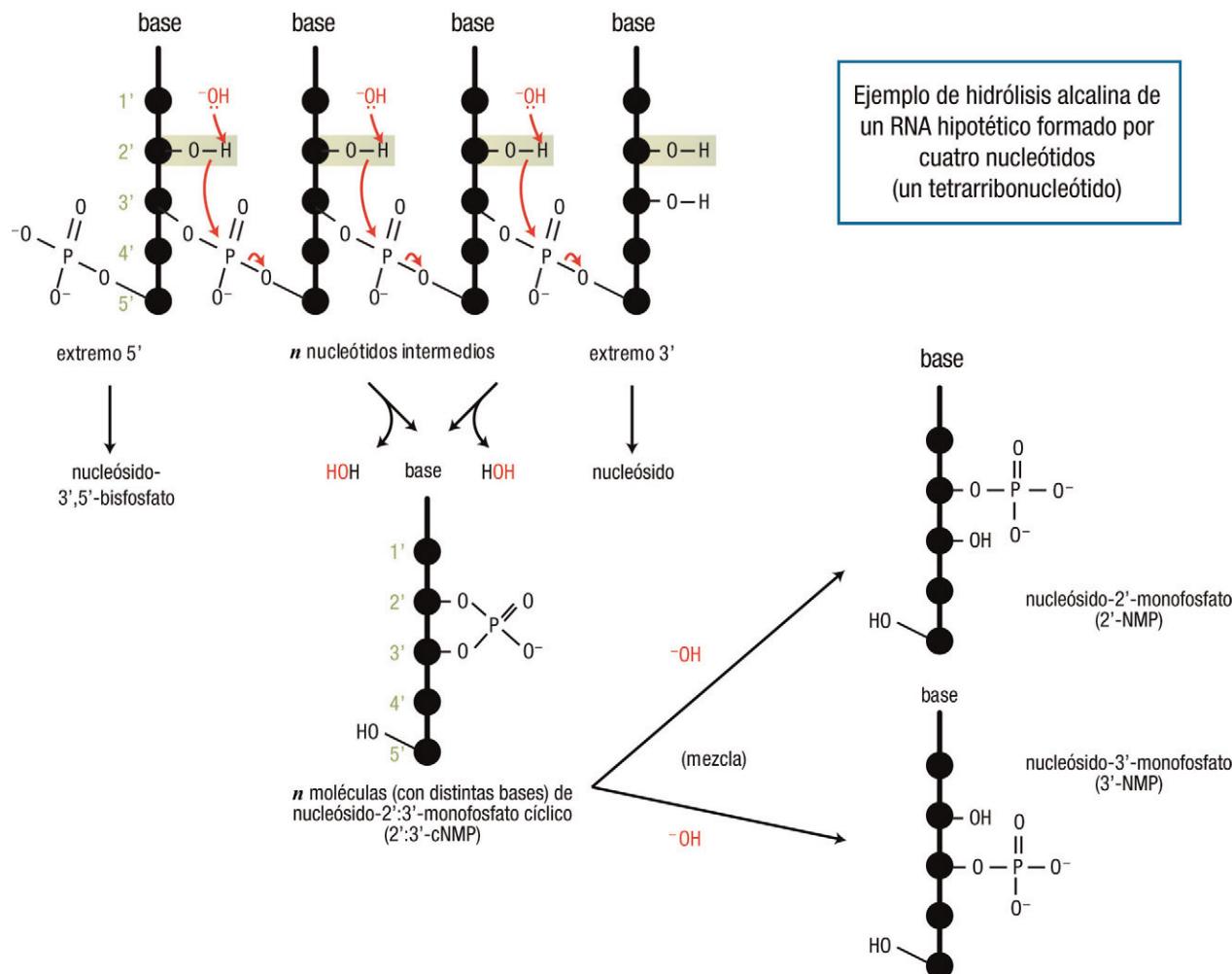
Los **ácidos débiles** respetan los enlaces fosfodiéster, pero tienden a romper la unión *N*-glicosídica. Este enlace es más lábil con purinas que con pirimidinas, por lo que se puede conseguir (por ejemplo, a pH = 3) escindir de manera selectiva las bases purínicas, dando lugar a los denominados ácidos *apurínicos*. Esta hidrólisis tuvo cierto interés históricamente para el análisis parcial de la secuencia del DNA.

¿Se hidrolizan?	DNA		RNA	
	Enlaces fosfodiéster (entre nucleósidos)	Enlaces <i>N</i> -glicosídicos (en cada nucleótido)	Enlaces fosfodiéster (entre nucleósidos)	Enlaces <i>N</i> -glicosídicos (en cada nucleótido)
Medio ácido fuerte (p. ej., HCl 1M)	Sí	Sí	Sí	Sí
Medio ácido débil	No	Sí (todos o purinas)	No	Sí (todos o purinas)
Medio alcalino	No	No	Sí	No

3.5.3.2 Hidrólisis alcalina

El medio básico afecta de distinta manera a los enlaces fosfodiéster, hidrolizando los del RNA, pero no los del DNA; los enlaces *N*-glicosídicos son resistentes.

El RNA se hidroliza rápidamente, liberando sus nucleótidos, en disoluciones diluidas de NaOH (pH = 11), sin afectar a los enlaces N-glicosídicos. La hidrólisis de los enlaces entre azúcar y fosfato se debe a la implicación directa de los grupos 2'-OH de la ribosa en la catálisis química. Se forman así NMP 2':3'-cíclicos como productos intermedios, que se hidrolizan rápidamente en el medio alcalino, al azar a uno u otro lado del enlace 2':3'-fosfodiéster, para dar una mezcla de los 2'-NMP y 3'-NMP.



3.4

El DNA, al contrario que el RNA, es resistente a la hidrólisis en medio alcalino. Ello se debe a la ausencia de grupos 2'-OH en las unidades de desoxirribosa, lo que impide el mecanismo expuesto anteriormente. Ésta es una de las pocas diferencias en las propiedades de ambos tipos de ácido nucleico y, por ello, es esencial en el diseño de los métodos de purificación o de análisis de DNA.

3.5.4 Absorción en el ultravioleta

La similitud en el valor máximo de absorbancia de las distintas bases (pág. 16) puede utilizarse no sólo para detectar y cuantificar las bases, los nucleósidos y los nucleótidos, sino también y especialmente para determinar la concentración de los ácidos nucleicos. De hecho, es el procedimiento que se utiliza comúnmente, a pesar de no ser el más exacto. Basta con medir el valor de A_{260} en una disolución diluida de DNA o RNA para poder calcular su concentración. Sin embargo, es conveniente la realización de algunos controles, teniendo en cuenta sobre todo la posible contaminación con proteínas, cuya presencia contribuye a la absorción y, por tanto, interfiere en esta determinación.

Efectivamente, aunque los máximos de absorción son característicos de cada uno de estos compuestos (260 nm para ácidos nucleicos y 280 nm para proteínas), ambos absorben en cierta medida a las dos longitudes de onda. Por ejemplo, la absorbancia de proteínas a 260 nm es, para igual concentración, unas 20 o 30 veces inferior a la de los ácidos nucleicos. En la práctica, teniendo en cuenta esta situación, se deben emplear ambos valores de absorbancia (A_{260} y A_{280}) para poder corregir la interferencia mutua y así determinar de forma correcta las concentraciones de ácido nucleico y de proteína.



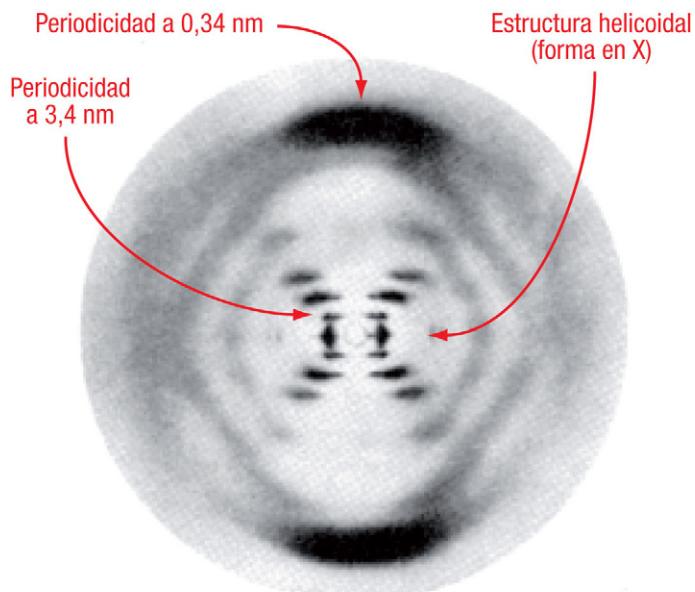
Web 3.1. Absorción de la luz por los ácidos nucleicos.



Estructura secundaria del B-DNA

4.1 ANTECEDENTES EXPERIMENTALES: ESTEREOQUÍMICA Y DIFRACCIÓN DE RAYOS X	35
4.2 PROPORCIÓN DE BASES NITROGENADAS: REGLAS DE CHARGAFF	36
4.2.1 Algunas relaciones entre bases son iguales en todas las especies	37
4.2.1.1 Relación entre bases	37
4.2.1.2 Relación entre purinas y pirimidinas	37
4.2.1.3 Relación entre aminobases y oxobases	37
4.2.2 Otras relaciones son diferentes según la especie	37
4.3 MODELO DE WATSON Y CRICK: FORMA B DEL DNA	38
4.3.1 Complementariedad de las bases nitrogenadas	38
4.3.1.1 Consecuencias de la complementariedad	40
4.3.1.2 Efecto del pH sobre el emparejamiento de bases	41
4.3.2 Geometría de los pares de bases	42
4.3.3 Antiparalelismo de las dos hebras	42
4.3.4 Diferencias en la secuencia de ambas hebras	42
4.3.5 Helicidad y carácter dextrorso	43
4.3.6 Parámetros cuantitativos de la doble hélice	43
4.3.7 Carácter anfipático y estabilidad de la doble hélice	44
4.3.7.1 Situación de los fosfatos y carácter hidrófilo	44
4.3.7.2 Coplanariedad, apilamiento de las bases y carácter hidrófobo	44
4.3.8 Superficie de la molécula: surcos en el DNA	45
4.3.9 Significado biológico: idoneidad para la replicación	45

4.1 ANTECEDENTES EXPERIMENTALES: ESTEREOQUÍMICA Y DIFRACCIÓN DE RAYOS X



4.1

El análisis por difracción de rayos X ha influido notablemente sobre el conocimiento de la estructura secundaria del DNA. Desde sus inicios se observaron, incluso en muestras de DNA sin purificar por completo, señales que correspondían a un espaciado regular de 0,34 nm a lo largo de las fibras de DNA. En la década de 1950 se obtuvieron datos más precisos (principalmente en los laboratorios de Maurice Wilkins y Rosalind Franklin),

con muestras de DNA mejor purificado y haciendo uso de los métodos de difracción de rayos X, que ya se habían desarrollado para el estudio estructural de proteínas. A partir de los correspondientes patrones de difracción, donde la distribución de las manchas en forma de aspa indica una estructura en hélice y las bandas intensas en las partes superior e inferior indican la repetición periódica de estructuras, se concluyó que las moléculas de DNA adoptaban una disposición helicoidal con una doble periodicidad a lo largo del eje de la hélice, cada 0,34 nm y cada 3,4 nm.

Por otra parte, la química de nucleótidos permitía conocer la existencia de interacciones por apilamiento, así como la asociación mediante puentes de hidrógeno (“emparejamiento” o “apareamiento”). Todos estos datos experimentales, junto con el análisis de la proporción de las bases constituyentes de la molécula, cuyas conclusiones se exponen a continuación, y la construcción de modelos moleculares, sirvieron de base para la propuesta por Watson y Crick de la estructura de doble hélice, cuya descripción es el objeto principal de este capítulo.

4.2 PROPORCIÓN DE BASES NITROGENADAS: REGLAS DE CHARGAFF

A pesar de que la composición de bases puede parecer un aspecto ligado a la estructura primaria, se estudia en este capítulo por estar directamente relacionada con la estructura secundaria del DNA. En un principio se creyó que el DNA estaba formado por cantidades iguales de los cuatro nucleótidos A, G, C y T, e incluso que la unidad fundamental era el tetranucleótido, hasta que Erwin Chargaff, entre otros, observó distintas proporciones de los 4 nucleótidos en diversos organismos. Concretamente, los DNA aislados a partir de células o tejidos diferentes de distintos individuos, dentro de una misma especie, presentan idéntica composición de bases, que además no cambia con la edad, estado nutricional o variaciones del entorno. La composición del DNA es, por tanto, característica de cada especie. A partir de esos datos se formularon unas **reglas o generalizaciones cuantitativas sobre la composición de bases del DNA**, que fueron la clave para deducir su estructura bicatenaria.

	Composición de bases (% moles)				Relación entre bases			
	Adenina	Timina	Guanina	Citosina	$\frac{A}{T}$	$\frac{G}{C}$	Purinas/ Pirimidinas $\frac{A+G}{T+C}$	Relación de asimetría $\frac{A+T}{G+C}$
Animales:								
Hombre	30,9	29,4	19,9	19,8	1,05	1,01	1,03	1,52
Oveja	29,3	28,3	21,4	21,0	1,04	1,02	1,03	1,36
Rata	28,6	28,4	21,4	20,4	1,01	1,05	1,02	1,36
Gallina	28,8	29,2	20,5	21,5	0,99	0,95	0,97	1,38
Tortuga	29,7	27,9	22,0	21,3	1,06	1,03	1,05	1,33
Salmón	29,7	29,1	20,8	20,4	1,02	1,02	1,02	1,43
Erizo de mar	32,8	32,1	17,7	17,3	1,02	1,02	1,02	1,85
Bacterias:								
<i>Escherichia coli</i>	24,7	23,6	26,0	25,6	1,05	1,02	1,03	0,94
<i>Staphylococcus aureus</i>	30,8	29,2	21,0	19,0	1,05	1,11	1,07	1,50
<i>Chlostridium perfringens</i>	36,9	36,3	14,0	12,8	1,02	1,09	1,04	2,73
<i>Brucella abortus</i>	21,0	21,1	29,0	28,9	1,00	1,00	1,00	0,73
<i>Sarcina lutea</i>	13,4	12,4	37,1	37,1	1,08	1,00	1,02	0,35
Virus:								
Fago ϕ 174 (monofilar)	24,3	32,7	24,1	18,5	0,74	1,30	0,95	1,34
Fago ϕ 174, forma replicativa (bifilar)	26,3	26,4	22,3	22,3	1,00	1,00	1,00	1,18

4.2.1 Algunas relaciones entre bases son iguales en todas las especies

A pesar de que la composición varía entre especies, en todas ellas se cumplen ciertas relaciones entre las bases, formuladas como las tres primeras reglas de Chargaff, que se exponen a continuación.

4.2.1.1 Relación entre bases

En prácticamente todas las especies, el contenido de A se aproxima al de T, y el contenido de G al de C:

$$\begin{array}{l} \text{A} \cong \text{T} \quad \longrightarrow \quad \text{Razón} \quad \text{A/T} \cong \text{T/A} \cong 1 \\ \text{G} \cong \text{C} \quad \longrightarrow \quad \text{Razón} \quad \text{G/C} \cong \text{C/G} \cong 1 \end{array}$$

4.2

Esta observación pudo explicarse una vez que Watson y Crick postularon la naturaleza bicanalicular del DNA, definida mediante la complementariedad de bases en los pares AT y GC. Hay que destacar que esta regla no se cumple para las pocas especies cuyo material genético está formado por RNA o DNA monocatenarios (compárense en la tabla los ejemplos del virus φ174, en sus formas monocatenaria y bicanalicular).

4.2.1.2 Relación entre purinas y pirimidinas

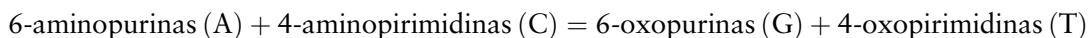
Como consecuencia de lo anterior, el contenido de bases purínicas (A + G) se aproxima al de bases pirimidínicas (T + C), es decir, siempre hay un 50% de purinas y un 50% de pirimidinas:

$$\begin{array}{l} \text{A+G} \cong \text{T+C} \cong 50\% \quad \longrightarrow \quad \text{Razón} \quad (\text{A+G}) / (\text{T+C}) \cong 1 \\ \text{purinas} \cong \text{pirimidinas} \cong 50\% \quad \longrightarrow \quad \text{Razón} \quad \text{purinas} / \text{pirimidinas} \cong 1 \end{array}$$

4.3

4.2.1.3 Relación entre aminobases y oxobases

Al observarse que A/T = C/G = 1, también se ha de cumplir (por razones estrictamente matemáticas) que (A + C) / (T + G) = 1. Por tanto, puede concluirse, como relación adicional (sin un significado especial), que:



4.2.2 Otras relaciones son diferentes según la especie

Otras relaciones entre bases no se mantienen constantes entre especies; de entre ellas, resulta útil la denominada **razón de asimetría**:

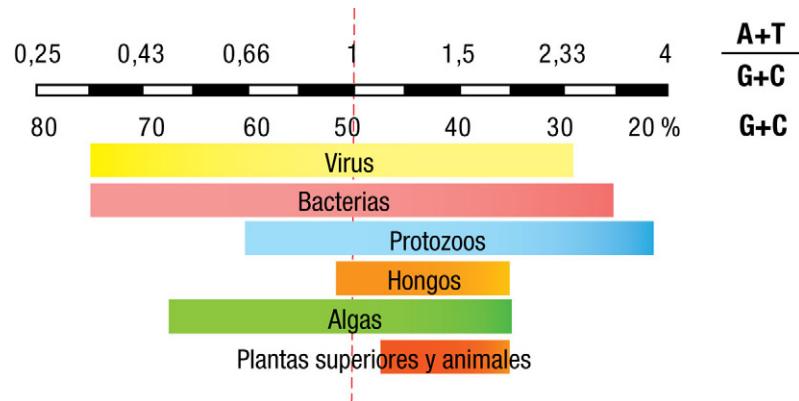
$$\frac{\text{A} + \text{T}}{\text{G} + \text{C}}$$

que, además de tener un valor diferente para cada especie, no se aproxima a la unidad como las razones anteriores (v. tabla). Por ejemplo, en animales la razón de asimetría es siempre superior a uno; lo mismo ocurre en plantas superiores y hongos. Por contraste, en bacterias y virus los valores son mucho más variados y superiores o inferiores a la unidad.

La razón de asimetría puede también calcularse directamente a partir del contenido en G+C, parámetro habitual en otros tipos de estudio, como la desnaturación del DNA (pág. 167 y web 12.1):

$$\text{razón de asimetría} = \frac{A + T}{G + C} = \frac{\%(A + T)}{\%(G + C)} = \frac{100 - \%(G + C)}{\%(G + C)}$$

De acuerdo con ello, los animales, plantas superiores y hongos muestran un exceso de A+T (de 50 a 65%) sobre G + C (de 50 a 35%), mientras que en bacterias y virus se aprecia la gran variabilidad en G+C (de 25 a 75% dependiendo de la especie).



4.4

En resumen, con estos estudios se ha observado que los DNA de especies evolutivamente relacionadas poseen una razón de asimetría similar, mientras que las especies muy alejadas tienden a tener una razón diferente. A pesar de ser ésta la menos conocida de las reglas de Chargaff, el hecho de que permita distinguir entre especies hace que pueda utilizarse como uno de los criterios en las relaciones evolutivas y en las clasificaciones taxonómicas de los organismos.

4.3 MODELO DE WATSON Y CRICK: FORMA B DEL DNA

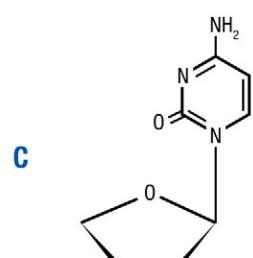
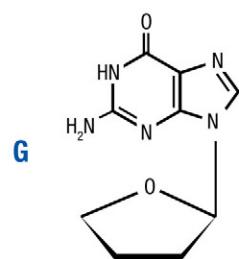
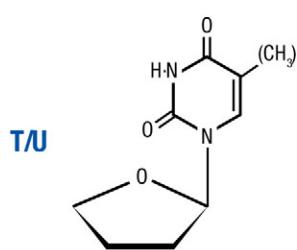
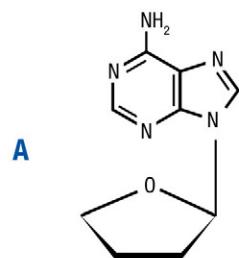
El **modelo tridimensional**, como estructura de referencia para las propiedades del DNA, fue postulado por James Watson y Francis Crick en 1953. Corresponde a la actualmente denominada **forma B del DNA**, la más común en las células y la más estable de todas las que puede adoptar un DNA bajo condiciones fisiológicas. Este modelo se basó tanto en las observaciones por rayos X como en las reglas de Chargaff, así como en otros estudios realizados mediante técnicas diversas, y permite justificar todos esos resultados experimentales. Simultáneamente se propuso la adecuación del modelo a la funcionalidad de la molécula de DNA para la conservación y transmisión de la información genética (pág. 45-46). A continuación se relacionan las características que conforman el modelo estructural de Watson y Crick.

4.3.1 Complementariedad de las bases nitrogenadas

La estructura química de las bases permite la formación de enlaces por puente de hidrógeno (dos átomos electronegativos, como N y O, comparten un átomo de hidrógeno cuando los 3 están en línea y la distancia es cercana a 3 Å). En concreto, dos nucleótidos pueden interaccionar de forma específica mediante dichos enlaces y se habla de “**bases complementarias**”, “**pares de bases**” o “**emparejamiento de bases**”.

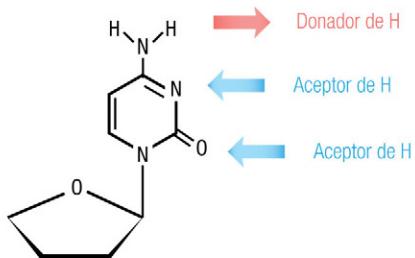
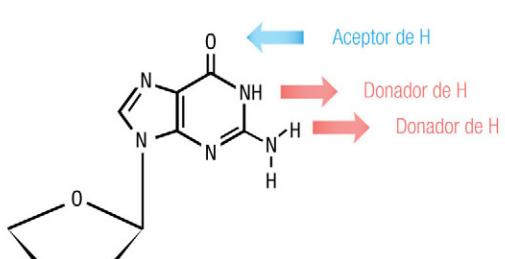
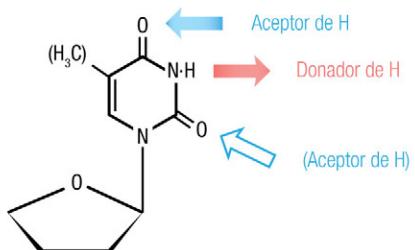
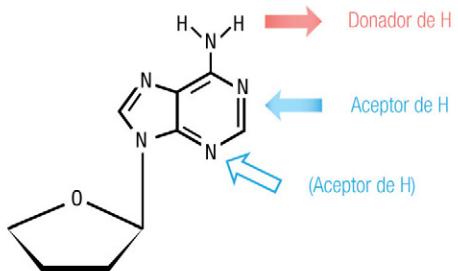
Para una mejor comprensión de la complementariedad, es conveniente establecer unos criterios previos en cuanto a la escritura de los nucleótidos en cada pareja de bases. Se trata con ello, simplemente, de facilitar la descripción de unas estructuras relativamente complejas:

1 Fórmulas de partida (las usadas anteriormente):



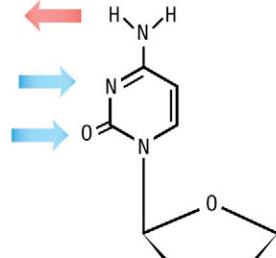
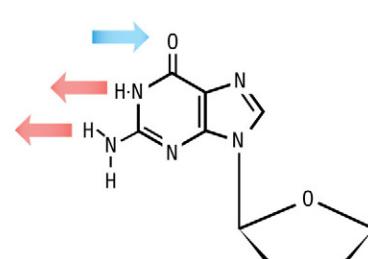
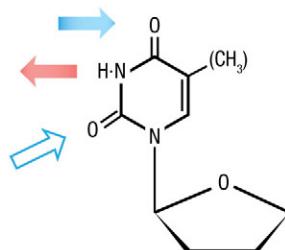
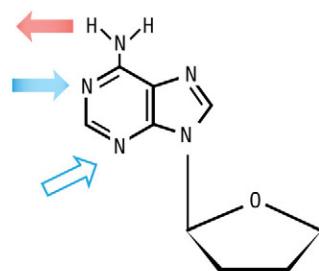
Esta disposición, con la base sobre el azúcar, se denomina conformación *sin*

2 Fórmulas con la base girada 180° alrededor del enlace N-glicosídico, con el fin de destacar el carácter donador o aceptor para la formación de puentes de hidrógeno:



Esta disposición, con la base alejada del azúcar, se denomina conformación *anti*

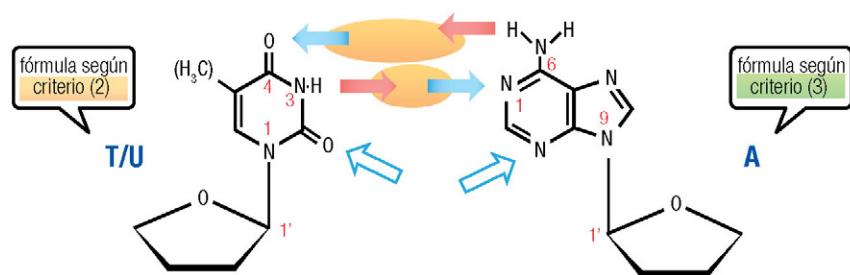
3 Fórmulas giradas 180° sobre sí mismas:



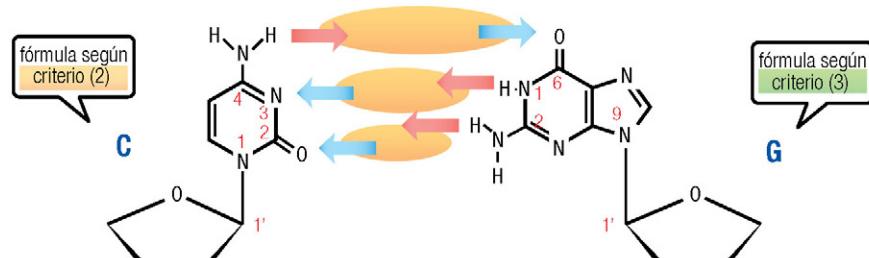
Sigue siendo conformación *anti*

- 4 Al representar un par de bases, se emplea la fórmula de tipo 2 para el nucleótido situado a la izquierda, y la de tipo 3 para el situado a la derecha (ambas conformación *anti*)

adenina y timina (o uracilo) pueden formar 2 enlaces por puente de hidrógeno:
 O del grupo C=O en C-4 de T con H del grupo NH₂ en C-6 de A
 H en N-3 de T con N-1 de A



guanina y citosina pueden formar 3 enlaces por puente de hidrógeno:
 H del grupo NH₂ en C-4 de C con O del grupo C=O en C-6 de G
 N-3 de C con H en N-1 de G
 O del grupo C=O en C-2 de C con H del grupo amino en C-2 de G



4.5b

Las interacciones óptimas se forman entre G y C (tres enlaces de hidrógeno) y entre A y T, o bien A y U (dos enlaces de hidrógeno). El par G≡C establece así una interacción más fuerte que el par A=T (en el DNA) o el par A=U (en el RNA). Obsérvese que siempre se forma un enlace de hidrógeno entre los dos átomos de N nucleares, mientras que los restantes se forman entre un grupo amino exocíclico y un grupo carbonilo.

Las interacciones óptimas mostradas corresponden a los tautómeros más abundantes (pág. 16); en los casos en que por alguna razón se estabilice otra forma tautomérica, los emparejamientos de bases serán diferentes.



Web 4.1. Diferencias en el emparejamiento de bases.

También son posibles otros emparejamientos entre bases, entre ellos los denominados “de tipo Hoogsteen”, que aparecen menos frecuentemente y no forman parte del B-DNA. Pueden encontrarse, por ejemplo, en el RNA (pág. 66), en los puentes entre 3 bases en DNA o RNA (págs. 55-56) y en la interacción codón-anticodón (pág. 315).

4.3.1.1 Consecuencias de la complementariedad

Gracias a la complementariedad descrita, dos cadenas polinucleotídicas se pueden asociar a todo lo largo siempre y cuando todas sus bases sean complementarias. Este emparejamiento completo de las dos cadenas da lugar a un ácido nucleico bicatenario, que es la estructura habitual del DNA. Esto quiere decir que siempre que en un DNA exista A, G, C o T en una cadena, se encontrarán respectivamente T, C, G y A en la otra.

Es importante destacar que la complementariedad entre las dos hebras del DNA es fundamental para su papel como portador de la información genética. Por un lado, porque se reduce la posibilidad de que las bases sufran modificaciones químicas accidentalmente. Por otro, porque la información está en cierto modo duplicada, aunque las secuencias de ambas hebras son diferentes entre sí (págs. 42-43); esa duplicidad facilita la corrección de mutaciones y de errores producidos durante la replicación del DNA.

Además, la existencia de la complementariedad explica las reglas de Chargaff (pág. 36), y justifica dos características estructurales importantes, la helicidad y el antiparalelismo, que se estudian más adelante.

Ilustración, mediante un ejemplo hipotético basado en la composición del DNA humano, de cómo la doble hebra del DNA permite explicar las reglas de Chargaff:

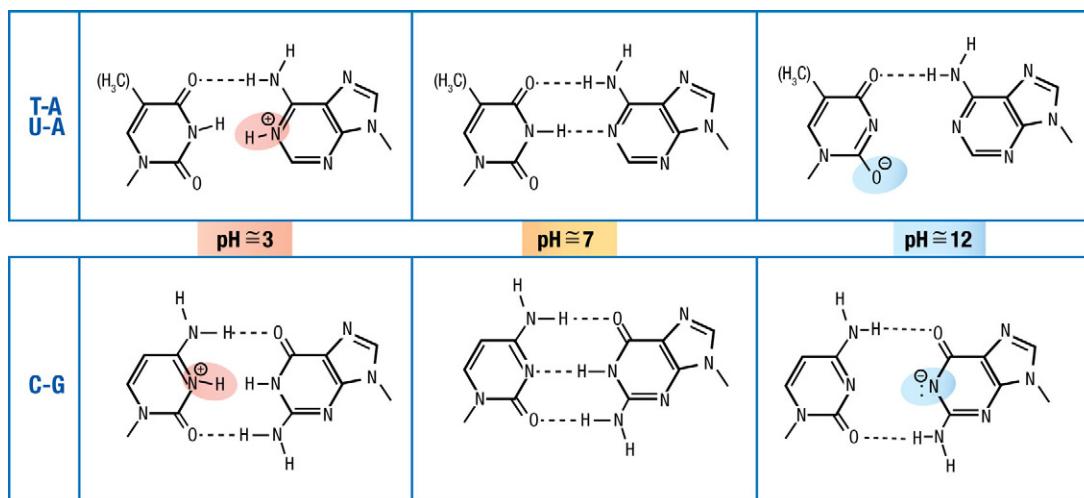
bases	Composición de bases			Regla de Chargaff correspondiente:
	% moles en hebra 1 y sus cocientes	% moles en hebra 2 y sus cocientes	% moles en la doble hebra y sus cocientes	
A	32	28	30	Relación entre bases individuales: A = T, C = G
T	28	32	30	
G	25	15	20	
C	15	25	20	
A + G = purinas C + T = pirimidinas purinas / pirimidinas	57 43 1,33	43 57 0,75	50 50 1,0	Relación entre purinas y pirimidinas: $A + G / C + T = 1$
A + C = aminobases G + T = oxobases amino / oxobases	47 53 1,08	53 47 0,92	50 50 1,0	

El emparejamiento implica que el contenido de una base en una hebra ha de ser igual al de su base complementaria en la otra hebra. Por ello, las reglas de Chargaff se cumplen cuando se considera la **composición total del DNA de doble hebra**, pero no para **cada hebra por separado**, ya que en este caso, A ≠ T y C ≠ G. Tampoco se cumplen las reglas en los ácidos nucleicos monocatenarios (p. ej., RNA y DNA vírico), donde A ≠ U, A ≠ T y C ≠ G

4.6

4.3.1.2 Efecto del pH sobre el emparejamiento de bases

Los cambios de pH, al afectar a la ionización de los grupos funcionales de las bases nitrogenadas y al equilibrio de tautómeros (pág. 16), modifican las posibilidades de formación de enlaces de hidrógeno y, por tanto, el emparejamiento de las bases.



4.7

Como consecuencia, a valores de pH inferiores o superiores al fisiológico la reducción en el número de enlaces de hidrógeno provoca la separación de las hebras componentes del DNA, fenómeno denominado desnaturalización (pág. 162). No debe confundirse este efecto con la hidrólisis ácida o alcalina, bajo condiciones más extremas, de las cadenas de ácido nucleico (pág. 31).

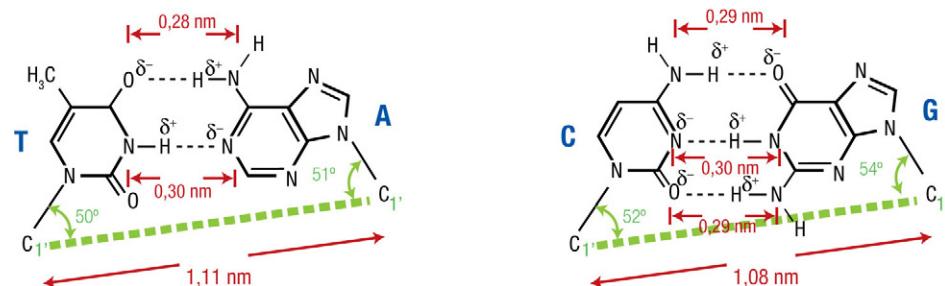
4.3.2 Geometría de los pares de bases

A pesar de la dificultad de visualizar la geometría de un par de bases en ausencia de un modelo cercano a la realidad (modelo tridimensional, sea material o informático), se describe a continuación de la forma más fiel posible.

La formación de un enlace de hidrógeno requiere que los tres átomos implicados (dos electronegativos y un H) se sitúen sobre una línea recta. Esta condición sólo se puede cumplir cuando los dos enlaces N-glicosídicos de los nucleósidos complementarios forman un ángulo, inferior a 90° , con la línea imaginaria que une los C-1' de ambas pentosas. Esta geometría del par de bases determina la disposición espacial relativa de las dos hebras en el DNA (helicidad) y es la causa molecular de la presencia de surcos menor y mayor en su estructura tridimensional; ambos aspectos se analizan en detalle más adelante (págs. 43 y 48).

Una vez establecidos los enlaces de hidrógeno múltiples, las distancias entre los átomos que los forman son prácticamente iguales para todos los casos. Asimismo, al emparejarse siempre una purina con una pirimidina, la separación entre las pentosas de ambas hebras también se mantiene constante (por ejemplo, la distancia entre los C1' o entre los P). Sólo de esta manera (distancia constante entre las hebras, independientemente de su secuencia) es posible la asociación continua de las dos cadenas para formar una molécula de DNA bicatenario.

representación bidimensional del par de bases



4.8



Web 4.2. Geometría de los pares de bases.

4.3.3 Antiparalelismo de las dos hebras

Ya se ha indicado que el emparejamiento completo entre las dos hebras exige que sus secuencias sean complementarias. Debe añadirse ahora que, estructuralmente, esto sólo se puede conseguir si ambas hebras se disponen en sentidos opuestos: si una avanza de 5' a 3', la otra debe hacerlo de 3' a 5'. Se dice que las hebras tienen polaridad opuesta, o que son **antiparalelas**. El concepto de antiparalelismo va, pues, estrechamente ligado al de complementariedad.



Web 4.3. Antiparalelismo de las hebras del DNA.

4.3.4 Diferencias en la secuencia de ambas hebras

Debe destacarse que las dos hebras de un DNA, complementarias y antiparalelas, poseen secuencias totalmente diferentes, como se puede ver en el último ejemplo mostrado. Es decir, son **hebras no idénticas**. Este hecho, aparentemente simple pero que a menudo pasa inadvertido, es de gran trascendencia; en concreto, en el proceso de transcripción, la molécula de RNA que se sintetiza es completamente diferente según cuál de las dos hebras del DNA se transcriba.

En ocasiones se utiliza el término *hebras equilibradas* refiriéndose a aquellas regiones del DNA en las que una de las hebras contiene igual proporción de purinas y pirimidinas. Como consecuencia, la hebra complementaria es también equilibrada:



Web 4.4. Relaciones de secuencia y composición entre ambas hebras.

4.3.5 Helicidad y carácter dextrorso

Las restricciones estéricas impuestas por la geometría de los pares de bases (que incluye los enlaces entre base, pentosa y fosfato, pág. 42) hacen que el emparejamiento completo de las dos hebras sólo se pueda establecer si los pares de bases sucesivos van girando unos con respecto a otros (concretamente, $34,6^\circ$ en el B-DNA). Como consecuencia de la acumulación de estos giros, las hebras de DNA adoptan una disposición helicoidal (es decir, cada hebra describe una hélice) y la molécula bicanalicular es una “doble hélice”.

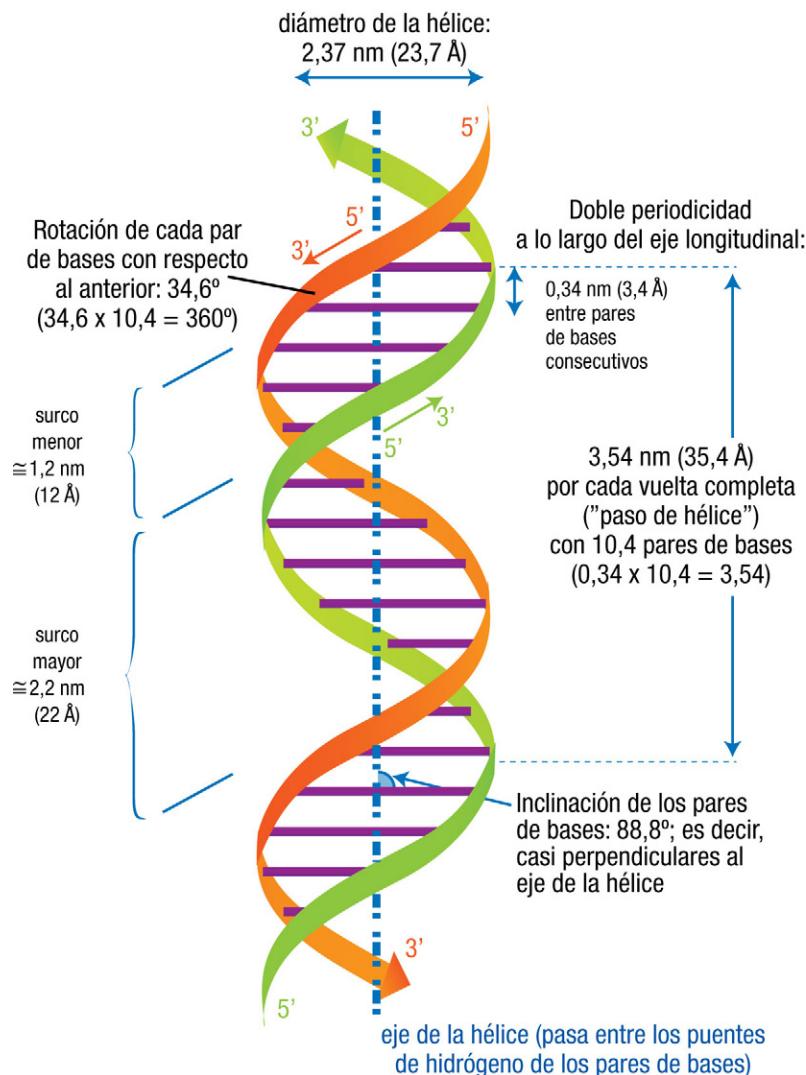
Geométricamente, una hélice (genérica, no de DNA), sea sencilla o doble, puede ser de dos tipos: dextrorso o sinistrorso, dependiendo de que gire hacia la derecha o hacia la izquierda.

En el caso de la forma B del DNA, ambas hebras se enrollan alrededor de un eje común, longitudinal, formando una **doble hélice dextrorso**. La observación de modelos moleculares tridimensionales permite una mejor comprensión de este carácter helicoidal.



Web 4.5. Helicidad y carácter dextrorso.

4.3.6 Parámetros cuantitativos de la doble hélice



Conocida la separación entre cada dos pares de bases sucesivos, es posible calcular el número total de pares de bases de una molécula de B-DNA de cualquier longitud, y viceversa.

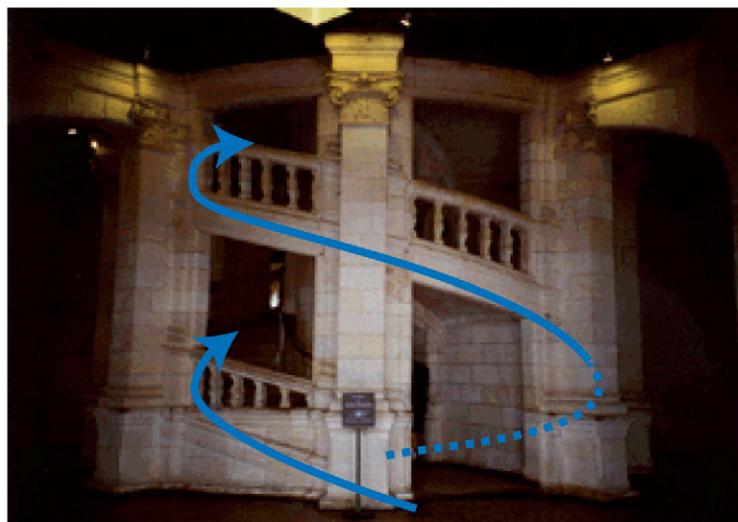


Web 4.6. Dimensiones de la molécula en doble hélice.

La escalera Francisco I del castillo de Chambord (valle del Loira, Francia), cuyo diseño se atribuye a Leonardo da Vinci, ha sido considerada como modelo de la doble hélice del DNA, con la salvedad de que su giro es a izquierdas, sinistrorso.

Se trata de una escalera helicoidal (de tipo caracol), con la característica atípica de poseer dos rampas independientes, cada una con su entrada (en la foto, una al frente y otra detrás). Por tanto, dos personas pueden utilizarla al mismo tiempo sin encontrarse.

En el caso del B-DNA, el esqueleto de las dos hebras correspondería a las barandillas externas de las dos rampas.



4.10

4.3.7 Carácter anfipático y estabilidad de la doble hélice

4.3.7.1 Situación de los fosfatos y carácter hidrófilo

Debido a la interacción entre las bases de las dos hebras, los esqueletos azúcar-fosfato de ambas deben disponerse hacia el exterior de la doble hélice, estableciendo así interacciones favorables con el medio acuoso circundante; por ello, el DNA bicatenario muestra un marcado carácter hidrófilo y es muy soluble en agua. Al mismo tiempo, la distribución relativa de los fosfatos es la que permite una menor repulsión electrostática entre ellos.



Web 4.7. Superficie de la molécula de DNA.

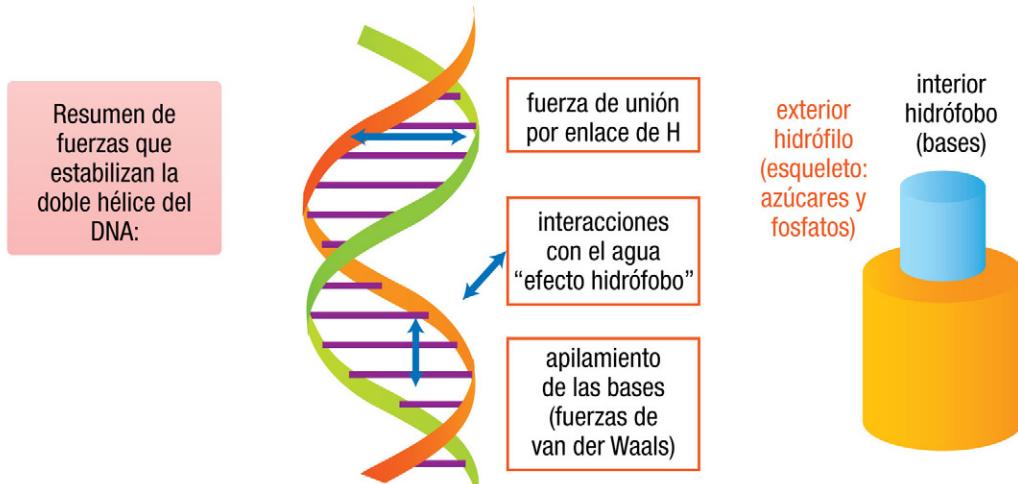
4.3.7.2 Coplanariedad, apilamiento de las bases y carácter hidrófobo



Web 4.8. Interior de la molécula de DNA.

Las bases purínicas y pirimidínicas de ambas cadenas se encuentran dirigidas hacia el interior de la doble hélice, superpuestas, formando un apilamiento similar a un montón de monedas. Como consecuencia de la proximidad y paralelismo de los anillos aromáticos, se originan entre ellos fuerzas de atracción, de tipo van der Waals, llamadas *interacciones de apilamiento*. Asimismo, esa disposición evita la interacción –entrópicamente desfavorable– de las bases con el medio acuoso; este *efecto hidrófobo* contribuye, junto con el apilamiento, a la estabilización de la molécula y a la protección de la información genética (constituida por la secuencia de bases).

Todas estas interacciones son mayoritariamente inespecíficas (es decir, son independientes del tipo de base implicada), a diferencia de los enlaces de hidrógeno entre bases complementarias, que son muy específicos. Aunque la atracción entre dos bases en un par no es muy fuerte de forma individual, la acumulación para todos los pares de bases de la molécula supone un aporte total muy elevado para la estabilidad de ésta.



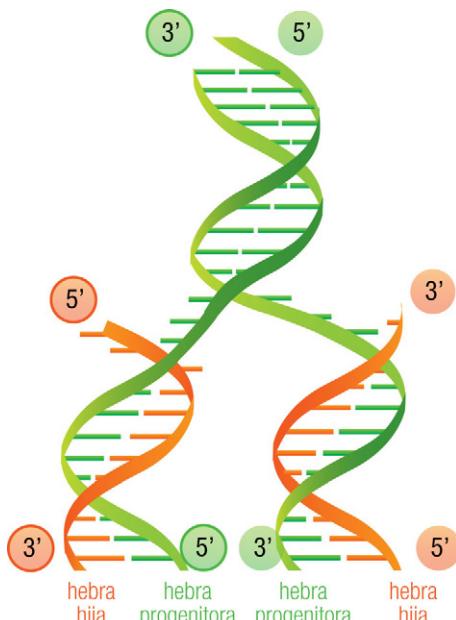
4.3.8 Superficie de la molécula: surcos en el DNA

Como se ha observado, la geometría de los pares de bases (pág. 42) y el antiparalelismo de las hebras (pág. 42), necesarios para la formación de los puentes de hidrógeno, hacen que los C-1' de los azúcares de los dos nucleótidos complementarios no se dispongan simétricamente con respecto a las bases y, por ello, que los esqueletos desoxirribosa-fosfato de las dos cadenas no se sitúen en puntos diametralmente opuestos de la doble hélice. Como consecuencia, a todo lo largo de la estructura en doble hélice se forman dos surcos de diferentes dimensiones, uno mayor (o hendidura profunda) y otro menor (o hendidura pequeña), que van girando de forma helicoidal, al igual que los dos esqueletos. Los surcos, especialmente el mayor, dejan espacio suficiente para que moléculas externas puedan contactar con las bases; éste es el fundamento de la interacción del DNA con numerosas proteínas capaces de reconocer secuencias de bases (págs. 56-59), con funciones tan importantes como la regulación de la expresión génica (por ejemplo, factores de transcripción). En cierto modo, es como si las proteínas "leyeran" la secuencia a través de los grupos funcionales de las bases que "asoman" dentro del surco mayor del DNA. Las dimensiones del surco mayor son adecuadas para que entre en ellas una hélice α de una proteína (pág. 57) o una tercera hebra de DNA (págs. 55-56).



Web 4.9. Los surcos en la molécula de DNA.

4.3.9 Significado biológico: idoneidad para la replicación



4.12

El modelo de doble hélice de Watson y Crick permite explicar que se mantenga una estructura regular a pesar de la variación de secuencia y que se produzca, mediante un mecanismo muy simple, la duplicación del material genético. En resumen, la replicación procede de una manera lógica, con dos procesos: *a)* la separación de las dos hebras, y *b)* la síntesis, a partir de ellas como molde, de otras dos cadenas complementarias (los detalles de este proceso se estudiarán posteriormente, en el capítulo 11). También permite la reparación de mutaciones y daños en una hebra, al disponerse de la otra como referencia de la secuencia correcta.



Variaciones en la estructura del DNA

5.1 VARIANTES BICATENARIAS: FORMAS A Y Z	47	5.3 MOTIVOS ESTRUCTURALES RESPONSABLES DE LA UNIÓN DEL DNA CON PROTEÍNAS	56
5.1.1 Forma A del DNA, en comparación con la forma B	48	5.3.1 Descripción de los motivos estructurales más frecuentes	57
5.1.2 Forma Z del DNA, en comparación con las formas B y A	49	5.3.1.1 Motivo hélice-giro-hélice o hélice-vuelta-hélice	57
5.2 VARIANTES LOCALES DE LA ESTRUCTURA SECUNDARIA DEL B-DNA	51	5.3.1.2 Motivo hélice-bucle-hélice	58
5.2.1 Curvatura de la doble hélice	51	5.3.1.3 Motivo homeodominio	58
5.2.2 Palíndromos	51	5.3.1.4 Motivo dedo de zinc	58
5.2.2.1 Concepto y características	51	5.3.1.5 Motivo cremallera de leucina	59
5.2.2.2 Palíndromos en moléculas monocatenarias	52		
5.2.2.3 Repeticiones de secuencia	53		
5.2.2.4 Consecuencias estructurales de los palíndromos	54		
a) En ácidos nucleicos monocatenarios	54		
b) En DNA bicatenario	54		
5.2.2.5 Funcionalidad de los palíndromos	55		
5.2.3 H-DNA: triple hélice	55		

La molécula de DNA no es una estructura estática, sino flexible y dinámica. Gracias a la capacidad de rotación alrededor de los enlaces de los nucleótidos, el DNA puede adoptar *in vivo* diversas formas, distintas de la forma B de Watson y Crick. Se estudian en este capítulo:

- Aquellas que, manteniendo la estructura helicoidal bicatenaria, presentan una conformación diferente: formas A y Z.
- Las variaciones conformacionales de carácter local en torno a la estructura B.
- La interacción de regiones específicas del DNA con estructuras características de algunas proteínas.

En general, estos tres tipos de variaciones estructurales se relacionan con funciones importantes en el metabolismo del DNA, e intervienen, por ejemplo, en el inicio y en la regulación de los procesos de replicación, recombinación, transcripción, etc. De ahí el interés de su descripción conjunta, como capítulo independiente, para ampliar la estructura secundaria y sentar las bases de dichas funciones, que serán objeto de estudio en capítulos posteriores.

5.1 VARIANTES BICATENARIAS: FORMAS A Y Z

El análisis por difracción de rayos X a resolución atómica reveló la existencia de otras dos formas estructurales o conformaciones, diferentes a la B-DNA de referencia, en las que se mantienen las características básicas de ésta (complementariedad, antiparalelismo, helicidad, etc.), pero con algunas diferencias:

Comparación entre las tres principales conformaciones del DNA bicanalario

	Tipo de hélice		
	A (deshidratada)	B (Watson-Crick)	Z (Rich-Dickerson)
Sentido de giro de la hélice	Dextrorso	Dextrorso	Sinistrorso
Forma y tamaño	La más ancha y corta	Intermedia	La más estrecha y larga
Diámetro de la hélice	2,55 nm (25,5 Å)	2,37 nm (23,7 Å)	1,84 nm (18,4 Å)
Distancia entre bases ¹ (<i>d</i>)	0,23 nm (2,3 Å)	0,34 nm (3,4 Å)	0,38 nm (3,8 Å)
Pares de bases por vuelta (<i>n</i>)	11	10,4	12
Rotación por cada base ² (360°/ <i>n</i>)	+ 32,7°	+34,6°	-30°
Longitud vuelta de hélice ³ (<i>n</i> × <i>d</i>)	2,53 nm (25,3 Å)	3,54 nm (35,4 Å)	4,56 nm (45,6 Å)
Inclinación del plano de los pares de bases ⁴	19° (gran inclinación)	1,2° (casi perpendicular al eje de la hélice)	9° (ligera inclinación)
Surco mayor o grande	Estrecho, profundo	Ancho, profundidad media	Plano, sin profundidad
Surco menor o pequeño	Amplio, no profundo	Estrecho, profundidad media	Estrecho, profundo
Enlace N-glicosídico	<i>anti</i>	<i>anti</i>	<i>sin</i> (G,A) y <i>anti</i> (C,T)

¹ Distancia entre pares de bases, medida a lo largo del eje de la hélice.

² Rotación por cada par de bases, o giro respecto al par anterior (+ dextrorso, - sinistrorso).

³ Longitud de la vuelta de hélice, medida a lo largo del eje; habitualmente se denomina “paso de hélice”.

⁴ Inclinación respecto al plano perpendicular al eje de la hélice.

5.1.1 Forma A del DNA, en comparación con la forma B

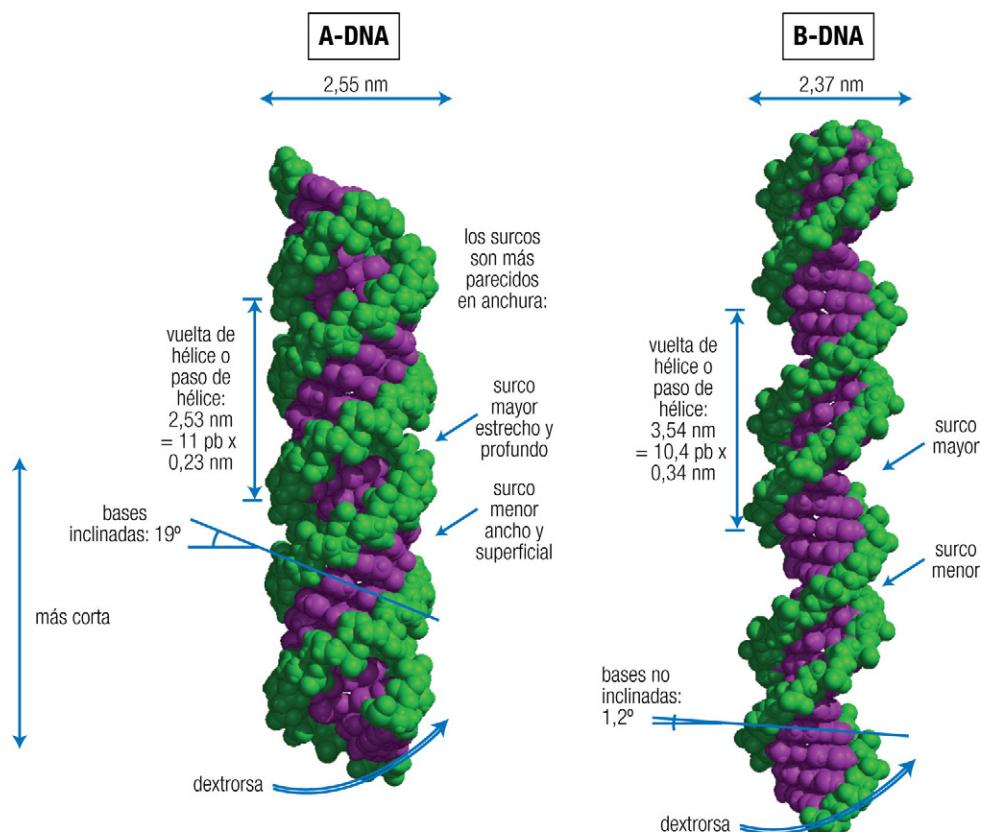
Una de las formas de estudiar la estructura del DNA es su preparación en forma de fibras para analizarlo mediante difracción de rayos X (pág. 35). Cuando las fibras se preparan en un ambiente de elevada humedad relativa se observa la estructura B, pero si la humedad es inferior al 75% el DNA adopta otra conformación, denominada A-DNA. Aunque esta forma no se ha encontrado para el DNA bajo condiciones fisiológicas, es un ejemplo de la flexibilidad de la molécula y su estudio es de interés por ser la que adopta el RNA cuando posee regiones bicanalarias, así como en los híbridos DNA:RNA.

Estructuralmente, el A-DNA es similar a la forma B: es igualmente una doble hélice dextrorsa, con dos hebras complementarias y antiparalelas unidas por enlaces de hidrógeno. Sin embargo, difiere en sus dimensiones: la hélice A es **más ancha** (mayor diámetro), **más corta** para un mismo número de nucleótidos (menor distancia entre pares de bases) y tiene **mayor número de bases por vuelta**. Como consecuencia, son menores la rotación por cada base y el paso de hélice. Además, las bases no se disponen casi perpendicularmente al eje de la hélice, sino claramente inclinadas, y se sitúan más alejadas del eje de la hélice. La superficie de la molécula es también diferente: el surco mayor es más estrecho que en la forma B y muy profundo, mientras que el surco menor es relativamente ancho y poco profundo. Mientras que la forma B puede acomodar en su surco menor una columna de moléculas de agua, el surco superficial de la forma A no ofrece espacio suficiente; esta diferencia explica, al menos en parte, por qué la deshidratación favorece la forma A.



Web 5.1. Las hélices A y B en el DNA.

Una visión complementaria a la anterior, a la vez que más cercana a la realidad, es la del **modelo espacial compacto**.
Las bases están coloreadas en rosa y el esqueleto de pentosas y fosfatos en verde



5.1

Las diferencias, recién comentadas, en la estructura de las dobles hélices A y B vienen originadas en última instancia por la diferente disposición relativa de los pares de bases, debida concretamente a una **distinta conformación de la desoxirribosa** en todos los nucleótidos: en el A-DNA, el C3' sobresale del plano formado por los otros cuatro carbonos de la ribofuranosa (conformación C3'-endo), mientras que en el B-DNA es el C2' el que no es coplanar (conformación C2'-endo).



Web 5.2. Conformaciones 2'-y 3'-endo en los nucleótidos.

Esta distinta conformación del anillo de ribosa hace que en el A-DNA los dos grupos fosfato consecutivos estén más próximos entre sí, lo que explica que los pares de bases sucesivos se acerquen, acortando la cadena. De manera similar, el cambio en la geometría conduce a los diferentes parámetros geométricos ya comentados.

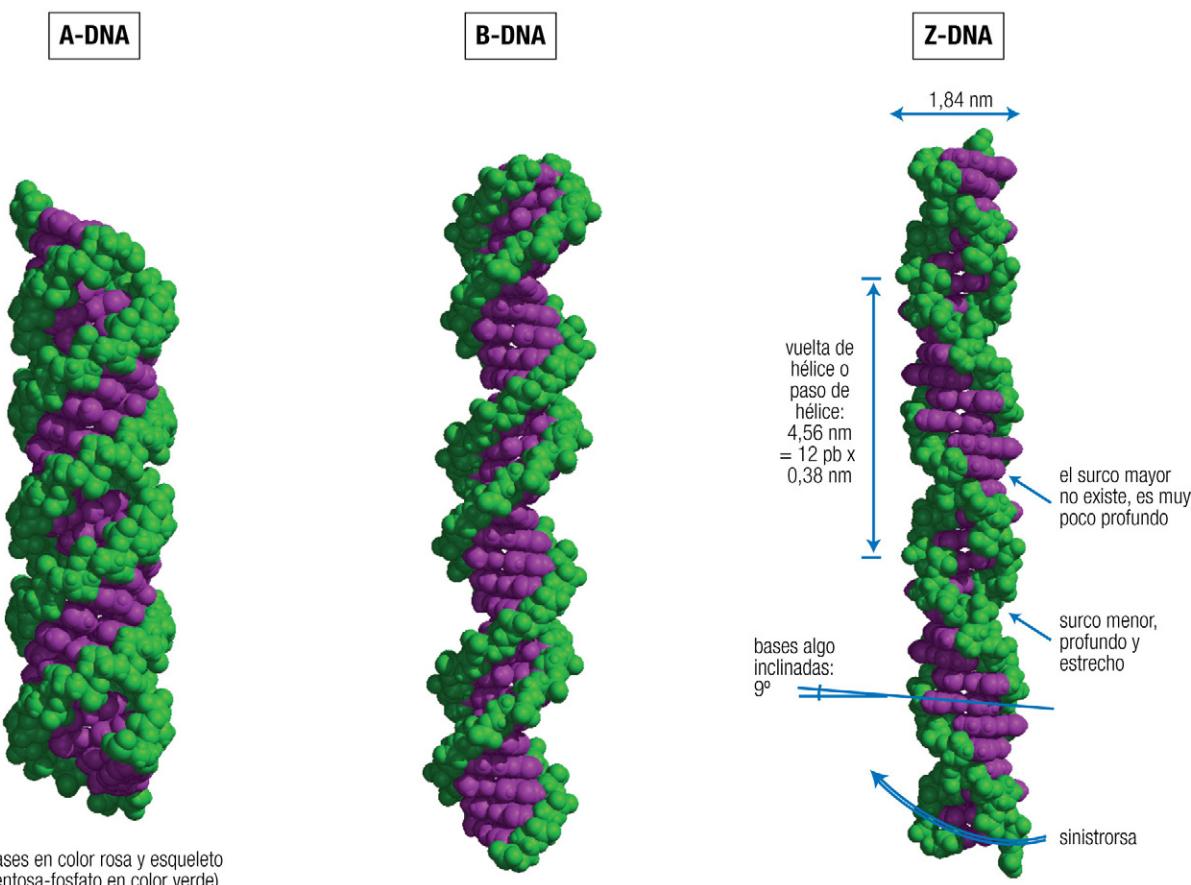
5.1.2 Forma Z del DNA, en comparación con las formas B y A

La existencia de la forma Z surge por primera vez de la observación, mediante difracción de rayos X, de una doble hélice del hexanucleótido CGCGCG, con **características distintas** a las formas B y A; entre ellas, como muy significativa, su giro a izquierdas o sinistrorso. Posteriormente, se ha encontrado esta estructura en otros oligonucleótidos sintéticos con secuencias de purinas y pirimidinas alternas. Aunque *in vivo* predomina el B-DNA, se ha podido observar la presencia de algunas regiones Z en el genoma de células eucarióticas, empleando anticuerpos contra esta última conformación. Su función, aún por confirmar, parece estar relacionada con el control de la expresión génica y con la recombinación, a través de sus efectos sobre el superenrollamiento.

En cuanto a sus dimensiones, observadas ahora comparativamente con las otras conformaciones, se observa que la hélice Z es la **más estrecha** (1,84 nm de diámetro), la **más larga** (mayor distancia entre bases: 0,38 nm) y

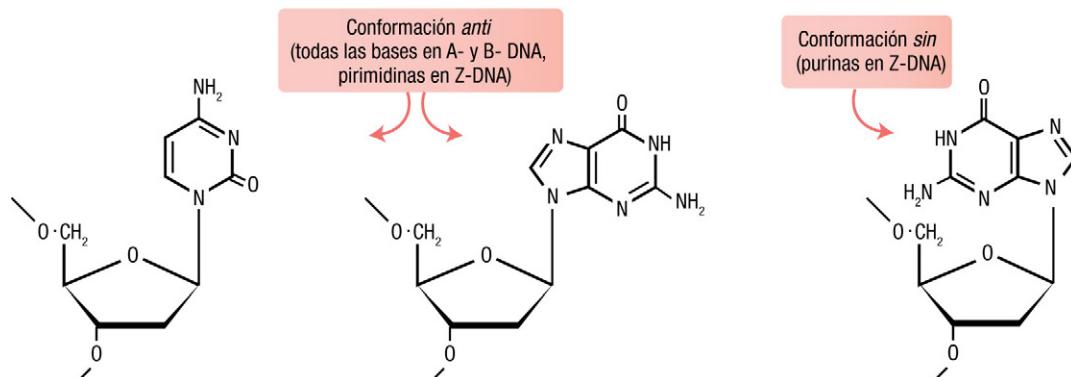
la que tiene **mayor número de pares de bases por vuelta** (12), lo cual supone una menor rotación por base (30°) y un mayor paso o longitud de una vuelta de hélice (4,56 nm). Es, en resumen, una hélice muy estirada. En su superficie forma **un solo surco, profundo** (correspondiente a la posición del surco menor en A- y B-DNA).

Modelos espaciales compactos:



5.2

La consideración de esta estructura a nivel molecular muestra que sus diferencias respecto a las formas A y B se deben también a una **distinta conformación**, en este caso alrededor del enlace N-glicosídico. En las formas A y B todas las bases nitrogenadas se orientan alejadas del anillo de pentosa, en la llamada conformación *anti*, mientras que en el Z-DNA sólo lo hacen las pirimidinas; las purinas tienden a situarse sobre la pentosa (conformación *sin*). Por ello, en las secuencias de purinas y pirimidinas alternantes (por ejemplo, CGCGCG), habituales del Z-DNA, se favorece la sucesión de ambas conformaciones, *sin* (G) y *anti* (C), lo que obliga al esqueleto azúcar-fosfato a plegarse en zig-zag para poder establecer los enlaces de hidrógeno.





Web 5.3. Conformaciones anti- y sin- en los nucleótidos.

Web 5.4. Disposición en zig-zag en el Z-DNA.

5.2 VARIANTES LOCALES DE LA ESTRUCTURA SECUNDARIA DEL B-DNA

Los estudios de difracción de rayos X también revelaron la existencia de variaciones en la conformación en determinadas secuencias del B-DNA. Aunque dentro de este apartado también podría incluirse el superenrollamiento del DNA, se ha preferido describir este aspecto como una estructura de orden superior de los ácidos nucleicos, dada su implicación en la organización del cromosoma.

5.2.1 Curvatura de la doble hélice

A pesar de la regularidad del modelo de Watson y Crick, existen pequeñas diferencias locales en la estructura de los DNA dependiendo de los nucleótidos integrantes. Por ejemplo, en determinadas secuencias los pares de bases no son exactamente coplanares; por otro lado, los pares AT son un poco más abiertos que los pares GC; todo ello puede ocasionar una desviación del eje de la doble hélice. Así, se ha comprobado que la repetición de 3 a 5 residuos de adenosina cada 10 pb (es decir, cada vuelta de hélice) provoca flexiones en la molécula de DNA bícatenaria. En otros casos, frecuentes, la curvatura no se origina por una secuencia particular de nucleótidos, sino que es forzada por la unión de una proteína a la cadena de DNA (como, por ejemplo, la mostrada en la web 5.9).

5.2.2 Palíndromos

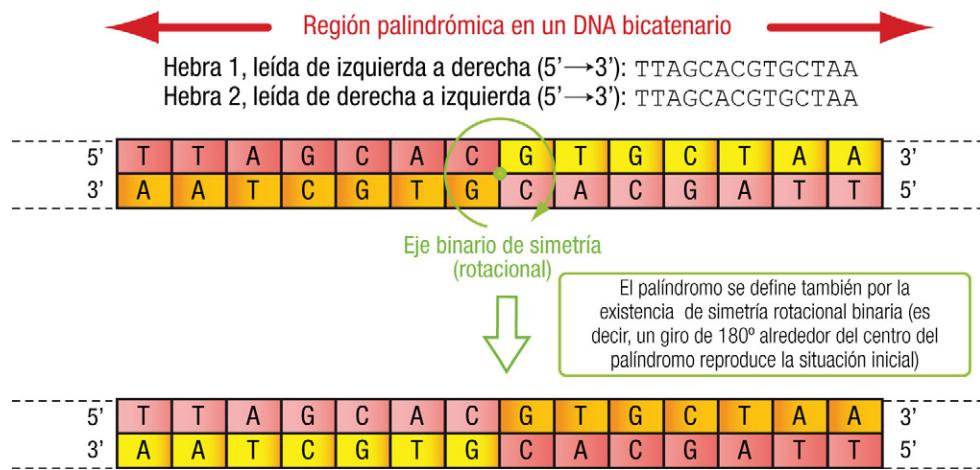
5.2.2.1 Concepto y características

Se denominan secuencias palindrómicas, o simplemente palíndromos, aquellas regiones de un DNA cuya secuencia es la misma en ambas hebras. Son bastante comunes, y tienen un significado especial por varias razones, principalmente: *a)* ser dianas de las enzimas de restricción (herramientas experimentales para escindir ácidos nucleicos en la tecnología del DNA recombinante); *b)* intervenir en la regulación de la expresión génica, pues son puntos de unión de los factores de transcripción, y *c)* generar estructuras secundarias peculiares en DNA y RNA (horquillas y cruciformes, pág. 54).

Gramaticalmente, un **palíndromo** (del griego *palin* = de nuevo, *dromos* = carrera) es toda palabra o frase que se deletrea igual de izquierda a derecha y de derecha a izquierda. Por ejemplo: RADAR, ANILINA, DÁBALE ARROZ A LA ZORRA EL ABAD. En los palíndromos hay un centro de simetría (a la vez eje y plano de simetría), que es la letra central o la posición entre las dos letras centrales.

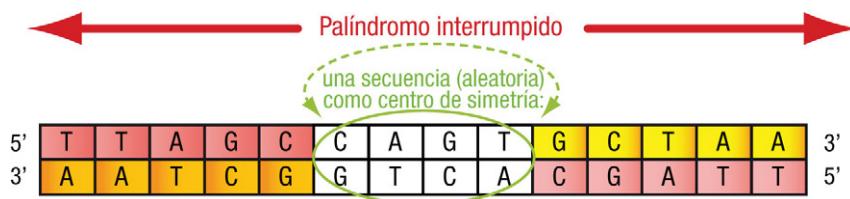
Sin embargo, la extensión de este término al ámbito de la biología molecular no es directa: se dice que una región de DNA es palindrómica cuando la secuencia de **una hebra**, leída de izquierda a derecha, es igual que la de la **otra hebra**, leída de derecha a izquierda. Por tanto, los palíndromos se observan en las dos cadenas conjuntamente, y no en una sola. Por ejemplo, la secuencia TTAGCAC-CACGATT sería un palíndromo gramatical, pero no es una secuencia palindrómica. Teniendo en cuenta que, según el convenio de escritura de los DNA, la secuencia de ambas hebras debe leerse de 5' a 3', se concluye que **las dos hebras de un palíndromo tienen la misma secuencia**; ésta es una de las definiciones de palíndromo.

5.4



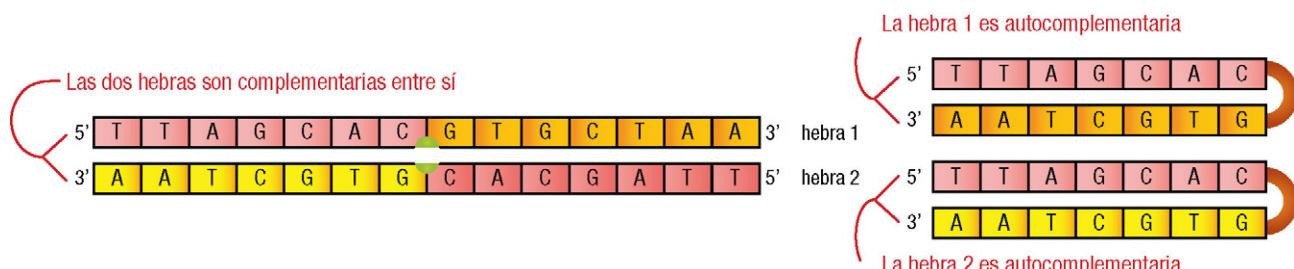
5.5

Como una extensión del concepto de palíndromo, se llaman **palíndromos interrumpidos** aquellos en los cuales una pequeña región del ácido nucleico separa las dos mitades del palíndromo y actúa, por tanto, como centro de simetría no puntual (en lugar del espacio entre dos nucleótidos de un palíndromo estricto). En cualquier caso, todo palíndromo está formado por dos “mitades” de secuencia idéntica, separadas por un centro de simetría, puntual o no puntual.



5.6

Al cumplirse tanto la simetría propia del palíndromo como la complementariedad de las dos hebras resulta que las secuencias de una misma hebra que quedan a ambos lados del centro de simetría son complementarias entre sí; se afirma, por ello, que las secuencias palindrómicas son **autocomplementarias** (otra forma de definir un palíndromo, aplicable no sólo a moléculas bicatenarias sino también a las monocatenarias, como se verá a continuación).



Web 5.5. Simetría y autocomplementariedad en los palíndromos.

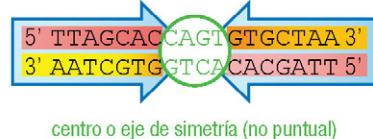
5.2.2.2 Palíndromos en moléculas monocatenarias

Para el caso de ácidos nucleicos monocatenarios, las dos primeras definiciones de palíndromo no se pueden aplicar de forma estricta. Sin embargo, dada la relación de complementariedad entre la secuencia de un RNA y la de una de las hebras del DNA que lo codifica, las moléculas de RNA monocatenario codificado por un DNA palindrómico son también autocomplementarias y se consideran como palíndromos. Bajo este criterio, cada una de las dos hebras de un DNA palindrómico también es una secuencia palindrómica monocatenaria.

5.2.2.3 Repeticiones de secuencia

palíndromos = repeticiones invertidas

(ejemplo: palíndromo discontinuo, con un centro de simetría de 4 pb)



centro o eje de simetría (no puntual)

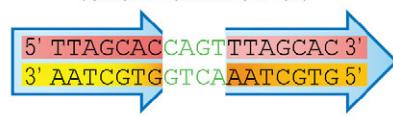
Las secuencias son iguales si se leen en sentidos opuestos:

Secuencia 1^a repetición(→): Secuencia 2^a repetición(←):

$$\begin{matrix} 5' & \text{TTAGCAC} & 3' \\ & \text{AATCGTC} & \end{matrix} = \begin{matrix} 5' & \text{TTAGCAC} & 3' \\ & \text{AATCGTC} & \end{matrix}$$

repeticiones directas:

(ejemplo: separadas por 4pb)



no hay eje ni plano de simetría

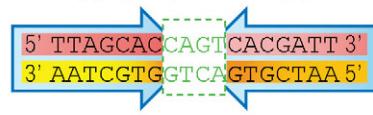
Las secuencias son iguales si se leen en el mismo sentido:

Secuencia 1^a repetición(→): Secuencia 2^a repetición(→):

$$\begin{matrix} 5' & \text{TTAGCAC} & 3' \\ & \text{AATCGTC} & \end{matrix} = \begin{matrix} 5' & \text{TTAGCAC} & 3' \\ & \text{AATCGTC} & \end{matrix}$$

repeticiones especulares:

(ejemplo: separadas por 4pb)



plano de simetría (no puntual)

Las secuencias son diferentes:

Secuencia 1^a repetición(→): Secuencia 2^a repetición(←):

$$\begin{matrix} 5' & \text{TTAGCAC} & 3' \\ & \text{AATCGTC} & \end{matrix} \neq \begin{matrix} 5' & \text{AATCGTC} & 3' \\ & \text{TTAGCAC} & \end{matrix}$$

Sólo son iguales si la segunda se lee en sentido no convencional, de 3' a 5':

$$\begin{matrix} 5' & \text{TTAGCAC} & 3' \\ & \text{AATCGTC} & \end{matrix} = \begin{matrix} 3' & \text{TTAGCAC} & 5' \\ & \text{AATCGTC} & \end{matrix}$$

5.7

En ocasiones se utiliza como sinónimo de palíndromo el término *repeticiones invertidas*, puesto que si se leen en sentidos opuestos, las dos mitades de la molécula bicatenaria tienen la misma secuencia (a ambos lados del centro de simetría, puntual o no).

Otra expresión relacionada es la de *repeticiones directas*, que se emplea cuando en un DNA hay repetición de una secuencia (unidad de repetición), leída siempre en el mismo sentido. No son palíndromos

porque no se cumple el criterio de igualdad de secuencias leídas en sentido opuesto, ni hay centro de simetría, por lo que no se habla de “mitades”. Estas repeticiones resultan de interés porque aparecen con una gran frecuencia en el genoma (pág. 111), tanto en posiciones contiguas o adyacentes (repeticiones *en tandem*) como separadas por un número elevado de pares de bases (repeticiones *dispersas*).

Finalmente, se habla de *repeticiones especulares* cuando la secuencia se repite, de forma invertida, dentro de la misma hebra. De nuevo, no son palíndromos, pues no hay centro de simetría rotacional, sino plano de simetría especular.

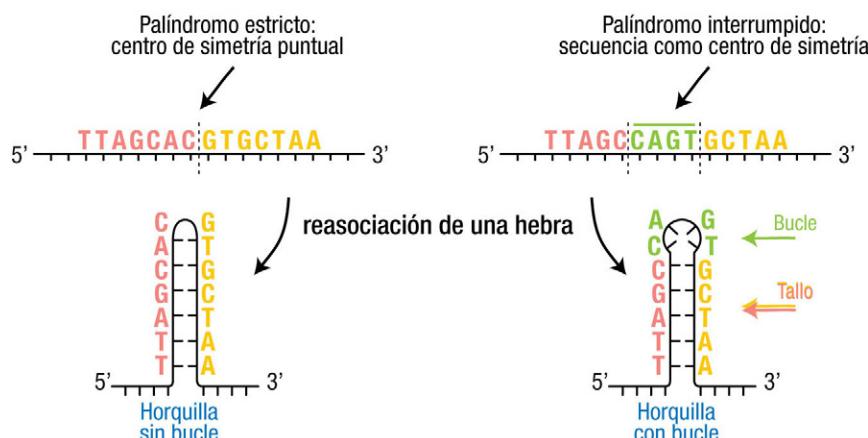
Estas dos últimas posibilidades, aunque aparecen con frecuencia en los DNA, ocupando una parte importante de su secuencia (pág. 108), no tienen consecuencias en cuanto a la estructura secundaria de los ácidos nucleicos, por lo que a partir de este momento sólo se tratarán los palíndromos.

5.2.2.4 Consecuencias estructurales de los palíndromos

a) En ácidos nucleicos monocatenarios

La existencia de palíndromos, por ser secuencias autocomplementarias, puede afectar de forma importante a la estructura secundaria. Así, los DNA o RNA monocatenarios a menudo se pliegan sobre sí mismos para formar una estructura *en horquilla*, siempre que posean regiones complementarias (las dos mitades del palíndromo). Si el centro de simetría no es puntual (palíndromos interrumpidos), esa porción forma un *bucle* en el extremo de la horquilla.

Monocatenario (DNA o RNA)



5.8

La formación de horquillas posee gran relevancia en la estructura tridimensional de algunos RNA. El ejemplo más significativo lo constituyen quizás los *brazos* de los tRNA, regiones con estructura de *tallo* y *bucle* (pág. 66), formados por secuencias palindrómicas. Por otro lado, este tipo de plegamiento en el mRNA contribuye, en algunos casos, a la terminación de la transcripción; concretamente, en el mecanismo de terminación independiente de ρ en procariotas (v. web 17.8).

En el caso de DNA monocatenarios también pueden encontrarse horquillas, por ejemplo en el DNA desnaturizado, que carece de una estructura secundaria regular y definida. También se propone su formación en situaciones particulares, por ejemplo, durante la síntesis en el laboratorio del cDNA (DNA complementario) a partir de un mRNA, donde el extremo 3' del DNA monocatenario se repliega sobre sí mismo y sirve así de cebador para la polimerasa (pág. 223).

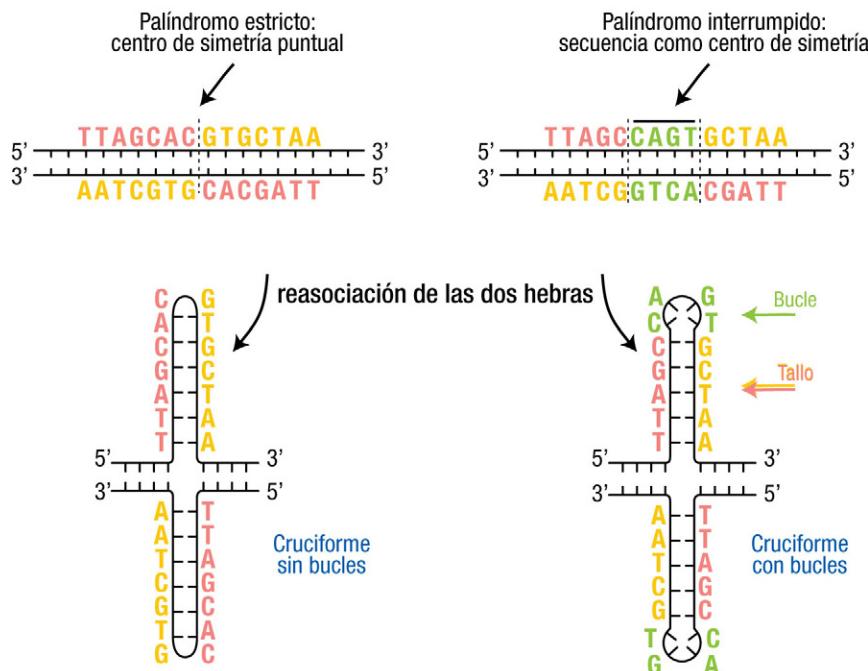


Web 5.6. Horquillas en los ácidos nucleicos.

b) En DNA bicanenario

En el caso de un DNA de doble hebra, las hebras del palíndromo pueden separarse y reasociarse de otra forma, generando estructuras de doble horquilla o *cruciformes* (en forma de cruz). De forma similar a lo que ocurre en el caso anterior, los palíndromos interrumpidos forman cruciformes con dos *bucles*.

Bicanenario (sólo DNA)



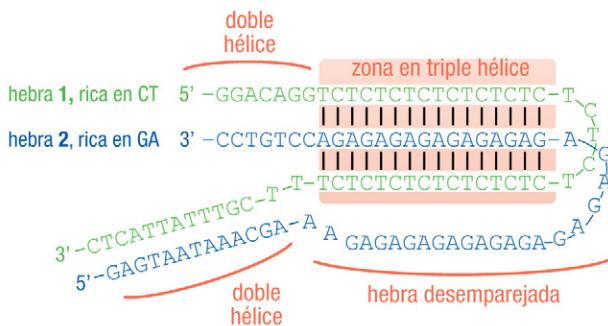
5.2.2.5 Funcionalidad de los palíndromos

Además de las alteraciones de estructura recién descritas, los palíndromos en DNA bicanenario actúan como señales de reconocimiento a las que se unen numerosas proteínas. Éstas poseen habitualmente estructuras diméricas, por lo que se unen de forma simétrica a ambas mitades del palíndromo (pág. 57). Destacan entre ellas los factores de transcripción, que regulan la expresión génica (pág. 288) y las enzimas de restricción, que fragmentan la molécula de DNA (pág. 212).

5.2.3 H-DNA: triple hélice

Se trata de una estructura poco habitual, formada por 3 hebras: una **triple hélice**, también llamada **triplex** o **tríplice**. Se descubrió por primera vez en un RNA, aunque es en el DNA donde se encuentra con más frecuencia. Se forma en regiones ricas en pirimidinas en una hebra (secuencia $(CT)_n$) y ricas en purinas (secuencia $(AG)_n$) en la hebra complementaria. En condiciones normales, dichas secuencias se encuentran totalmente emparejadas sin alteración alguna en la doble hélice. Sin embargo, puede ocurrir que parte de la doble hélice se abra y la hebra rica en C y T se repliegue y se empareje con la otra hebra (AG), mediante una nueva clase de enlaces de hidrógeno (conocidos como enlaces de Hoogsteen), en una zona donde ésta aún forma doble hélice. Aparece así una triple hélice $(CT)_n/(AG)_n/(CT)_n$ en dicho segmento del DNA. La otra hebra sencilla (rica en A y G), desemparejada, forma una especie de bucle.

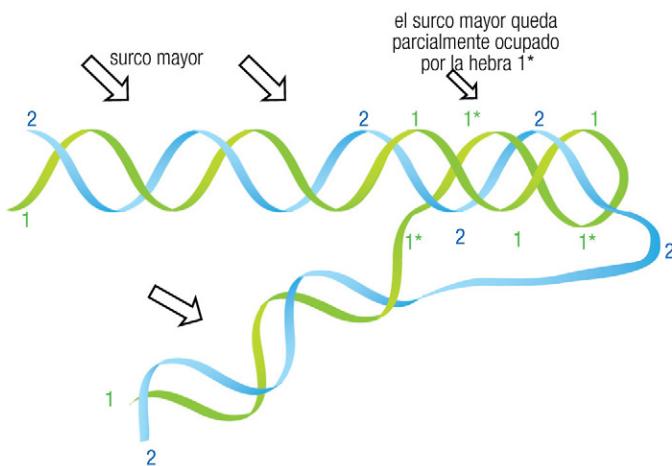
Organización de las cadenas en el plano:



Esquemáticamente:

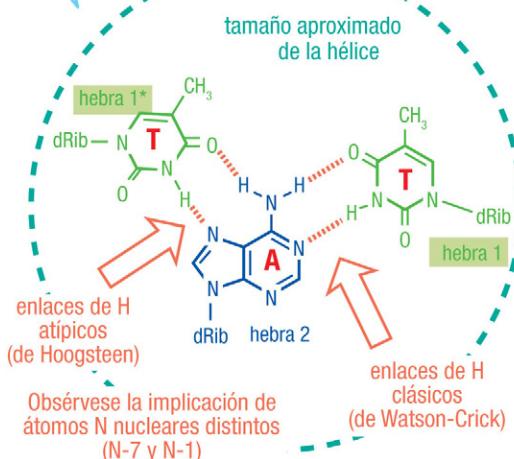
- 1) Las dos hebras (1 y 2) forman una doble hélice.
- 2) Una de las hebras gira volviendo hacia atrás (1^*) y se empareja con las dos hebras anteriores, formando el segmento de triple hélice.
- 3) La otra hebra (2) queda desemparejada, hasta que se reúne de nuevo con la hebra 1 para formar un segundo tramo de doble hélice.

Organización espacial de las cadenas:



5.10

Ejemplo de emparejamiento de 3 bases que da lugar a la triple hélice en el H-DNA (TAT). El otro emparejamiento es entre citosina, guanina y citosina protonada: CGC⁺. En el caso de RNA, la interacción TAT vendría sustituida por UAU.



Este tipo de estructura ha despertado gran interés por su aplicación en terapias de inhibición de la expresión génica, mediante la unión de un oligonucleótido sintético a la doble hélice del DNA en la región promotora de un gen, formando una triple hélice que impide la expresión del gen.

5.3 MOTIVOS ESTRUCTURALES RESPONSABLES DE LA UNIÓN DE PROTEÍNAS AL DNA

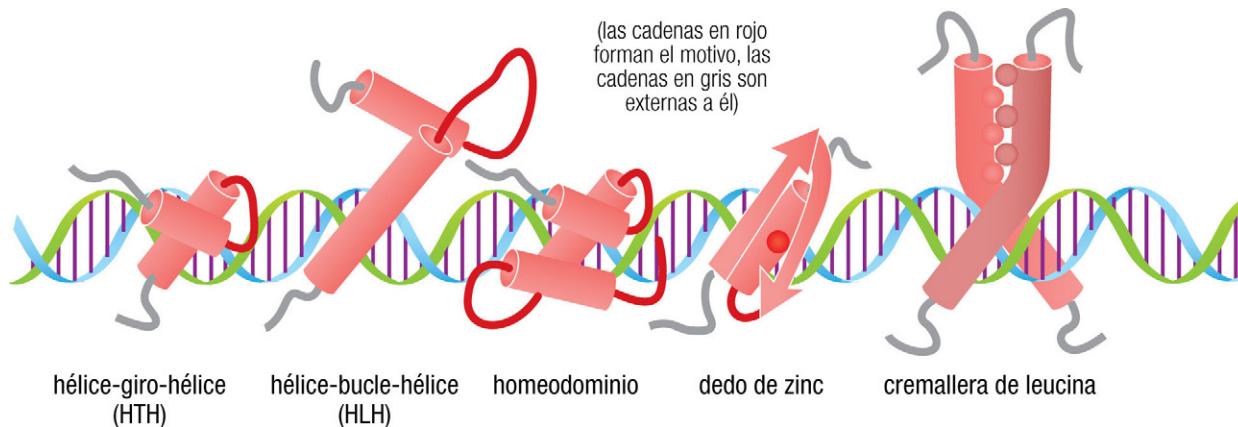
Son numerosos los procesos biológicos en los que la función del DNA depende de su interacción específica con proteínas; destaca entre ellos el control de la expresión génica, ejercido por proteínas denominadas genéricamente *factores de transcripción* (cuyo papel se estudia más adelante, pág. 288). El estudio de estas proteínas tiene interés dentro del contexto de las variantes estructurales del DNA, precisamente por su capacidad de **reconocer e interaccionar con secuencias concretas del DNA**. Este reconocimiento tiene lugar gracias a la posibilidad de variaciones locales en la estructura tridimensional del DNA, así como a la interacción directa de la proteína con las bases, sin necesidad de que se separen las hebras de la doble hélice. En la mayor parte de los casos, los numerosos contactos establecidos (de tipo iónico, hidrófobo y enlaces de hidrógeno), aunque individualmente débiles, se combinan para dar una interacción específica y muy fuerte entre DNA y proteína.

El reconocimiento se ejerce a través de regiones o elementos estructurales, que son sólo una parte de la estructura terciaria completa de la proteína, y que se pueden encontrar, con una estructura espacial muy similar, en proteínas distintas, interviniendo en una misma función en todas ellas. Si se cumplen estos requisitos,

reciben la denominación genérica de “motivos” estructurales proteicos (en inglés, *motifs*), en este caso “motivos de unión a DNA” o “motivos ligantes de DNA”.

La interacción tiene lugar con la superficie de la doble hélice, en especial sobre el surco mayor. Como característica importante, una misma proteína puede contactar simultáneamente con varios puntos diferentes del DNA, forzando así en éste una modificación conformacional que hace que secuencias lejanas puedan aproximarse. Dicho cambio conformacional puede inducir incluso la separación parcial de las hebras (desnaturalización local), con lo que se vería favorecido el acceso de otras proteínas. Tal es el caso, por ejemplo, de la RNA-polimerasa en el proceso de transcripción.

5.3.1 Descripción de los motivos estructurales más frecuentes



5.11

5.3.1.1 Motivo hélice-giro-hélice o hélice-vuelta-hélice

Se abrevia como HTH, del inglés *helix-turn-helix*. Fue el primer motivo proteico de unión al DNA bien estudiado. Se ha encontrado en proteínas reguladoras de la expresión génica, tanto en virus como en prokariotas y eucariotas. Está formado por dos segmentos peptídicos en hélice α , de estructura rígida, separados por un corto tramo peptídico que permite que las dos hélices se aproximen entre sí. Una de las hélices α (llamada “hélice de reconocimiento”) se encaja en el surco mayor del DNA, de forma que los residuos aminoácidos de un lado de la hélice interaccionan mediante puentes de hidrógeno o enlaces de van der Waals con las bases nitrogenadas expuestas en ese surco. El diámetro de la hélice α posee la dimensión adecuada (1,2 nm) para entrar en el surco. Algunos aminoácidos, situados en la cara opuesta de la misma hélice, establecen interacciones con otros de la segunda hélice, fijando la posición relativa de ambas en un ángulo casi perpendicular (v. el esquema inicial). La segunda hélice α queda así cruzada sobre la de reconocimiento y situada en el exterior del DNA.



Web 5.7. Estructura de los motivos hélice-giro-hélice.

Es frecuente que una proteína presente dos motivos HTH o bien que se asocien dos moléculas de proteína iguales (es decir, un homodímero), cada una con un motivo HTH. En ambos casos, las dos hélices de reconocimiento (una de cada motivo) se sitúan a una distancia mutua de unos 3,5 nm, lo que coincide con el paso de hélice del B-DNA (3,54 nm) y, por tanto, con dos posiciones consecutivas de su surco mayor. Es decir, ambos motivos interaccionan de manera simultánea con el DNA, contribuyendo a estabilizar la unión de la proteína.

Los ejemplos mejor conocidos de proteínas con motivos HTH actúan regulando la transcripción en bacterias y virus, respectivamente. El primero, la *proteína receptora de AMP cíclico* (CRP, también llamada CAP) es un dímero –por tanto, con 2 motivos HTH– que cuando une cAMP aumenta su afinidad por secuencias específicas del DNA, desencadenando la expresión de los genes del operón lactosa (éste fue el primer ejemplo conocido de control de la transcripción). El segundo ejemplo (en la web 5.7) es la proteína Cro

del fago lambda, codificada por un gen del virus y expresada en la bacteria anfitriona, que interacciona a través de su motivo HTH con el gen de la proteína vírica denominada *repressor del fago λ*, bloqueando como consecuencia su expresión (transcripción del gen y posterior síntesis de la proteína).

5.3.1.2 Motivo hélice-bucle-hélice

Se abrevia como HLH, del inglés *helix-loop-helix*. No debe confundirse con el anterior. Consta también de dos segmentos en hélice α , pero en este caso el péptido que los une es más largo, lo que lo dota de mayor flexibilidad y, en consecuencia, hay más posibilidades de orientación mutua de las dos hélices. Generalmente, al igual que en HTH, dos motivos HLH se asocian, formándose proteínas diméricas, aunque no siempre interaccionan en posiciones consecutivas del surco mayor. Un ejemplo de motivo HLH bien conocido es el de la proteína *Max* de ratón.



Web 5.8. Estructura de los motivos hélice-bucle-hélice.

5.3.1.3 Motivo homeodominio

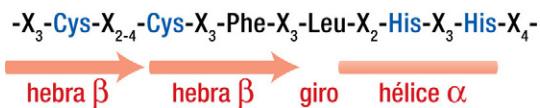
Se puede considerar como una ampliación del motivo hélice-giro-hélice (HTH), pero adquiere entidad propia por aparecer repetidamente, con idéntica estructura, en distintas proteínas. Además de la disposición de dos hélices cruzadas típica del HTH, existe un tercer tramo en hélice α , del que sale una región sin estructura secundaria definida cuyos aminoácidos interaccionan de manera directa con el DNA. Este motivo (en inglés, *homeodomain*) es importante por aparecer en las proteínas que regulan el desarrollo embrionario, estudiadas especialmente en *Drosophila*.



Web 5.9. Estructura de los motivos homeodominio.

5.3.1.4 Motivo dedo de zinc

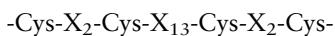
Se ha encontrado este tipo de motivo estructural (en inglés, *zinc finger*) en multitud de proteínas eucarióticas que se unen al DNA, entre ellas el factor de transcripción TFIIIA y el receptor de estrógenos. Al igual que en las estructuras anteriores, es frecuente que una misma proteína posea más de un motivo; en este caso, se observan normalmente múltiples dedos de zinc consecutivos. Éste es un elemento estructural formado por unos 30 aminoácidos, de los cuales 2 cisteínas y 2 histidinas aparecen en posiciones constantes, coordinando tetraédricamente un ion Zn^{2+} . Esta unión obliga a un plegamiento de la cadena peptídica, de forma que el motivo se caracteriza por una conformación tridimensional alargada, en forma de “dedo”, que le da nombre. La variante más frecuente muestra dos hebras β antiparalelas, cada una con un residuo Cys, seguidas de un giro y de una estructura helicoidal α con los dos residuos His. Este motivo se denomina C_2H_2 (abreviaturas de una letra para los aminoácidos que unen el Zn^{2+}) y su secuencia consenso es:



5.12

donde los subíndices representan el número de residuos (aminoácidos variables, X) que separa los aminoácidos consenso (constantes). Estas secuencias adquieren espontáneamente gracias al Zn^{2+} la conformación apropiada para que el segmento α -helicoidal interaccione con tres bases expuestas en el surco mayor del DNA. Los residuos variables permiten que se reconozcan distintas secuencias del DNA.

En otras proteínas se observa una variante de dedo de zinc llamada C4, donde el ion metálico se coordina con cuatro residuos Cys en la secuencia consenso



que no forma hebras β , sino dos segmentos en hélice α . Esta estructura aparece, por ejemplo, en los receptores de hormonas esteroides.



Web 5.10. Estructura de los motivos dedo de zinc.

5.3.1.5 Motivo cremallera de leucina

Se trata de una región de la proteína cuya secuencia tiene un residuo de leucina cada 7 aminoácidos. Al adoptar la conformación α -helicoidal se concentran en un lado los aminoácidos hidrófobos, pues Leu se repite en la misma cara cada dos vueltas de hélice (2 vueltas de 3,6 residuos \cong 7 aminoácidos). Dos cadenas peptídicas de este tipo pueden asociarse debido al efecto hidrófobo, intercalando sus restos de Leu como si se tratara de los dientes de una cremallera (en inglés, *leucine zipper*). En el otro extremo del motivo las dos hélices α encajan en el surco mayor del DNA y presentan en la cara exterior abundantes aminoácidos básicos que interaccionan favorablemente con los fosfatos. De alguna manera, este modelo se asemeja a una “Y”, en la que el tallo correspondería a la cremallera de leucina y los brazos a los tramos de las hélices que contactan con el DNA. Estructuras de este tipo aparecen, por ejemplo, en las proteínas reguladoras de la transcripción llamadas *bZIP*.



Web 5.11. Estructura de los motivos cremallera de leucina.



Estructura y función de los RNA

6.1 ESTRUCTURA DE LOS RNA	61
6.1.1 Composición de bases y estructuras primaria y secundaria del RNA	61
6.1.2 Estructura tridimensional de los RNA en general	62
6.1.3 Tipos de RNA: propiedades y estructuras particulares	63
6.1.3.1 RNA mensajero (mRNA)	64
6.1.3.2 RNA transferente (tRNA)	64
a) Especificidad de los tRNA por los aminoácidos	64
b) Presencia de nucleótidos infrecuentes	65
c) Emparejamiento de bases intracatenario	66
d) Estructura secundaria en trébol	66
e) Estructura terciaria en "L"	67
6.1.3.3 RNA ribosómico (rRNA)	68
6.1.3.4 RNA interferente y otros RNA pequeños	69
6.2 FUNCIONES DE LOS RNA	70
6.2.1 RNA como material genético	70
6.2.1.1 Depositario de la información genética: genoma de RNA	70
6.2.1.2 Portador transitorio de la información genética	70
6.2.2 Síntesis de proteínas	70
6.2.3 Maduración del RNA	71
6.2.4 Regulación de la expresión de los genes	71
6.3 RNA CATALÍTICO: RIBOZIMAS	71
6.3.1 Acción catalítica en la maduración de RNA	71
6.3.2 Acción autocatalítica en la eliminación de intrones	72
6.3.3 Acción catalítica y autocatalítica en viroides: ribozimas de posible aplicación terapéutica	73
6.3.4 Acción catalítica transpeptidasa en la síntesis de proteínas	74
6.4 LOS RIBOSOMAS	74
6.4.1 Ubicación subcelular	75
6.4.2 Origen	75
6.4.3 Composición	75
6.4.4 Estructura	77
6.4.5 Separación de los ribosomas, subunidades y componentes	78

Los ácidos nucleicos no se encuentran en las células en las formas extendidas correspondientes a su estructura primaria y secundaria, sino molecularmente más compactados, lo que hemos denominado (pág. 28) nivel estructural “de orden superior”. En el caso del ácido ribonucleico, RNA, esto se plasma en estructuras tridimensionales variadas que, en algunos casos, se complementan con su asociación estrecha con proteínas formando ribonucleoproteínas (por ejemplo, el RNA ribosómico, que al unirse a proteínas forma los ribosomas). Todo ello se relaciona con una serie de funciones ejercidas específicamente por cada tipo de RNA, siempre relacionadas con la expresión de la información contenida en el DNA genómico.

6.1 ESTRUCTURA DE LOS RNA

Como ya se ha indicado (pág. 28), los RNA son polirribonucleótidos (polinucleótidos con ribosa, y U en lugar de T) monocatenarios, formados por cadenas lineales de varias decenas o millares de unidades, pero de longitud muy inferior a la del DNA. No se presentan como moléculas bicatenarias con dos hebras complementarias (salvo en algunos virus que tienen como genoma un RNA bicatenario (pág. 70).

6.1.1 Composición de bases y estructuras primaria y secundaria del RNA

La estructura primaria de los RNA se ha estudiado anteriormente (v. Capítulo 3). En cuanto a la secundaria, es característica la ausencia de una estructura repetitiva y regular, lo que los diferencia del DNA; en realidad sólo existe estructura secundaria en una parte de la molécula de algunos tipos de RNA.

Al no existir un emparejamiento sistemático de dos hebras, la composición de bases en los RNA es mucho más variada que en el DNA y no se cumplen las reglas de Chargaff (pág. 36). Por otro lado, en algunos RNA hay cierta abundancia de nucleótidos distintos de los cuatro principales (A, G, C y U), aunque en proporción minoritaria con respecto a la de éstos.

Composición de bases de varios RNA (%)

Organismo	Tipo de RNA	A	G	C	U
Rata (hígado)	nuclear	20,2	25,7	29,5	24,6
	mitocondrial	17,8	31,8	28,4	20,9
	ribosómico	20,0	30,5	31,6	20,2
	transferente	20,9	30,9	28,5	20,4
	mensajero	23,8	28,3	27,4	20,5
Levaduras	ribosómico	24,9	27,7	19,4	26,7
	transferente	18,5	29,2	28,4	20,0
	mensajero	27,5	25,1	20,3	27,1
<i>E. coli</i>	ribosómico	~ 25	~ 31	~ 22	~ 21
	transferente	19,3	32,0	28,3	16,0
	mensajero	24,1	27,7	24,7	23,5

Obsérvese cómo, a diferencia de lo que ocurre con el DNA,

- la composición de bases varía dentro de una misma especie (según el RNA considerado)
- no existe una equivalencia entre bases ($A \neq U$, $G \neq C$, purinas \neq pirimidinas)

Algunos RNA adoptan una estructura secundaria en regiones limitadas de su molécula, gracias al plegamiento de su única hebra sobre sí misma, cuando esto permite el emparejamiento *intracatenario* de dos zonas cuyas secuencias sean complementarias, a pesar de estar separadas en la estructura primaria; se forma así una doble hélice local, de carácter antiparalelo. De manera similar se forman también dobles hélices entre una molécula de RNA y una hebra sencilla de DNA (híbridos RNA-DNA).

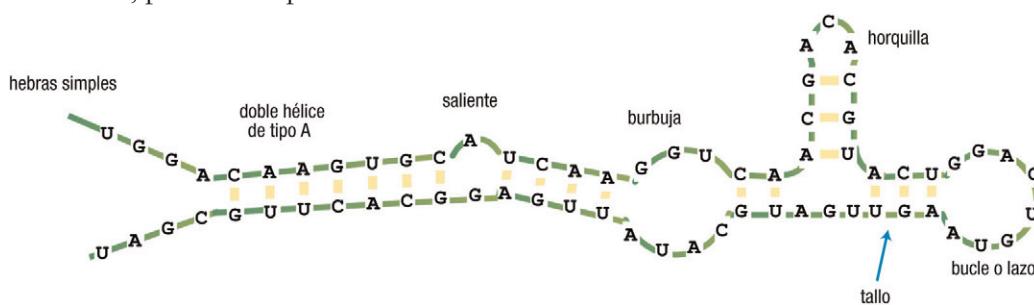
En estos casos, si el grupo 2'-OH de la ribosa adopta la conformación C2'-*endo* presenta impedimentos estéricos con otros átomos de la cadena de RNA, lo que impide que se forme la doble hélice (que sería de tipo B). En concreto, el 2'-OH estaría muy cerca de los átomos del grupo fosfato contiguo y del C8 de la base adyacente. En cambio, en la conformación C3'-*endo* el grupo 2'-OH se proyecta hacia fuera, lejos de otros átomos del esqueleto y no genera impedimento estérico alguno, por lo que es posible la formación de la doble hélice, en este caso de tipo A. En resumen, la única doble hélice en la que puede participar un RNA es la de tipo A (el DNA, al carecer del OH en 2', puede adoptar tanto la A como la B, como ya se ha estudiado en el Capítulo 5).



Web 6.1. Conformaciones de la ribosa en el RNA.

6.1.2 Estructura tridimensional de los RNA en general

Existe una gran diversidad en la estructura tridimensional de las moléculas de RNA. Ésta es compleja, única para cada tipo de RNA, y resulta de la combinación de estructuras secundarias locales, estabilizadas por enlaces de hidrógeno, por el efecto hidrófobo y por interacciones de apilamiento de bases. Su estudio concreto se hará más adelante, para cada tipo de RNA.



6.1

A pesar de la inexistencia de una estructura secundaria propiamente dicha para toda la molécula, la cadena sencilla de cualquier RNA adopta localmente variadas estructuras: horquillas, bucles, lazos, etc. Algunas de ellas se deben al emparejamiento de bases *intracatenario* formando cortos tramos de doble hélice de tipo A, como se acaba de indicar. La proporción de zonas helicoidales varía ampliamente entre distintos RNA, pero en algunos casos supera la mitad de la estructura total.

6.1.3 Tipos de RNA: propiedades y estructuras particulares

Tanto en procariotas como en eucariotas existen tres tipos principales de RNA, denominados mensajero (mRNA), transferente (tRNA) y ribosómico (rRNA). Además de ellos, en eucariotas existen RNA nuclear heterogéneo (hnRNA), nuclear pequeño (snRNA) y citoplásmico pequeño (scRNA), y RNA de orgánulos (mitocondrias y cloroplastos), así como otros RNA pequeños con actividad reguladora (miRNA y siRNA, principalmente). Todos ellos tienen en común que sus funciones están relacionadas, directa o indirectamente, con la expresión génica (maduración postranscripcional y traducción). Adicionalmente, en algunos virus el genoma está constituido por RNA, que desempeña así el papel de portador de la información genética, el mismo que en el resto de organismos corresponde al DNA.

Las diferencias fundamentales entre los distintos tipos de RNA se resumen en la tabla.

Clase	Ubicación subcelular	Cantidad	Tamaño		Características particulares
			s	nº nucleótidos	
mRNA	citoplasma (P y E)	5%		600 a 3.000	Estructura sencilla lineal, sin plegamiento
tRNA	citoplasma (P y E)	20%		75 a 95	Hay de 50 a 60 diferentes, específicos para cada aminoácido
rRNA	citoplasma (P y E)	75%			
	3 tipos, en ribosomas de procariotas:		16S	1.500	Forma parte de la subunidad pequeña del ribosoma
			23S	2.900	Forman parte de la subunidad grande del ribosoma
			5S	120	
	4 tipos, en ribosomas de eucariotas:		18S	1.900	Forma parte de la subunidad pequeña del ribosoma
			28S	4.700	Forman parte de la subunidad grande del ribosoma
			5,8S	160	
			5S	120	
hnRNA	núcleo (E)	minoritario		200 a 30.000	Transcrito primario, precursor del mRNA
snRNA	núcleo (E)	minoritario		100 a 300	Forman parte de ribonucleoproteínas nucleares, snRNP ¹
scRNA	citoplasma (E)	minoritario			Forman parte de ribonucleoproteínas citoplásMICAS, scRNP ²
mtRNA	mitocondrias (E)	minoritario			De los 3 tipos: mensajero, transferente y ribosómico
cpRNA	cloroplastos (E)	minoritario			Sólo en vegetales, los 3 tipos: mensajero, transferente y ribosómico
miRNA	citoplasma (E)	minoritario		21 a 25	RNA pequeños que regulan la expresión génica
siRNA	citoplasma (E)	minoritario		21 a 25	
otros	núcleo/citoplasma	minoritario			

P = procariotas

E = eucariotas

s = coeficiente de sedimentación (medido en svedbergs, web 10.2)

¹Son ejemplos de ribonucleoproteínas nucleares las proteínas U1, U2, U4, U5 y U6 que forman el ayustosoma (pág. 304), que interviene en la maduración del mRNA eucariótico.

²Un ejemplo de scRNA es el RNA 7SL, componente clave de la partícula de reconocimiento de señal (pág. 353).

6.1.3.1 RNA mensajero (mRNA)

La propuesta de la existencia de una molécula que transfiera el mensaje genético desde el núcleo al citoplasma se debe a François Jacob y Jacques Monod (premios Nobel en 1965). Dicha molécula es el RNA mensajero que, al ser una copia de la información contenida en la secuencia del DNA, actúa posteriormente en el ribosoma como molde o plantilla para la síntesis proteica.

Las características principales del mRNA son las siguientes:

- Se encuentra en el citoplasma de todas las células, en baja proporción (5%) respecto al RNA total y con una vida media corta dada su rápida degradación. Aparece asociado transitoriamente a los ribosomas.
- Sus moléculas son de tamaño muy diverso (relacionado con el tamaño de las proteínas que codifican).
- Es el tipo de RNA de estructura más sencilla, con una cadena lineal formada exclusivamente por los ribonucleósidos A, U, C y G. La composición de bases refleja directamente la del DNA molde del que procede. La molécula carece de una conformación definida: la cadena se mantiene extendida, sin adoptar estructuras secundarias ni terciarias significativas, en parte gracias a la participación de proteínas que se asocian a ella.

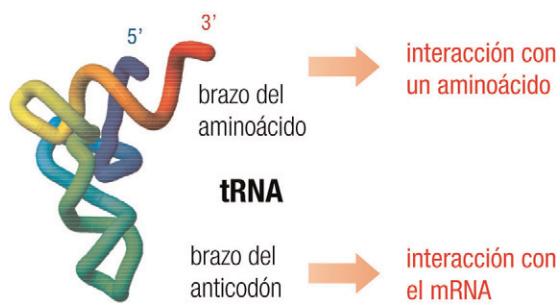
6.1.3.2 RNA transferente (tRNA)

Fue Francis Crick quien sugirió que los aminoácidos se unirían a una molécula intermedia, que llamó “adaptador”, para poder interaccionar con el mRNA. Efectivamente, los tRNA actúan así en la síntesis de proteínas. Esta función depende de una doble interacción:

- Por un extremo de la molécula unen el aminoácido para dar aminoacil-tRNA (pág. 326), la forma activa necesaria para la reacción de formación del enlace peptídico.
- Por el otro extremo se emparejan con el mRNA, en presencia del ribosoma.

La combinación de ambas interacciones permite la transferencia del aminoácido al péptido creciente, generando el orden correcto de los aminoácidos en la proteína, de acuerdo con la secuencia “leída” por los sucesivos tRNA sobre el mRNA (pág. 337).

El tRNA es el mediador entre el mensaje del mRNA y la secuencia de aminoácidos del polipéptido que se está sintetizando



6.2

A pesar de ser los RNA más pequeños, los tRNA presentan una gran complejidad estructural. Se han aislado y estudiado muchos, de distintos organismos, mostrando una estructura común a pesar de sus distintas secuencias; estas últimas reflejan la necesidad de adaptarse a su especificidad por un único aminoácido.

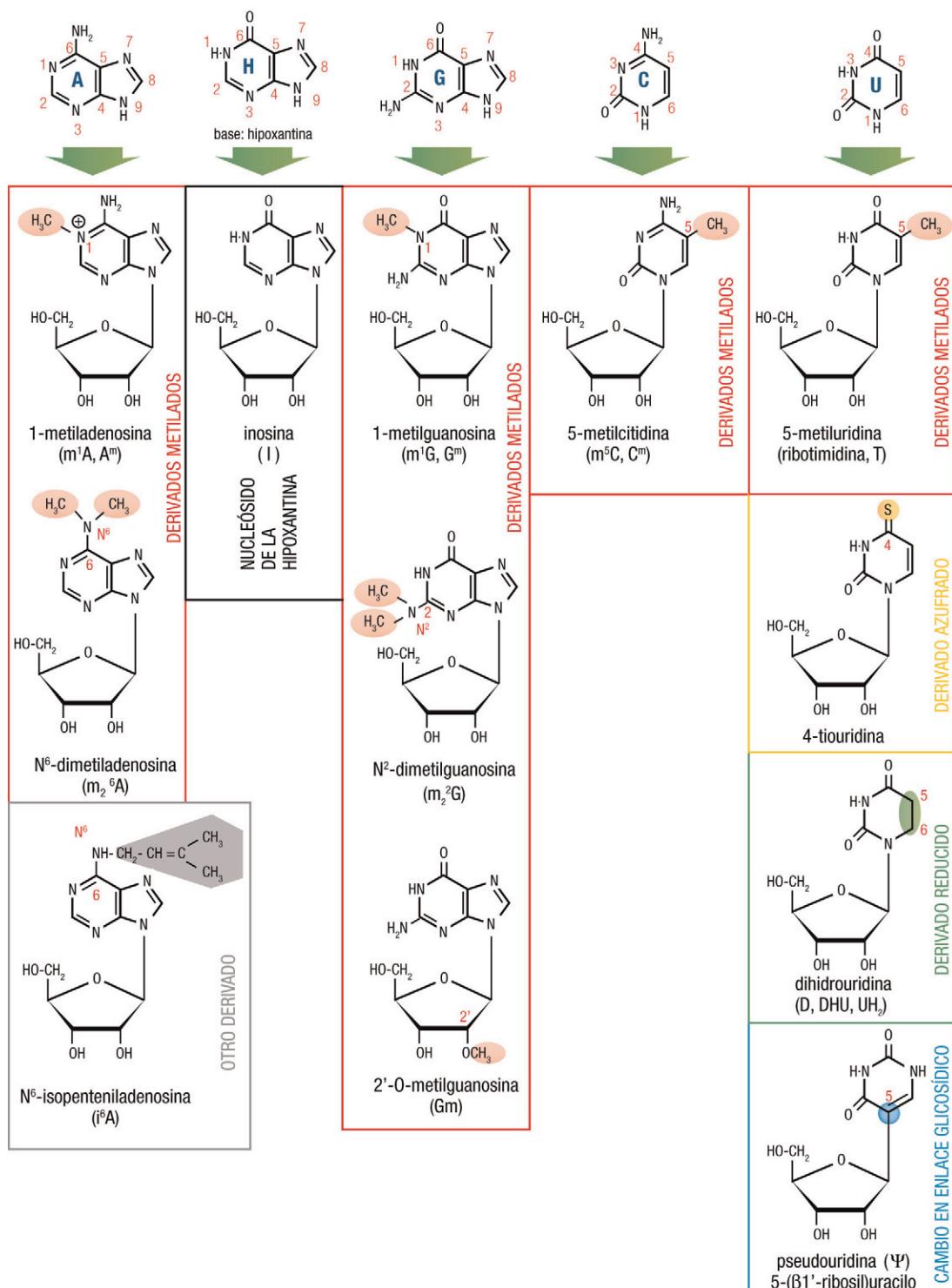
a) Especificidad de los tRNA por los aminoácidos

Los tRNA constituyen una fracción importante del RNA total presente en la célula (20%) y se encuentran disueltos en el citosol. Una célula puede contener hasta 60 tRNA diferentes, es decir, en número superior al de aminoácidos (20), puesto que cada tRNA interacciona específicamente con un solo aminoácido, pero algunos aminoácidos son reconocidos por más de un tRNA (se habla en este caso de tRNA “sinónimos”, pág. 319).

La especificidad se indica con el nombre del aminoácido en superíndice (por ejemplo, tRNA^{Ala}); los tRNA sinónimos llevan subíndices numéricos (por ejemplo, tRNA^{Trp₁} y tRNA^{Trp₂}). Cuando un tRNA está unido con el aminoácido se dice que está “activado” o “cargado”, se denomina aminoacil-tRNA y se indica con el aminoácido como prefijo (por ejemplo, Ala-tRNA^{Ala}, Trp-tRNA^{Trp₁}, Trp-tRNA^{Trp₂}, etc.).

b) Presencia de nucleósidos infrecuentes

Una de las características de los tRNA es que su composición incluye, además de los cuatro principales, otros nucleósidos no corrientes o **infrecuentes**, que pueden suponer hasta un 10% del total. Generalmente se trata de derivados de los cuatro principales, A, G, C, U, o bien de la inosina (nucleósido de la base hipoxantina), a menudo con grupos metilo en la base o en la desoxirribosa (las bases infrecuentes y los nucleósidos infrecuentes se han presentado en las págs. 15 y 19). Además, es característica la presencia de ribotimidina. La participación de todos estos nucleósidos está relacionada con la estructura terciaria y tiene por ello una clara influencia sobre la función.



c) Emparejamiento de bases intracatenario

A pesar de la abundancia de nucleótidos no usuales, los tRNA muestran un grado notable de complementariedad, es decir, de **regiones con estructura secundaria en doble hélice** basada en emparejamientos de bases **intracatenarios**. Con frecuencia se observan emparejamientos distintos de los de Watson y Crick, llamados genéricamente *de Hoogsteen*: por ejemplo, G puede emparejarse con U mediante un enlace menos fuerte que la unión G≡C. Incluso, en algunos casos, en el enlace de hidrógeno participa el 2'-OH de la ribosa. También pueden existir bases enfrentadas pero no complementarias, así como uno o más nucleótidos formando un bucle en la cadena para facilitar el emparejamiento de sus vecinos.

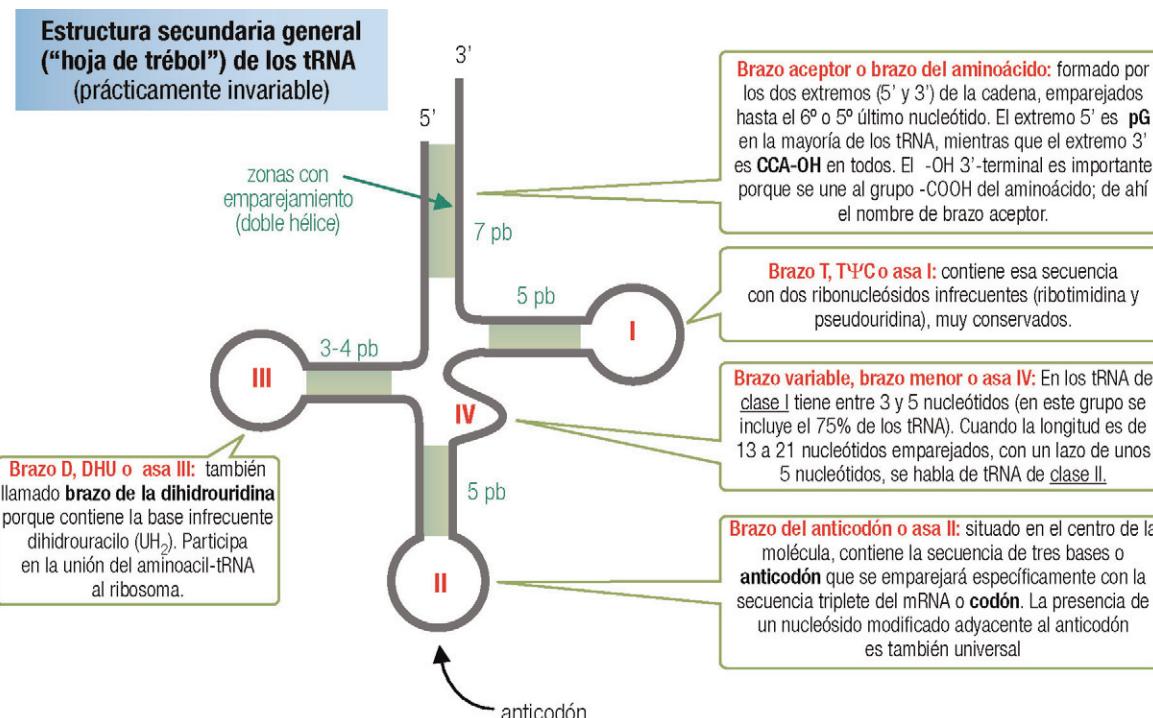


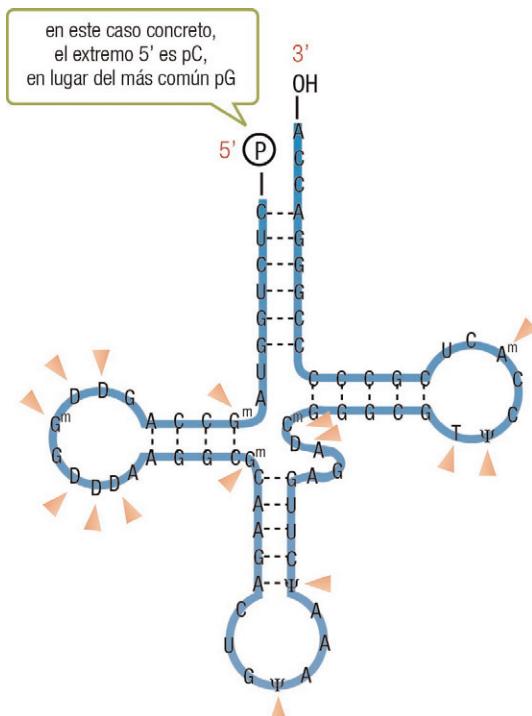
Web 6.2. Emparejamientos de bases atípicos en los RNA.

d) Estructura secundaria en trébol

Al comparar la estructura de distintos tRNA se observa una notable semejanza. El plegamiento más probable es la estructura secundaria denominada *en trébol* (por su analogía con tres hojas y un tallo, o para otros con cuatro hojas), donde el emparejamiento entre bases es máximo. En ella se distinguen:

- Tres **brazos o asas principales**; una central, II, y dos periféricas, I y III (cada una con tallo y bucle).
- Un **brazo o asa menor** (IV), de dimensión variable en diferentes tRNA.
- Un brazo abierto: **brazo acceptor** o **brazo del aminoácido**.





6.5

Ejemplo:
Detalles de la estructura primaria y secundaria del tRNA^{Tyr} de levadura

Las puntas de flecha señalan los nucleótidos infrecuentes:

D = dihidrouridina
(también DHU o UH₂)

G^m = 1-metilguanosina
(también m¹G)

A^m = 1-metiladenosina
(también m¹A)

C^m = 5-metilcitosina
(también m⁵C)

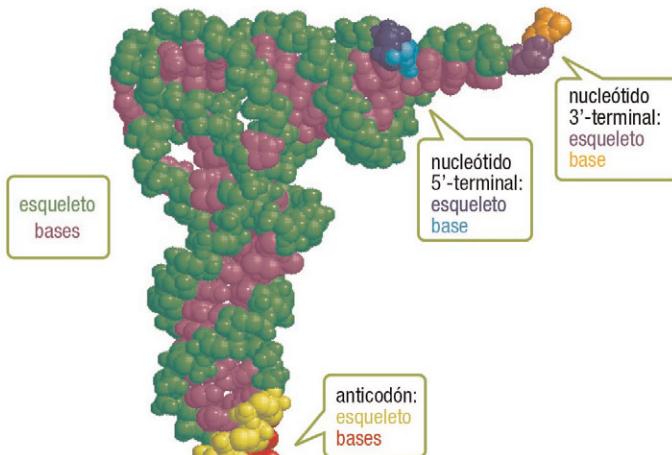
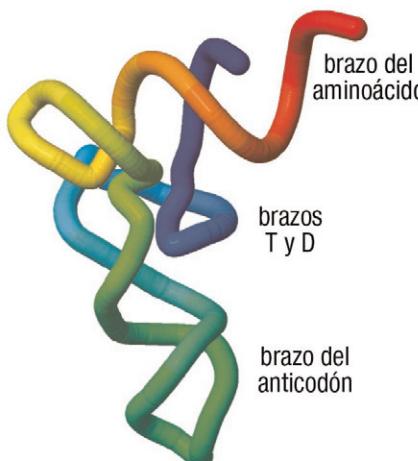
Ψ = pseudouridina

T = ribotimidina

En otros tRNA aparecen otros nucleótidos, como la inosina en el anticodón.

e) Estructura terciaria en "L"

Mediante análisis por difracción de rayos X de cristales de tRNA se ha podido deducir una estructura terciaria compacta, en forma de codo o "L". Ésta se origina por un retorcimiento en el espacio de la estructura secundaria en trébol, impuesto por las dobles hélices locales y por puentes de hidrógeno que acercan los brazos D y T. Estos puentes son de tipos infrecuentes, incluyendo en algunos casos tríos de bases (web 6.2). La estructura en trébol es, por tanto, sólo una representación bidimensional simplificada de la verdadera estructura tridimensional.



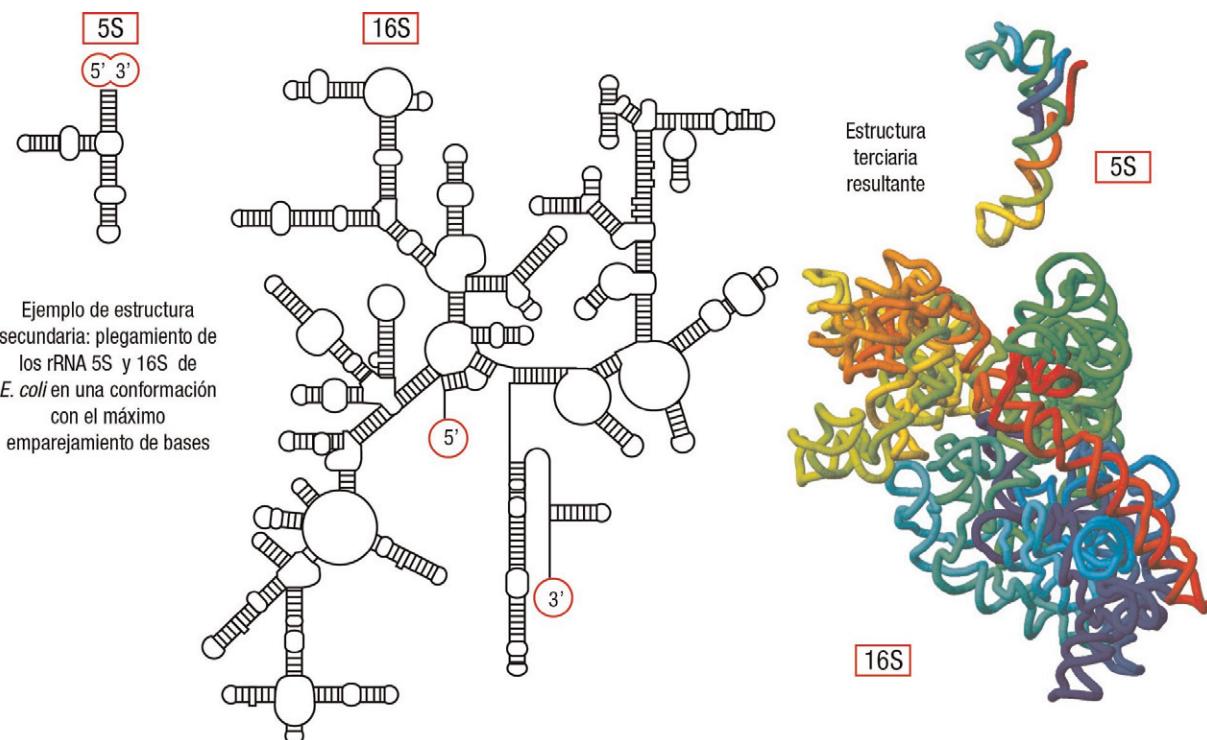
Web 6.3. Estructura tridimensional del RNA transferente.

El conocimiento detallado de esta estructura permite comprender una peculiaridad de los tRNA: todos son distintos (por tener distinta secuencia e interaccionar con un codón de mRNA y un aminoácido específicos), pero al tiempo son de algún modo similares (por adoptar una estructura terciaria común, unirse a ribosomas y ser reconocidos por los factores que intervienen en la traducción).

6.1.3.3 RNA ribosómico (rRNA)

Es el soporte estructural y componente principal de los ribosomas. En todas las células, sean procarióticas o eucarióticas, existen varios tipos de rRNA, constituyendo la mayor parte del RNA total (75%). Al igual que los tRNA, presentan un contenido relativamente elevado de nucleósidos infrecuentes (fundamentalmente Ψ , T y los que tienen bases metiladas).

Los RNA ribosómicos adoptan una conformación compleja. Están formados por una sola hebra con numerosas regiones helicoidales cortas resultantes del emparejamiento intracatenario de bases. Son moléculas con muchos repliegues sobre sí mismas, formando una estructura tridimensional bastante compacta. Su estructura constituye un armazón sobre el que se ensamblan varias proteínas para dar lugar al ribosoma (pág. 75-76).



Web 6.4. Estructura tridimensional de los RNA ribosómicos.

Conviene hacer en este punto una breve mención a la relación entre estas moléculas de rRNA funcionales y sus moléculas precursoras, pues éstas se encuentran también presentes tanto en las células procarióticas como en el núcleo de las eucarióticas. Obviamente, la comprensión completa de estas interrelaciones sólo se alcanzará cuando se describa el proceso de maduración postranscripcional (v. Capítulo 19). Por otra parte, debe indicarse que la “S” en los nombres de los rRNA aquí presentados corresponde a la unidad de coeficiente de sedimentación, el svedberg (v. web 10.2).

En eucariotas, la mayor parte del RNA presente en el núcleo corresponde a los *transcritos primarios*, moléculas obtenidas directamente de la transcripción del DNA y que se convertirán en RNA funcional o maduro tras sufrir el citado proceso de maduración. Hay dos transcriptos primarios precursores de los rRNA; el mayor tiene un tamaño que varía entre 6 y 14 kb, según la especie (con un coeficiente de sedimentación comprendido entre 35 y 47S). Poco después de ser sintetizado, ese pre-rRNA grande es escindido para dar tres rRNA maduros, de 28S, 18S y 5,8S (v. tabla en pág. 63). Estos tres, junto al rRNA 5S (producto del otro pre-RNA, y por ello de un gen diferente), son los RNA que constituyen el núcleo estructural de los ribosomas eucarióticos (pág. 76).

En procariotas, un único pre-rRNA da lugar, también por procesamiento postranscripcional, a rRNA maduros de 23S, 16S y 5S (pág. 63), que forman el ribosoma procariótico (pág. 76). Además, a partir del mismo precursor se generan algunos RNA transferentes.

6.1.3.4 RNA interferente y otros RNA pequeños

Entre los RNA minoritarios en cuanto a su abundancia en la célula se encuentran los **RNA nucleares pequeños** y **citoplásmicos pequeños**. Se trata de RNA monocatenarios con menos de 300 nucleótidos cuya estructura no ofrece características de relevancia que merezcan su discusión en detalle y cuyas funciones, particulares para cada uno, se estudiarán más adelante (v. las citas en la tabla de pág. 63).

Recientemente se han ido descubriendo múltiples moléculas pequeñas de RNA, mono- o bicatenario, que ejercen un papel regulador en la expresión de ciertos genes. Han alcanzado rápidamente relevancia debido a la gran utilidad de su aplicación en investigación y potencialmente en terapia (básicamente, la posibilidad de bloquear a voluntad la expresión de un gen concreto). A menudo se habla genéricamente de RNA pequeños o RNA reguladores, pero se han identificado dos grupos de moléculas con estructura, biosíntesis y mecanismo de actuación definidos: los **microRNA** (o miRNA) y los **RNA interferentes pequeños** (siRNA, de *small interfering RNA*).

Los **microRNA** son moléculas con entre 21 y 25 residuos nucleótidos, monocatenarias en su forma madura, bicatenarias en la forma precursora, cuya secuencia es complementaria al extremo 3' de un mRNA, de modo que se emparejan con él y bloquean su traducción. Desempeñan así un papel regulador en la traducción de algunos genes de proteínas (que se estudiará con más detalle en el Capítulo 21). Bajo la categoría microRNA se incluyen los **RNA temporales pequeños** (stRNA).

Los **RNA interferentes pequeños** son moléculas bicatenarias, igualmente de 21 a 25 pb. Ambos extremos 3' presentan dos nucleótidos sin emparejar (extremos protuberantes, equivalente a los extremos cohesivos, pág. 215). Las dos hebras se separan dando la forma monocatenaria, que se asocia con un complejo proteico (RISC) y lo dirige hacia un mRNA específico que tenga secuencia complementaria a la del siRNA. Como consecuencia, RISC degrada el mRNA diana. Se trata, de nuevo, de un mecanismo de control postranscripcional y pretraduccional, en este caso no por bloqueo del mRNA, sino mediante su eliminación (v. Capítulo 21).

La **ribointerferencia** (*RNA interference*, RNAi, interferencia por RNA) es el fenómeno desencadenado por los **RNA interferentes** (*interfering RNA*, iRNA), incluidos los microRNA y los siRNA. El nombre se utiliza para designar tanto al fenómeno biológico de regulación como a la técnica de laboratorio en la que se usan RNA interferentes sintetizados ex profeso con el fin de silenciar un gen determinado, con propósitos de investigación o aplicados.

Los **ribointerruptores** o **riborreguladores** (en inglés, *riboswitches*) son un tipo más de RNA con actividad reguladora. En este caso, se trata de una parte de la molécula de un mRNA que regula la expresión del propio mRNA (el final de su transcripción, su maduración postranscripcional o su traducción). Estas regiones del mRNA poseen cierto grado de estructura secundaria que genera una estructura terciaria capaz de reconocer y unir de forma específica un metabolito. Por ejemplo, se han descubierto ribointerruptores que unen adenina, guanina, xantina, hipoxantina, glutamina, glicina, lisina, pirofosfato de tiamina, riboflavina, cobalamina, Mg²⁺, FMN, S-adenosilmetionina, o glucosamina-6P. Todas estas biomoléculas pequeñas actúan como señal para la regulación de la expresión de algún gen, habitualmente relacionado con su metabolismo. El cambio conformacional inducido en el RNA al unirsele el ligando conduce a la actuación sobre la expresión génica. La mayoría de ribointerruptores se han detectado en bacterias, pero al menos el de tiamina está también presente en algunos hongos y plantas.

Web 6.5. Estructura de RNA pequeños con función reguladora.

El descubrimiento de estas nuevas funciones para moléculas de RNA añade una pincelada a la percepción anterior de que gran parte del genoma humano –y, en general, de organismos superiores en la escala evolutiva– sea no codificante y pudiera ser inútil (el denominado *DNA basura*). La búsqueda tradicional de genes se centró en la codificación de proteínas; cada vez está más claro que el panorama es más complejo y al menos una parte de ese DNA presumiblemente inútil tiene realmente una función codificante. Parece hoy probable que existan miles de genes en el genoma humano codificando RNA pequeños con función reguladora.

6.2 FUNCIONES DE LOS RNA

En directa relación con la clasificación anterior, basada en su estructura, se puede abordar el estudio de las funciones ejercidas por los RNA, un catálogo que está sufriendo una ampliación continua en los últimos años, según se van descubriendo nuevas estructuras y acciones de moléculas pequeñas de RNA. Se desarrolla aquí tan solo una clasificación y breve descripción, pues la función de cada uno debe entenderse en el contexto de cada proceso que forma parte de la expresión génica (control de la transcripción del DNA en un RNA, procesamiento de los RNA transcritos primarios para dar los RNA maduros, síntesis de proteínas, control de la traducción del mRNA para sintetizar proteínas), procesos abordados en capítulos posteriores.

6.2.1 RNA como material genético

6.2.1.1 Depositario de la información genética: genoma de RNA

Algunas especies de virus poseen un genoma DNA (al igual que todos los procariotas y eucariotas), mientras que otras tienen un genoma constituido por RNA. De éstas, algunas emplean un RNA monocatenario y otras uno bicatenario. Esto incluye dos categorías de virus: los **retrovirus** (cuyo genoma es RNA pero necesita retrotranscribirse en un DNA para que el virus pueda propagarse) y los **ribovirus o virus RNA**, que no utilizan DNA intermedio en su propagación.

En todos estos casos, el papel del RNA ya no es el de una molécula auxiliar, sino que es el propio material genético de la especie, su genoma, la molécula depositaria de la información genética que proporciona al virus su identidad como especie y le permite reproducirse, multiplicarse.

Dado el enfoque de este texto, no se desarrollará en más detalle la biología molecular de este tipo de virus; el lector interesado debe acudir a textos de virología. Aun así, hay que destacar que el conocimiento de estos virus es relevante para las ciencias de la salud debido a su implicación en infecciones y enfermedades y la necesidad de combatirlas, habitualmente interfiriendo en el mecanismo de propagación del virus. Como ejemplos, este tipo de virus causan la gripe, la hepatitis C, el síndrome respiratorio agudo grave (SARS, neumonía coronavírica o neumonía atípica asiática), el sida, la hepatitis B, la leucemia de células T y el linfoma de células T.

6.2.1.2 Portador transitorio de la información genética

En la gran mayoría de organismos hay moléculas de RNA (el mensajero, mRNA) que actúan como una copia temporal de la información contenida en el DNA genómico y se utilizan como fuente de información, como plantilla, para la síntesis de proteínas.

6.2.2 Síntesis de proteínas

El RNA también participa en la síntesis de proteínas de otras formas, sin aportar un mensaje genético. (Esta función es conceptualmente independiente del papel citado del RNA mensajero.) Éste es el caso de los RNA transferentes y ribosómicos, que intervienen en dicho proceso como moléculas accesorias (aunque imprescindibles). Como se ha indicado –y se estudiará con más detalle en capítulos posteriores (20 y 21)–, los tRNA ejercen de **portadores** que llevan los aminoácidos hasta el ribosoma para que reaccionen de manera ordenada formando el polipéptido y de **activadores químicos** para que esa reacción sea termodinámicamente favorable. Por su parte, los rRNA son constituyentes estructurales del ribosoma, responsable de todo el proceso de síntesis de las proteínas.

En segundo lugar, una molécula concreta, el más grande de los RNA ribosómicos (23S en procariotas, 28S en eucariotas) es la **transpeptidasa**, la enzima que cataliza la unión de los aminoácidos en enlace peptídico para formar las proteínas (v. a continuación la descripción de ribozimas).

6.2.3 Maduración del RNA

Se denomina maduración o modificación postranscripcional (v. capítulo 19) al conjunto de transformaciones que experimenta un RNA recién sintetizado (transcrito primario) hasta convertirse en un RNA funcional (RNA maduro). En este proceso intervienen también –no exclusivamente– algunas moléculas de RNA. Por ejemplo, la RNasa P es una ribozima que cataliza una de las reacciones de hidrólisis en la maduración de los tRNA procarióticos. Por otro lado, el ayuste (pág. 303), parte de la maduración de los mRNA eucarióticos, depende de ribonucleoproteínas que incluyen algunos RNA nucleares pequeños (snRNA).

6.2.4 Regulación de la expresión de los genes

Como se ha indicado (pág. 9), los microRNA, los siRNA y los ribointerruptores, entre otros, ejercen una función reguladora sobre diversas etapas de la expresión de algunos genes, bien bloqueando el fin de la transcripción o, más comúnmente, provocando la degradación del mRNA o bloqueando su uso en la traducción. La mayor parte de estos mecanismos forman parte del proceso conocido como ribointerferencia, o también silenciamiento génico postranscripcional. Todo ello tiene lugar de forma selectiva sobre un gen determinado en cada caso.

6.3 RNA CATALÍTICO: RIBOZIMAS

Algunas moléculas de RNA son capaces de actuar como catalizadores de reacciones bioquímicas; de ahí que el concepto clásico de enzima haya sido ampliado más allá de las proteínas. A estos RNA con actividad catalítica, es decir, con capacidad de disminuir la energía de activación de determinadas reacciones, se les llama ribozimas, para distinguirlos de las enzimas proteicas. Su capacidad catalítica estriba también en la gran complejidad y diversidad de estructura tridimensional que pueden adoptar los RNA, lo que les permite adaptarse al sustrato, así como en su reactividad gracias a los grupos funcionales de bases y esqueleto. De hecho, se han podido realizar estudios cinéticos con algunas ribozimas, verificando que cumplen la cinética de Michaelis y Menten, típica de enzimas proteicas. Se cree que las ribozimas fueron los primeros biocatalizadores que aparecieron en la evolución, siendo desplazados posteriormente por las enzimas proteicas, más versátiles. Los RNA serían, así, las moléculas primordiales de la vida, antes de que evolucionasen los DNA como portadores de información genética y las proteínas como catalizadores.

A continuación se describen algunos ejemplos de ribozimas, que actúan en diversos tipos de reacciones de interés en biología molecular o en las técnicas de ingeniería genética.

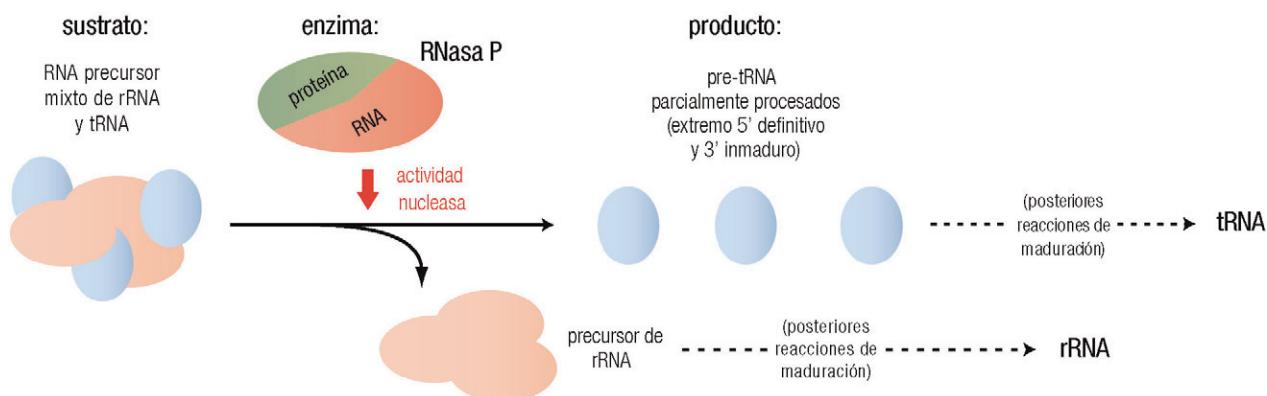


Web 6.6. Estructura tridimensional de algunas ribozimas y reconocimiento de su sustrato.

6.3.1 Acción catalítica en la maduración de RNA

El primer ejemplo conocido (1981) de molécula de RNA con actividad catalítica fue la RNasa P (se ha estudiado principalmente en procariotas, pero existe una actividad similar en eucariotas). Esta enzima está formada por moléculas de RNA y de proteína (una en la RNasa P bacteriana, 9 o 10 en la del núcleo eucariótico), pero se ha comprobado que la actividad catalítica reside en el componente RNA. Reconoce como sustrato al precursor mixto de rRNA y tRNA en procariotas (un transcrito primario) y cataliza una de las reacciones de su maduración postranscripcional (que genera los rRNA y tRNA funcionales), en concreto la generación de los extremos 5' de los tRNA (pág. 307).

6.8

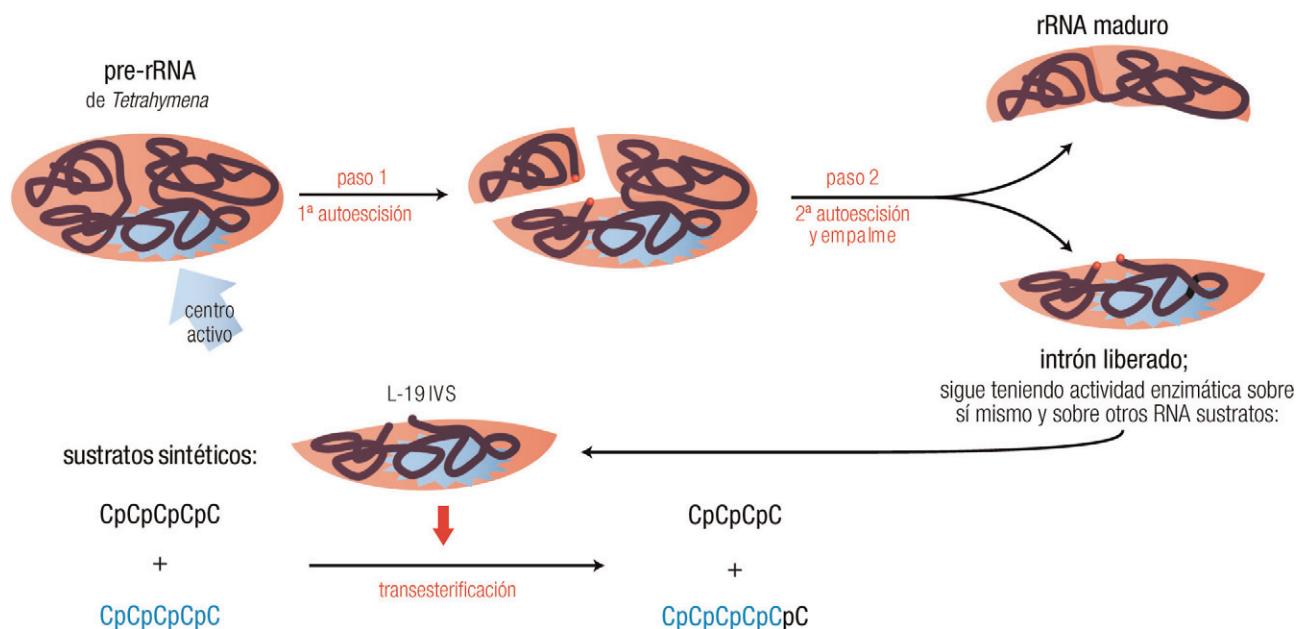


6.3.2 Acción autocatalítica en la eliminación de intrones

La siguiente “actividad ribozímática” encontrada fue la del precursor de rRNA en el protozoario *Tetrahymena* (un eucariota unicelular). A diferencia del caso anterior, esta ribozima actúa de forma autocatalítica, es decir, el sustrato es su propia molécula. Mediante un mecanismo similar actúan como ribozimas otros pre-RNA procarióticos y eucarióticos, aquellos que poseen **intrones de los tipos I y II**. Los intrones (pág. 264) son partes existentes en el RNA precursor, pero no en el RNA maduro, pues se eliminan durante la maduración; el pre-RNA comprende, pues, las regiones del RNA maduro y las de los intrones. Los tipos de intrón y los procesos de maduración del RNA precursor por corte y empalme se estudian como parte del procesamiento postranscripcional del RNA (pág. 303).

En este caso es cuestionable hablar de catálisis, pues uno de los requisitos de la definición de catalizador es que éste no se consuma en el proceso. Sin embargo, el producto final de una cadena de reacciones de autohidrólisis del intrón de *Tetrahymena* es una molécula (derivada del propio intrón y denominada L-19 IVS) que es capaz de catalizar reacciones de hidrólisis y de esterificación de oligonucleótidos externos a ella. Por tanto, si bien no lo son el pre-RNA o el intrón originales, L-19 IVS sí es un verdadero biocatalizador.

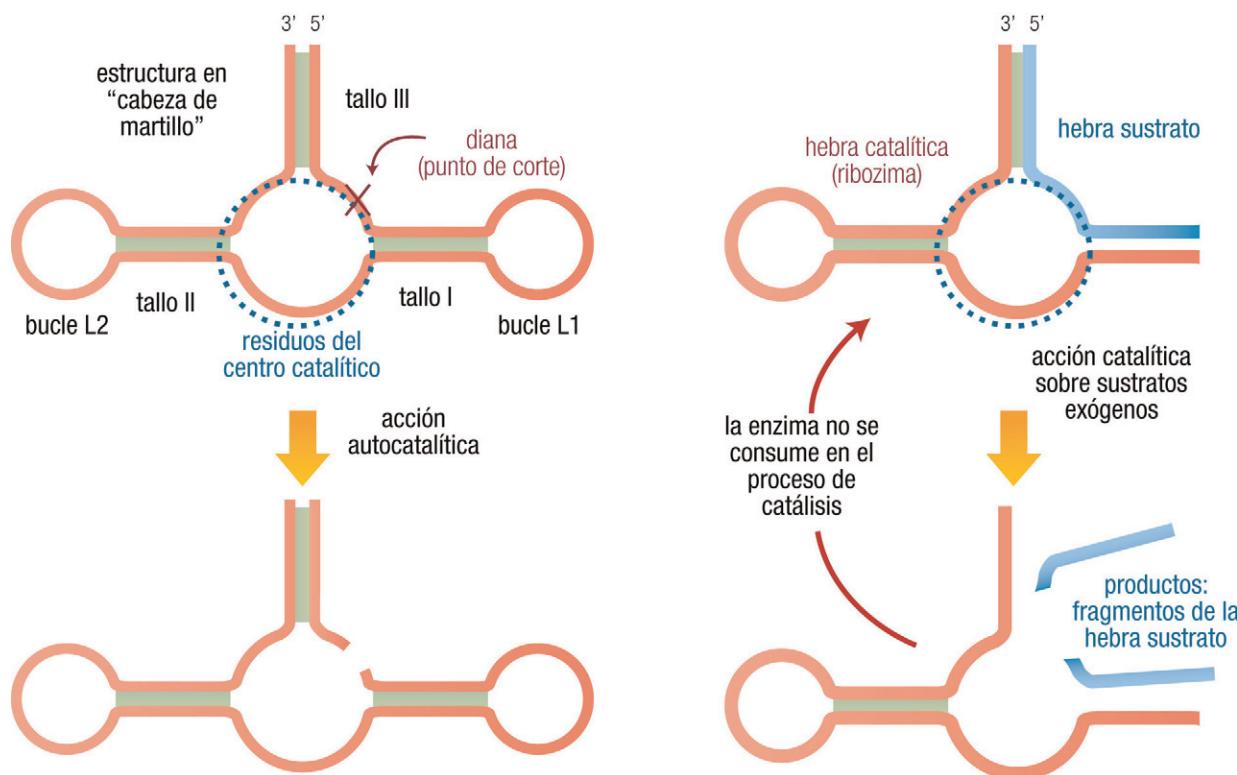
6.9



6.3.3 Acción catalítica y autocatalítica en viroides: ribozimas de posible aplicación terapéutica

Los *viroides* son agentes infecciosos formados por una molécula circular de RNA monocatenario (de 250 a 400 nt), sin cubierta proteica, con capacidad para infectar plantas. No codifican proteína alguna, simplemente portan señales para replicarse a costa de la maquinaria molecular de las células infectadas. Poseen actividad como ribozimas, de modo que escinden uno de sus propios enlaces fosfodiéster, como parte de su mecanismo de replicación. Se ha descubierto –y aquí radica su interés aplicado– que son también capaces de actuar sobre otras moléculas de RNA.

La estructura secundaria de estas ribozimas les ha merecido el nombre de “cabeza de martillo” (*hammerhead*). La molécula posee dos características esenciales: por un lado, las secuencias de reconocimiento del RNA sustrato (mediante emparejamiento de bases complementarias) y, por otro, el centro catalítico, que produce la escisión.

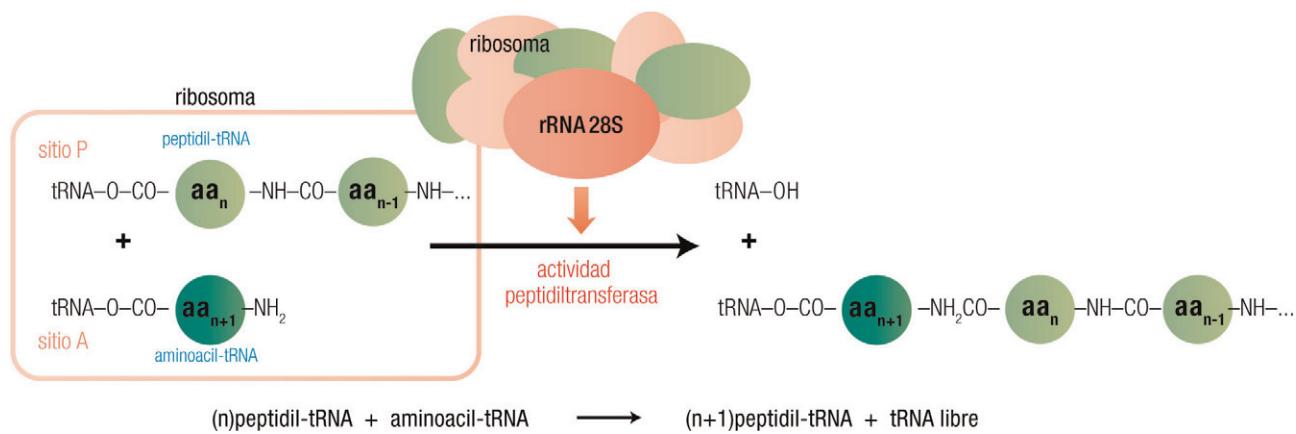


6.10

Se han ensayado análogos sintéticos de alguna de estas *ribozimas en cabeza de martillo* (obtenidos mediante ingeniería genética) con el objeto de destruir de forma selectiva moléculas de RNA con secuencias determinadas. Éste es un planteamiento de **terapia génica**, dentro de la estrategia de “inhibición dirigida de la expresión génica”. Se trata de preparar ribozimas artificiales con secuencias de reconocimiento que les permitan unirse en orientación “antisentido”, antiparalela, a un mRNA determinado, de forma que el centro catalítico lo escinda, suprimiendo en consecuencia la síntesis de la proteína correspondiente. (El concepto de cadena *antisentido* se introducirá en el capítulo de transcripción, pág. 267) Otro tipo de planteamiento de terapia, dentro de la estrategia de “corrección dirigida de mutaciones”, consiste en el corte mediante ribozimas de un transcripto precursor de RNA cuya secuencia estuviese mutada (por proceder de un gen mutado), de forma que la escisión ocurra alrededor de la mutación y resulte un transcripto corregido.

6.3.4 Acción catalítica transpeptidasa en la síntesis de proteínas

El cuarto ejemplo –y enormemente importante para todos los organismos– lo constituye la ribozima que ejerce la actividad **transpeptidasa** o **peptidiltransferasa**, responsable de la formación del enlace peptídico en la biosíntesis proteica, entre los residuos aminoácidos situados en los sitios A y P del ribosoma (pág. 339). En este caso, la “acción ribozimática” radica en el rRNA 23S, uno de los componentes de la subunidad mayor (50S) del ribosoma procariótico, y en el rRNA homólogo de eucariotas (rRNA 28S de la subunidad 60S).



6.11

Fisiológicamente, la RNasa P y la peptidiltransferasa ribosómica son posiblemente los que mejor reflejan el parecido de las ribozimas con las enzimas proteicas, a saber: actúan sobre un sustrato específico, aceleran la reacción varios órdenes de magnitud, no se destruyen durante la catálisis y al parecer se comportan de acuerdo con la cinética michaeliana.

6.4 LOS RIBOSOMAS

Estructuralmente, los ribosomas son **asociaciones supramoleculares de proteínas y rRNA** (35 y 65% en peso, respectivamente). Es decir, son nucleoproteídos. (Otros ejemplos de nucleoproteídos son el nucleosoma –pág. 84– y los virus, que se abordan en una materia propia, la virología.) Los ribosomas se pueden considerar también partículas subcelulares (con un tamaño de 20 a 23 nm) y orgánulos, aunque no delimitados por una membrana.

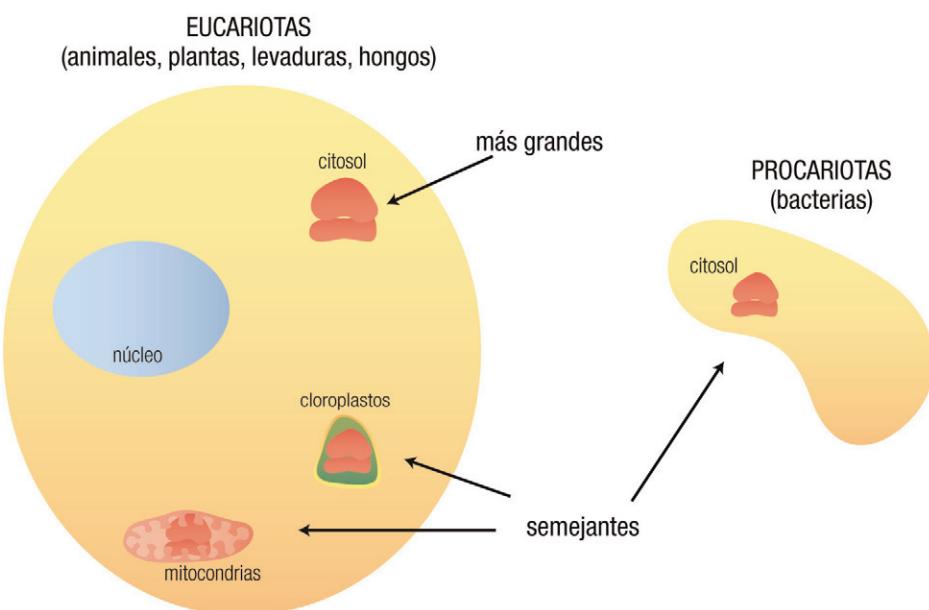
Los ribosomas poseen una importancia crucial para la transmisión de la información genética, pues es en ellos donde se realiza la síntesis de proteínas. Se los ha calificado de estructuras químico-mecánicas pues, por un lado, se desplazan sobre el mRNA ordenando la interacción de sus codones con los correspondientes anticodones de los tRNA y, por otro, proporcionan el entorno necesario para que los aminoácidos formen los enlaces peptídicos del polipéptido en crecimiento.

La función de los ribosomas se estudiará en el Capítulo 21, que describe la síntesis de proteínas. Aquí se aborda sólo su descripción estructural.

6.4.1 Ubicación subcelular

En cada célula existen varios miles de ribosomas (por ejemplo, más de 15.000 en un *E. coli*), que llegan a suponer un cuarto del peso seco de la célula. Se encuentran en el **citosol**, así como en la **matriz mitocondrial** y en el **estroma de cloroplastos**. Pueden estar en forma libre, como partículas discretas, pero a menudo se observan en forma de hilera, agrupados sobre un mRNA; en este caso se denominan **polisomas** o **polirribosomas**; su grado de agrupación o polimerización depende de la actividad metabólica celular (si la síntesis de proteínas es muy activa, habrá mayor número de ribosomas asociados a cada mensajero).

Los ribosomas del citosol de eucariotas son más complejos en composición y mucho mayores en tamaño que los ribosomas bacterianos. En contraste, los ribosomas de mitocondrias y cloroplastos eucarióticos son semejantes a los bacterianos. Evolutivamente para todos los tipos se ha conservado una misma estructura tridimensional, pero no la secuencia que la produce.



6.12

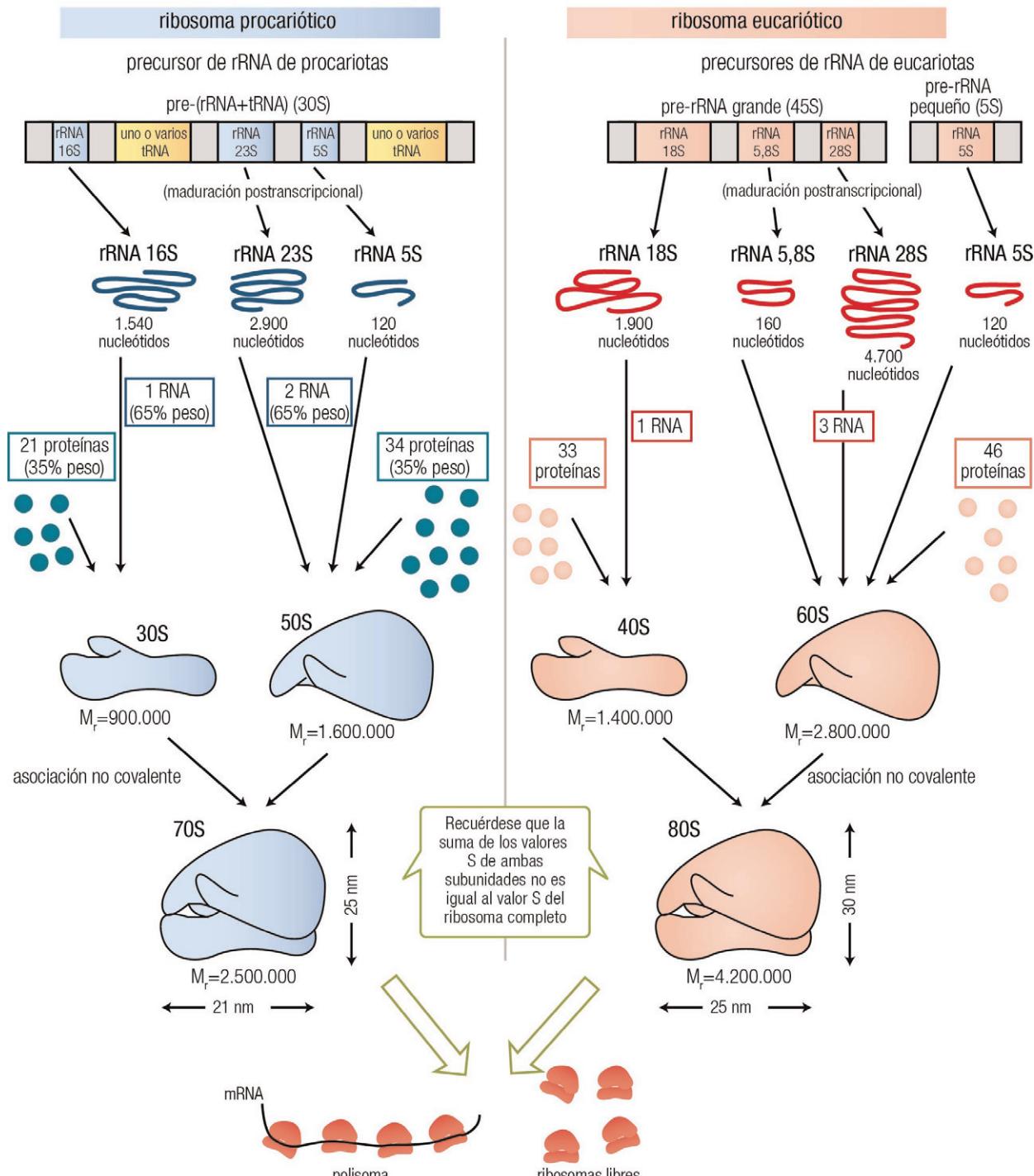
6.4.2 Origen

Los precursores del rRNA muestran una gran diversidad en su secuencia, que se manifiesta en la aparición (durante el proceso de maduración postranscripcional, capítulo 19) de distintos tipos de rRNA, caracterizados especialmente por su diferente coeficiente de sedimentación, *s* (pág. 63 y 67) (v. web 10.2).

6.4.3 Composición

Durante su actividad en las células, los ribosomas se separan de forma reversible en dos partes, que se denominan **subunidades**. Éstas tienen tamaño diferente, ambas una forma irregular, y cada una está constituida por una o varias moléculas de rRNA y numerosas proteínas de diversa masa molecular (entre 6 y 75 kDa). Tanto el ribosoma completo como sus subunidades y los rRNA componentes se nombran de acuerdo con su coeficiente de sedimentación. Las proteínas, por su parte, se designan mediante números precedidos por L o S, según pertenezcan a la subunidad grande o pequeña, respectivamente (*large/small*).

Diferencias en origen, composición y estructura de los ribosomas procarióticos y eucarióticos



6.13

Como se ha comentado, los ribosomas de orgánulos son más parecidos a los de bacterias que a los citosólicos eucarióticos. Por ejemplo, el ribosoma mitocondrial, con 55S, está formado por una subunidad mayor 39S con rRNA 16S (1560 nt) y 48 proteínas, más una subunidad menor 28S con rRNA 12S (950 nt) y 29 proteínas.

6.4.4 Estructura

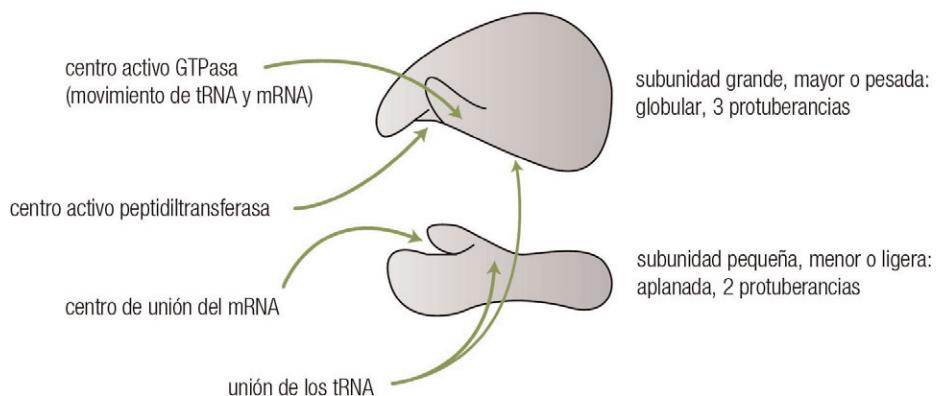
La estructura de los ribosomas, igual en todas las células, es muy característica; se ha dilucidado empleando difracción de rayos X, difracción de neutrones, microscopía electrónica en combinación con métodos inmunológicos, etc. Debido a su enorme tamaño (MDa) y la complejidad de su asociación interna, el estudio de la estructura, dinámica y función de los ribosomas constituyó un importante reto investigador.



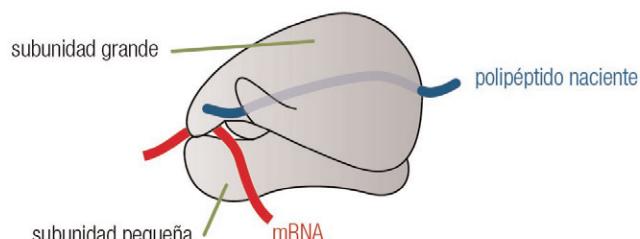
Web 6.7. Estructura tridimensional del ribosoma.

Las dos subunidades se asocian mediante interacciones no covalentes, de una forma peculiar: la subunidad pequeña cubre una oquedad de la grande, dejando entre ambas espacio para acomodar 3 moléculas de tRNA y también para el paso del mRNA a medida que el ribosoma se desplaza durante el proceso de traducción. Además, la subunidad mayor está atravesada por un túnel por el cual emerge la cadena polipeptídica que se está sintetizando.

Estructura en dos subunidades de los ribosomas y
situación aproximada de sus centros activos

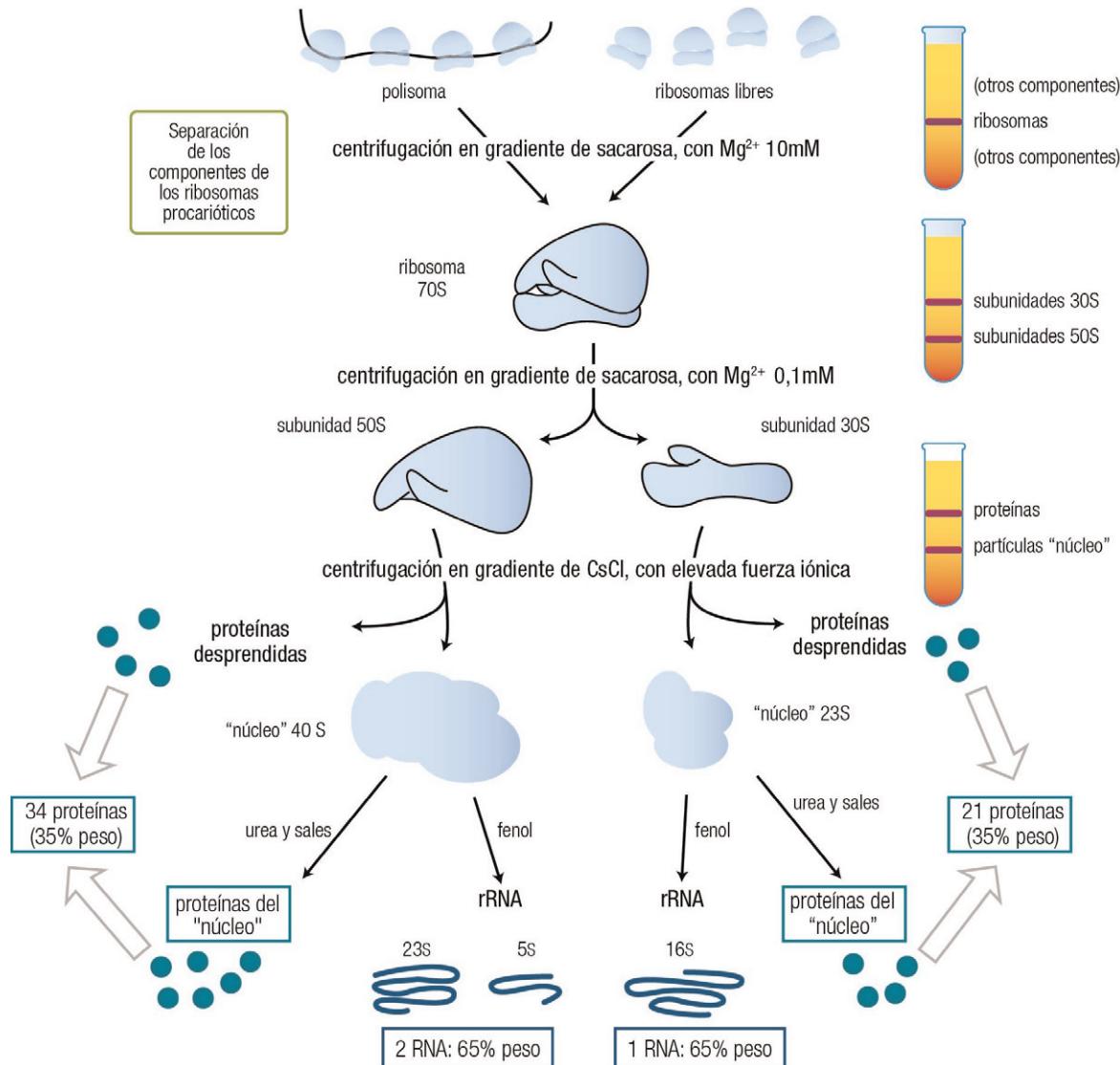


Disposición sobre el ribosoma del mRNA y
de la cadena polipeptídica naciente



6.4.5 Separación de los ribosomas, subunidades y componentes

Tanto los ribosomas individuales como sus subunidades y sus componentes (varios RNA y diferentes proteínas) se pueden separar mediante centrifugación en gradientes (pág. 126) de sacarosa o de cloruro de cesio, en presencia de concentraciones definidas de Mg^{2+} . Por otro lado, si se mezclan los componentes (rRNA y proteínas) a pH y fuerza iónica correctos puede observarse su asociación espontánea en ribosomas funcionales.



Condensación del DNA y cromosomas



7.1 SUPERENROLLAMIENTO DEL DNA	79		
7.1.1 Superenrollamiento del DNA circular	80	7.2.2.1 Disposición en nucleosomas y fibra de 10 nm	86
7.1.1.1 Conformación relajada	80	7.2.2.2 Formación de la fibra de 30 nm	87
7.1.1.2 Conformaciones superenrolladas	80	7.2.2.3 Cromatina o cromosoma interfásico	87
7.1.1.3 Demostración experimental del superenrollamiento	81	7.2.2.4 Cromosoma metafásico	89
7.1.2 Análisis teórico del superenrollamiento	82		
7.1.3 Superenrollamiento del DNA lineal	83		
7.2 CONDENSACIÓN DEL DNA EN EUCA RIOTAS	83	7.3 EL CROMOSOMA METAFÁSICO	89
7.2.1 Proteínas componentes de la cromatina	84	7.3.1 Aspectos citogenéticos	89
7.2.1.1 Histonas	84	7.3.1.1 Morfología de los cromosomas	89
7.2.1.2 Proteínas no histónicas	84	7.3.1.2 Cariotipo y su análisis	90
7.2.2 Niveles de condensación del DNA nuclear eucariótico	85	7.3.1.3 Bandeo	92
		7.3.2 Regiones con significado funcional	93
		7.3.2.1 Centrómero	93
		7.3.2.2 Telómeros	94

Los ácidos nucleicos no se encuentran en las células en las formas extendidas correspondientes a su estructura primaria y secundaria, sino molecularmente más compactados, lo que hemos denominado (pág. 28) nivel estructural “de orden superior”. En el caso del DNA, parte de esta compactación o *condensación* viene representada por el denominado *superenrollamiento*, o retorcimiento de la cadena sobre sí misma. La condensación se completa mediante la asociación estrecha con proteínas que inducen el plegamiento de la doble hélice del DNA. La contribución de esta asociación con proteínas posee una relevancia especial en el caso del DNA nuclear eucariótico.

7.1 SUPERENROLLAMIENTO DEL DNA

El superenrollamiento, como parte de la condensación del DNA en general, posee un especial significado para entender la organización estructural del cromosoma. En concreto, contribuye a explicar cómo un DNA de tan gran magnitud puede alojarse en el interior de la célula (procariotas) o del núcleo (eucariotas), cuyas dimensiones son muy inferiores a la longitud teórica del DNA en doble hélice. Por ejemplo:

DNA	Tamaño	Longitud en forma de B-DNA	Dimensiones en las que se compacta
<i>E. coli</i>	$4,7 \times 10^6$ pb	1,6 mm = 1.600 μm	2 μm (célula)
Cada cromosoma humano	$50 - 250 \times 10^6$ pb	17-85 mm	4-6 μm (núcleo)
Genoma humano diploide	$6,4 \times 10^9$ pb	2,1 m	



superenrollamiento:
la doble hélice
forma una
superhélice, en dos
sentidos opuestos



Se entiende por **superenrollamiento** (enrollar algo que ya está enrollado) del DNA el retorcimiento o giro sobre sí misma de la doble hélice (que ya lleva implícito el **enrollamiento mutuo** de dos hebras), de forma que el eje de la doble hélice no sigue una línea recta, sino otra hélice (una “superhélice”). Su origen, como se verá a continuación, es la introducción de una tensión estructural en la propia molécula. Aparece de algún modo en todos los DNA celulares y se encuentra estrechamente regulado *in vivo* por la acción de las *topoisomerasas* (pág. 155). En resumen, las dos hebras del DNA sufren un enrollamiento, mientras que la doble hélice sufre un superenrollamiento.

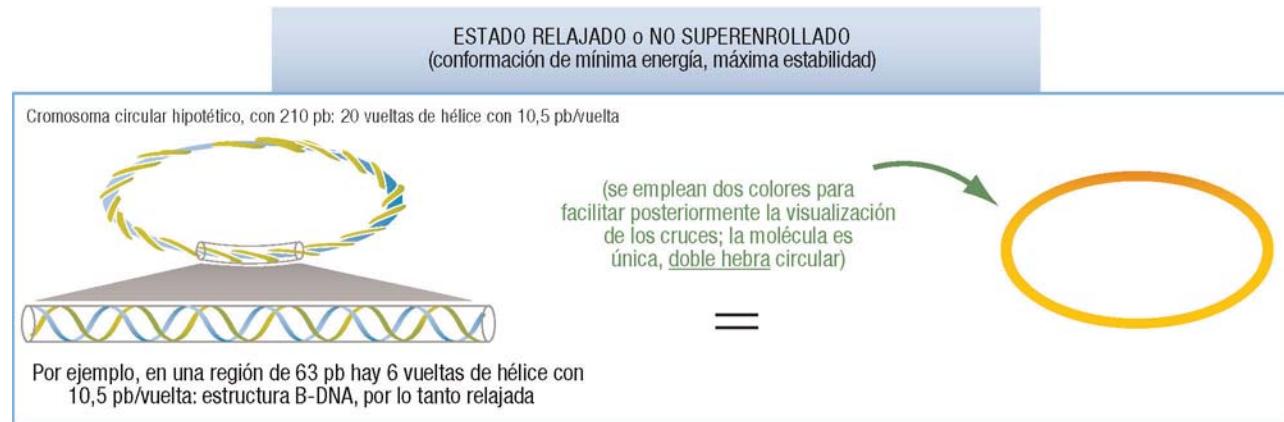
Un ejemplo intuitivo del concepto de superenrollamiento es el cordón del teléfono, que de por sí presenta un enrollamiento helicoidal y, con el uso cotidiano, termina por adoptar un enrollamiento adicional (cordón retorcido). Este símil se utilizó para explicar muchas propiedades del DNA vírico (circular y pequeño). Un modelo más próximo a la estructura del DNA es una cuerda formada por dos hebras arrolladas entre sí. Aunque la cuerda tiende a mantener ese arrollamiento, puede ser forzada manualmente a un mayor o menor número de vueltas entre sus hebras (torsión), con lo que se genera una tensión que se libera mediante la formación de superenrollamientos (v. web 7.1).

7.1.1 Superenrollamiento del DNA circular

Aunque afecta tanto a procariotas como a eucariotas, el superenrollamiento es más conocido y de comprensión más sencilla en los primeros, por lo que así se abordará a continuación. Estas consideraciones también son válidas para el DNA de plásmidos procarióticos y el de cromosomas mitocondriales y cloroplásticos de eucariotas, dado que en todos estos casos la molécula de DNA muestra una estructura *circular* o cerrada (una doble hélice con sus extremos unidos entre sí por enlace covalente), en la que es fácil apreciar los fenómenos de superenrollamiento.

7.1.1.1 Conformación relajada

La conformación adoptada por el DNA en su forma B (aproximadamente 10,5 pb/vuelta) corresponde a un estado estable, de mínima energía, que se considera **no superenrollado** o **relajado** cuando el eje de la doble hélice se puede disponer por completo sobre un plano.

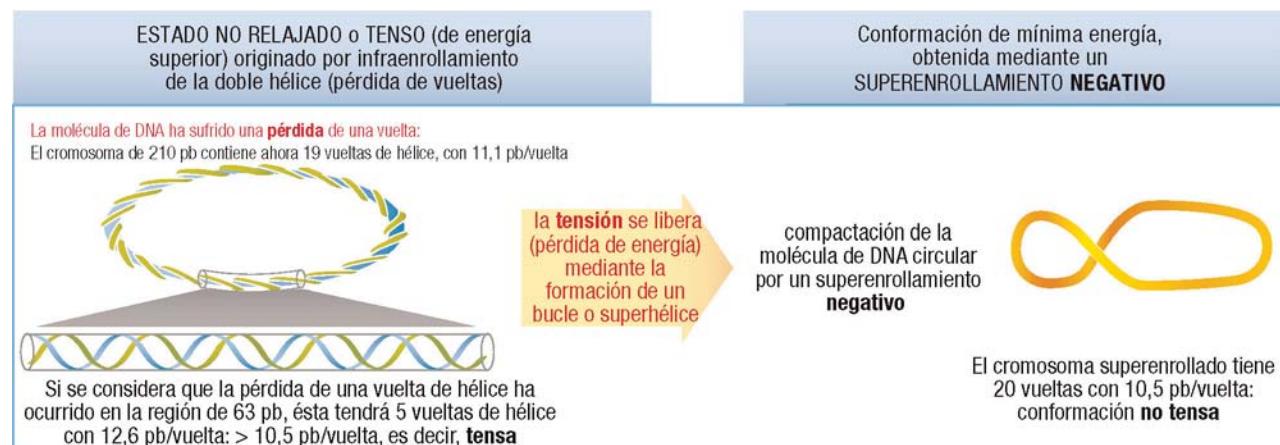


7.1.1.2 Conformaciones superenrolladas

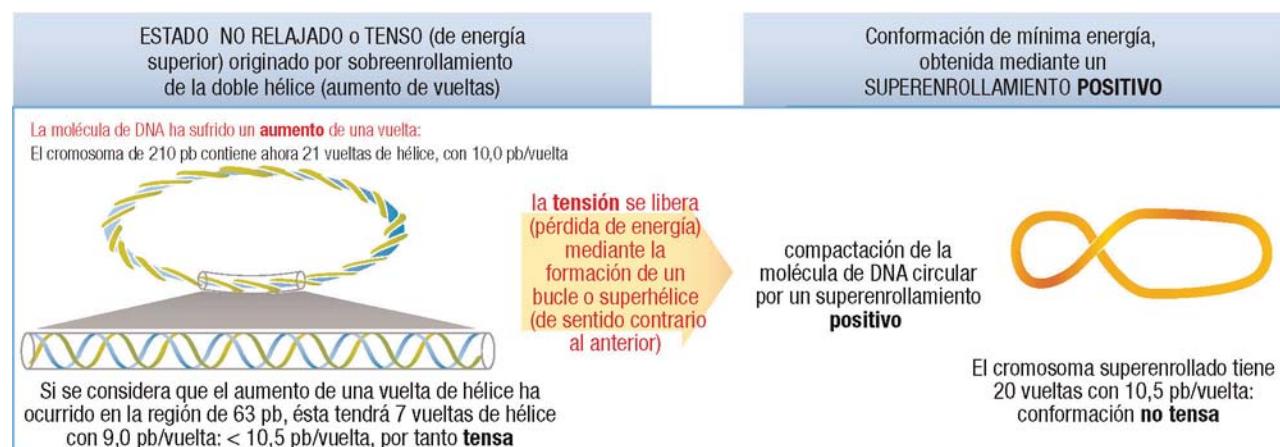
En las células, el DNA generalmente no se encuentra en la conformación relajada, sino en un estado **no relajado**, **superenrollado** y más compacto. Éste lo producen topoisomerasas que, mediante el corte transitorio de una o ambas hebras, introducen un aumento o una disminución en el número de vueltas de hélice (pág. 155).

En ambos casos se produce una tensión estructural, debida a un plegamiento local de las hebras y a una posición relativa de los nucleótidos diferentes de su geometría energéticamente más estable, la del B-DNA. Esta tensión se contrarresta en la molécula mediante la formación de un superenrollamiento: la cadena entera del DNA gira sobre sí misma de forma que las bases puedan disponerse de nuevo del modo más próximo a la conformación B-DNA.

En primer lugar, si se **reduce** el número de vueltas de hélice, la tensión se libera formando una superhélice a la que se le asigna un valor **negativo** de superenrollamiento.



Si, por el contrario, se **aumenta** el número de vueltas de hélice, la tensión se libera por formación de una superhélice de sentido opuesto, con un superenrollamiento **positivo**.



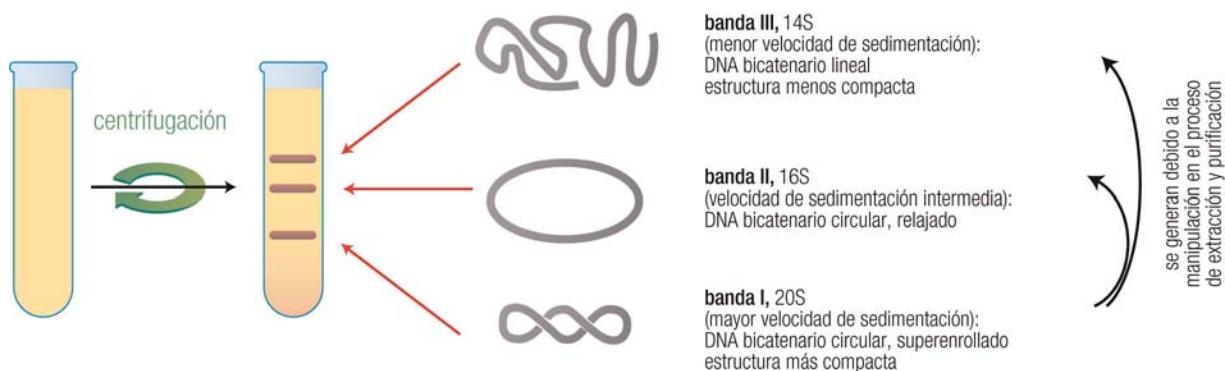
Web 7.1. Generación de superenrollamientos positivo y negativo.

7.1.1.3 Demostración experimental del superenrollamiento

De estas dos posibilidades, el DNA se encuentra habitualmente en las células en un estado de **superenrollamiento negativo**. El primer indicio experimental que llevó a proponer la hipótesis del superenrollamiento de las cadenas se obtuvo gracias a la aplicación de la técnica de *ultracentrifugación isopícnica* (basada en el equilibrio de sedimentación en un gradiente de densidad de cloruro de cesio, pág. 126).

Se observó que el DNA del *polioma*, un pequeño virus que causa cáncer a ratones, se separaba formando tres bandas en el gradiente. Las tres fracciones, sin embargo, tenían la misma masa molecular, por lo que la diferencia debía estar en su forma o estructura. Se pudo comprobar que las moléculas menos densas eran lineales, mientras que las demás eran circulares. Sólo tras la hipótesis del superenrollamiento se pudo explicar la diferencia entre esas dos fracciones más densas.

7.4

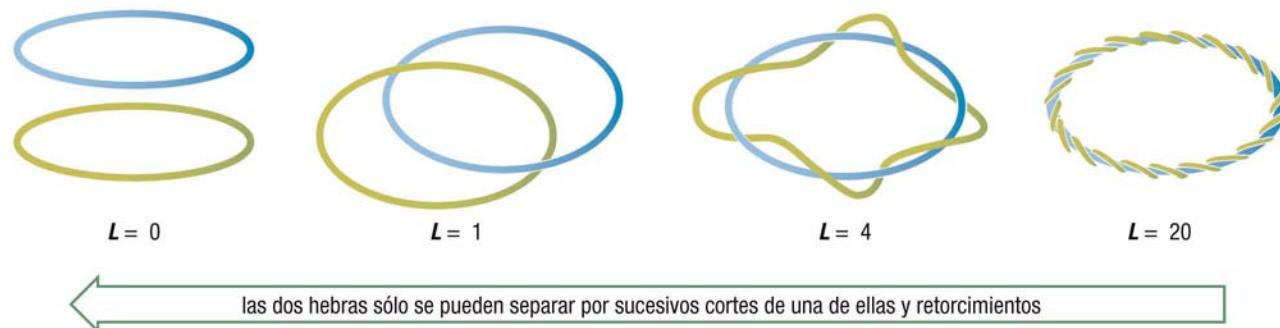


7.1.2 Análisis teórico del superenrollamiento

El empleo de la **topología** (rama de las matemáticas que estudia las propiedades de posición relativa de las partes de un objeto, propiedades que no cambian al ser sometido a deformación) ha permitido aclarar la estructura del DNA superenrollado. El signo y magnitud del superenrollamiento se asignan a partir de dicho estudio topológico, del cual sólo se comentan aquí los aspectos más fundamentales.

Posibles disposiciones teóricas de dos hebras (por ejemplo, dos cuerdas) con diferente grado de enrollamiento entre sí (diferente *número de enlace*). L = número de enlace = nº de veces que las hebras se cruzan.
(El modelo de la derecha, con $L = 20$, correspondería al DNA relajado de 210 pb considerado anteriormente)

7.5



Se define el **índice o número de enlace topológico** (L o Lk , de *linking number*) como el número de veces que las dos hebras se cruzan entre sí; más específicamente, el número de veces que una hebra cruza el plano definido por la otra (como anillo o curva cerrada). Se le asigna valor positivo cuando el cruce se produce girando hacia la derecha.

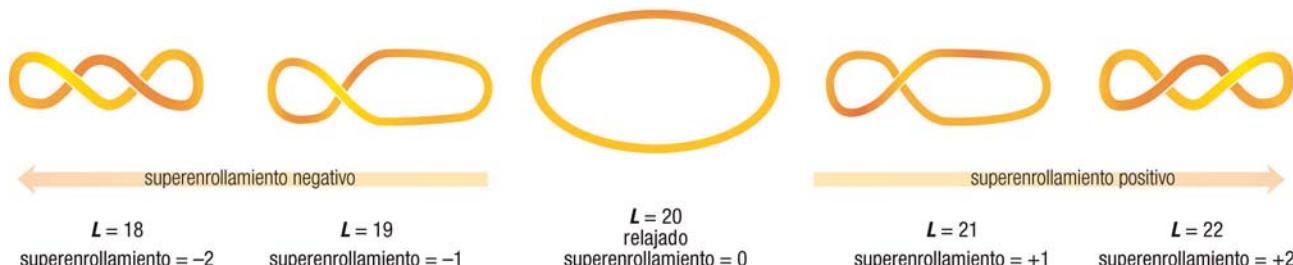
Se habla de **enlace topológico** porque cuando $L \neq 0$ las dos hebras están físicamente ligadas (es decir, no se pueden separar), a pesar de que no hay enlaces covalentes entre ellas. Sólo se podrían separar (y conseguir así $L = 0$) rompiendo de forma transitoria algún enlace covalente (es decir, cortando una hebra). Lo mismo es aplicable para cualquier cambio del valor L (por ejemplo, pasar de $L = 9$ a $L = 10$ requiere la ruptura de un enlace fosfodiéster). Por la misma razón, la deformación, plegamiento, etc., de las hebras, sin cortarlas, no alteran el número de enlace ni el número de superenrollamiento.

Para el caso del DNA, la conformación relajada posee, por tanto, un número de enlace que coincide con el número de vueltas de hélice, siempre positivo puesto que la doble hélice del DNA es dextrorsa. Como ya se ha estudiado, las situaciones de superenrollamiento surgen de cambios en el número de vueltas de hélice, es decir, en el número de enlace L .

Para discernir y cuantificar los estados superenrollados se refiere su valor de L al del estado relajado, definiendo un nuevo valor numérico de **superenrollamiento** como la diferencia (ΔL) entre el número de enlace de una conformación concreta del DNA y el número de enlace de su forma relajada. Este valor de superenrollamiento puede, lógicamente, ser positivo o negativo y coincide con el número de vueltas de superhélice formadas para que la molécula libere la tensión y alcance de nuevo localmente la conformación más estable de doble hélice B.

Para un DNA bicatenario circular de 210 pb de longitud, la estructura B tiene $210 \text{ pb} / 10,5 \text{ pb/vuelta} = 20$ vueltas. Por tanto, la conformación relajada de ese DNA sería la de $L=20$. Por encima y por debajo de $L=20$ se habla de un superenrollamiento positivo o negativo, respectivamente. (Todas las formas representadas son las de mínima energía para cada caso)

7.6



7.1.3 Superenrollamiento del DNA lineal

En eucariotas, el superenrollamiento del DNA es menos evidente y más difícil de estudiar debido a que, por la naturaleza lineal de las moléculas de DNA, las conformaciones superenrolladas no son estables por sí mismas. La libertad de giro de las cadenas con extremos libres permite que las superhélices se deshagan de manera espontánea, salvo que exista algún impedimento a su libre giro. Precisamente, es la unión de proteínas al DNA lineal lo que permite la estabilización de estados superenrollados, y debe considerarse éste como uno de los componentes que contribuyen a la condensación del genoma nuclear eucariótico (v. el ejemplo de las asas radiales en la cromatina, pág. 87).

Por otra parte, en los DNA eucarióticos se aprecia otra forma de superenrollamiento: la unión a proteínas permite el superenrollamiento *toroidal*, diferente al considerado hasta ahora, que se denomina *plectonémico*. El ejemplo principal del primero es el resultante del arrollamiento del DNA bicatenario en torno al octámero de histonas en la estructura de los nucleosomas (pág. 86).



Web 7.2. Superenrollamientos plectonémico y toroidal.

7.2 CONDENSACIÓN DEL DNA EN EUCA RIOTAS

Para completar la descripción de la estructura del DNA como material genético, y como requisito previo al estudio del genoma en el ciclo celular (Capítulo 8), es conveniente analizar el complejo problema de la condensación del DNA nuclear, que tiene lugar a lo largo de la interfase de dicho ciclo. Este aspecto, junto al proceso enzimático de replicación del DNA (Capítulo 11) durante la fase S, es de la mayor importancia desde el punto de vista básico y de sus aplicaciones.

Como ya se ha indicado, la *condensación* del DNA comprende dos aspectos, en teoría independientes pero funcionalmente relacionados y subordinados entre sí:

Condensación del DNA

- Superenrollamiento** propiamente dicho de la molécula de DNA bicatenario (doble hélice dextrorxa de tipo B).
 - Es análogo al superenrollamiento negativo ya descrito para procariotas; se conoce peor en eucariotas por su mayor complejidad, debida en parte a la gran magnitud del genoma.
- Empaquetamiento**, plegamiento o compactación de la molécula de DNA debido a la asociación con proteínas, para dar lugar a nucleoproteínas.
 - Es característico del genoma nuclear eucariótico.
 - Este aspecto apenas existe en procariotas y orgánulos de eucariotas, donde no hay histonas y la condensación se basa principalmente en el superenrollamiento de la doble hélice.

7.7

La condensación del DNA desde la doble hélice extendida hasta cromatina y luego cromosoma tiene lugar de forma gradual durante la fase G₂ del ciclo celular (Capítulo 8), mientras que el proceso inverso, la

descondensación del DNA, comienza después de la división celular (fase M) y continúa durante la fase G₁ hasta alcanzar de nuevo la condición difusa que permite la replicación (fase S). No se conocen con precisión los aspectos moleculares de ambos procesos, aunque se asume una gran analogía y se considera, por tanto, que ambos transcurren por los mismos mecanismos, pero en sentido inverso.

7.2.1 Proteínas componentes de la cromatina

7.2.1.1 Histonas

Las histonas son las principales proteínas que forman parte del material genético de eucariotas; su masa es equivalente a la del DNA. Su función fundamental es **estabilizar la estructura del DNA** y contribuyen de forma notable a compactarla. Las histonas constituyen una familia de proteínas semejantes, de tamaño relativamente pequeño y con un contenido muy elevado (casi 1/3 del total) de aminoácidos básicos (con pI de 9 a 10). Gracias a ello, muestran naturaleza poliacidónica a pH fisiológico (numerosas cargas positivas en cada molécula) y se asocian fuertemente, mediante interacciones electrostáticas, con los grupos fosfato del esqueleto polianiónico del DNA.

	% Lisina	% Arginina	% (Lys + Arg)	Nº de aminoácidos	Masa molecular
H1	24,8-29,5	1,3-2,6	27,4-30,8	215-244	21,1-23,0 kDa
H2A	10,9	9,3	20,2	129	14,0 kDa
H2B	16,0	6,4	22,4	125	13,8 kDa
H3	9,6	13,3	22,9	135	15,3 kDa
H4	10,8	13,7	24,5	102	11,2 kDa

Existen **cinco tipos** principales de histonas, bien caracterizadas, que se diferencian en su composición y, como se verá después, en su papel estructural en la organización del nucleosoma y la cromatina. La histona H1 se presenta en distintas especies o tejidos con mayor diversidad de secuencia y de tamaño, mientras que las otras cuatro están altamente conservadas. En algunos casos, se encuentran histonas alternativas; por ejemplo, H1 puede ser sustituida por una H5 que posee especial afinidad por la cromatina, lo que refuerza el empaquetamiento y la inactividad transcripcional.

El grado de condensación del DNA se regula en parte mediante la acetilación, fosforilación y metilación de las histonas (págs. 284 y 359-361), afectando así a la accesibilidad del DNA para la transcripción; de este modo las histonas se ven implicadas en uno de los mecanismos de control de la expresión génica.

7.2.1.2 Proteínas no histónicas

En la cromatina se encuentran, asimismo, otras proteínas, aunque menos abundantes, más variables entre tejidos y especies y en general peor caracterizadas:

a) Con función estructural

- Proteínas básicas:
 - **Protaminas:** Desempeñan la función equivalente a las histonas en tipos celulares concretos, particularmente el espermatozoide. Son proteínas pequeñas (50 aminoácidos, en algún caso) con elevada proporción de Arg, Ala y Ser, y estructura predominantemente en hélice α. Su menor tamaño permite la compactación del DNA en el espacio mínimo disponible en el espermatozoide.
- Proteínas ácidas:
 - **HMG:** Hasta un 5% de las proteínas nucleares son moléculas relativamente pequeñas que se incluyen en el grupo llamado HMG (*high mobility group*), proteínas de alta movilidad electroforética debido a su carácter ácido. De ellas, algunas se asocian al nucleosoma (al menos *in vitro*), mientras que otras interactúan con el DNA espaciador (págs. 86-87). Este grupo incluye también algunos factores de transcripción (Capítulo 18).
 - **Proteínas del esqueleto o matriz nuclear:** Esta estructura, que sirve de soporte o guía en el empaquetamiento de la cromatina (pág. 87), está formada por diversas proteínas, incluyendo algunas HMG y otras proteínas de carácter ácido.

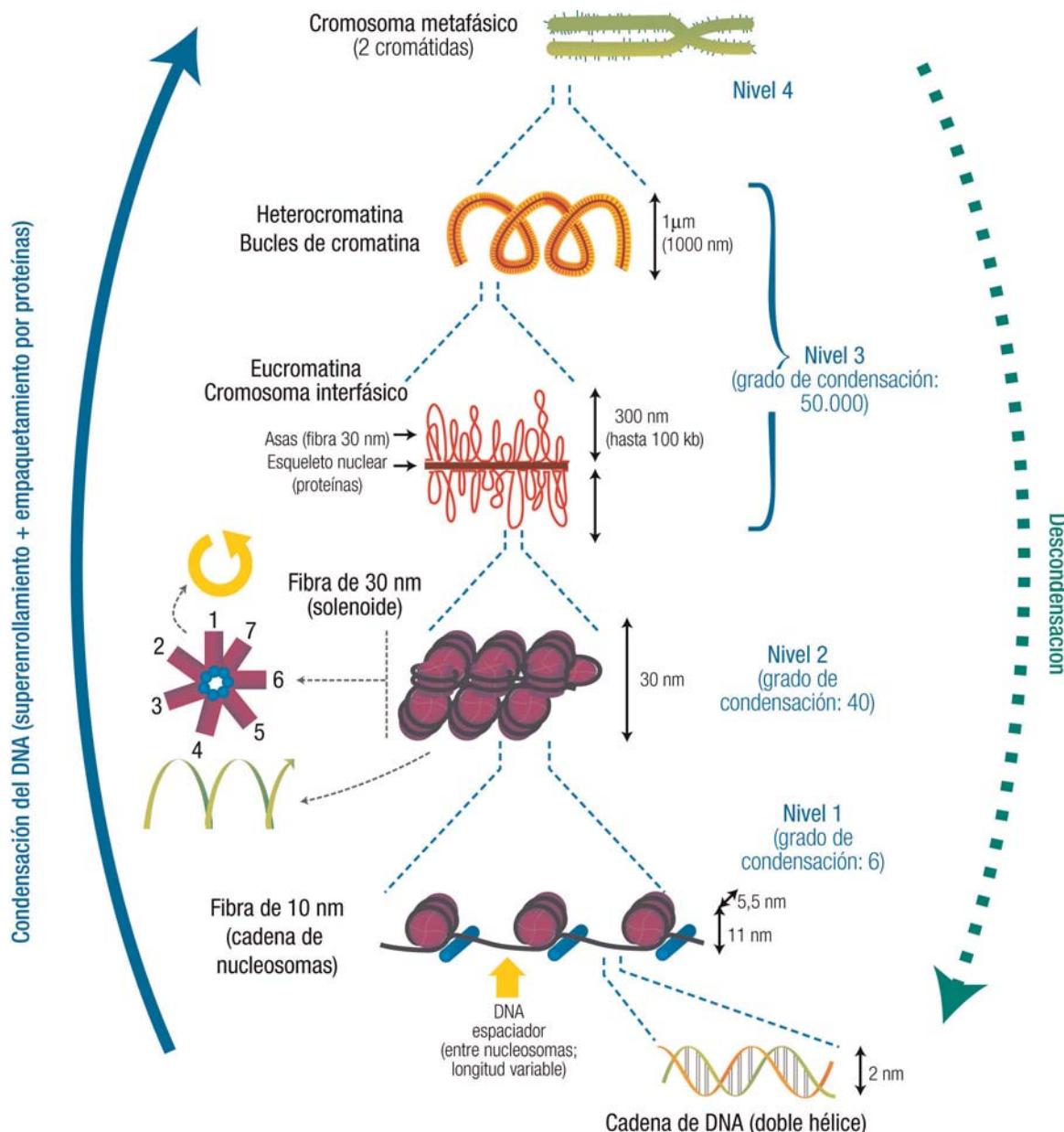
b) Con otras funciones

Asociadas a la cromatina, aunque en mínima cantidad y sin función estructural, se pueden encontrar numerosas proteínas, muy diferentes, que intervienen en la replicación y transcripción (Capítulos 11 y 17), así como en la regulación del grado de condensación de la cromatina: DNA- y RNA-polimerasas, sus proteínas auxiliares, factores de replicación, factores de transcripción, receptores de hormonas, topoisomerasas, etc.

7.2.2 Niveles de condensación del DNA nuclear eucariótico

Se puede resaltar el elevado grado de condensación que alcanza el DNA en las células recordando que la longitud de la molécula en doble hélice es varios órdenes de magnitud superior a las dimensiones de las estructuras celulares que la albergan (pág. 79). Se denomina *grado de condensación o de empaquetamiento* al cociente entre la longitud del B-DNA bícatenario y el tamaño del mismo una vez condensado. Este parámetro se utilizará para comparar los distintos niveles de compactación.

Para facilitar su comprensión, la condensación del DNA se describe aquí en cuatro niveles, aunque esta división es relativa, ya que no hay límites precisos y la comprensión con detalle de este fenómeno a nivel molecular es aún limitada. Es importante indicar que la condensación y la descondensación a través de esos niveles están estrechamente asociadas a la progresión por las distintas fases del ciclo celular (págs. 102-103).



7.2.2.1 Disposición en nucleosomas y fibra de 10 nm

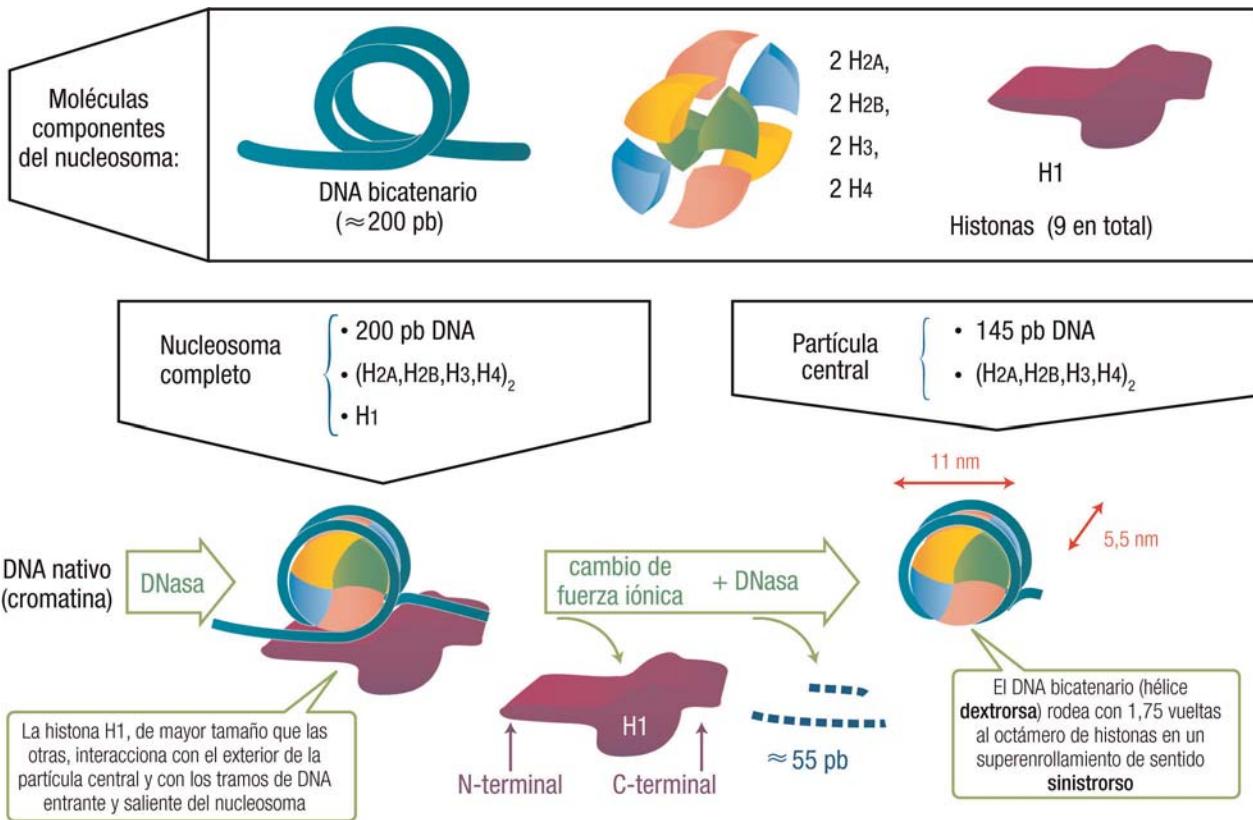
El nucleosoma es un complejo plurimolecular resultante de la asociación estrecha del DNA con las histonas; se trata, pues, de un ejemplo de nucleoproteido. En términos estructurales, se considera la unidad elemental constituyente de la cromatina y de los cromosomas, siendo así responsable de la organización estructural del material genético en todos los organismos eucarióticos. Tanto el nucleosoma como sus componentes y las fibras en las que se integra se han obtenido y analizado gracias a técnicas diversas, como digestión con DNAsas, entrecruzamiento covalente, hibridación, difracción de rayos X, microscopía electrónica, difracción de neutrones, dicroísmo circular y ultracentrifugación analítica.

La propuesta de la existencia del nucleosoma surgió del estudio de la acción de DNAsas: el tamaño de los fragmentos producidos a partir de un DNA nativo es siempre múltiplo de unos 200 pb, mientras que si previamente se le trata para retirar las histonas, los fragmentos de DNA producidos por la DNasa son de tamaños aleatorios. Esto sugirió que las histonas protegen regiones de DNA con una extensión próxima a 200 pb. Este fenómeno se pudo comprender al averiguar los detalles de la estructura molecular, descrita a continuación.

www

Web 7.3. Acción de DNAsas sobre el DNA genómico.

Cada nucleosoma está formado por 9 moléculas de histonas y un tramo de la molécula de DNA con aproximadamente 200 pb. En el centro se ubica un octámero de histonas, alrededor del cual se enrolla el DNA bicatenario, y más externamente se une la novena histona.



www

Web 7.4. Estructura del nucleosoma.

Bajo ciertas condiciones experimentales se puede liberar la histona H1, localizada externamente, lo que hace accesible a la DNasa una parte adicional del DNA. El nucleosoma queda así reducido a la llamada **partícula central**, formada por el octámero de histonas rodeado por DNA.

Una misma molécula de DNA envuelve sucesivamente a distintos octámeros de histonas, de modo que los nucleosomas están conectados entre sí. La porción presente entre dos nucleosomas sucesivos se denomina **DNA espaciador**, ligador o internucleosómico; su longitud puede variar entre 10 y 80 pb, dependiendo del organismo.

o incluso del tejido dentro de un mismo organismo. Esta estructura de nucleosomas espaciados a lo largo de la molécula de DNA se denomina **fibra básica** de cromatina o **fibra de 10 nm**, pues su grosor corresponde al diámetro del nucleosoma. El grado de empaquetamiento alcanzado es unas 6 veces superior al de la doble hélice del DNA extendida. La fibra de 10 nm también se denomina “estructura en cuentas de rosario”, por su aspecto característico al microscopio electrónico, que en general sólo se observa cuando se prepara *in vitro* a baja fuerza iónica (por ejemplo, 1 mM), mientras que dentro de la célula la cromatina se encuentra mayoritariamente en formas más condensadas, y su disposición en este estado relativamente extendido se limita, en todo caso, a pequeñas zonas.

También participan en la estructura otras proteínas no histónicas: HMG-14 y HMG-17 se asocian, al menos *in vitro*, con cada nucleosoma, mientras que HMG-1 y HMG-2 interaccionan con el DNA espaciador.

El arrollamiento sinistrorso del DNA sobre el octámero de histonas supone un superenrollamiento toroidal, de signo negativo, de la cadena. En contraste, en procariotas no existen nucleosomas y el superenrollamiento –plectonémico– se consigue por simple retorcimiento del DNA circular, sin una dependencia tan directa de la unión de proteínas.

7.2.2.2 Formación de la fibra de 30 nm

En las preparaciones de cromatina obtenidas *in vitro* a mayor fuerza iónica (por ejemplo, hasta 100 mM), en lugar de la fibra de 10 nm se observa una forma más condensada, denominada **fibra de 30 nm**, que supone un grado de empaquetamiento al menos 40 veces superior al de la doble hélice original. Aún no se ha podido confirmar su estructura exacta *in vivo*; uno de los primeros modelos propuestos consiste en el arrollamiento de la fibra de 10 nm sobre sí misma en una espiral o **solenoide**, de sentido sinistrorso, con algo más de 6 nucleosomas por cada vuelta y con la histona H1 situada en el interior del solenoide, acercando los nucleosomas entre sí; el DNA espaciador se flexiona para conectar los sucesivos nucleosomas a lo largo de la espiral. En otros modelos se propone que la cadena de DNA describe un zig-zag para que el DNA espaciador, que se mantiene rígido, conecte nucleosomas dispuestos en caras opuestas de la fibra. Se postulan también actualmente modelos en los que las vueltas de solenoide se intercalan consiguiendo un empaquetamiento aún más denso, y se cree que la disposición exacta puede ser variable, para acomodar espaciadores de distinta longitud mientras se mantienen constantes los puntos de contacto entre nucleosomas adyacentes en la espiral de la fibra de 30 nm. Estas formas más densas de la fibra podrían ya corresponder a los estados de compactación superiores que se describen someramente a continuación (eucromatina y heterocromatina). La fibra de 30 nm es probablemente la forma fisiológica más descondensada de la cromatina, adoptada por la mayor parte del DNA durante la interfase.



Web 7.5. Estructura de la fibra de 30 nm.

En todo caso, de forma experimental se ha confirmado que la histona H1 es un requisito necesario para la formación de la fibra de 30 nm, mientras que no lo es para la de 10 nm. Además, los extremos amino-terminales o “colas” de las histonas centrales, que sobresalen del octámero, parecen también desempeñar un papel en la organización de la fibra de 30 nm, estableciendo contactos con los tramos de DNA espaciador y con los nucleosomas vecinos. Dichas colas contienen precisamente los residuos que sufren acetilación, fosforilación y metilación como mecanismo de regulación de la condensación y, en consecuencia, de la actividad transcripcional de la cromatina (págs. 284 y 359-361).

7.2.2.3 Cromatina o cromosoma interfásico

Este nivel de condensación supone la presencia de lazos y superenrollamientos de la estructura anterior. Los **bucles o asas radiales**, formados por entre 50 y 100 kb de fibra de 30 nm, emergen de un armazón de proteínas no histónicas denominado **andamiaje, matriz o esqueleto nuclear**. Estos bucles sufren superenrollamiento, regulado por topoisomerasas asociadas al esqueleto. Parece que, además de contribuir al empaquetamiento del DNA, los bucles constituyen unidades de transcripción, habiéndose observado la presencia dentro de un solo bucle de un gen completo o bien de un grupo de genes regulados de forma conjunta. Es típica la observación de asas radiales en la interfase de cromosomas polítenicos gigantes de larvas de insectos y en los cromosomas plumosos meióticos de oocitos de anfibios.

A partir de esta estructura se forman enrollamientos superiores, que hacen que la cromatina posea durante la interfase un grado de condensación variable; se calcula que se multiplica el grado de condensación del DNA por 1.000 o 2.000. Esta diversidad local en la densidad de condensación se observa en forma de la presencia simultánea de **eucromatina** y **heterocromatina** en una célula en interfase.

	Heterocromatina	Eucromatina*
Aspecto al microscopio bajo tinción (definición original)	Se tiñe intensamente con el colorante May-Grünwald-Giemsa	Se tiñe débilmente y con aspecto difuso
Grado de condensación del DNA	Máxima condensación de la cromatina (próxima, aunque menor, a la del cromosoma metafásico)	Forma menos condensada
Contribución al total de la cromatina	Minoritaria; sólo algunas porciones del material genético se encuentran en esta forma:	Mayoritaria durante la interfase
	<p>Heterocromatina constitutiva: regiones del genoma que se encuentran permanentemente en forma de heterocromatina (durante toda la interfase) y en todas las células.</p> <p>Se descondensa el tiempo mínimo necesario para replicarse durante la fase S del ciclo celular.</p> <p>Incluye el <i>DNA satélite</i> (pág. 116), concentrado especialmente en las regiones teloméricas y centromérica de los cromosomas (págs. 93-94)</p>	
	<p>Heterocromatina facultativa: se encuentra como heterocromatina sólo en algunas células de un mismo organismo o en algunos individuos de una especie; puede pasar a eucromatina, en respuesta a determinadas señales.</p> <p>Uno de los ejemplos es la inactivación de uno de los 2 cromosomas X en la mujer, puesto que se puede inactivar tanto el X materno como el paterno. Uno permanece condensado como heterocromatina, mientras el otro está en forma de eucromatina</p>	
	Al final de la fase G ₂ , previa a la división por mitosis o meiosis, toda la cromatina de la célula adopta la forma de heterocromatina (antes de condensarse aún más para dar el cromosoma)	
Accesibilidad de la molécula de DNA para su interacción con proteínas (DNA polimerasas, RNA polimerasas, factores de transcripción, etc.; pág. 268)	No es accesible, debido a su elevada condensación. Se habla de cromatina <i>transcripcionalmente inactiva</i> . Se replica al final de la fase S. Los genes que contiene no se expresan	Es accesible: cromatina <i>transcripcionalmente activa</i> . Se replica al principio de la fase S. Sus genes se expresan

* El nombre “eucromatina” significa etimológicamente “cromatina verdadera”.

7.2.2.4 Cromosoma metafásico

Aunque al final de la profase cada cromosoma ya está condensado como heterocromatina, durante la metafase se condensa aún más, adquiriendo el aspecto típico de los **cromosomas metafásicos**, corpúsculos observables al microscopio óptico que poseen, en general, aspecto de bastoncillos con dos cromátidas más o menos separadas entre sí.

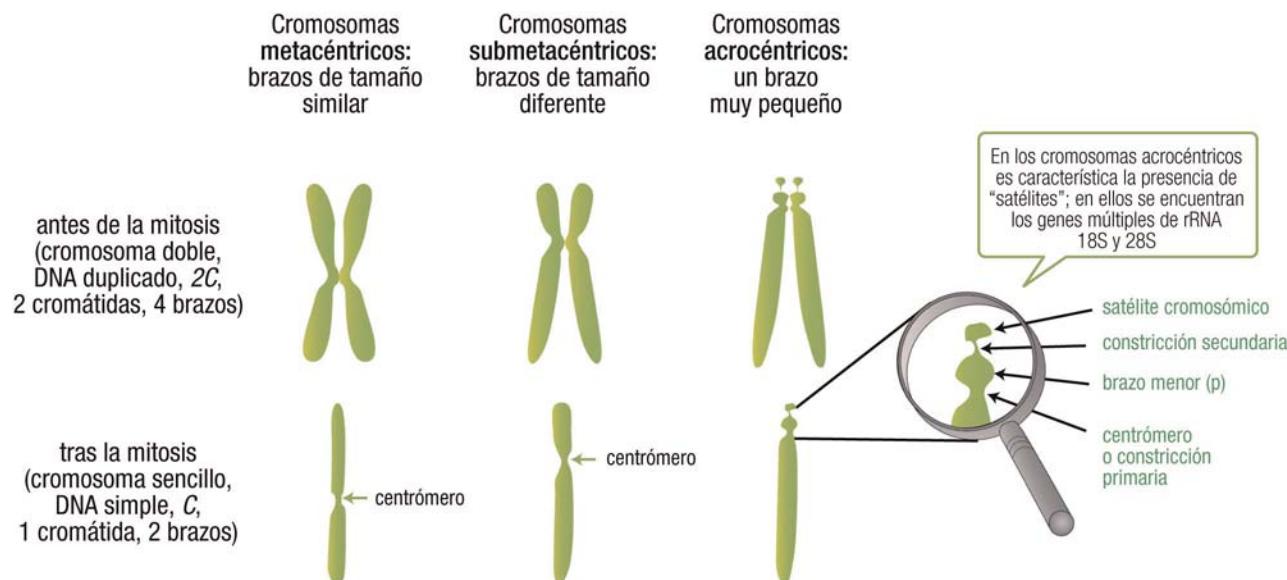
Las fibras de heterocromatina se arrollan en una primera espiralización para formar una **espiral menor** de paso muy pequeño, que se arrolla a su vez sobre sí misma (en unas 10 a 30 vueltas) para dar la **espiral somática**. Con esa doble espiralización se acorta la cromátida a una vigésima parte, sin aumentar apenas su diámetro. Esto la hace más visible al microscopio óptico. Algunas regiones de la cromátida sólo alcanzan el arrollamiento en espiral menor, como, por ejemplo, la zona del centrómero.

7.3 EL CROMOSOMA METAFÁSICO

7.3.1 Aspectos citogenéticos

7.3.1.1 Morfología de los cromosomas

Se llama **centrómero, restricción primaria o central** a la región más estrecha del cromosoma metafásico, por la que permanecen unidas las dos cromátidas hermanas. El centrómero delimita además los **brazos cromosómicos** (cuatro antes de la mitosis, dos tras ella). Los brazos cortos se designan con la letra *p* y los largos con *q* (la *p* procede del francés *petit*, pequeño y la *q*, dependiendo de las fuentes, de *queue*, cola, o bien simplemente por ser la letra siguiente a la *p*). Cada brazo se nombra, pues, por el número del cromosoma al que pertenece, seguido de la letra *p* o *q*. Por ejemplo, *6p* es el brazo corto del cromosoma 6, *8q* el brazo largo del cromosoma 8, etc. Los cromosomas se clasifican atendiendo a la posición del centrómero y, por tanto, según el tamaño relativo de los brazos:

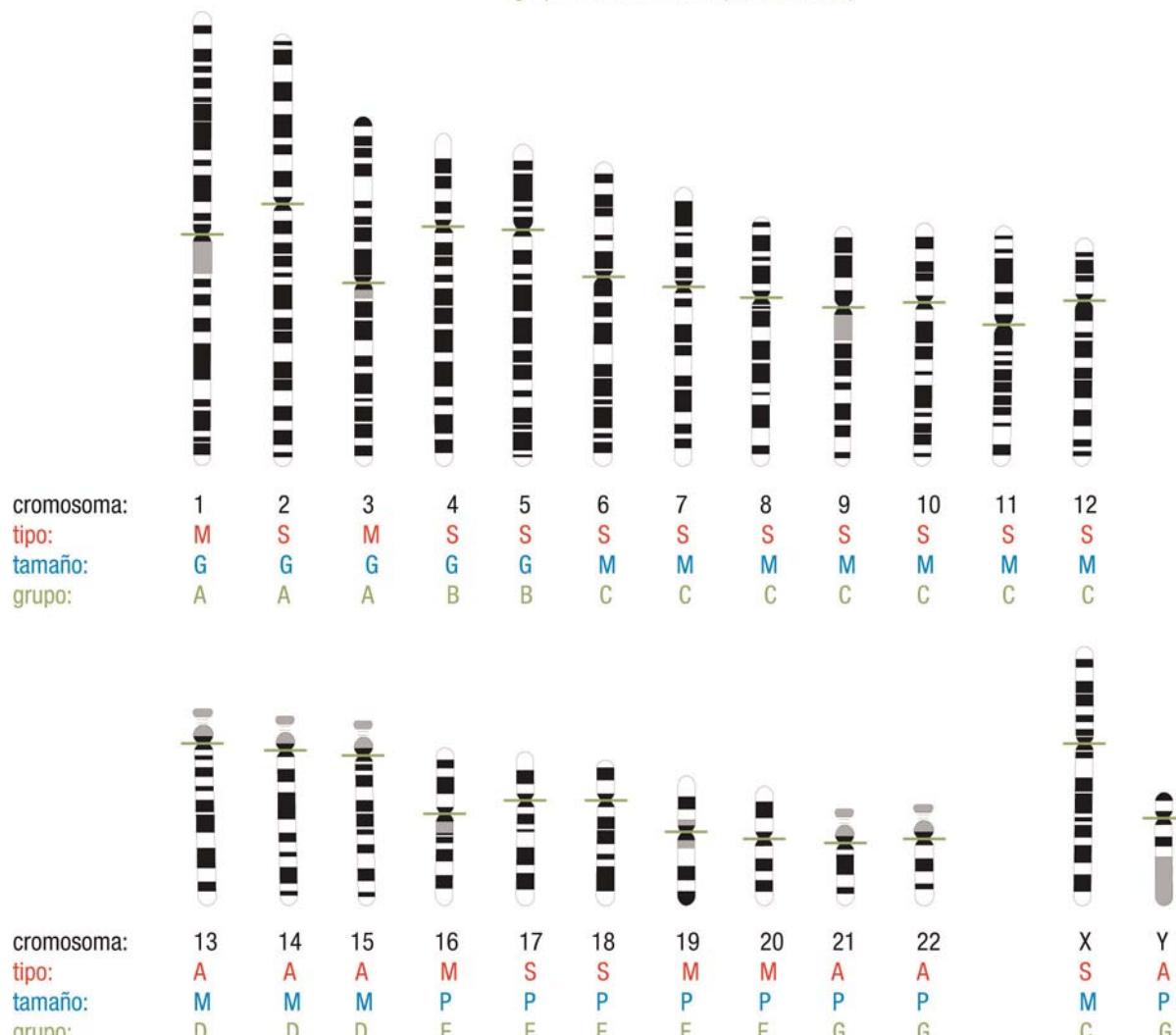


7.10

En los brazos cortos de los cromosomas acrocéntricos (excepto el Y) existen, además, otros estrechamientos de la cromátida, denominados **restricciones secundarias**, zonas en las que la espiral somática presenta un diámetro más reducido. Delimitan una sección terminal en el cromosoma, a modo de una pequeña esfera, a la que se llama **satélite cromosómico**.

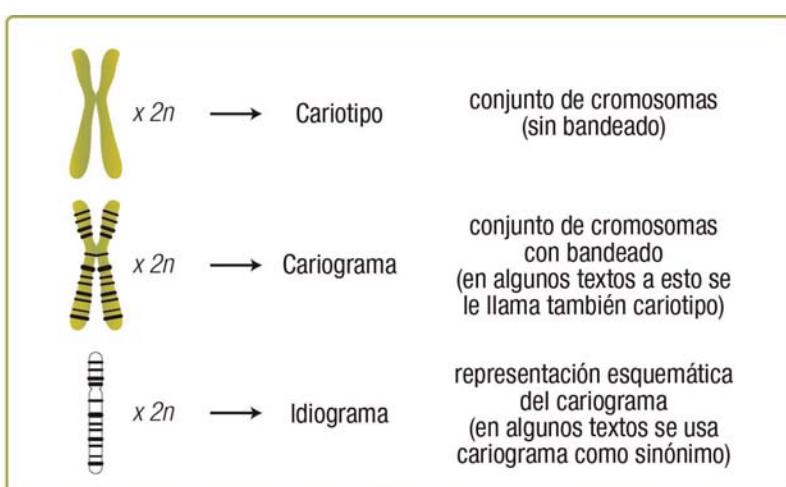
Clasificación de los cromosomas humanos, según:

- tipo (Metacéntrico, Submetacéntrico o Acrocéntrico)
- tamaño (Grande, Mediano, Pequeño)
- grupo de clasificación (A,B,C,D,E,F,G)



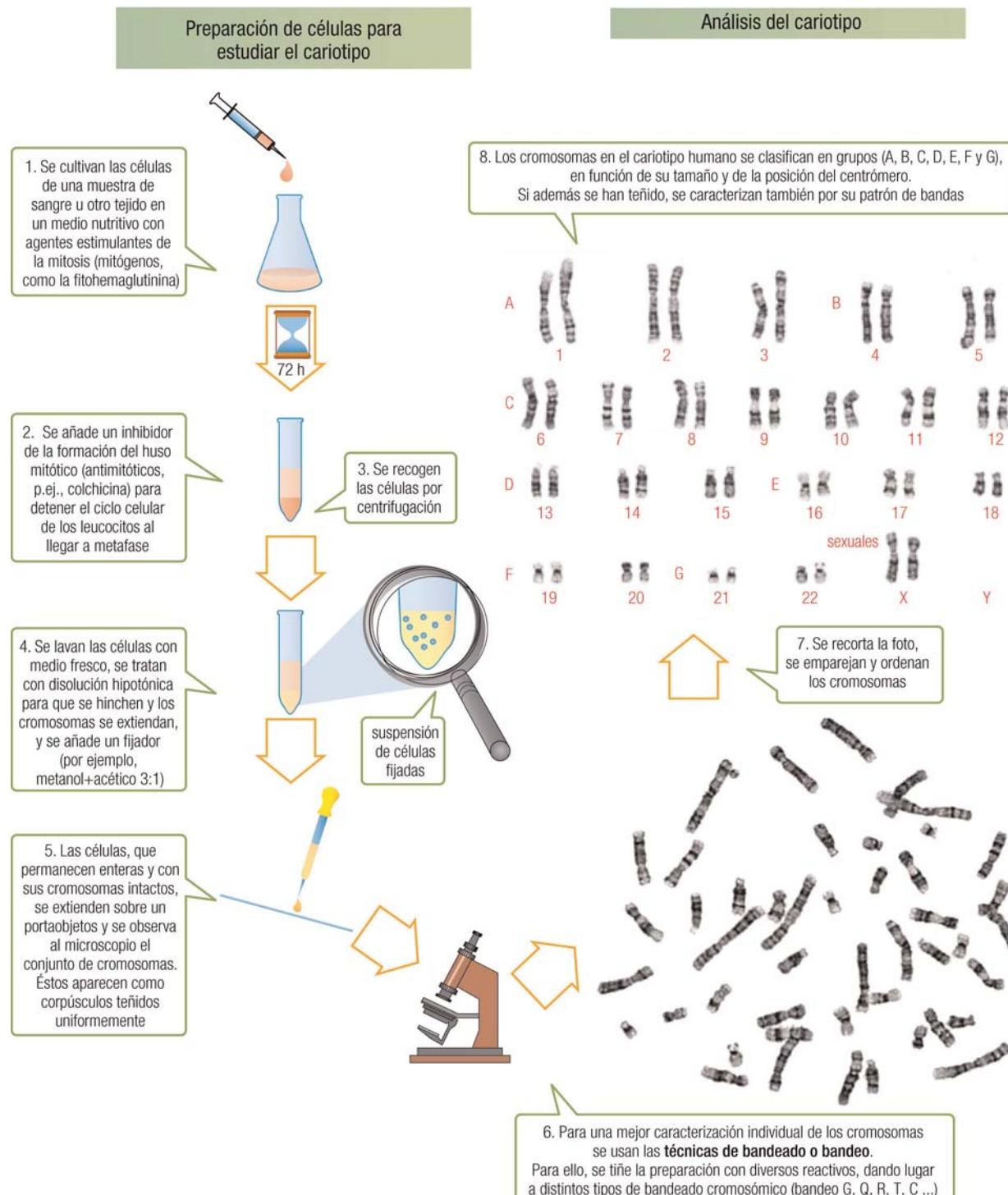
7.3.1.2 Cariotipo y su análisis

Cariotipo =
constitución cromosómica o *complemento* cromosómico de un individuo.
Es un rasgo característico de cada especie: todos los organismos de una especie tienen el mismo cariotipo.
Sin embargo, especies muy similares pueden tener cariotipos muy diferentes.



Son varios los métodos desarrollados para el análisis del cariotipo. Habitualmente, este estudio se realiza sobre cromosomas en el estado de prometafase o metáfase. Las preparaciones de cromosomas metafásicos se emplean, además, para analizar anomalías cromosómicas y determinar alguna ausencia o la presencia de cromosomas extra.

A este fin, el método original y más sencillo de obtención de cromosomas parte de leucocitos, como células nucleadas del organismo más asequibles.



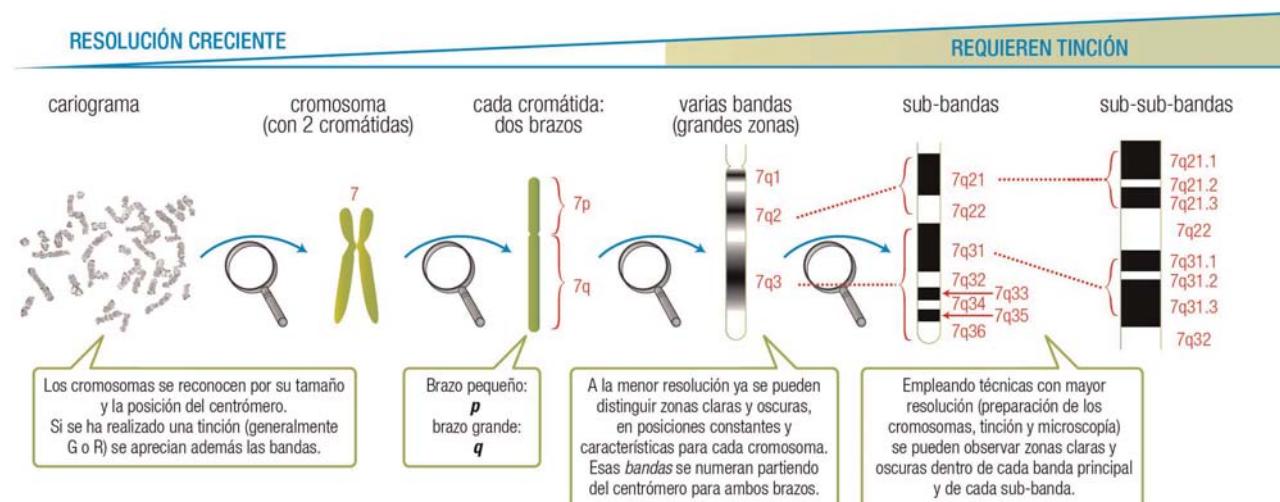
Actualmente, se parte ya no sólo de cultivos de leucocitos, con la desventaja de su vida corta, de 3 a 4 días, sino también de cultivos a largo plazo de células de diversos tipos de tejidos (fibroblastos, médula ósea, líquido amniótico, biopsias, etc.). Las más usadas son las de mayor potencial de división o mayor facilidad

de cultivo. Asimismo, existe una variedad de tratamientos de desnaturización o degradación enzimática de la cromatina, y de tinciones con colorantes específicos para el DNA. Por último, se han desarrollado sistemas de cariotipificación que pueden ser informatizados de forma que seleccionen automáticamente las células a analizar, el tipo de tinción y el análisis de los cromosomas.

7.3.1.3 Bandeado

Aplicando distintas técnicas de tinción, utilizadas de forma sistemática en laboratorios de citogenética, se pueden observar en los cromosomas bandas pálidas y oscuras alternativamente, que definen grandes regiones cromosómicas llamadas **isocoros**. Este **bandeo** o **bandeado** de cromosomas no es consecuencia de agrupamientos fortuitos, sino que está relacionado con la organización estructural del genoma, reflejando variaciones tales como composición de bases, grado de condensación, conformación de la cromatina, secuencias repetitivas y no transcritas, etc.

Los patrones de bandas de cada cromosoma son prácticamente idénticos en células diferentes, y en casi todos los tejidos, pero pueden diferir entre individuos; por ejemplo, al menos 12 cromosomas muestran variaciones en la longitud de ciertos segmentos en distintas personas. Se pueden llegar a apreciar hasta 500 bandas en un cromosoma; en función de la resolución microscópica alcanzada, las **bandas** principales se subdividen sucesivamente en **subbandas** y **subsubbandas**. A todas ellas se les asigna un identificador, que incluye el cromosoma, el brazo y un número correlativo desde el centrómero hacia los extremos.



7.14

Para el bandeado de cromosomas se emplean diversos métodos de tinción del DNA con colorantes específicos:

a) Bandeo G

Es la técnica más utilizada. Se basa en una desnaturización controlada de las proteínas cromosómicas (en general por digestión con tripsina), seguida de tinción con Giemsa y observación al microscopio (el reactivo de Giemsa es una mezcla compleja de colorantes; de él procede el nombre de bandeado G). Cada par de cromosomas muestra así patrones característicos de tinción, con bandas oscuras (G-positivas o bandas G) y bandas claras (G-negativas).

b) Bandeo Q

Los cromosomas se tiñen con un compuesto fluorescente, como mostaza de quinacrina (hidrocloruro de quinacrina), 4',6-diamidino-2-fenilindol (DAPI) o Hoechst 33258, que se intercalan en el DNA bicatenario. Se requiere, por tanto, un examen mediante microscopía de fluorescencia. Aparece un patrón específico de bandas brillantes (bandas Q o Q-positivas), que se corresponden casi exactamente con las bandas G (G-positivas, G-oscuras).

Las bandas G y Q contienen DNA que, en general, es algo más rico en pares AT (55-60%), pues la quinacrina se inserta preferentemente entre estos pares. Corresponden a cromatina altamente condensada, con DNA de replicación tardía dentro de la fase S del ciclo celular, relativamente inactivo transcripcionalmente.

c) Bandeo R

Se basa en un tratamiento de los cromosomas con calor (para desnaturalizar el DNA rico en AT), antes de la tinción con Giemsa. Resultan así bandas oscuras (bandas R) que coinciden con las bandas G o Q claras (el nombre deriva de patrón *reverso*). También se consigue el mismo patrón de bandas empleando olivomicina, un fluorocromo con afinidad por los pares GC.

Las bandas R contienen DNA rico en GC (50-60%), de baja condensación cromatínica, que se replica en etapas tempranas de la fase S y es relativamente activo transcripcionalmente.

d) Bandeo T

Identifica un subconjunto de bandas R especialmente concentradas en los telómeros, las de tinción más intensa, y se visualizan empleando un tratamiento térmico particularmente severo antes de teñir los cromosomas con Giemsa o una combinación de tinciones y fluorocromos.

e) Métodos especiales

Otros métodos de cultivo cromosómico y de tinción se utilizan para situaciones particulares: **bandeo C**, para regiones heterocromáticas; **tinción NOR**, para la *región organizadora del núcleo*, que contiene los genes de rRNA 18S y 28S; **bandeo de profase y prometafase, o de alta resolución**, sobre cromosomas detenidos en una etapa temprana de la mitosis, con un estado relativamente descondensado (su mayor longitud permite apreciar más detalles, llegándose a observar hasta 850 bandas en un solo cromosoma), etc.

Finalmente, la **citogenética molecular** emplea sondas de DNA clonado (preparadas por técnicas de DNA recombinante, pág. 180) para examinar cromosomas, de forma similar a los ensayos de hibridación molecular pero sobre preparaciones de cromosomas completos. Ésta es la técnica denominada **hibridación *in situ* con fluorescencia (FISH, fluorescence *in situ* hybridization**, pág. 177). Se dispone de sondas específicas para cromosomas individuales y para regiones cromosómicas, que permiten, por ejemplo, identificar reordenamientos cromosómicos particulares u obtener un diagnóstico rápido de la existencia de un número anómalo de cromosomas.

7.3.2 Regiones con significado funcional

Los cromosomas eucarióticos muestran tres elementos esenciales para una función correcta del ciclo celular y de la división, imprescindibles para la expresión, duplicación y segregación de los cromosomas: el **centrómero** (lugar de unión del cromosoma a las fibras del huso acromático), los **telómeros** (regiones que constituyen los extremos de los cromosomas) y los **orígenes de replicación** (puntos numerosos en cada cromosoma donde se inicia la copia de las dos hebras de DNA). Se tratan a continuación los aspectos morfológicos relacionados con los dos primeros; el análisis de los orígenes de replicación se pospone hasta que se estudie el proceso de replicación (pág. 153).

7.3.2.1 Centrómero

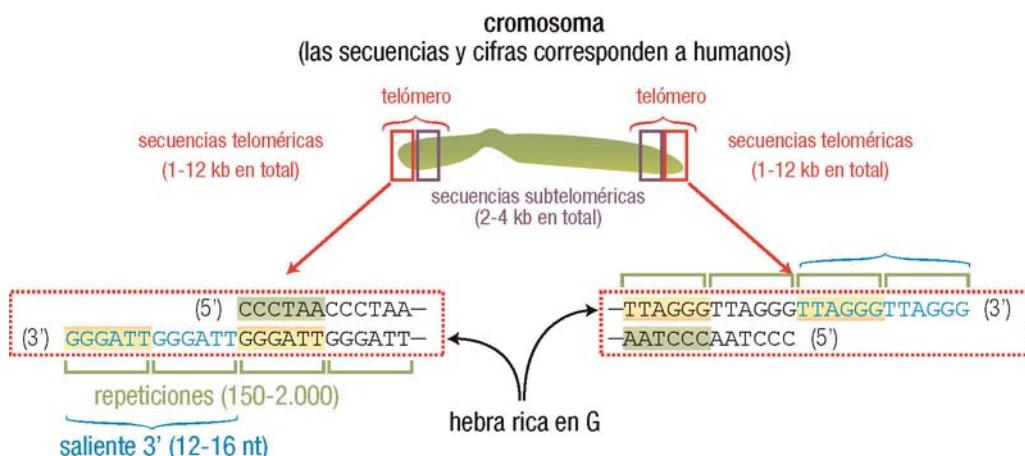
Además de ser el punto de contacto de las dos cromátidas y la frontera que delimita los dos brazos del cromosoma, funcionalmente la importancia del centrómero radica en que sobre él se sitúan los **cinetocoros**, estructuras proteicas en las que se anclan las fibras que constituyen el huso acromático o huso mitótico. Estas fibras, filamentos contráctiles o microtúbulos, de naturaleza proteica, ejercen la tracción necesaria para separar las dos cromátidas durante la división celular. De esta forma se produce la segregación ordenada de los cromosomas y cada célula hija recibe una cromátida de cada cromosoma, es decir, idéntica dotación genética. La ausencia de centrómero (cromosoma **acéntrico**) impide que el cromosoma se una al huso mitótico y, por tanto, que se reparta de manera correcta entre las células hijas.

Las secuencias esenciales para la función de los centrómeros son muy ricas en pares AT, tienen unos 170 pb en humanos y se repiten entre 2.000 y 30.000 veces en cada centrómero. Forman, pues, parte del llamado DNA repetitivo (pág. 116) y además suponen una parte importante de la fracción de heterocromatina constitutiva, la que sólo llega a descondensarse un tiempo mínimo en el ciclo celular, el preciso para su replicación.

7.3.2.2 Telómeros

Telómero es una palabra de etimología griega: *telos*, extremo, y *meros*, parte o región, que hace referencia a regiones situadas en los extremos de los cromosomas eucarióticos. Están constituidos por secuencias especializadas de DNA asociado a proteínas, y con características estructurales y funcionales propias que las diferencian de otras regiones cromosómicas.

Los telómeros de la mayoría de eucariotas están formados por dos tipos de secuencias de DNA. Una de ellas, denominada secuencia telomérica, repetición telomérica o repetición terminal, constituye el extremo de la cadena de DNA del cromosoma, mediante la repetición en tandem de un oligonucleótido corto (variable según especies; en humanos es TTAGGG). En general esta secuencia telomérica es más rica en G en una de las hebras, la que forma el extremo 3', que sobresale unos 12-16 nucleótidos sobre la hebra que aporta el extremo 5'. Ese extremo saliente es reconocido por **proteínas ligantes del telómero** (*TBP, telomere binding proteins*), que actúan a modo de "caperuza" protectora de dicho DNA terminal. Existen también unas secuencias más complejas adyacentes a las anteriores (es decir, más internas en el cromosoma), que se denominan **asociadas a telómeros** o **secuencias subteloméricas**.



7.15

Se han relacionado los telómeros con varias funciones:

- Participar en la estabilización y mantenimiento de la integridad estructural del cromosoma: en ausencia de telómeros, el extremo del cromosoma tiende a unirse a otros y aumentan las posibilidades de que sufra recombinación y degradación. Por otra parte, los telómeros permiten a las enzimas encargadas de la reparación del DNA diferenciar entre el extremo natural del cromosoma y uno que resulte de la fragmentación accidental de la cadena de DNA; en este último caso se detiene el ciclo celular, evitando la replicación hasta que se haya reparado la lesión.
- La función más importante consiste en asegurar la replicación completa de los extremos del cromosoma. Este aspecto se estudiará una vez considerado con detalle el proceso de la replicación del DNA (pág. 158).
- Se especula también con la implicación del telómero en la arquitectura tridimensional del núcleo o la del emparejamiento cromosómico.



El genoma en el ciclo celular

8.1 DOTACIÓN GENÉTICA EN EUCAΡΙOTAS	95	8.5.1 Gametogénesis	102
8.1.1 Número de cromosomas: valor <i>n</i> , haploidía y diploidía	95	8.5.1.1 Espermatogénesis	103
8.1.2 Contenido de DNA: valor <i>C</i>	97	8.5.1.2 Oogénesis	104
8.1.3 Locus y alelo	97	8.5.2 Reparto del material genético en la meiosis	105
8.1.4 Genotipo, fenotipo y dominancia	98	8.5.2.1 Carácter reductor de la meiosis	105
8.2 INTRODUCCIÓN AL CICLO CELULAR	99	8.5.2.2 Comparación entre meiosis y mitosis	106
8.3 LA CROMATINA EN LA INTERFASE	100	8.5.2.3 Segregación de cromosomas y diversidad	
8.3.1 Variaciones en condensación y contenido		genética	106
en DNA de la cromatina	101	8.5.2.4 Recombinación meiótica	108
8.4 DIVISIÓN DE CÉLULAS SOMÁTICAS POR MITOSIS	101	8.6 UNIÓN DE LOS GAMETOS HAPLOIDES DURANTE	
8.5 DIVISIÓN DE CÉLULAS GERMINALES PRIMARIAS		LA FECUNDACIÓN	108
POR MEIOSIS	101	8.7 RESUMEN INTEGRADO	108

8.1 DOTACIÓN GENÉTICA EN EUCAΡΙOTAS

Es conveniente completar el conocimiento de la organización del material genético considerando el contenido total de este material presente en las células eucarióticas y su relación con el tipo de célula y el estadio en el que se encuentra. Para ello, es preciso describir el significado de algunos términos y conceptos genéticos básicos, necesarios también para la comprensión del ciclo celular y de los dos tipos de división, mitosis y meiosis.

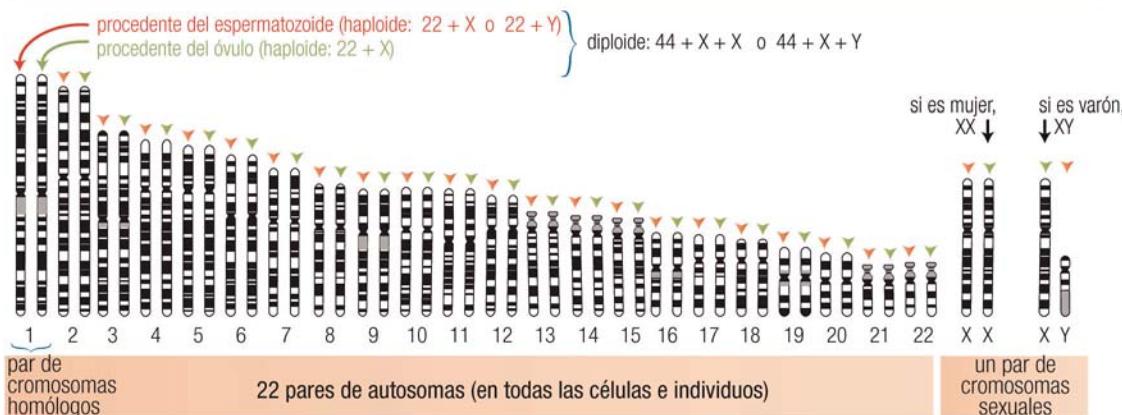
8.1.1 Número de cromosomas: valor *n*, haploidía y diploidía

El número de cromosomas en el núcleo es característico de cada especie, y varía mucho de unas a otras. Se habla de células **haploides** (y de organismos haploides, si están formados sólo por ellas) cuando cada cromosoma se presenta en forma individual, de modo que cada célula contiene un *juego* de cromosomas, todos distintos entre sí. El número de cromosomas se identifica con el número *n*.



Por el contrario, se denominan células **diploides** (y organismos diploides) a aquellas en las que se observan parejas o *pares de cromosomas* morfológicamente iguales entre sí; los dos miembros de la pareja se denominan **cromosomas homólogos**. Por tanto, cada célula diploide tiene dos *juegos* de cromosomas, y este carácter diploide se identifica con el número $2n$ (n parejas de cromosomas). En cada par, un cromosoma se ha heredado del padre y otro de la madre.

Célula diploide humana:
células somáticas (hepatocito, neurona...), células germinales precursoras y cigoto.
Tiene 46 cromosomas ($2n$): 46 moléculas lineales de DNA bicatenario.
Forman 2 conjuntos de 23 cromosomas distintos, o 23 pares de cromosomas
(a veces denominados 23 "cromosomas diploides").
De ellos, 22 pares, que son homólogos y además similares en ambos sexos, se llaman **autosomas**;
los 2 cromosomas restantes difieren dependiendo del sexo y sólo son homólogos en la mujer.



8.2

cromosomas sexuales	mujer:	varón:
Homología	X X, homólogos El par nº 23 está formado por dos cromosomas homólogos, igual que los autosómicos (pares n.º 1 a 22)	X Y, muy distintos El "par" nº 23 está formado por dos cromosomas no homólogos Sólo una pequeña región del Y y el X son homólogas (región <i>pseudoautosómica</i>)
Gametos producidos	Se dice que la mujer es <i>homogamética</i> : todos sus óvulos llevan un mismo tipo de cromosoma, X	Se dice que el varón es <i>heterogamético</i> : la mitad de sus espermatozoides son portadores de X, la otra mitad de Y
Procedencia de sus cromosomas	un X de la madre, el otro X del padre	el X siempre de la madre, el Y siempre del padre
Transmisión de los cromosomas	cromosoma X a las hijas y a los hijos, indistintamente (por fertilización con un espermatozoide X o Y, respectivamente)	el X sólo a las hijas, el Y sólo a los hijos (al fertilizar cualquier óvulo, que siempre será X)
En algunas especies, a diferencia de los humanos, no existe cromosoma Y, sino que los machos tienen un solo cromosoma sexual X, sin pareja, mientras las hembras tienen dos cromosomas X homólogos.		

8.3

En el caso de células (u organismos) con más de $2n$ cromosomas ($4n$, $8n$, etc.) se habla de **poliploidía**. Hay células que son poliploides de forma natural, porque han formado copias adicionales de su juego de cromosomas inicial mediante una replicación sin división celular (endomitosis). Por ejemplo, las células del hígado en regeneración son tetraploides ($4n$), y los megacariocitos gigantes de la médula ósea contienen usualmente $8n$,

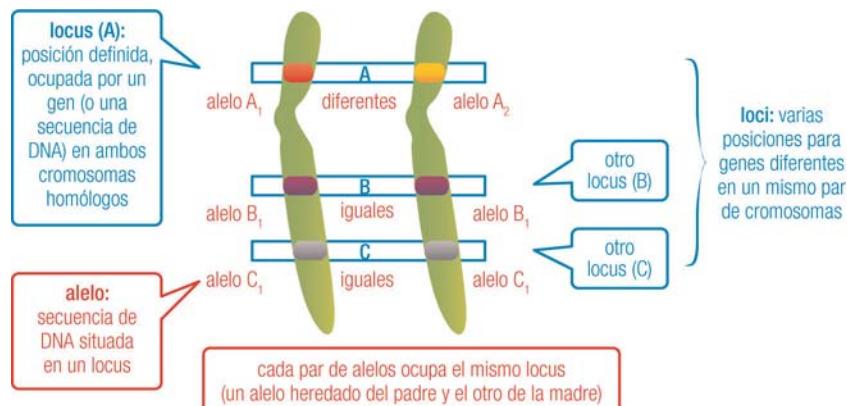
$16n$ o $32n$. Hay también situaciones patológicas caracterizadas por poliploidía: las más frecuentes son la triploidía ($3n$) y la tetraploidía ($4n$). Estas anomalías cromosómicas se estudian en el Capítulo 25.

8.1.2 Contenido de DNA: valor C

Durante el ciclo celular, la formación de gametos y la fertilización tienen lugar variaciones en el número de cromosomas de la célula, indicadas mediante los múltiplos de n . En humanos y animales, estas variaciones transcurren entre n (haploide) y $2n$ (diploide). No debe confundirse esa variación con la del contenido en DNA de cada célula, aunque éste también varía durante los procesos citados. Para indicar este contenido se emplea una notación basada en denominar C a la cantidad de DNA (bicatenario) correspondiente a un juego haploide de cromosomas (es decir, n) antes de su replicación. Se pueden encontrar células, tanto haploides como diploides, cuyo contenido C es diferente:

Célula haploide	n, C	gameto maduro (óvulo y espermatocito)
	$n, 2C$	gametocito secundario (precursor de los gametos maduros)
Célula diploide	$2n, 2C$	célula somática cuando no está en división, antes de sufrir la replicación de su DNA
	$2n, 4C$	célula somática tras sufrir la replicación, a punto de dividirse por mitosis gametocito primario, antes de dividirse por meiosis dando 2 gametocitos secundarios

8.1.3 Locus y alelo



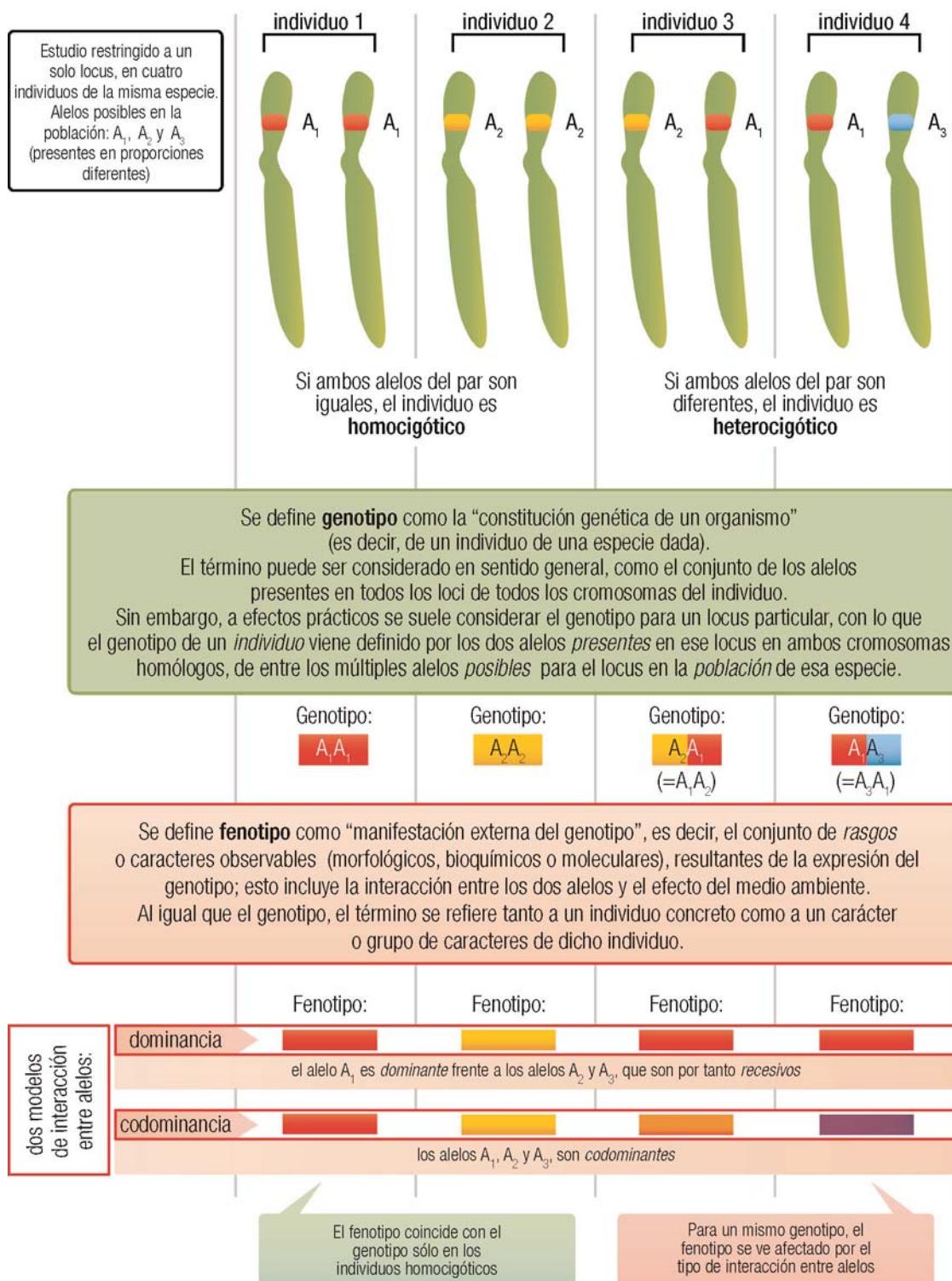
8.4

En cualquier célula de una misma especie, un gen ocupa siempre el mismo lugar en un cromosoma determinado, y la posición equivalente en su homólogo. (Por ahora, basta con que se entienda por gen una secuencia de DNA que codifica un producto funcional, bien proteína o bien RNA. Más adelante se define con mayor precisión [pág. 263].) Para referirse de forma genérica a esa posición se utiliza el término *locus*. Por tanto, genes distintos ocuparán diferentes *loci* (plural de *locus*), ya sea en el mismo par de cromosomas o en otro. Estos términos también se utilizan para la posición de regiones del DNA que no sean genes.

Cada porción de secuencia concreta de DNA (génica o no) situada en un locus se denomina *alelo*. Puesto que los cromosomas se presentan en pares, en cada individuo hay dos alelos en el mismo locus, que pueden ser idénticos o diferentes.

En general, los dos alelos que presenta un individuo en un locus determinado son diferentes de los de otro individuo de la misma especie (diversidad genética o *polimorfismo*, pág. 405). Como consecuencia, en el conjunto de la población existen más de dos alelos diferentes para un locus. El estudio de enfermedades ha permitido identificar numerosos alelos, aquellos menos frecuentes en la población, que sirven hoy como marcadores para identificar dichas patologías.

8.1.4 Genotipo, fenotipo y dominancia



Relación entre el número de alelos existentes en la población para un solo locus y los genotipos diferentes a los que pueden dar lugar				Probabilidad de los genotipos en la población (asumiendo que todos los alelos son igualmente frecuentes)	
5 situaciones posibles en la población: Nº de alelos existentes	Nº de genotipos homocigóticos	Nº de genotipos heterocigóticos	Nº de genotipos posibles en total	Relación de homocigóticos a heterocigóticos	Frecuencia de heterocigosis
1: A ₁	1: A ₁ A ₁	0	1	1: 0	0%
2: A ₁ , A ₂	2: A ₁ A ₁ , A ₂ A ₂	1: A ₁ A ₂	3	2: 2 (*)	50%
3: A ₁ , A ₂ , A ₃	3: A ₁ A ₁ , A ₂ A ₂ , A ₃ A ₃	3: A ₁ A ₂ , A ₁ A ₃ , A ₂ A ₃	6	3: 6 (*)	66%
4: A ₁ , A ₂ , A ₃ , A ₄	4: A ₁ A ₁ , A ₂ A ₂ , A ₃ A ₃ , A ₄ A ₄	6: A ₁ A ₂ , A ₁ A ₃ , A ₁ A ₄ , A ₂ A ₃ , A ₂ A ₄ , A ₃ A ₄	10	4: 12 (*)	75%
5: A ₁ , A ₂ , A ₃ , A ₄ , A ₅	5: A ₁ A ₁ , A ₂ A ₂ , A ₃ A ₃ , A ₄ A ₄ , A ₅ A ₅	10: A ₁ A ₂ , A ₁ A ₃ , A ₁ A ₄ , A ₁ A ₅ , A ₂ A ₃ , A ₂ A ₄ , A ₂ A ₅ , A ₃ A ₄ , A ₃ A ₅ , A ₄ A ₅	15	5: 20 (*)	80%
X	X	X(X-1)/2	X(X+1)/2	X : X(X-1)	(X-1)X · 100%

(*) Un mismo genotipo heterocigótico es el resultado de dos combinaciones de cada dos alelos (por ejemplo, A₁A₂ = A₂A₁, en la columna 3). Como consecuencia (columna 5), la probabilidad de existencia de cada uno de los genotipos heterocigóticos en la población es doble que la de cada genotipo homocigótico.

8.6

En las situaciones en las que existen dos o más alelos en la población se dice que hay **polimorfismo**, obviamente referido al locus considerado. Los polimorfismos del genoma humano son objeto de atención especial en el Capítulo 24.

8.2 INTRODUCCIÓN AL CICLO CELULAR

Es importante destacar que los aspectos estructurales considerados hasta ahora para la molécula de DNA deben estudiarse bajo la situación fisiológica en la que realmente se encuentra el material genético en el núcleo eucariótico, es decir, durante las etapas de reposo y división que se alternan en la vida de la célula. Con frecuencia, estas dos cuestiones se relacionan poco entre sí, o se analizan bajo planteamientos que carecen de una base molecular suficiente. Se pretende en este texto resaltar la conexión entre los cambios moleculares de la cromatina y los procesos del ciclo celular.

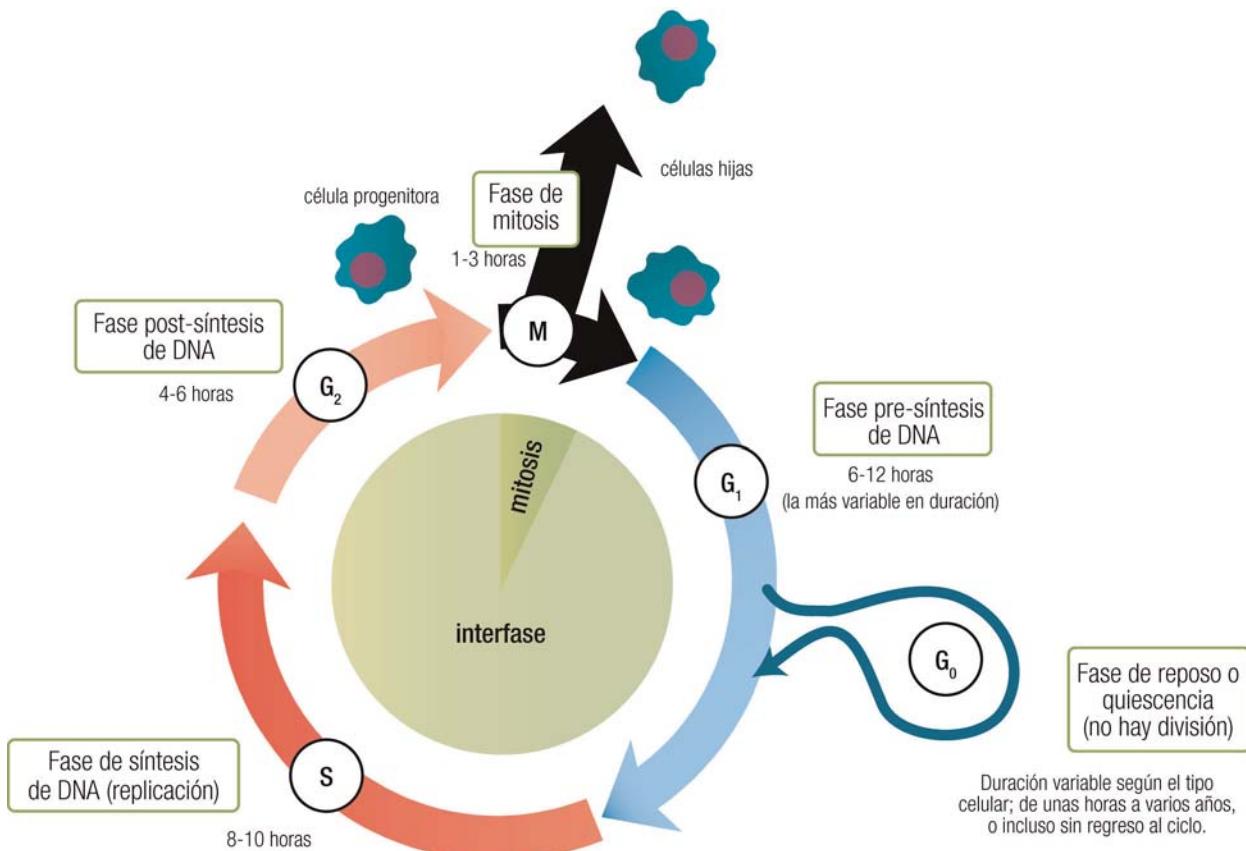
Mediante la división celular se generan tanto los gametos (división meiótica) como, a partir del cigoto, todas las células necesarias para que un organismo crezca y se desarrolle durante los períodos embrionario, fetal y de crecimiento hasta el estado adulto, así como para la renovación continua de las células que mueren durante dichos períodos (división mitótica).

La división por mitosis tiene lugar normalmente durante toda la existencia del organismo, desde la primera división del cigoto hasta que el individuo muere. Se calcula que durante la vida humana se producen alrededor de 10^{17} divisiones. La transición entre dos divisiones sucesivas (interfase) constituye, junto a la propia división, el **ciclo de división celular** (o, simplemente, **ciclo celular**). Su duración varía ampliamente de acuerdo con el tipo de célula, así como en función de determinados estímulos. Algunas células experimentan una sucesión continua de ciclos de división (por ejemplo, células epiteliales del intestino y precursores hematopoyéticos en la médula ósea), otras poseen ciclos más espaciados (una vez al mes en células hepáticas, una vez cada 3 años en células endoteliales de vasos sanguíneos, etc.) y, como caso extremo, algunas células detienen su progresión en el ciclo, reanudándolo tras un período prolongado (quiescencia) o incluso no dividiéndose más (caso de las neuronas).

Existen unas pocas células especializadas en la producción de gametos o células sexuales, que permitirán la formación de un nuevo individuo. Dichas células germinales se dividen mediante **meiosis**, un mecanismo diferente al del resto (células somáticas), aunque también está precedido de un ciclo celular similar.

Una visión global del ciclo celular permite observar cómo la célula progresa de manera ordenada y sin discontinuidad por dos grandes etapas bien diferenciadas: la **división celular** propiamente dicha (por mitosis para células somáticas y por meiosis para células germinales) y la **interfase** (común a ambas). Durante esta última, la célula se prepara para la siguiente división duplicando su material genético (DNA) y todo su

contenido (proteínas, RNA, orgánulos, membranas), de modo que se duplica su tamaño antes de dividirse en dos células hijas. Clásicamente, el ciclo se plantea en torno al proceso de replicación, o duplicación del DNA nuclear, con lo que se definen dentro de la interfase tres etapas: la fase S (durante la cual tiene lugar la replicación), la fase G₁ (previa) y la fase G₂ (posterior). Mientras que algunas células completan rápidamente cada interfase para entrar de nuevo en división, otras pasan a un estadio modificado de G₁, llamado fase G₀ o quiescente, en el que permanecen días o incluso años hasta emprender un nuevo ciclo, o del que nunca salen (células que no se dividen, como es el caso de las neuronas).



Siguiendo el planteamiento general de este texto, interesa relacionar los aspectos genéticos clásicos del ciclo con aquellos otros de índole más bioquímica y molecular. Para ello, se abordarán en forma detallada la interfase y los dos tipos de división celular, con estos objetivos:

- Poner de manifiesto la enorme variación estructural que supone, durante la interfase, la condensación de la molécula de DNA desde cromatina a cromosoma y su descondensación en sentido inverso.
- Resaltar el proceso de la replicación, que tiene lugar igualmente durante la interfase pero en un momento diferente del anterior. Sus características moleculares y enzimáticas se estudiarán con detalle en el Capítulo 11.
- Establecer las analogías y diferencias entre la división por mitosis de las células somáticas y la división por meiosis de las células germinales primarias para dar lugar a células germinales maduras o gametos.

8.3 LA CROMATINA EN LA INTERFASE

La **interfase**, o lapso que transcurre entre dos divisiones celulares, es normalmente el período más prolongado del ciclo celular, con sus tres fases G₁, S y G₂. Esta etapa termina cuando una célula somática entra en división por mitosis (para formar células somáticas hijas), y cuando una célula germinal primaria se divide por meiosis (para formar las células germinales maduras). Se va a estudiar cómo durante la interfase varía el grado de condensación del material genético y también cómo varía el contenido de DNA (2C en G₁, 4C en G₂) sin que se modifique el número de cromosomas (la célula se mantiene diploide, 2n).

8.3.1 Variaciones en condensación y contenido en DNA de la cromatina

Durante la mayor parte de la vida celular (bien interfase o bien quiescencia, según los casos), la molécula de DNA no está presente en la forma extendida de doble hélice ni tampoco en la forma compacta de cromosomas, sino en el estado parcialmente condensado conocido como **cromatina** (págs. 85 y 87).

Al progresar en el ciclo celular, la cromatina se descondensa de manera gradual durante la fase G₁ hasta adoptar localmente una conformación totalmente extendida (correspondiente a la doble hélice), necesaria para la separación de las dos hebras durante la replicación (fase S). Para ello, la eucromatina ha de pasar antes por los estados de fibra de 30 y 10 nm (pág. 103).

Durante la fase G₁, las células, tanto somáticas como germinales, son diploides ($2n$) y su contenido en DNA es 2C. En la fase S o de síntesis se produce la replicación del DNA de los $2n$ cromosomas individuales. Cada hebra de este DNA sirve de molde para la síntesis de otra nueva, que permanece asociada por emparejamiento de bases. Las dos moléculas de DNA bicatenario resultantes permanecen unidas por el **centrómero**, dando así lugar a cromosomas con 4 hebras de DNA; cada doble hebra constituye una **cromátida**. De este modo, el número de cromosomas permanece constante (diploide, $2n$), pero se duplica el contenido de DNA, desde 2C a 4C. Una vez finalizada la replicación de todo el genoma, la célula entra en la fase G₂ donde se inicia la condensación gradual de la cromatina (pág. 102, en secuencia inversa a la descondensación ocurrida durante la fase G₁, pág. 103), que se completa en las primeras etapas de la mitosis dando lugar a cromosomas visibles al microscopio, con su aspecto típico de 2 cromátidas y 4 brazos. La célula sigue siendo $2n$, 4C.

Aunque se puede considerar que cada molécula de DNA es un cromosoma en cualquier momento del ciclo, el término **cromosoma** corresponde estrictamente a esta forma condensada al máximo. A veces se usan los términos *cromosoma metafásico* y *cromosoma interfásico* para las formas condensada y descondensada (cromatina), respectivamente. Conviene, finalmente, recordar que sólo la cromatina suficientemente descondensada puede sufrir transcripción, o expresión de sus genes (págs. 88, 268 y 281).

8.4 DIVISIÓN DE CÉLULAS SOMÁTICAS POR MITOSIS

La mitosis, o fase M, es habitualmente el período más corto del ciclo celular, durante el que tiene lugar la división de la célula, que ya ha duplicado su material genético y citoplasmático, en dos células hijas idénticas. Clásicamente se subdivide en cinco etapas: 1) **Profase**: la cromatina se condensa hasta formar los cromosomas, dejan de observarse los nucléolos y aparece el huso mitótico. 2) **Prometafase**: comienza a romperse la membrana nuclear, se ven por primera vez las 2 cromátidas de cada cromosoma y los $2n$ cromosomas dobles migran hacia el plano ecuatorial de la célula. 3) **Metafase**: los $2n$ cromosomas, totalmente condensados, aparecen alineados en el plano ecuatorial, independientes unos de otros. 4) **Anafase**: los cromosomas se dividen por sus centrómeros, separándose en dos cromátidas, que son desplazadas hacia polos opuestos de la célula. Cada cromátida es ya un cromosoma hijo independiente. 5) **Telofase**: comienzan a descondensarse los cromosomas y se forman membranas nucleares para delimitar dos núcleos. Paralelamente, la membrana celular separa el citoplasma en dos partes, cada una con un núcleo y la mitad de los orgánulos, formando las dos células hijas, en el proceso llamado **citocinesis** (o citoquinesis).

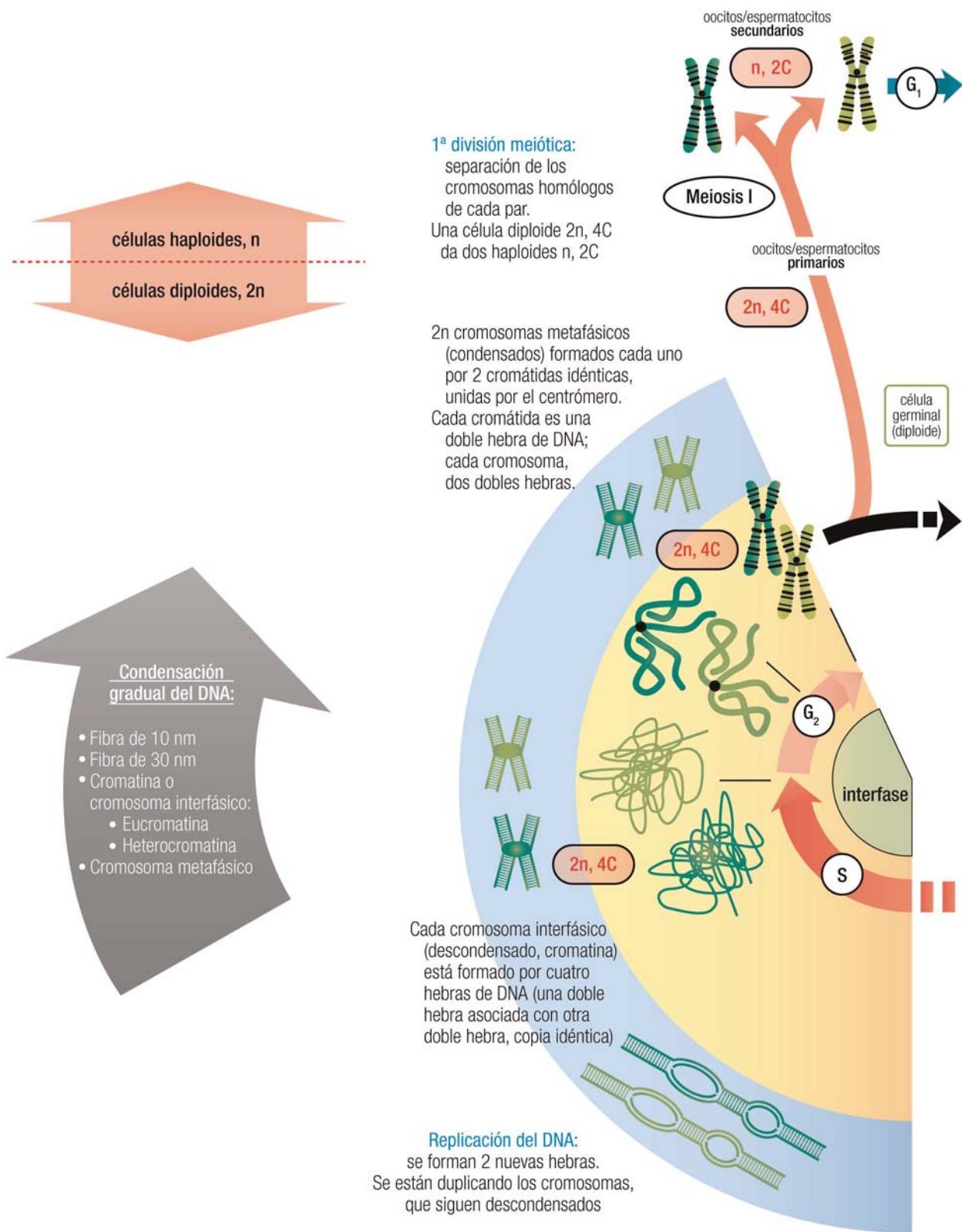


Web 8.1. Separación de las cromátidas durante la mitosis.

Como consecuencia de la mitosis, cada célula hija recibe una cromátida de cada cromosoma de la célula progenitora, es decir, la mitad de DNA, pero constituyendo un conjunto cromosómico completo (23 pares de cromosomas) con toda la información genética de la célula progenitora. Por ello, se dice que la **mitosis es un proceso de división no reductora** (pág. 105). La única diferencia es que la célula justo antes de dividirse tiene $2n$ cromosomas de 4 hebras y 2 cromátidas cada uno (4C) y tras la división cada célula hija tiene $2n$ cromosomas de 2 hebras y una cromátida cada uno (2C) (págs. 102-103).

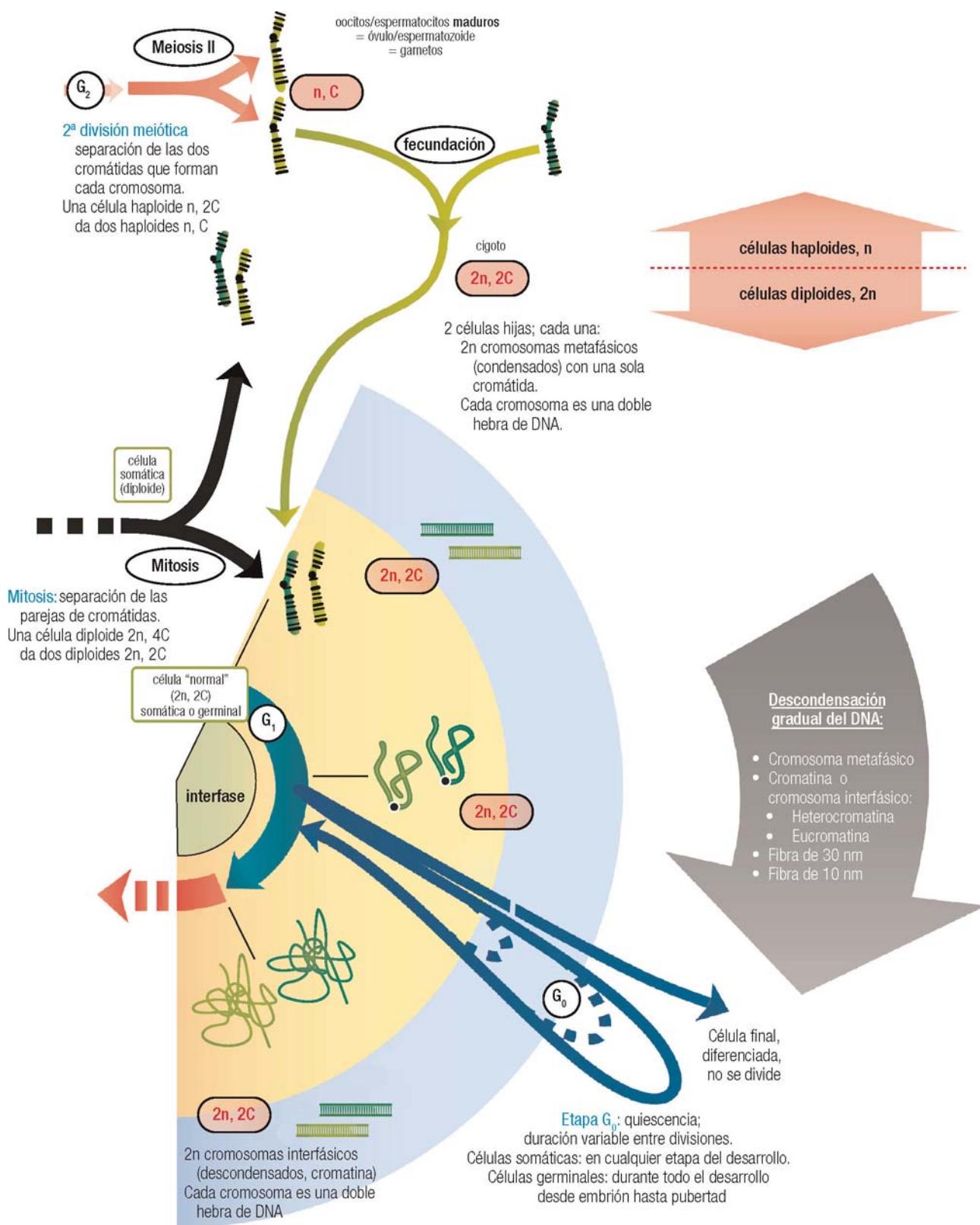
8.5 DIVISIÓN DE CÉLULAS GERMINALES PRIMARIAS POR MEIOSIS

La división meiótica forma parte del proceso de **gametogénesis**, es decir, formación de **gametos** o células sexuales haploides a partir de células germinales primordiales (que son diploides, al igual que las células somáticas).



8.5.1 Gametogénesis

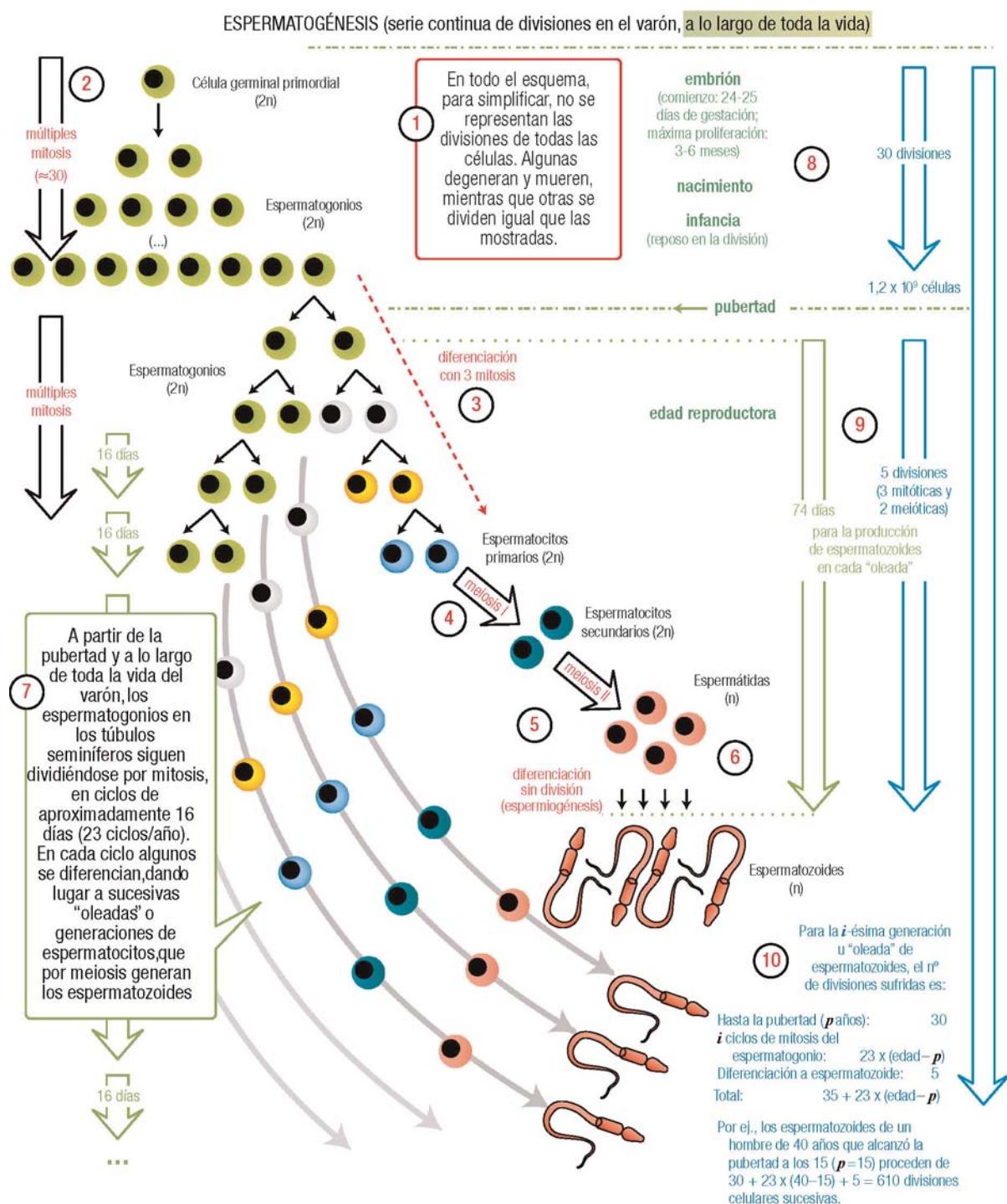
La oogénesis y la espermatogénesis se inicián a partir de **células germinales primordiales**, células especializadas del óvulo y testículo, respectivamente, mediante una serie de divisiones mitóticas que, con el consiguiente proceso de crecimiento y diferenciación celular, producen sucesivas generaciones de células llamadas **oogonios** y **espermatogonios** (también oogonia y espermatogonia; págs. 103 y 104). A continuación éstas, también por mitosis, dan lugar a **oocitos y espermatocitos primarios**. Todas estas células son diploides. Finalmente, a partir



de cada oocito o espermatocito primario, y mediante dos divisiones sucesivas diferentes (meiosis I y meiosis II), se produce la reducción del estado diploide al haploide. Por ello, el conjunto de esas dos divisiones recibe el nombre de **meiosis** o **división celular reductora**.

8.5.1.1 Espermatogénesis

En el caso del varón, el proceso, que tiene lugar de forma repetida a lo largo de toda su vida, culmina con la producción de espermatozoides o gametos masculinos.



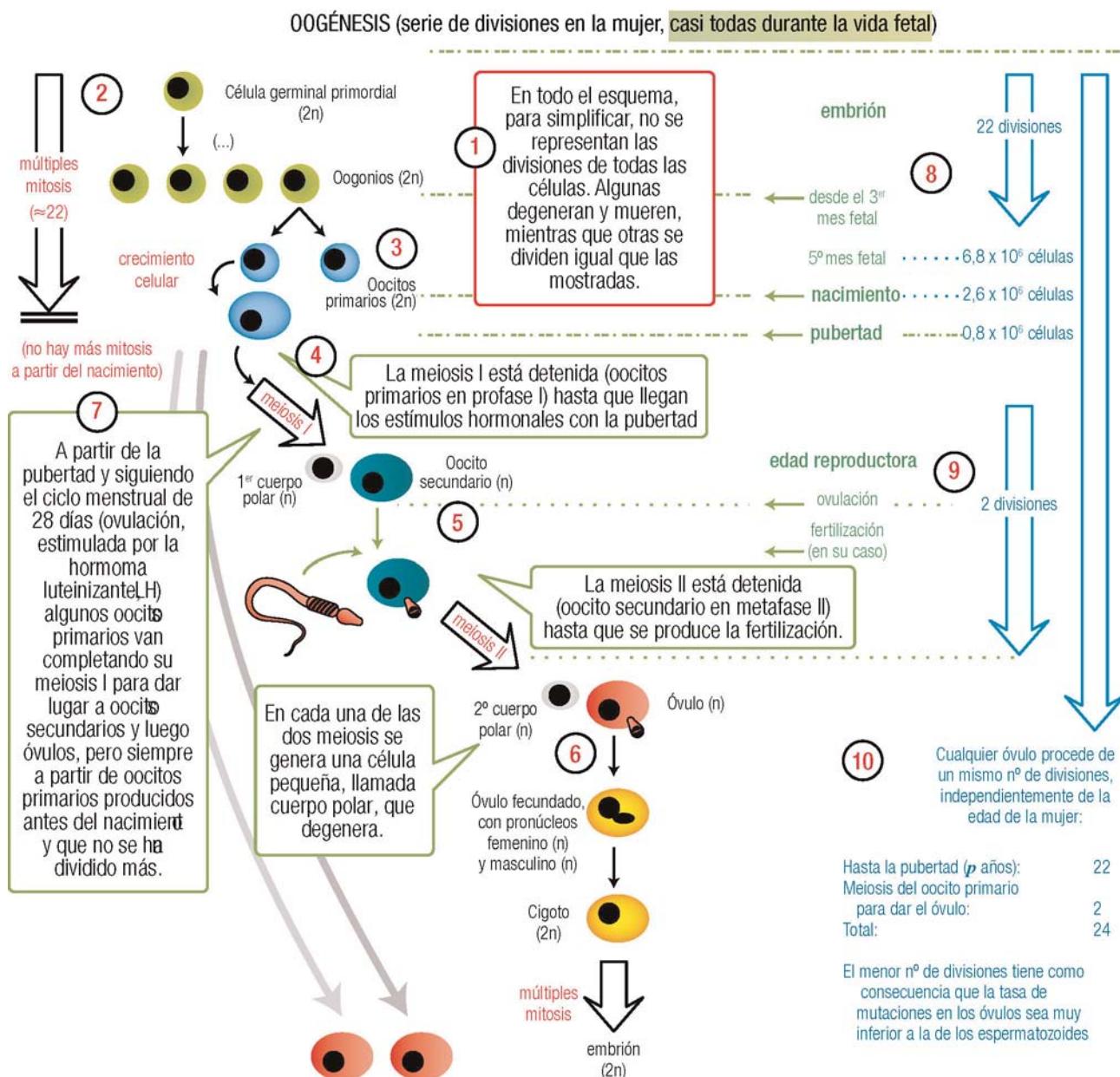
8.9

Cada uno de los espermatoxenitos primarios ha dado lugar a un reparto simétrico del citoplasma en cuatro espermatozoides iguales. El número total de divisiones alcanza valores sorprendentes, variables según la edad, ya que dependen de los años transcurridos desde la pubertad. El número total de espermatozoides realmente producidos es obviamente variable; baste indicar que un hombre sano puede formar unos 200 millones diarios durante toda su vida fértil y que un eyaculado, con unos 2 ml, contiene unos 60 millones de espermatozoides.

8.5.1.2 Oogénesis

Por el contrario, en la mujer la oogénesis ocurre en su mayor parte durante la vida fetal, a partir de la célula germinal primordial originada en el endodermo del saco vitelino del embrión. El mayor número de oocitos primarios se alcanza hacia el quinto mes de vida fetal (compárese con la pubertad para el varón) y es muy inferior

al de espermatocitos primarios. A partir de ese momento, el número de oocitos se reduce continuamente hasta la pubertad, se mantiene durante la vida fértil de la mujer y desaparece casi totalmente al llegar la menopausia.



8.10

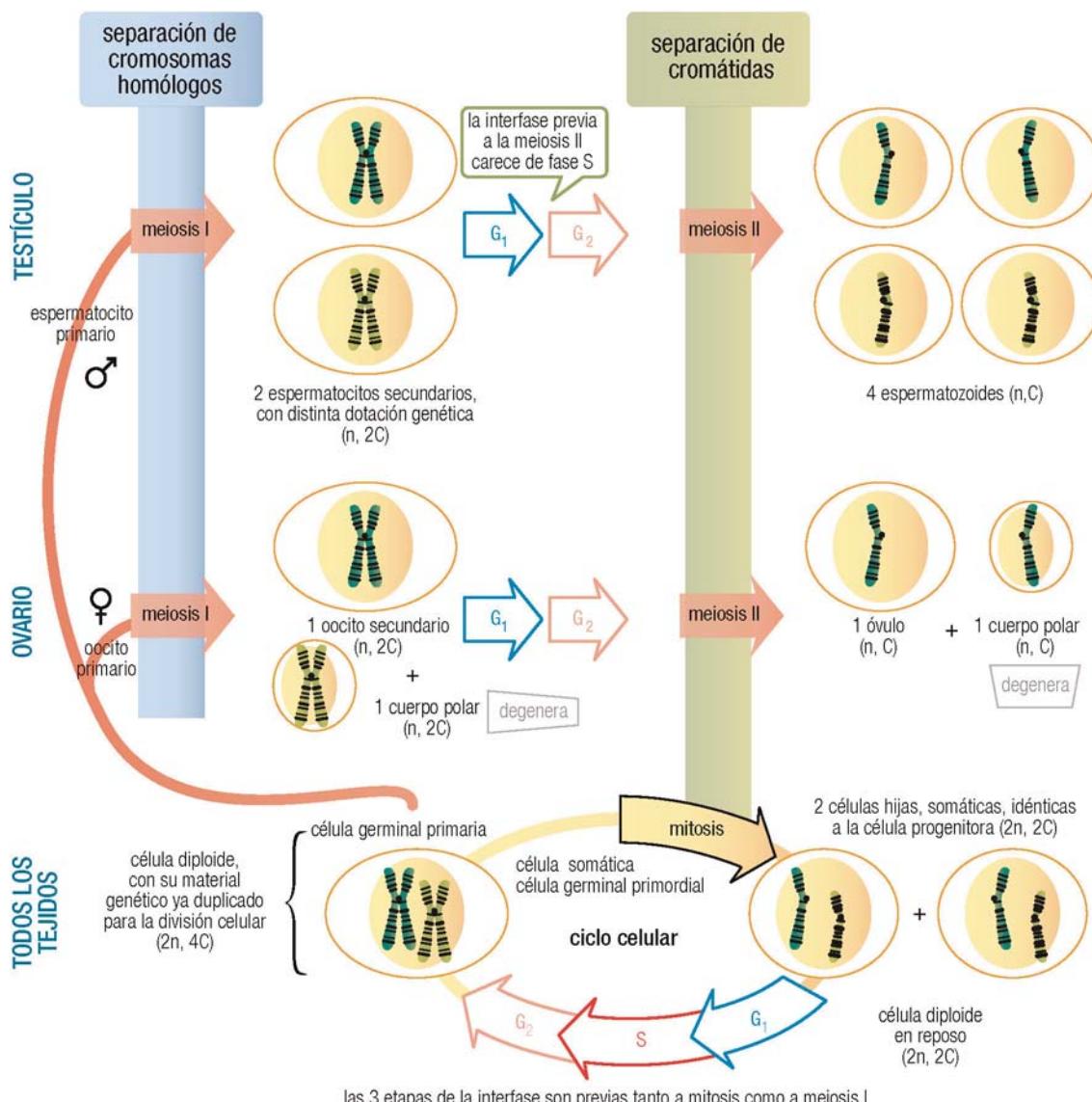
Otra peculiaridad que diferencia espermatogénesis y oogénesis es que en las dos divisiones meióticas de esta última el reparto del citoplasma es asimétrico: en la meiosis I se forma una célula pequeña que recibe el nombre de *cuerpo polar* y un oocito secundario grande. La división de éste en la meiosis II origina un *segundo cuerpo polar* y un óvulo maduro. Ambos cuerpos polares degeneran posteriormente. Por tanto, un oocito primario sólo genera un óvulo, mientras que un espermatocito primario produce cuatro espermatozoides.

8.5.2 Reparto del material genético en la meiosis

8.5.2.1 Carácter reductor de la meiosis

En la meiosis I, el gametocito (oocito o espermatocito) primario separa sus 23 pares de cromosomas homólogos ($2n, 4C$, con dos cromátidas cada uno), de forma que pasan 23 cromosomas diferentes (uno de cada par; $n, 2C$, con 2 cromátidas) a cada gametocito secundario (primera ley de Mendel). Se dice que se produce la *segregación* de los cromosomas, o que cada par ha sufrido *disyunción*; se reduce así a la mitad el número de cromosomas.

Por ello, se afirma que la meiosis es una división *reductora*, debido a la meiosis I. La meiosis II es *no reductora* y en cuanto al reparto de material genético se asemeja a la mitosis, como se discute a continuación.



8.11

8.5.2.2 Comparación entre meiosis y mitosis

La diferencia esencial entre la mitosis y las dos etapas de la meiosis radica en la distribución del material genético entre las células hijas (somáticas diploides en mitosis; gametos haploides en meiosis). El mecanismo de la meiosis II es equivalente al de la mitosis, puesto que se **separan las dos cromátidas** y se reduce la cantidad de DNA ($2C \rightarrow C$) sin cambiar el carácter haploide (n). De este modo, cada célula hija recibe **copias idénticas** de DNA. En contraste, en la meiosis I se **separan los miembros de cada pareja de cromosomas**, con lo que se reducen a la vez la cantidad de DNA ($4C \rightarrow 2C$) y el número de cromosomas ($2n \rightarrow n$), y las células hijas reciben **diferente material genético** (pues cada cromosoma de un par puede contener alelos diferentes en cada locus, pág. 97).

8.5.2.3 Segregación de cromosomas y diversidad genética

En ambas divisiones meióticas el reparto de los cromosomas homólogos de cada par (meiosis I) y de las cromátidas de cada cromosoma (meiosis II) se hace al azar, de forma totalmente independiente para cada cromosoma (segunda ley de Mendel). En definitiva, se asegura con creces que cada gameto recibe diferente material genético; esta variabilidad es el fundamento del polimorfismo genético, cuyos detalles y consecuencias se estudian más adelante (Capítulo 24).

GAMETOCITO PRIMARIO

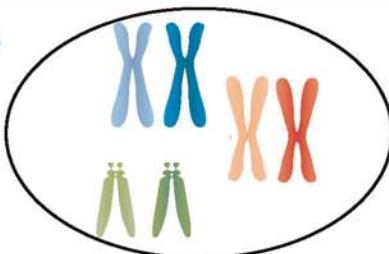
Segregación independiente de los cromosomas en la meiosis

Ilustrada mediante un ejemplo de una célula con 3 pares de cromosomas homólogos (de los 23 realmente presentes)

1 gametocito primario

2n, 4C

En la meiosis I, cada pareja de cromosomas homólogos se separa, repartiéndose entre las dos células hijas de forma independiente de los demás pares de cromosomas (esto fue ya propuesto por Mendel en su Ley de la Segregación)

genotipo
AaBbCc

Por ej.: letra mayúscula y color oscuro para los cromosomas de origen paterno, minúscula y claro para el materno

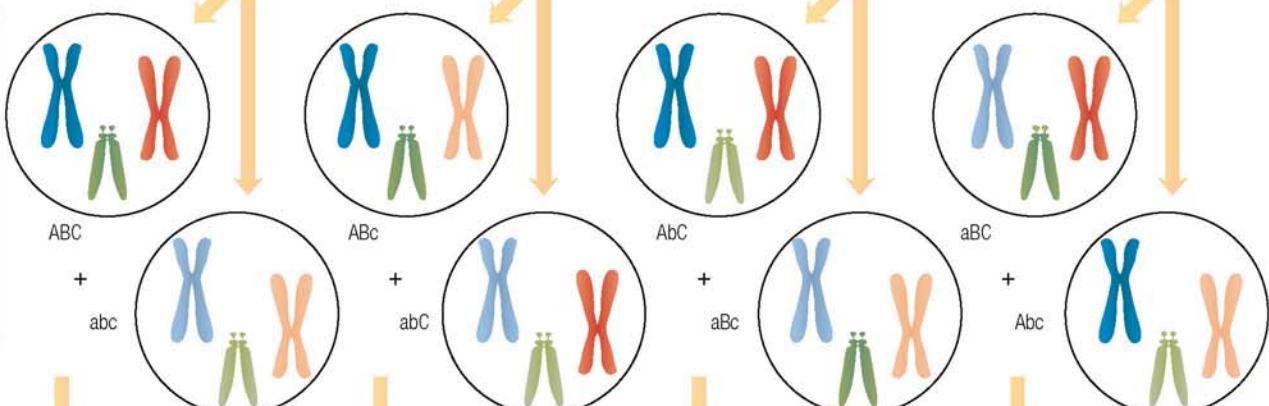
meiosis

de cada gametocito primario:

2 gametocitos secundarios x 4 posibilidades = 8 posibles gametocitos secundarios

n, 2C

GAMETOCITOS SECUNDARIOS

La segregación independiente produce 2^n posibles combinaciones de cromosomas (en el ejemplo, $2^3 = 8$).Para los humanos ($n=23$) hay $2^{23} = 8,4$ millones de gametos posibles a partir de una misma célula germinal precursora.

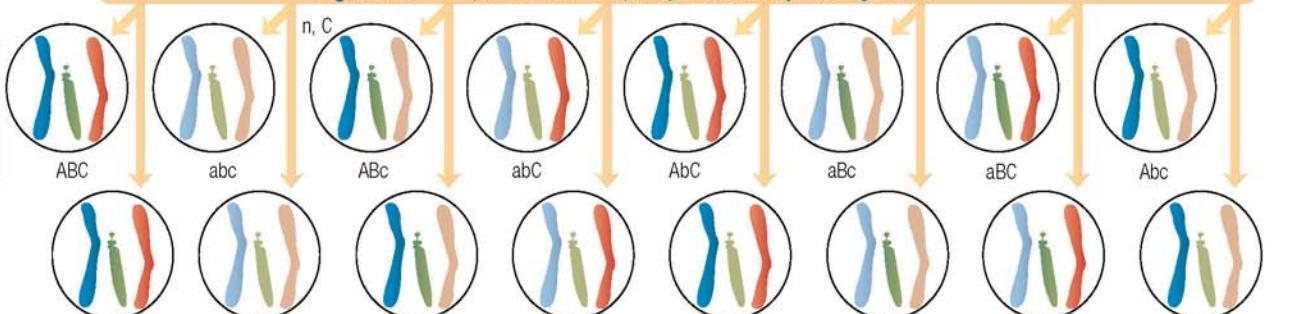
A esta diversidad se le suma la resultante de la recombinación de los cromosomas en la profase de la meiosis I (no mostrada).

meiosis II

de cada gametocito secundario:

2 gametos x 8 posibilidades = 16, pero iguales 2 a 2: 8 posibles gametos

GAMETOS



GAMETO

n, C

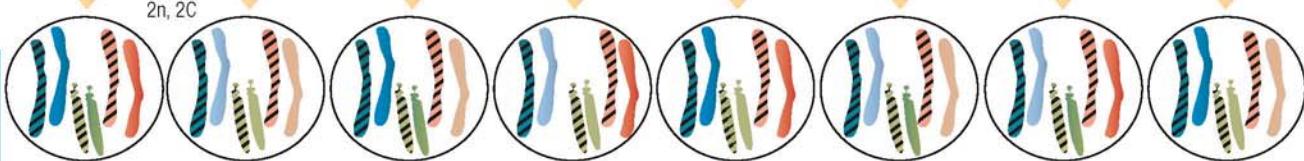
Fertilización (con un mismo gameto)

por un mismo espermatozoide, si el anterior es un óvulo;
de un mismo óvulo, si el anterior es un espermatozoide

cigotos : 8 posibilidades (1 de cada tipo de gameto)

CIGOTOS

2n, 2C



Estos 8 cigotos proceden de la fertilización por un mismo gameto. Si se considera el total de posibilidades, de unos mismos padres y madre hay 8 posibles gametos masculinos y 8 posibles gametos femeninos, lo que da un total de 64 combinaciones en el cigoto. Por tanto, en humanos, 2^{23} posibles gametos masculinos y 2^{23} posibles gametos femeninos dan 70,4 billones de posibles cigotos.

8.5.2.4 Recombinación meiótica

Durante la profase de la meiosis I (profase I) tiene lugar un proceso de gran trascendencia en cuanto a la herencia y a la generación de la diversidad genética de cada individuo: se trata de la **recombinación meiótica**, o simplemente **recombinación**. Los detalles de este proceso no se tratan aquí, pero básicamente consiste en el *sobrecrezamiento*, o intercambio de regiones de DNA entre cromátidas de los dos cromosomas homólogos (es decir, nunca hay recombinación entre cromátidas de distinto par). De ahí que se llame también **recombinación homóloga**.



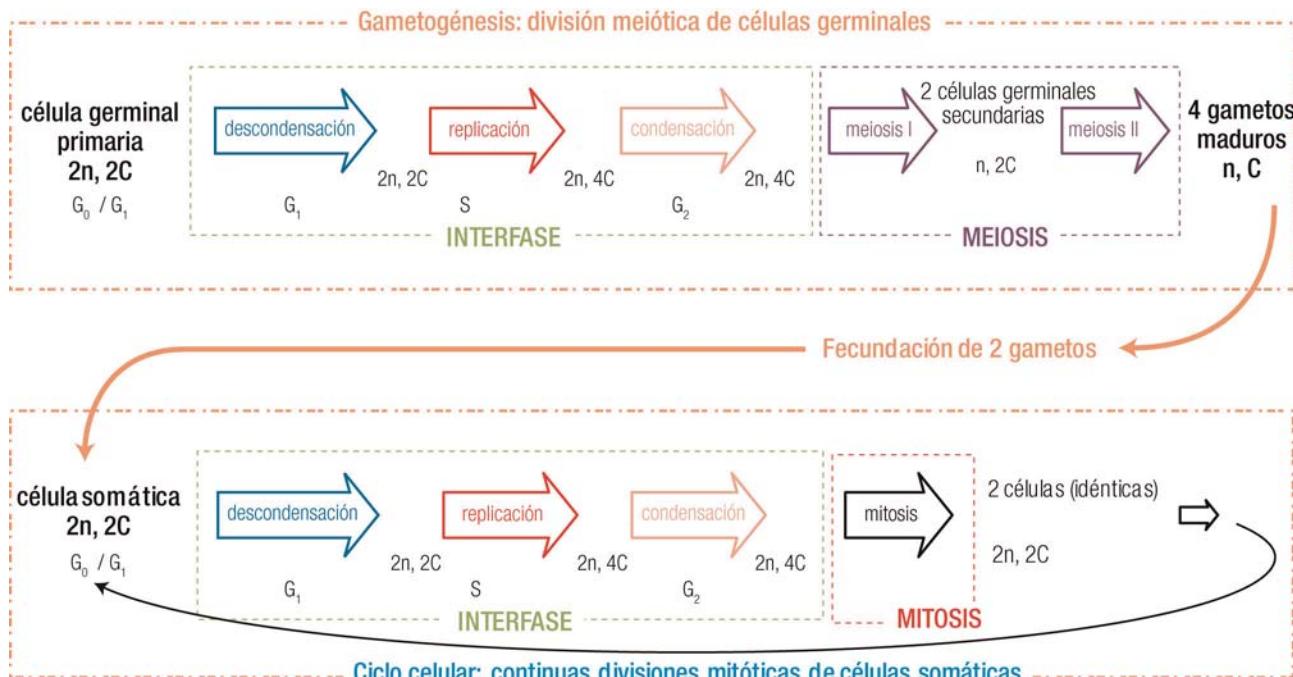
Web 8.2. Alteración de genotipos como resultado de la recombinación.

8.6 UNIÓN DE LOS GAMETOS HAPLOIDES DURANTE LA FECUNDACIÓN

Cuando los dos gametos (n, C) se unen en la **fecundación**, el pronúcleo macho del espermatozoide se funde con el pronúcleo hembra del óvulo, combinando sus dos genomas haploides para dar un genoma $2n, 2C$ en el núcleo de la célula huevo o cigoto. Ésta es, así, la primera célula somática propiamente dicha. Como tal, entra en la fase G_1 y posteriores de la interfase. A partir del cigoto se forman en definitiva todas las células somáticas (y también las germinales primordiales). El mecanismo que rige la transformación de una simple célula, el cigoto, en un individuo completo constituye un tema de gran interés biológico y poco conocido aún desde el punto de vista molecular. Tras múltiples divisiones celulares, se origina un agregado globular de células embrionarias, las cuales se autoorganizan y especializan hasta formar un fantástico conjunto de células (estimado entre 50 y 100 billones para un ser humano adulto), repartidas en alrededor de 250 tipos celulares diferentes, cada uno con una función y disposición concretas, en las distintas estructuras del organismo completo (cerebro, ojos, nariz, corazón, etc.). La biología del desarrollo es un ámbito de activa investigación y se han podido identificar genes que controlan, por ejemplo, la disposición dorsal/ventral o derecha/izquierda de algunos órganos; sin embargo, queda aún mucho por conocer sobre los mecanismos moleculares que gobiernan la organogénesis.

8.7 RESUMEN INTEGRADO

Para terminar, el siguiente esquema resume los distintos procesos estudiados, recogiendo de una forma unificada los cambios en la cantidad de material genético que sufre la célula durante los procesos de fecundación, división meiótica y mitótica y ciclo celular.





Organización del genoma eucariótico

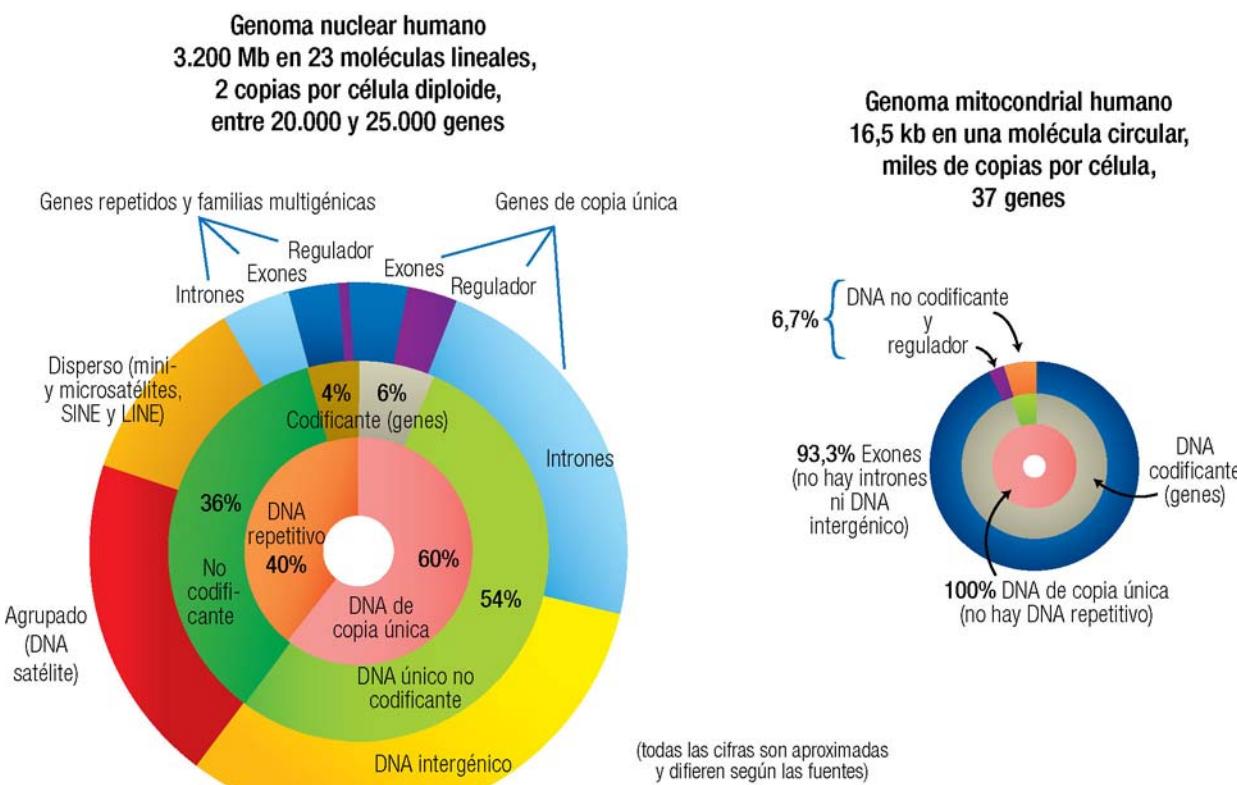
9.1 ESTRUCTURACIÓN DEL GENOMA EUCARÍÓTICO	109	9.4.2 DNA repetitivo no codificante	116
9.2 DNA CODIFICANTE: INTRODUCCIÓN AL CONCEPTO DE GEN	110	9.4.2.1 DNA altamente repetitivo y agrupado: DNA satélite	116
9.3 DNA DE COPIA ÚNICA, SIMPLE O NO REPETITIVO	111	9.4.2.2 DNA moderadamente repetitivo y disperso	117
9.4 DNA REPETITIVO	111	a) Bloques dispersos de repeticiones en tandem: DNA minisatélite y DNA microsatélite	117
9.4.1 DNA repetitivo codificante	112	b) Repeticiones dispersas: SINE y LINE	118
9.4.1.1 DNA repetitivo codificante agrupado	112		
a) Familias génicas clásicas	112		
b) Familias multigénicas de genes agrupados	113		
9.4.1.2 DNA repetitivo codificante disperso: familias multigénicas de genes dispersos	115		
		9.5 ESTUDIO EXPERIMENTAL DE LA COMPLEJIDAD EN EL GENOMA EUCARÍÓTICO	118

En los procariotas prácticamente todo el DNA existe en forma de copia única y codifica productos génicos (proteínas o RNA). Sin embargo, esto no se cumple en el genoma nuclear de eucariotas; por ejemplo, se calcula que sólo 100 Mb del total de 3.200 Mb del genoma humano haploide son codificantes. Además, el genoma eucariótico se caracteriza por la repetición de secuencias. El objetivo de este capítulo es plantear esta distinción entre *DNA de copia única* y *DNA repetitivo*, así como estudiar con detalle este último en el genoma humano. El carácter *codificante*, aunque se incluye también aquí, se comprenderá mejor en capítulos posteriores, al estudiar la expresión génica (transcripción, maduración y traducción). Los aspectos experimentales que han permitido la demostración de la presencia de DNA repetitivo en el genoma eucariótico se describen al final del capítulo. Basándose en toda esta información, se podrá abordar más adelante su aplicación a problemas biotecnológicos en general y clínicos en particular.

9.1 ESTRUCTURACIÓN DEL GENOMA EUCARIÓNICO

Para el estudio de la distribución de secuencias dentro del genoma humano éste puede subdividirse en distintas categorías de DNA en función de su repetitividad y de su carácter codificante:

DNA de copia única	Codificante: genes	Estructural: exones (codifica proteínas o RNA) Regulador (controla la expresión del gen)
	No codificante (función desconocida)	Intrones o DNA intragénico (intercalado dentro de los genes) DNA intergénico (separa los genes)
DNA repetitivo	Codificante	Agrupado Genes repetidos en tandem (familias génicas clásicas) Familias multigénicas agrupadas Disperso: Familias multigénicas dispersas
	No codificante	Agrupado (en regiones heterocromáticas) y altamente repetitivo Disperso (por todo el genoma) y moderadamente repetitivo Bloques dispersos de repeticiones en tandem Repeticiones dispersas Minisatélites Microsatélites SINE: secuencias <i>Alu</i> y otras LINE: secuencias <i>Kpn</i> y otras



9.2

A diferencia del nuclear, el genoma de la mitocondria, de tamaño muy pequeño (16.569 pb en humanos, por tanto 200.000 veces inferior al genoma nuclear haploide), está formado casi en su totalidad por DNA codificante.

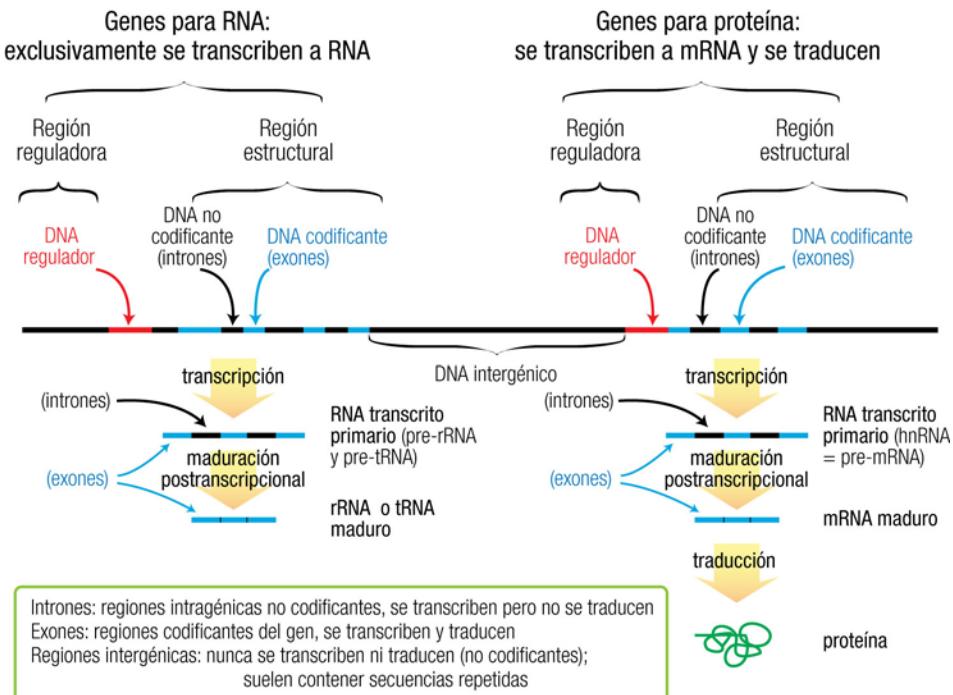
Todas estas cifras que cuantifican la proporción de DNA codificante y los distintos subtipos de DNA repetitivo varían de forma notable entre diferentes textos. La razón de esta discrepancia puede proceder, entre otros factores, de si se cuentan como codificantes los intrones o las regiones de control de la expresión génica. A todo ello se suma la incapacidad, al menos en este momento, para asignar fehacientemente todas las secuencias del genoma a un producto génico o función concretos. Respecto a las secuencias repetitivas, la clasificación aquí planteada en términos de tamaño, número de las repeticiones y ubicación, posee, obviamente, fronteras no siempre precisas.

Por tanto, las cifras aquí recogidas deben tomarse como una orientación con un interés didáctico más que como valores certeros. La relevancia de conocer esta realidad de los genomas eucarióticos no depende de la frecuencia precisa de cada tipo de secuencia.

9.2 DNA CODIFICANTE: INTRODUCCIÓN AL CONCEPTO DE GEN

El concepto de gen, aunque ha sufrido una gran evolución en eucariotas, haciéndose más complicado y con múltiples componentes, se sigue definiendo bajo un criterio estrictamente funcional, siempre en relación con la transmisión de la información desde DNA a RNA y, en su caso, desde mRNA a proteínas. El estudio detallado de las características de los genes eucarióticos se deja para más adelante, cuando se analice su expresión (Capítulo 17).

Brevemente, un gen es aquella región del DNA necesaria para que pueda sintetizarse una macromolécula funcional, un **producto génico**; éste puede ser bien una proteína, bien un RNA. Más en concreto, un gen comprende tanto *secuencias estructurales* como *secuencias reguladoras*. Las primeras –comúnmente consideradas como la parte **codificante** del gen– se transcriben de forma exclusiva en RNA (región estructural de genes de tRNA, rRNA o RNA pequeños), o bien se transcriben en un mRNA que, a su vez, se traduce a proteína. Las segundas actúan en ambos casos controlando la transcripción de las regiones estructurales.



9.3

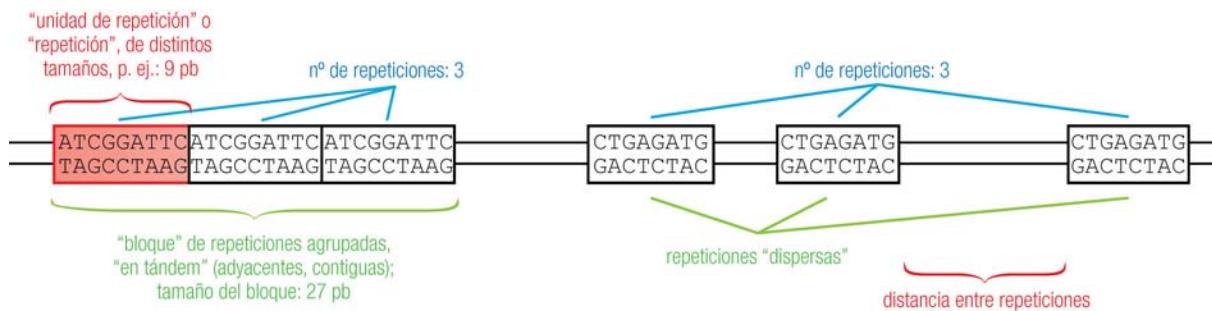
9.3 DNA DE COPIA ÚNICA, SIMPLE O NO REPETITIVO

El DNA no repetitivo constituye la mayor parte del genoma, aunque en una proporción variada dependiendo del tipo de organismo (100% en procariotas, 80% en eucariotas inferiores, 50-70% en animales superiores). Parte de este DNA ($\approx 5\%$) constituye las secuencias de genes, que codifican los RNA y proteínas celulares; otra parte ($\approx 5\%$) es responsable del control de la expresión de esas secuencias, mientras que el resto, mayoritario ($\approx 50\%$), es DNA no codificante, cuya función, si existe, apenas se conoce.

El DNA de copia única se asoció durante mucho tiempo al concepto de gen; hoy se sabe que también una pequeña parte del DNA repetitivo forma algunos genes eucarióticos, como se estudiará a continuación.

9.4 DNA REPETITIVO

Las secuencias de DNA que aparecen en copias múltiples, o DNA repetitivo, constituyen entre un 30 y un 50% del total del genoma humano, dependiendo de las fuentes. Esta abundancia es una característica determinante de la gran complejidad del genoma eucariótico. Las secuencias repetidas, llamadas **unidades de repetición** o simplemente **repeticiones**, poseen tamaños diversos y cada una aparece de forma idéntica o casi idéntica muchas veces en el genoma. Su distribución puede darse en **forma dispersa** por todo el genoma, entremezcladas con las secuencias de copia única, o bien en **forma agrupada**, localizada en regiones concretas de un cromosoma. En ambos casos, el número de copias de la unidad de repetición varía desde unas centenas o miles (DNA **moderadamente repetitivo**) hasta cientos de miles (DNA **altamente repetitivo**). Esto coincide en muchos casos con una distribución en forma dispersa y agrupada, respectivamente.



Una parte del DNA repetitivo tiene carácter **codificante**, es decir, contiene la información para expresar un producto funcional (RNA o proteína). Para el resto del DNA repetitivo no se conoce una función clara; alguno posiblemente contribuya a mantener la estructura de los cromosomas, mientras que se ha llegado a proponer que una parte sea *DNA chatarra* o *DNA basura*, un vestigio evolutivo sin función actual. Finalmente, debe destacarse que algunas de las secuencias repetidas no codificantes (en concreto, las denominadas minisatélites y microsatélites), aun con función desconocida, tienen gran relevancia aplicada en pruebas de identificación y en estudios familiares, debido a su variabilidad entre individuos (polimorfismo, Capítulo 24).

La repetitividad del DNA posee probablemente un origen evolutivo. Por un lado, las repeticiones agrupadas pueden surgir debido a errores en la replicación o en la recombinación genética. Por otro, la separación entre repeticiones dispersas puede provenir de translocaciones o transposiciones cromosómicas.

9.4.1 DNA repetitivo codificante

Se le llama también **DNA repetitivo funcional**, pues forma genes que se expresan (aunque no todos ellos, como se verá). Su tamaño se estima alrededor del 10% del genoma (esta cifra incluye probablemente las regiones de control e intrones respectivos, por lo que la fracción realmente codificante será inferior). Aparece en forma de “familias de genes”, cuyos miembros se caracterizan por su homología, al haberse originado mediante duplicaciones (pág. 407) y variaciones de un gen ancestral.

Tipos de DNA repetitivo codificante y ejemplos característicos:	
1. Agrupado:	
a) Familias génicas clásicas o conservadas, con genes repetidos en tandem (contiguos)	Histonas: Ej. 1 rRNA: Ej. 2 tRNA snRNA
Muestran un alto grado de homología de secuencia, es decir, son copias prácticamente idénticas, lo que indica su relación evolutiva y funcional. Se expresan todas las repeticiones.	
b) Familias multigénicas con genes agrupados (no contiguos)	Globinas: Ej. 3 HLA-1 (y otras muchas)
Situados en localizaciones específicas del genoma (pero sin relación con las regiones centroméricas y teloméricas). Sus repeticiones presentan menor homología: variantes, pseudogenes, genes truncados y fragmentos de genes. Sólo se expresan algunas de las repeticiones.	
2. Disperso: Familias multigénicas con genes dispersos	Aldolasa Actina Ferritina GA3PDH
Normalmente, cada familia está formada por un número pequeño de repeticiones repartidas por todo el genoma (incluso en cromosomas distintos), que casi siempre son todas funcionales.	

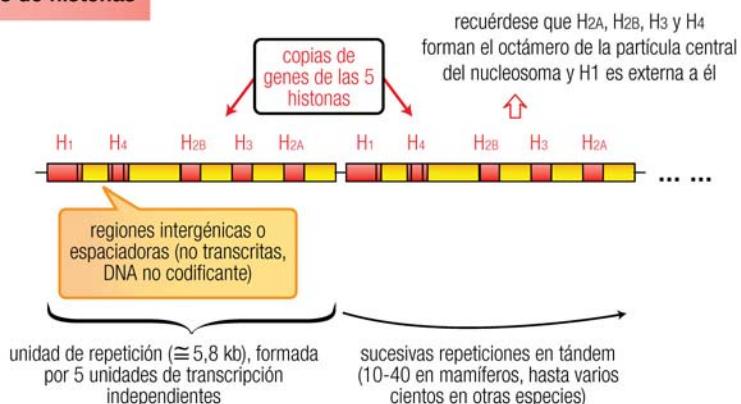
9.4.1.1 DNA repetitivo codificante agrupado

a) Familias génicas clásicas

Contienen secuencias repetidas que codifican un mismo RNA o proteína, que puede, así, sintetizarse con rapidez y en gran cantidad. Son ejemplos característicos los genes de histonas y los de RNA ribosómico,

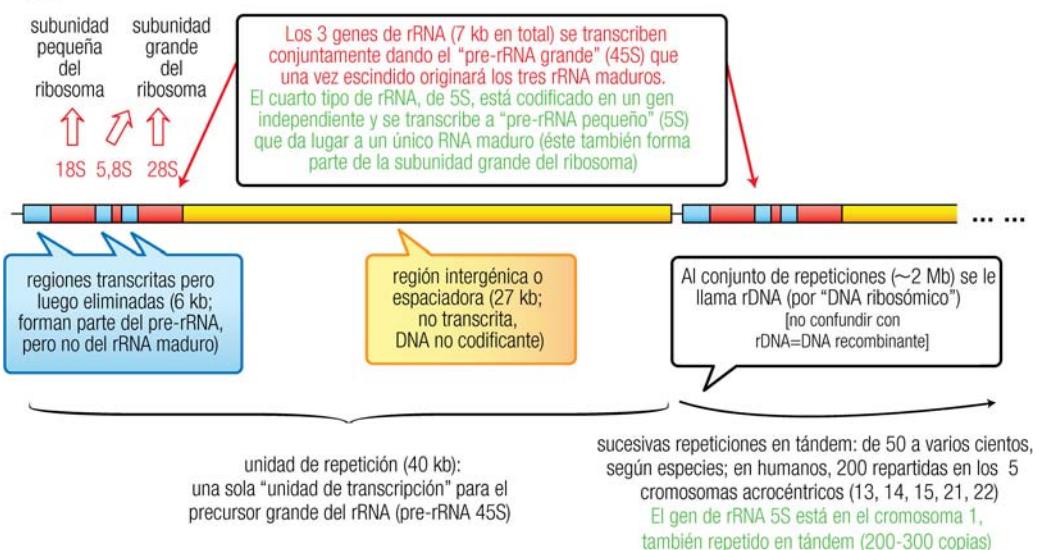
productos ambos que la célula necesita en abundancia (las histonas –pág. 84– deben sintetizarse rápidamente en la fase S del ciclo celular, para asociarse al nuevo DNA procedente de la replicación, mientras que el rRNA –capítulo 6– supone el 75% del total de RNA celular).

Ej. 1 Genes de histonas



9.5

Ej. 2 Genes de RNA ribosómico



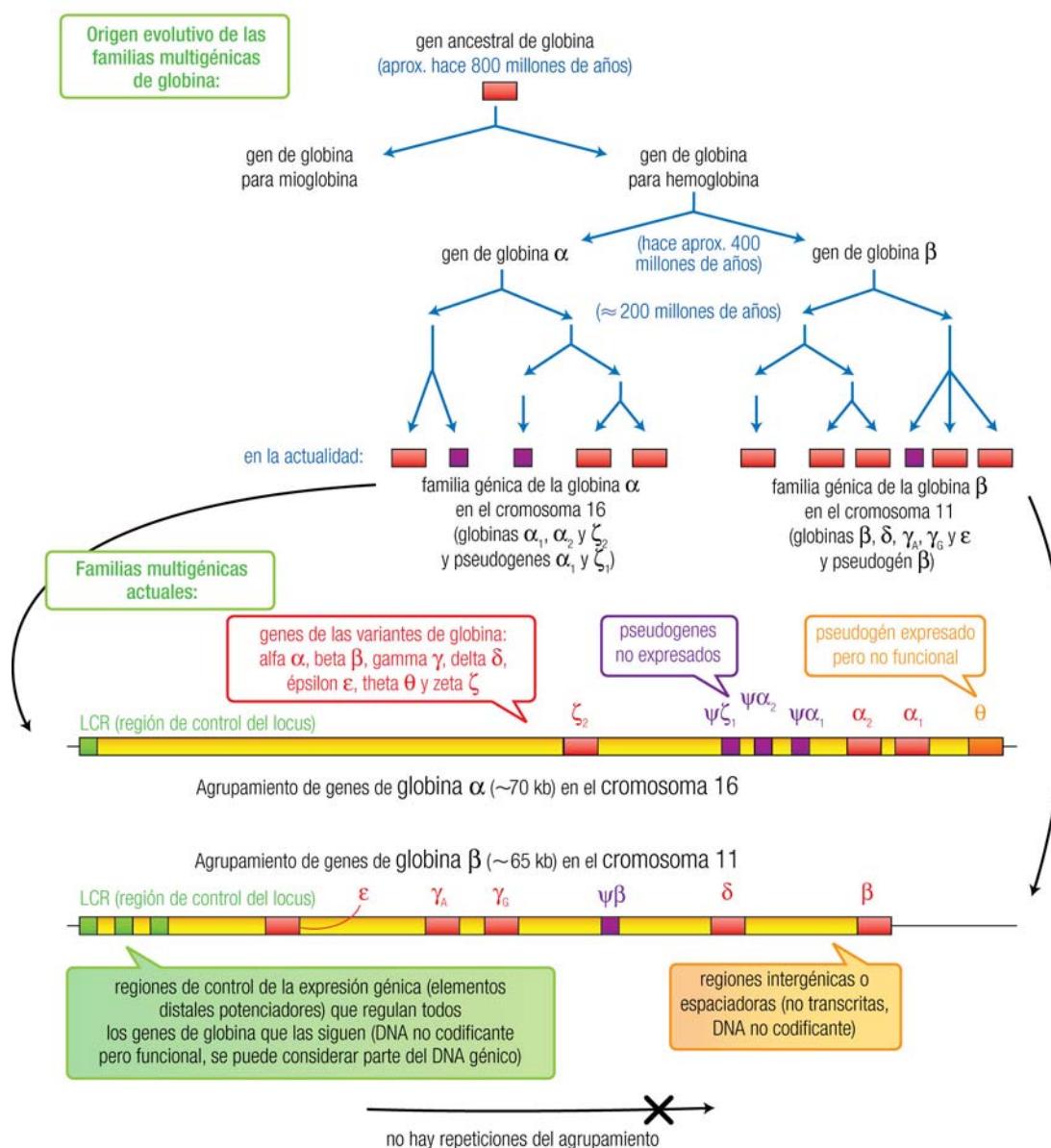
9.6

b) Familias multigénicas de genes agrupados

A diferencia de las anteriores, los miembros de cada familia son genes con cierta diversidad de secuencia. Se pueden incluir en este grupo tres tipos de familias:

- Las que presentan homología sólo en parte de la secuencia de DNA, que da lugar a proteínas con grandes regiones de secuencia y estructura comunes (*dominios*).
- Familias con homología escasa en las secuencias del DNA, lo que genera proteínas con pequeños *motivos* comunes, cada uno de pocos aminoácidos.
- Grandes “superfamilias”, con muy débil homología en la secuencia de DNA y, por tanto, en la de las proteínas, a pesar de lo cual éstas mantienen una similitud estructural (en parte de su molécula, *dominios* comunes) y una relación funcional. Los genes que forman cada superfamilia parecen estar relacionados evolutivamente, pero más distantes entre sí que los miembros que componen las familias anteriores.

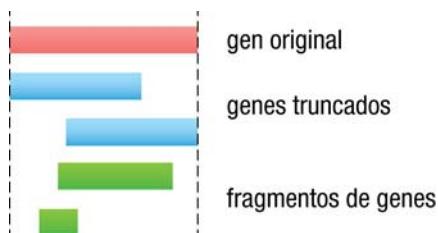
Ej. 3 Familias de genes de las cadenas α y β de globina:



9.7

Como consecuencia de los procesos evolutivos que dan lugar a las familias multigénicas, aparecen con frecuencia en ellas miembros con ciertas características diferenciales:

- **Copias de genes (variantes)** con pequeñas diferencias de secuencia, que codifican productos génicos funcionales pero de características ligeramente distintas (por ejemplo, diferencias en actividad enzimática, parámetros cinéticos, especificidad por sustrato, afinidad por su ligando, etc.). Es común que se active de forma alternativa la expresión de una u otra variante en función del estadio de desarrollo o de diferenciación celular, el tipo de tejido, etc.



9.8

- **Pseudogenes:** se aplica este término (abreviado como ψ) a copias inactivas de un gen, es decir, que nunca se expresan, o que, aun expresándose, no son funcionales. En general, la inactivación se ha producido por acumulación de mutaciones durante la evolución en lo que inicialmente era un duplicado del gen. Existen tipos diversos de pseudogenes, en función de si se expresan y si sufren procesamiento postranscripcional, así como de su origen evolutivo.
- **Genes truncados:** se trata de copias incompletas del gen, por pérdida de una región situada en uno de los extremos (5' o 3').
- **Fragментos de genes:** en este caso se conserva sólo una porción del interior de la secuencia del gen (se han perdido ambos extremos).

Obviamente, sólo las copias idénticas (familias clásicas) y las variantes conducen a productos génicos funcionales, mientras que los pseudogenes, genes truncados y fragmentos de genes parecen ser vestigios evolutivos sin función actual.

Los genes de la globina constituyen uno de los ejemplos más representativos de familias multigénicas con genes agrupados, implicados en una función tan importante como es el transporte de oxígeno en la corriente sanguínea. Su estudio incluye consideraciones evolutivas, de desarrollo y diferenciación, de regulación de la expresión génica, etc.

Este diagrama ilustra la regulación y expresión de las familias multigénicas de genes de globina α y β a lo largo del desarrollo humano. Se muestra un cuadro resumen que enumera los genes expresados en cada etapa, las regiones de control (LCR) activadas y las hemoglobinas sintetizadas.

Etapa de desarrollo y órgano donde se sintetizan las globinas	Genes que se expresan en la familia génica de globina α (en el cromosoma 16)	Genes que se expresan en la familia génica de globina β (en el cromosoma 11)	Hemoglobinas sintetizadas
Embrión (saco vitelino)			 Mayoritaria: globina ε, Hb Gower I Mayoritaria: globina ζ, Hb Portland I
Feto (hígado y bazo)			 Mayoritaria: globina α, γ, Hb F Mayoritaria: globina α, β, Hb A
Adulto (médula ósea)	Majoritaria: globina α.		 Mayoritaria: globina α, β, Hb A Mayoritaria: globina α, δ, Hb A₂

En cada etapa del desarrollo humano es un órgano diferente el único que expresa los genes de globina. Además, en cada órgano se activan genes diferentes.

Las "regiones de control del locus" (LCR) de ambas familias multigénicas activan la expresión de uno u otro de los genes dependiendo de la etapa de desarrollo y órgano.

Cada etapa del desarrollo se caracteriza, por tanto, por la presencia de diferentes hemoglobinas. P.ej., HbF (fetal), HbA (adulta).

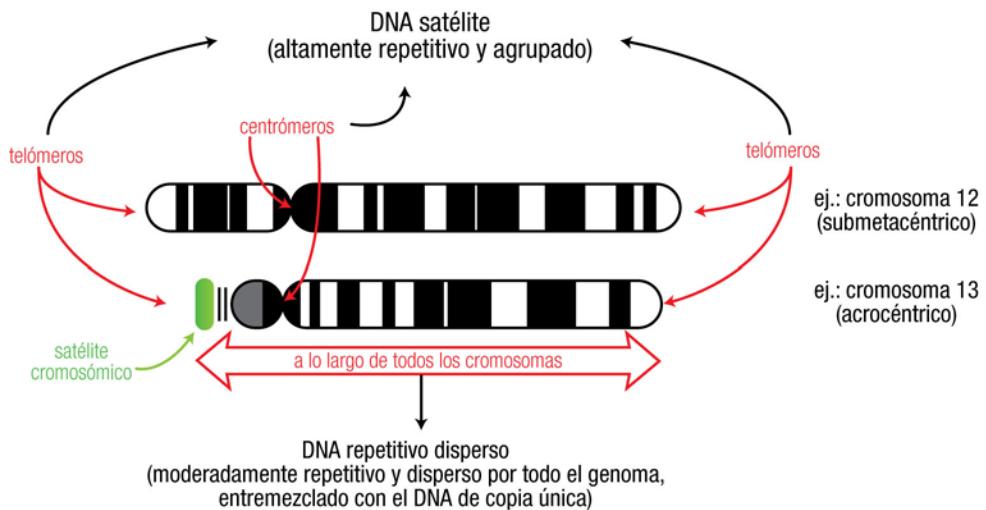
9.9

9.4.1.2 DNA repetitivo codificante disperso: familias multigénicas de genes dispersos

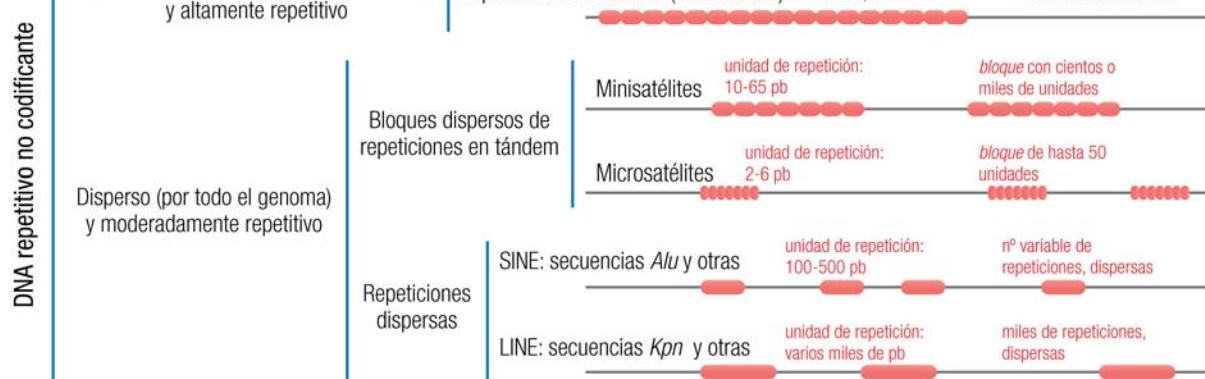
Mientras que en las familias multigénicas anteriores las repeticiones se disponen **agrupadas**, próximas en la secuencia lineal del genoma, en otras los genes que forman la familia (copias idénticas, variantes, pseudogenes, etc.) se localizan de forma **dispersa** por el genoma, a menudo incluso en cromosomas distintos. Entre estas familias se encuentran genes que codifican proteínas de funciones diversas: enzimáticas, reguladoras, de almacenamiento, estructurales, etc. Por ejemplo, la familia de aldolasa comprende 3 genes funcionales y 2 pseudogenes, repartidos en 5 cromosomas; los genes de actina forman otra familia con 6 genes y más de 20 pseudogenes en 12 cromosomas.

9.4.2 DNA repetitivo no codificante

Ésta es la fracción del genoma a la que el término *DNA repetitivo* se refiere de forma más habitual. Su tamaño se estima entre un 30 y un 40% del DNA genómico total. Como ya se ha indicado, se subdivide en dos categorías: el moderadamente repetitivo, que se ubica en forma dispersa, y el altamente repetitivo, que además es agrupado. Su distribución general en el genoma puede representarse de forma gráfica como se muestra en la figura, para dos cromosomas representativos.



9.10



9.11

9.4.2.1 DNA altamente repetitivo y agrupado: DNA satélite

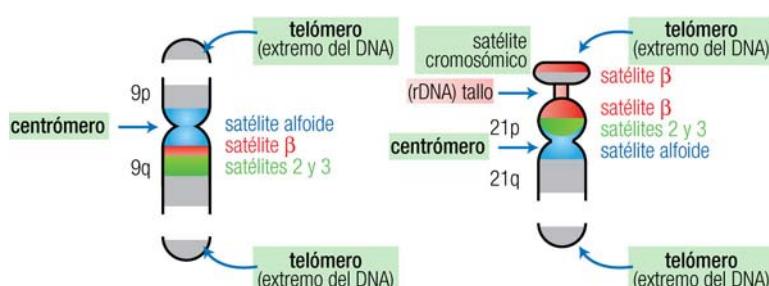
Supone entre un 10 y un 20% del genoma, y se denomina también **DNA repetitivo en tandem**. No debe confundirse con los DNA *minisatélite* y *microsatélite* (ambos corresponden al DNA moderadamente repetitivo y disperso, pág. 117) ni con los *satélites cromosómicos*, las pequeñas masas de cromatina que forman el extremo de los brazos cortos de un cromosoma acrocéntrico (pág. 89).

Este tipo de DNA está formado por unidades de pequeño tamaño (entre 2 y 50 pb) repetidas en tandem entre miles y un millón de veces, con lo que el bloque de repetición abarca cientos de kb o incluso varias Mb. También es característica su **localización agrupada** en lugares específicos de varios cromosomas, correspondientes a las **regiones heterocromáticas**, principalmente en torno al centrómero y en los telómeros. Se conoce el papel de este DNA satélite en las regiones teloméricas (págs. 94 y 158), pero aún no su función exacta en el centrómero.

El nombre de *DNA satélite* tiene un origen metodológico: debido a su alta tasa de repetición, las regiones de este tipo de DNA poseen una composición de bases diferente a la del resto del genoma, que es más promediada. Cuando el DNA total se fragmenta (por ejemplo, por fuerzas de cizalla) y se analiza en gradientes de densidad

de cloruro de cesio (págs. 137-138), los fragmentos procedentes de las regiones de DNA satélite, que poseen menor densidad debido a su menor contenido en G + C, se separan de la banda principal de DNA formando “*bandas satélites*”. Cada una de ellas corresponde a una subfamilia de DNA satélite (como 1, 2, 3, α y β), con distinta secuencia de repetición o número de repeticiones.

Ejemplos de ubicación de DNA satélite en centrómeros y telómeros



9.12

subfamilia o tipo de satélite	tamaño de la unidad de repetición
satélite 1	25-48 pb
satélites 2 y 3	5 pb
satélite α o alfoide	171 pb
satélite β	69 pb
telómero (humano)	6 pb

9.4.2.2 DNA moderadamente repetitivo y disperso

Como indica su nombre, este tipo de DNA se distribuye a lo largo de todos los cromosomas, con un número de repeticiones no muy elevado (entre 10^2 y 10^4). Se puede subdividir en dos categorías, en función de la forma en que se distribuyen las repeticiones: en primer lugar, algunas secuencias se repiten en tandem formando bloques, los cuales aparecen de forma dispersa por el genoma. El segundo tipo consiste en repeticiones igualmente dispersas por todo el genoma, pero que no se agrupan en bloques.

a) Bloques dispersos de repeticiones en tandem: DNA minisatélite y DNA microsatélite

Se distinguen estos dos subgrupos en función del tamaño de la unidad de repetición. Constituyen conjuntamente entre el 15 y el 20% del genoma. A pesar de la similitud de nombre, no deben confundirse con el *DNA satélite* (pág. 116), del que se diferencian por estar localizados de forma dispersa por el genoma, aunque se asemejan a él en la presencia de repeticiones en tandem.

La importancia práctica del DNA minisatélite y microsatélite radica en su gran variabilidad entre individuos, que permite utilizarlo como marcador molecular en medicina forense, pruebas de paternidad y diagnóstico de enfermedades moleculares; estas aplicaciones se describen más adelante al hablar de polimorfismo genético (pág. 424).

a1) DNA minisatélite:

Está formado por repeticiones de 10 a 65 pb, ricas en G + C, agrupadas en tandem formando bloques relativamente grandes, de cientos o miles de repeticiones; dichos bloques se encuentran repartidos por todo el genoma.

Algunas de estas repeticiones, denominadas “*DNA minisatélite hipervariabile*”, se caracterizan por su elevado polimorfismo (diferencias tanto en secuencia como en número de repeticiones), no sólo entre individuos de una especie, sino también incluso entre dos cromosomas homólogos de un individuo. Sus repeticiones, de 10 a 24 pb, presentan la secuencia consenso GGGCAGGANG (donde N representa cualquier nucleótido). Este DNA es importante por las aplicaciones derivadas de su polimorfismo, principalmente la obtención de **huellas de DNA**, donde una sonda para la secuencia consenso puede hibridar de manera simultánea con múltiples *loci* minisatélites diferentes distribuidos por todos los cromosomas, lo que da lugar a un esquema o patrón de hibridación distinto para cada individuo (pág. 425).

Algunos autores consideran al DNA telomérico como parte del minisatélite. No obstante, según los criterios de clasificación adoptados en este texto, ni el tamaño de su unidad de repetición ni su disposición agrupada en pocos lugares del cromosoma permiten su asignación a esta categoría.

a2) DNA microsatélite:

Se aplica este término cuando la unidad de repetición es inferior a 7 pb. Se presentan agrupadas en tandem en bloques de hasta 50 repeticiones, también distribuidos de forma dispersa por el genoma. Su polimorfismo se aplica igualmente para la obtención de huellas genéticas, para pruebas de identidad y para estudios familiares.

Aunque las secuencias teloméricas se han clasificado anteriormente como DNA repetitivo agrupado (DNA satélite), por su ubicación en puntos concretos del cromosoma, algunos autores las consideran DNA microsatélite por el hecho de compartir con éste características de tamaño, número y disposición en tandem de las repeticiones.

b) Repeticiones dispersas: SINE y LINE

De nuevo, la longitud de la unidad de repetición sirve de criterio para subdividirlas:

b1) Secuencias SINE:

Las siglas proceden de *Short Interspersed Nuclear Elements*, elementos nucleares dispersos cortos. Se habla comúnmente de unidades de entre 100 y 500 pb, repetidas hasta 20 veces. Sin embargo, el ejemplo más característico, el elemento *Alu*, de 300 pb, aparece en el genoma humano entre medio millón y un millón de veces. Su nombre se debe a que contiene un sitio diana para la enzima de restricción *Alu* I. Se trata en realidad de una familia de secuencias, con un 85% de homología, que supone por sí sola alrededor de un 5% del genoma, de modo que, en promedio, existe una secuencia *Alu* cada 4 o 5 kb.

Se cree que los elementos *Alu* se originaron evolutivamente por la integración en el genoma de múltiples copias (*retroposones*) formadas por transcripción inversa del gen del RNA 7SL, un RNA citoplásmico pequeño (scRNA) componente de la *partícula de reconocimiento de señal* (pág. 357). A pesar de la homología (parcial) de secuencia, las secuencias *Alu* no ejercen la función de este scRNA.

b2) Secuencias LINE:

En este caso, se trata de elementos nucleares dispersos largos (*Long Interspersed Nuclear Elements*), con unidades de mayor tamaño, generalmente varios miles de pb, que se repiten hasta 50.000 veces en forma dispersa. La principal familia LINE en humanos la constituyen las secuencias *Kpn*, también denominadas *L1* o *LINE-1*. Se encuentran en el genoma de 50.000 a 100.000 repeticiones dispersas, de tamaños comprendidos entre 1,4 y 6,1 kb cada una, debido a su diversidad.

9.5 ESTUDIO EXPERIMENTAL DE LA COMPLEJIDAD EN EL GENOMA EUCARÍÓTICO

El descubrimiento y análisis del DNA repetitivo, característico del genoma eucariótico, se ha basado fundamentalmente en las técnicas de reasociación del DNA desnaturizado.



Web 9.1. Estudio experimental de la complejidad en el genoma eucariótico.



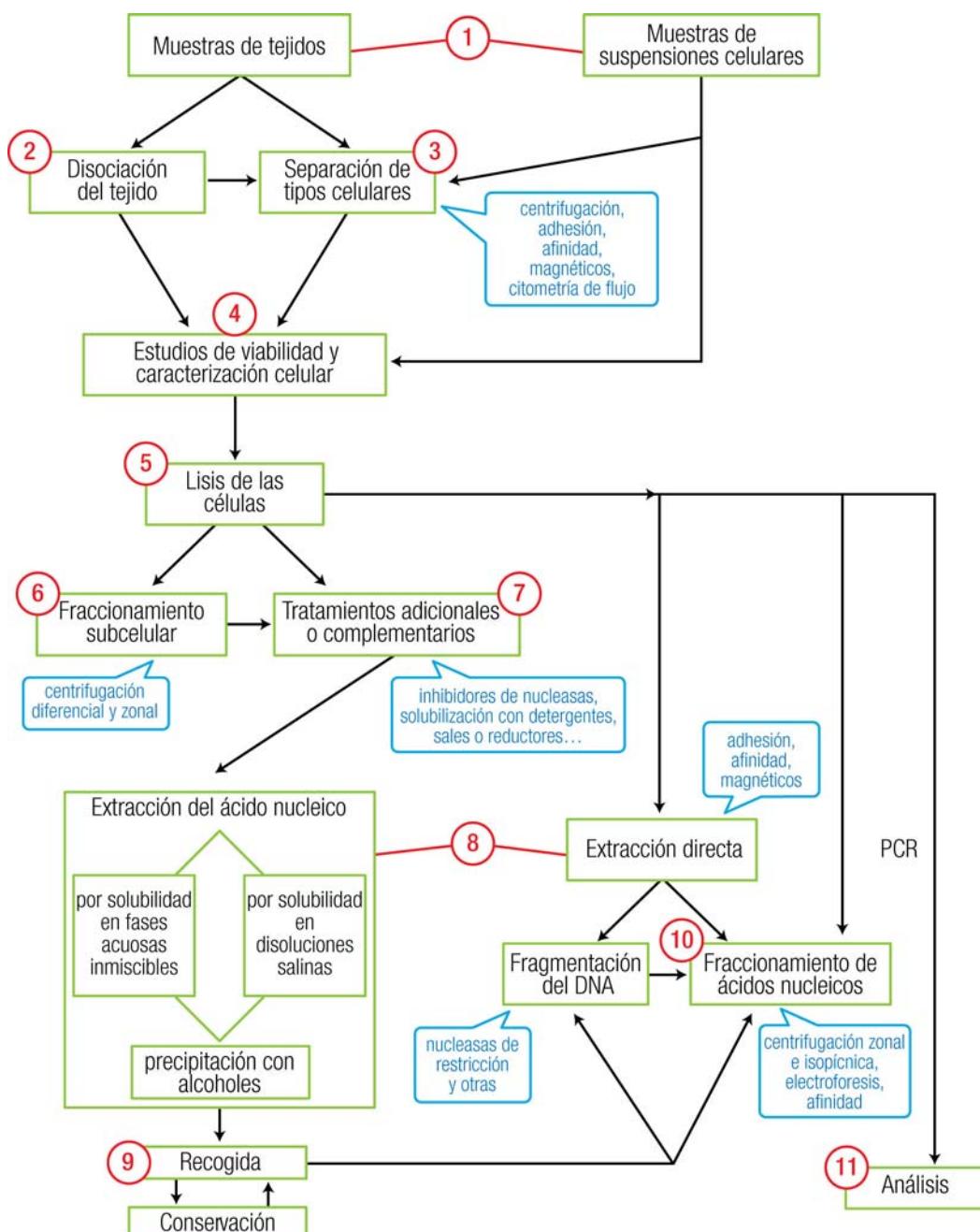
Preparación de muestras, extracción y análisis de ácidos nucleicos

10.1 OBTENCIÓN Y PREPARACIÓN PRELIMINAR DE LAS MUESTRAS	
10.1.1 Tejidos sólidos	121
10.1.2 Células embrionarias	121
10.1.2.1 Células del embrión no implantado	122
10.1.2.2 Líquido amniótico	122
10.1.2.3 Tejidos o sangre fetales	122
10.1.3 Células de sangre o de médula ósea adultas	122
10.1.4 Muestras bucales	122
10.1.5 Semen	123
10.1.6 Células en cultivo	123
10.1.7 Otros líquidos extravasculares	123
10.1.7.1 Líquido cefalorraquídeo	123
10.1.7.2 Líquidos o derrames serosos	123
10.1.7.3 Líquido articular o sinovial	124
10.1.7.4 Muestras biológicas menos habituales	124
10.2 DISOCIACIÓN DE LA MUESTRA TISULAR	124
10.3 SEPARACIÓN DE CÉLULAS	125
10.3.1 Separación de células por métodos de sedimentación y centrifugación	125
10.3.1.1 Centrifugación diferencial	125
10.3.1.2 Centrifugación en gradiente de densidad	126
a) Métodos de barrera: centrifugación a través de un medio de densidad constante	126
b) Centrifugación zonal o de velocidad de sedimentación	126
c) Centrifugación isopícnica o de equilibrio de sedimentación	126
10.3.1.3 Elutriación centrífuga	127
10.3.2 Caracterización y separación de células por citofluorimetría de flujo	127
10.3.3 Separación directa de tipos celulares	129
10.3.3.1 Separación por afinidad con partículas magnéticas	129
10.4 CARACTERIZACIÓN CELULAR Y MEDIDAS DE VIABILIDAD	131
10.4.1 Ensayos de exclusión	131
10.4.2 Recuento de células viables por incorporación de un marcador	132
10.4.3 Ensayos por medida de la función metabólica	133
10.5 LISIS DE LAS CÉLULAS	134
10.6 PREPARACIÓN DE FRACCIONES SUBCELULARES	134
10.7 TRATAMIENTOS ADICIONALES O COMPLEMENTARIOS	135
10.8 EXTRACCIÓN Y PURIFICACIÓN DE ÁCIDOS NUCLEICOS	135
10.8.1 Extracción de ácidos nucleicos por solubilidad	136
10.8.1.1 Precipitación con alcoholes	136
10.8.1.2 Extracción del DNA mediante fases inmiscibles	136
10.8.1.3 Extracción del RNA	137
10.8.1.4 Purificación de ácidos nucleicos por precipitación salina diferencial	137
10.8.2 Extracción directa de ácidos nucleicos	137
10.9 RECOGIDA Y CONSERVACIÓN DE LAS MUESTRAS	139
10.10 FRACCIONAMIENTO DE ÁCIDOS NUCLEICOS	139
10.10.1 Ultracentrifugación	139
10.10.1.1 Centrifugación zonal	139
10.10.1.2 Centrifugación isopícnica	139
10.10.2 Electroforesis	140
10.10.2.1 Soporte para la electroforesis	141
10.10.2.2 Detección del DNA tras su separación ("revelado")	141
10.10.2.3 Calibración de la electroforesis para el cálculo del tamaño de los fragmentos de DNA	141
10.10.2.4 Electroforesis de campo pulsante	142

Las técnicas derivadas de la biología molecular y la ingeniería genética se utilizan de forma creciente tanto con fines básicos como aplicados, fundamentalmente al diagnóstico clínico, partiendo de muestras biológicas muy diversas. Esto requiere la extracción y análisis de RNA y DNA y, en su caso, de proteínas u otras macromoléculas. Todos estos aspectos, como en todo proceso analítico, se desarrollan en tres fases:

- **Preanalítica**, desde la obtención y preparación preliminar de la muestra inicial hasta su transporte al laboratorio.
- **Analítica** o de aplicación de procedimientos para la extracción de los ácidos nucleicos, sobre los que se realiza luego el análisis bioquímico-genético propiamente dicho: toma y tratamiento de datos, y recogida de resultados en un informe.
- **Postanalítica** o de validación técnica del informe analítico. El conjunto de esta fase y la preanalítica se denomina “fase extraanalítica”.

Estas fases, hoy día condicionadas por los continuos avances tecnológicos, instrumentales, de automatización e informáticos, han llegado a tal desarrollo que pueden a veces enmascarar el carácter científico que subyace a todo el proceso analítico. Dentro del planteamiento de este libro, corresponde considerar solamente y de forma breve los aspectos esenciales de las fases preanalítica y analítica. Los detalles concretos, descritos en libros especializados, protocolos de laboratorio, etc., son objeto de estudio teórico-práctico durante los períodos de formación y de ejercicio profesional o científico en diversas áreas de laboratorio. Aunque de difícil delimitación y contenido (por ejemplo, a veces se llama *procesado de la muestra* a la preparación preliminar, pero otras en esta expresión se incluye la fase analítica), ambas fases se han integrado aquí, para su descripción, bajo un esquema unificado en el que se intenta reunir todas las posibilidades, aunque éstas puedan no darse de forma simultánea en la realidad. Se establecen así una serie de etapas más o menos consecutivas, conectadas entre sí, en las cuales se aplican técnicas muy diversas, en ocasiones sencillas y comunes, a veces muy complejas y específicas.



10.1 OBTENCIÓN Y PREPARACIÓN PRELIMINAR DE LAS MUESTRAS

La muestra celular se obtiene y prepara a partir de órganos, tejidos o fluidos biológicos, en función del objetivo perseguido (clínico o de identificación de individuos), tipo de célula (somática o germinal), tipo de genoma (nuclear o mitocondrial), tipo de estudio que debe realizarse (secuencia del genoma, genes, productos génicos), etc. En cualquier caso, deben emplearse condiciones que *garanticen la calidad* de la muestra frente a las diversas variables que afectan a la fase preanalítica (recogida, anticoagulante empleado –en su caso–, manipulaciones para la conservación, transporte, etc.). En general, la preparación preliminar se hace rápidamente tras la obtención (dentro de 1 hora, como máximo 24 h), o la muestra se mantiene congelada hasta que llega al laboratorio.

10.1.1 Tejidos sólidos

Habitualmente se obtienen mediante **biopsia** (método *invasivo* de recogida, bajo condiciones quirúrgicas, de células o porciones de tejido de un organismo vivo, con fines diagnósticos). Se liberan de restos de sangre, grasa o tejidos acompañantes mediante lavados con disolución salina fisiológica fría y se envían al laboratorio, según los casos, refrigeradas o congeladas, protegidas con hielo, nieve carbónica o nitrógeno líquido. La elección del tejido y tipo de extracción depende de numerosos factores, entre otros, contenido en DNA, finalidad del estudio, o necesidad de amplificación por PCR de pequeñas muestras. En ocasiones se parte de tejidos fijados o de DNA antiguo.

Características generales de las principales muestras de tejidos y órganos de adultos

Tejidos tumorales	Para analizar marcadores de membrana (glicoproteínas, receptores, etc.), de citosol (receptores hormonales, etc.) y de núcleo (oncogenes)
Tejido hepático	Rico en DNA (15 µg/mg); para enfermedades metabólicas
Tejido muscular	Pobre en DNA (3 µg/mg); para enzimopatías
Tejido intestinal	Para enzimopatías
Extensiones de piel	Pobre en DNA (3 µg /mg). Se obtiene de zonas no vellosas
Raíz de pelo	Cantidad mínima de DNA (0,5 µg/raíz); con fines forenses; difícil de estudiar
Huesos	Pulpa de los dientes o médula de otros huesos. El tejido óseo se tritura por métodos físicos
Tejidos fijados	Muestras anatomicopatológicas, fijadas con formaldehído (formol)
Tejidos incluidos en parafina	Se debe eliminar la parafina empleando disolventes, como xileno o tolueno
DNA antiguo (restos humanos y animales)	Interés arqueológico y antropológico. Debe evitarse la contaminación por DNA exógeno e inhibidores de la polimerasa Taq (que alteran el análisis por PCR)
DNA atrapado	Originado hace millones de años. La resina vegetal (ámbar) mantiene el DNA en perfecto estado

10.1.2 Células embrionarias

El óvulo fertilizado origina en el útero, por sucesivas mitosis, estructuras de 2, 4 y 8 células totipotentes (8 *blastómeros*, que forman la *mórula*) y luego de 16, 32 y 64 células pluripotentes. A partir de ahí se forma una estructura hueca denominada *blastocisto* o *blastodermo*, a partir de la cual comienza una especialización marcada. La cavidad interior del blastocisto, rellena de líquido, se denomina *blastocito*.



Web 10.1. Desarrollo embrionario y muestras celulares.

El blastocisto (formado por entre 70 y 100 células) pasa a diferenciarse en una capa externa de células o *trofoblasto* y una masa celular interna o *embrioblasto*. Las células de la masa interna del blastocisto originarán el embrión. Unos 5 días tras la fecundación se produce la implantación, cuando el trofoblasto invade la pared del endometrio (que reviste internamente el útero), dando lugar a la placenta. Como parte de este proceso, se forman las *vellosidades coriónicas* (v. a continuación), que penetran en la placenta para permitir la circulación fetal.

10.1.2.1 Células del embrión no implantado

Para observar si el embrión es portador de un trastorno genético se puede obtener DNA extrayendo 1 o 2 blastómeros de la mórula (lo que no afecta al desarrollo de las otras células). Esto permite el **diagnóstico preimplantatorio**.

De la masa celular interna del blastocisto pueden obtenerse células troncales pluripotentes, de elevado interés potencial para trasplantes, ya que pueden dar lugar en cultivo a células especializadas de distintos tipos. Las muestras de este origen se conocen comúnmente como células troncales (o células madre) embrionarias (págs. 196 y 198).

10.1.2.2 Líquido amniótico

A partir de la primera semana de gestación, el embrión queda protegido al estar inmerso en el *líquido amniótico* –o fluido amniótico–, contenido dentro de una membrana llamada saco amniótico o *amnios*. La composición del líquido amniótico, cuyo volumen alcanza los 1,1 litros a término, depende de un equilibrio dinámico de intercambios con el feto y la madre, y contiene células epiteliales que aportan unos 65 ng de DNA/ml. La obtención de una muestra, mediante punción abdominal a través de la pared del útero, se denomina *amniocentesis*. Ésta se usa a partir de las 14 semanas de gestación para el diagnóstico prenatal de anomalías cromosómicas, enfermedades metabólicas congénitas, defectos del desarrollo (como la inmadurez pulmonar del feto), determinación del sexo, etc. Por otro lado, las muestras del tercer trimestre sirven, por ejemplo, para valorar la gravedad de la eritroblastosis fetal y predecir el síndrome de dificultad respiratoria del neonato.

10.1.2.3 Tejidos o sangre fetales

Como alternativa al líquido amniótico se pueden utilizar muestras de las *vellosidades coriónicas*, tejido fetal procedente del *corion* (membrana más externa del embrión, que contacta con la placenta). Éstas constituyen la mejor fuente de DNA fetal. Se obtienen por biopsia entre las semanas 8^a y 12^a de gestación, y han de cultivarse.

Se pueden obtener también muestras de **sangre fetal**, dentro del seno materno, empleando dos procedimientos. La *fetoscopia* (método similar a la endoscopia, válido tanto para exploración como para toma de muestras) se utiliza entre las semanas 16^a y 21^a de gestación, para determinar hemoglobinopatías, coagulopatías, análisis de cariotipo, etc. Una técnica alternativa más sencilla, la *cordocentesis* o *funiculocentesis*, permite obtener sangre fetal directamente del cordón umbilical en el tercer trimestre de embarazo; sus eritrocitos (nucleados) y linfocitos se utilizan para diagnóstico prenatal sin interferencias con el feto.

10.1.3 Células de sangre o de médula ósea adultas

Las más empleadas son las muestras de **sangre periférica**, habitualmente de sangre venosa (unos 40 µg DNA/ml) recogida sobre anticoagulante. Se utiliza la fracción celular total o sus distintos tipos separados por alguno de los métodos disponibles, en especial la centrifugación según sus diferencias de tamaño o densidad (págs. 125-126).

Se usan también **manchas de sangre**, que se obtienen de sangre fisiológica (talón, dedo...) secada sobre un papel o cartón, o bien de origen accidental (prendas, vehículos, utensilios...). La mancha se raspa y se extrae mediante lavado con disolución salina. Se utilizan para metabolopatías, identificación, fines forenses, etc. Se pueden destacar las *tarjetas Guthrie*, para diagnóstico neonatal, y las tarjetas de *papel FTA*, material especialmente tratado en el que se seca la sangre, se lisan las células y se extraen sus componentes, quedando fijado exclusivamente el DNA. Estas tarjetas poseen la ventaja de su fácil transporte y conservación.

Las **muestras de médula ósea roja** (activa en hematopoyesis) se obtienen, bajo anestesia local, por aspiración mediante *punción medular*, en general en el esternón o en la cadera, y se usan para observaciones citológicas al microscopio y para estudios genéticos e histológicos.

10.1.4 Muestras bucales

Los **raspados y enjuagues bucales** se obtienen por métodos no invasivos, con cepillo suave o algodón y disolución salina. Contienen células epiteliales de descamación con suficiente DNA genómico para su detección, previa amplificación por PCR.

Se utiliza cada vez más la **saliva** por su facilidad de obtención (toma directa o presencia sobre la piel, sellos, colilla de cigarrillo, etc.). Las células epiteliales que incorpora, procedentes de descamación, pueden aportar cantidad suficiente de DNA (8-9 µg en una toma de muestra con hisopo) para su análisis con fines forenses o para el estudio de la paternidad.

10.1.5 Semen

Contiene millones de espermatozoides (5-7% del volumen total, con 1,3 pg DNA por célula) formados en los testículos, en suspensión en una fracción líquida de secreciones de las vesículas seminales (50% en volumen), la próstata (20% en volumen), las glándulas bulborrectales y el epidídimo. Se recoge no sólo directamente –para estudios de funcionamiento del tracto genital masculino, marcadores de fertilidad, etc.– sino también de la vagina por aspiración –para investigación criminal por agresión sexual–, o por lavado salino de manchas secas en ropa, piel, pelo, etc. Las muestras frescas deben analizarse antes de 1 h de su recogida, manteniéndolas a temperatura corporal durante su transporte. Debido a su especial resistencia, los espermatozoides se pueden separar del resto de células (que incluyen las de la víctima, en casos de asalto sexual) mediante lisis diferencial de éstas con SDS (dodecilsulfato sódico) y proteinasa K.

10.1.6 Células en cultivo

Las células aisladas o procedentes de tejidos u órganos disgregados se pueden cultivar en medios especiales. El cultivo celular constituye *per se* un área de estudio independiente, con requerimientos especializados en cuanto a equipo, procedimientos, sustratos, nutrientes y factores de crecimiento, etc. Pueden citarse como ejemplos líneas celulares clónicas (genética y metabólicamente uniformes con la célula inicial) y cultivos de células de tejidos, con un potencial de aplicación como alternativa al trasplante de órganos y tejidos o para la reparación de lesiones tisulares; por ejemplo, fibroblastos y queratinocitos para preparar “parches” de piel artificial (dermis y epidermis, respectivamente), hepatocitos para patologías hepáticas terminales, y células de miocardio para lesiones cardíacas.

10.1.7 Otros líquidos extravasculares

Se incluyen en este apartado algunas muestras celulares con interés diagnóstico por sí mismas o complementario de las anteriores.

10.1.7.1 Líquido cefalorraquídeo

Es un fluido resultante de la ultrafiltración del plasma (de ahí su analogía en composición) y que recoge los productos de desecho del metabolismo cerebral. Se encuentra en dos compartimentos que recubren el cerebro y la médula espinal, comunicados entre sí: el *sistema ventricular* (30 ml) y el *espacio subaracnoidal*, delimitado por las membranas piamadre y aracnoides (120 ml). Se produce y reabsorbe continuamente (con un recambio de unos 500 ml/día) en los *ventrículos* de los dos hemisferios cerebrales. Sirve como soporte protector del sistema nervioso central, para que “flote” el cerebro evitando su apoyo directo sobre la superficie ósea. Se puede obtener una muestra mediante *punción lumbar*. Aunque su aspecto normal es transparente, en condiciones patológicas aumenta su turbidez por la presencia de células (esencialmente leucocitos, eritrocitos, microorganismos, etc.), en función del tipo y gravedad de la enfermedad. Se utiliza para detectar enfermedades del sistema nervioso central, así como para detectar marcadores tumorales (pág. 476).

10.1.7.2 Líquidos o derrames serosos

Se localizan en las cavidades del cuerpo que rodean los órganos, delimitadas por dos membranas, la *parietal* (externa) y la *visceral* (interna, superficie del órgano). Estas membranas están formadas por tejido conjuntivo y una rica red de capilares sanguíneos y linfáticos, a través de cuya pared se filtra el plasma cuando se altera el balance hidrostático u osmótico; de ahí la analogía de composición entre plasma y líquidos serosos. En la

práctica clínica, se distinguen dos tipos de ultrafiltrado: *transudados*, de origen no inflamatorio, con pocas células (menos de 300/ml), y *exudados*, debidos a derrames inflamatorios y con numerosas células (más de 1.000/ml). Las muestras, que suelen ser transparentes, sólo se toman en situaciones anormales, presentando entonces un grado de turbidez variable en función de su contenido en leucocitos y eritrocitos, fundamentalmente.

	Volumen normal	Ubicación	Obtención por punción o paracentesis
Líquido peritoneal o ascítico	100 ml	Cavidad abdominal	Punción peritoneal
Líquido pleural	10-30 ml	Cavidad pleural (se acumula en la base de los pulmones)	<i>Toracocentesis o pleurocentesis</i> (punción de la pared torácica o del espacio pleural)
	Contiene células de muchos tipos y tiene interés clínico en enfermedades inflamatorias, infecciosas, marcadores tumorales, etc.		
Líquido pericárdico	20-50 ml	Espacio pericárdico	<i>Pericardiocentesis</i>
	Se usa con menor frecuencia, ya que presenta más complicaciones que los anteriores		

10.1.7.3 Líquido articular o sinovial

Es un ultrafiltrado del plasma, viscoso, con un alto contenido en proteínas y ácido hialurónico, localizado en la cavidad articular para conferir un carácter *lubricante* frente a la fricción propia de las articulaciones durante el movimiento. Con un volumen normal de 1 a 4 ml, su contenido celular tiene interés tanto en procesos mecánicos como en procesos inflamatorios patológicos articulares. Se obtiene por *artrocentesis*, es decir, aspiración tras punción en el espacio articular.

10.1.7.4 Muestras biológicas menos habituales

Pueden contener células y, por tanto, DNA, una serie de muestras de interés en casos concretos: humor acuoso (cámara anterior del ojo), humor vítreo (detrás del iris), lágrimas, endolinfa (tubo espiral del oído interno) y perilinfa (canales del equilibrio), moco cervical, líquidos folicular, oviductal y vaginal, heces, secreciones nasales, uñas, etc.

La **orina**, que en condiciones normales no contiene células, se utiliza para la cuantificación de metabolitos en enfermedades hereditarias relacionadas con el metabolismo de los aminoácidos (fenilcetonuria, alcaponuria, cetonuria de aminoácidos ramificados, cistinuria, etc.), de las bases purínicas (gota) y pirimidínicas (aciduria orótica), del grupo hemo (porfirias), etc.

10.2 DISOCIACIÓN DE LA MUESTRA TISULAR

En los tejidos, las células se encuentran atrapadas en una organización tridimensional, con interacciones o adhesiones tanto entre ellas como con la matriz extracelular (gel hidratado formado por proteínas fibrosas, como el colágeno, y polisacáridos complejos). El mejor rendimiento en cuanto a disociación de células viables se alcanza a partir de tejidos fetales o neonatales. Pueden señalarse **tres estrategias**:

Fragmentación mecánica: mediante corte, troceado o triturado de la muestra original en una disolución amortiguadora de pH, para obtener células individuales o fragmentos tisulares pequeños que faciliten su separación.

Disgregación química: la rotura de la matriz extracelular y de las uniones intercelulares se suele realizar por homogeneización isotónica o por sonicación (exposición a ultrasonidos, ondas sonoras o acústicas no percibidas por el oído humano por poseer una frecuencia superior a 20 kHz), en general en presencia de detergentes (iónicos, como el dodecilsulfato sódico o SDS, o no iónicos, como Tween-20 y Tritón X-100). La eliminación de cationes divalentes, en especial el Ca^{2+} (usando agentes quelantes como EDTA), también desestabiliza la adhesión de células entre sí y con la matriz extracelular.

Digestión enzimática: las proteínas de la matriz se digieren con proteasas como colagenasa, proteinasa K, pepsina o tripsina. Dependiendo del tipo de muestra, esta digestión puede requerir tiempos prolongados de incubación (por ejemplo, células en cultivo o raíces de pelo).

10.3 SEPARACIÓN DE CÉLULAS

La obtención de muestras celulares adecuadas para la fase analítica requiere variados procedimientos de separación, fraccionamiento y caracterización, que suelen basarse en las propiedades de cada tipo celular existente en la muestra. Su elección en distintas etapas de las fases preanalítica y analítica depende de numerosos factores: muestra de partida, grado de separación deseado, conservación de la viabilidad, operaciones y técnicas que van a aplicarse posteriormente, etc. Además de los métodos más comunes, que se exponen a continuación, existen otros procedimientos específicos, pero que no han alcanzado un uso general (por ejemplo, reparto en sistemas bifásicos acuosos, electroforesis de flujo libre, etc.).

10.3.1 Separación de células por métodos de sedimentación y centrifugación

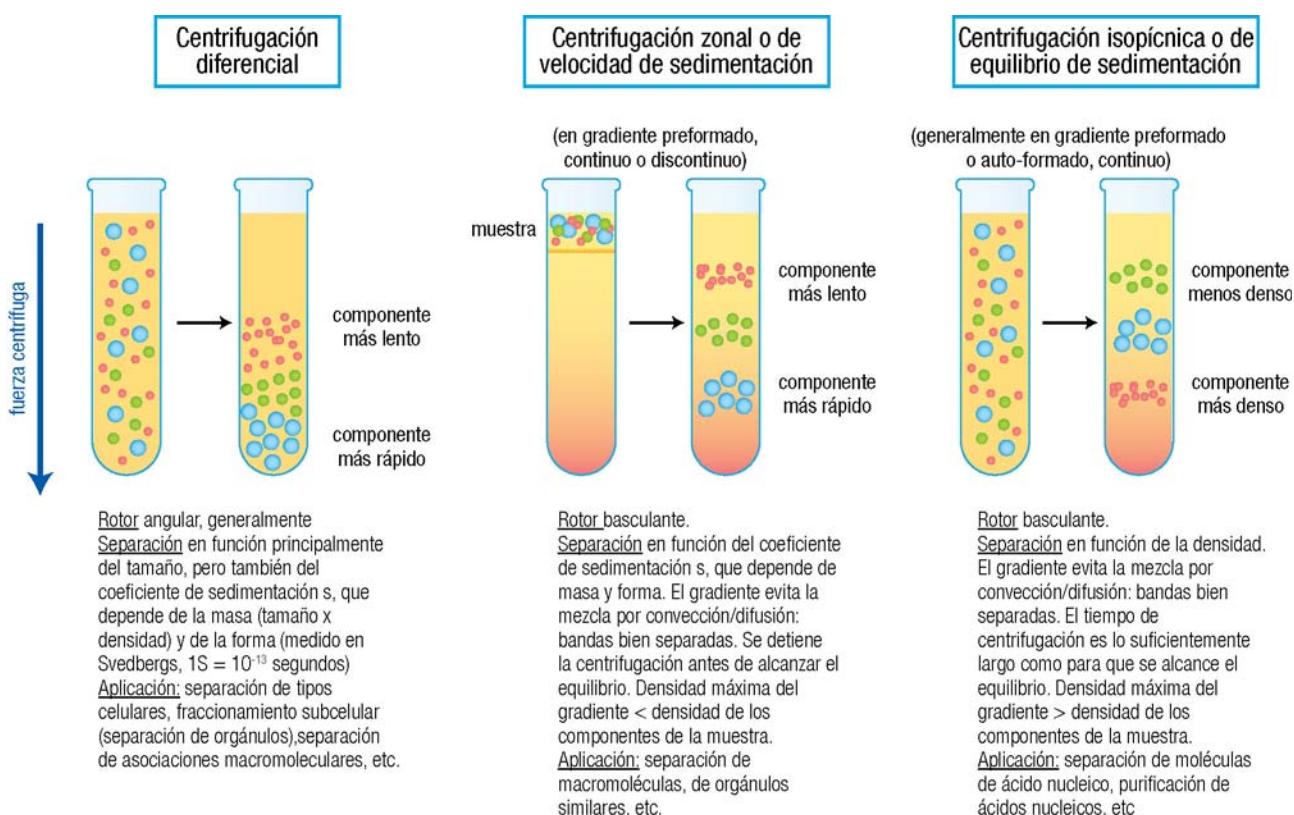
La separación celular puede hacerse por sedimentación *a gravedad unidad*, bajo la fuerza de la gravedad terrestre, pero lo habitual es acelerar el proceso aumentando muchas veces esta fuerza mediante la centrifugación. Al realizarse ésta en el seno de un medio líquido, la separación no sólo depende de la fuerza centrífuga, sino también de otros factores que modifican la sedimentación y que dependen de las características de la célula (o partícula en general) y del medio. El conjunto define la **velocidad de sedimentación** de la célula.



Web 10.2. Parámetros de la sedimentación y la centrifugación.

10.3.1.1 Centrifugación diferencial

Habitualmente, el planteamiento más sencillo es la separación por centrifugación, comenzando con velocidades bajas (*precentrifugación*). En la práctica, sólo se pueden separar así células que muestren grandes diferencias de tamaño, al menos de 10 veces (por ejemplo, las plaquetas, mucho más pequeñas que las restantes células



Web 10.3. Centrifugaciones diferencial, zonal e isopícnica.

sanguíneas). Por ejemplo, en el caso de sangre con anticoagulante, a $200 \times g$ los eritrocitos sedimentan en la zona inferior, los leucocitos aparecen en la interfase y queda como sobrenadante un plasma rico en plaquetas (*PRP*). Se puede preparar plasma pobre en plaquetas (*PPP*) empleando mayor velocidad (alrededor de $3.000 \times g$).

10.3.1.2 Centrifugación en gradiente de densidad

a) Métodos de barrera: centrifugación a través de un medio de densidad constante

Para conseguir separaciones más eficaces (por ejemplo, la de los leucocitos respecto al resto de las células sanguíneas) se centrifuga la muestra sobre un lecho de densidad intermedia entre aquellas de los tipos celulares que se quieren separar. Se emplean para ello medios comerciales como *Ficoll-Paque™*, *Lymphoprep™* y otros muchos, formados en general por mezclas de *Ficoll™* (un polisacárido sintético) y metrizamida (un compuesto sintético yodado). (Esta aplicación no emplea un gradiente en términos estrictos, pero utiliza principios comunes y puede considerarse un caso extremo de gradiente de un solo escalón.)

Se dispone de varios medios con densidades adecuadas para la separación de tipos celulares concretos; por ejemplo, *Nycoprep™-1.077* para células mononucleares, *Nycoprep™-1.068* para monocitos, *Polymorhprep™* para células polimorfonucleares, o *Nycoprep™-1.063* para plaquetas.



Web 10.4. Centrifugación en barrera: aplicación a la separación de células sanguíneas.

b) Centrifugación zonal o de velocidad de sedimentación

Para separar tipos celulares cuyo coeficiente de sedimentación difiere poco (por ejemplo, linfocitos y monocitos) es preciso acudir al uso de medios que forman un gradiente de densidad, lo que favorece la concentración de cada tipo celular en una banda o zona estrecha; por ello, esta técnica se denomina *centrifugación zonal*. El gradiente de densidad más simple se genera en forma escalonada, antes de añadir la muestra (gradiente *preformado*), utilizando varias disoluciones de distinta concentración de un compuesto adecuado (*Ficoll™*, sacarosa, albúmina, suero fetal bovino, etc.). La preparación puede hacerse de forma manual (lo que da lugar a un gradiente discontinuo) o empleando un aparato *formador de gradientes* (gradiente continuo).



Web 10.5. Preparación de gradientes de densidad.

Un pequeño volumen de muestra se deposita sobre el gradiente y se centrifuga a alta velocidad durante un tiempo corto. Habitualmente se diseña de modo que la densidad máxima del gradiente sea inferior a la de las células, con lo que éstas podrían alcanzar el fondo del tubo si se centrifugase durante el tiempo suficiente. Al emplear tiempos inferiores, el resultado depende del tiempo de centrifugación; dicho de otro modo, no se alcanza el equilibrio de sedimentación. El parámetro clave en la separación es el **coeficiente de sedimentación** de cada célula. Con esta técnica se consigue, por ejemplo, separar todos los tipos celulares sanguíneos, purificar espermatozoides viables, separar células viables y no viables de tejidos desagregados, muestras con células en suspensión, etc. Asimismo, se puede aplicar para separar orgánulos, estructuras supramoleculares (como las subunidades ribosómicas) o macromoléculas.

c) Centrifugación isopícnica o de equilibrio de sedimentación

Se pueden separar células en función exclusivamente de su **densidad** empleando otra variante de centrifugación, denominada *isopícnica* (palabra de origen griego que significa *de igual densidad*). Se realiza también sobre un gradiente de densidad, pero en este caso el tiempo de centrifugación es lo suficientemente largo como para que se alcance el equilibrio de sedimentación (entre la fuerza centrífuga, el empuje hidrostático –o fuerza de flotación– de la célula y su difusión). Para conseguirlo, se usan gradientes continuos que cubren todo el intervalo de densidades de las células presentes en la muestra: en el fondo del tubo la densidad del medio ha de ser superior a la de las células más densas. De esta forma, con independencia del tiempo de centrifugación, las células nunca sedimentarán al fondo, sino que alcanzan una posición estable intermedia en el gradiente.

También se utiliza esta técnica para la separación de moléculas cuya densidad es ligeramente diferente (por ejemplo, DNA, págs. 137-138 y 145).

Con una variante peculiar de esta técnica se consigue purificar linfocitos de su mezcla con monocitos, aprovechando la capacidad de estos últimos para fagocitar microesferas de oro coloidal, con lo que su densidad se hace artificialmente muy superior y sedimentan más abajo del tubo.

10.3.1.3 Elutriación centrífuga

La **elutriación** consiste en la separación de células (o partículas en general) de acuerdo con su velocidad de sedimentación. El proceso original, consistente en la realización de ciclos sucesivos de sedimentación y decantación, fue facilitado por la incorporación al sistema de un flujo del propio medio líquido de suspensión, en sentido contrario a la sedimentación (contra la gravedad). Éste fue finalmente combinado con el efecto de la fuerza centrífuga, dando lugar a la modalidad de elutriación de mayor uso actual, la **elutriación centrífuga**. Aquí, las células se ven sometidas a dos fuerzas opuestas: por una parte, la fuerza centrífuga y, por otra, la fuerza de arrastre por el flujo continuo del medio en sentido contrario.



Web 10.6. Elutriación centrífuga.

El resultado es un método versátil, rápido, eficaz y relativamente suave para separar subpoblaciones celulares según su tamaño a partir de una mezcla en suspensión. Admite números grandes de células e, incluso, permite obtener una mayor proporción de células viables respecto a la muestra original (al separarse las células dañadas, que poseen distinta densidad). Se aplica para obtener una gran diversidad de subpoblaciones enriquecidas, a partir de poblaciones celulares hepáticas, células en proliferación (sincronizadas en la fase M, G₁, S o G₂), células sanguíneas y otros tipos celulares.

10.3.2 Caracterización y separación de células por citofluorimetría de flujo

Es desde hace años un método de separación importante en investigación biomédica y en el laboratorio clínico, especialmente por su capacidad para el análisis automático de suspensiones celulares. El instrumento empleado se basa en el transporte de una suspensión celular (células sanguíneas, aspirado medular, tejidos disociados, etc.), impulsada por un flujo de disolución isotónica, al punto de medición o cámara de flujo. En su aplicación habitual, la citometría de flujo analiza y separa células previamente marcadas con uno o varios fluorocromos.



Web 10.7. Fundamento de la fluorimetría.

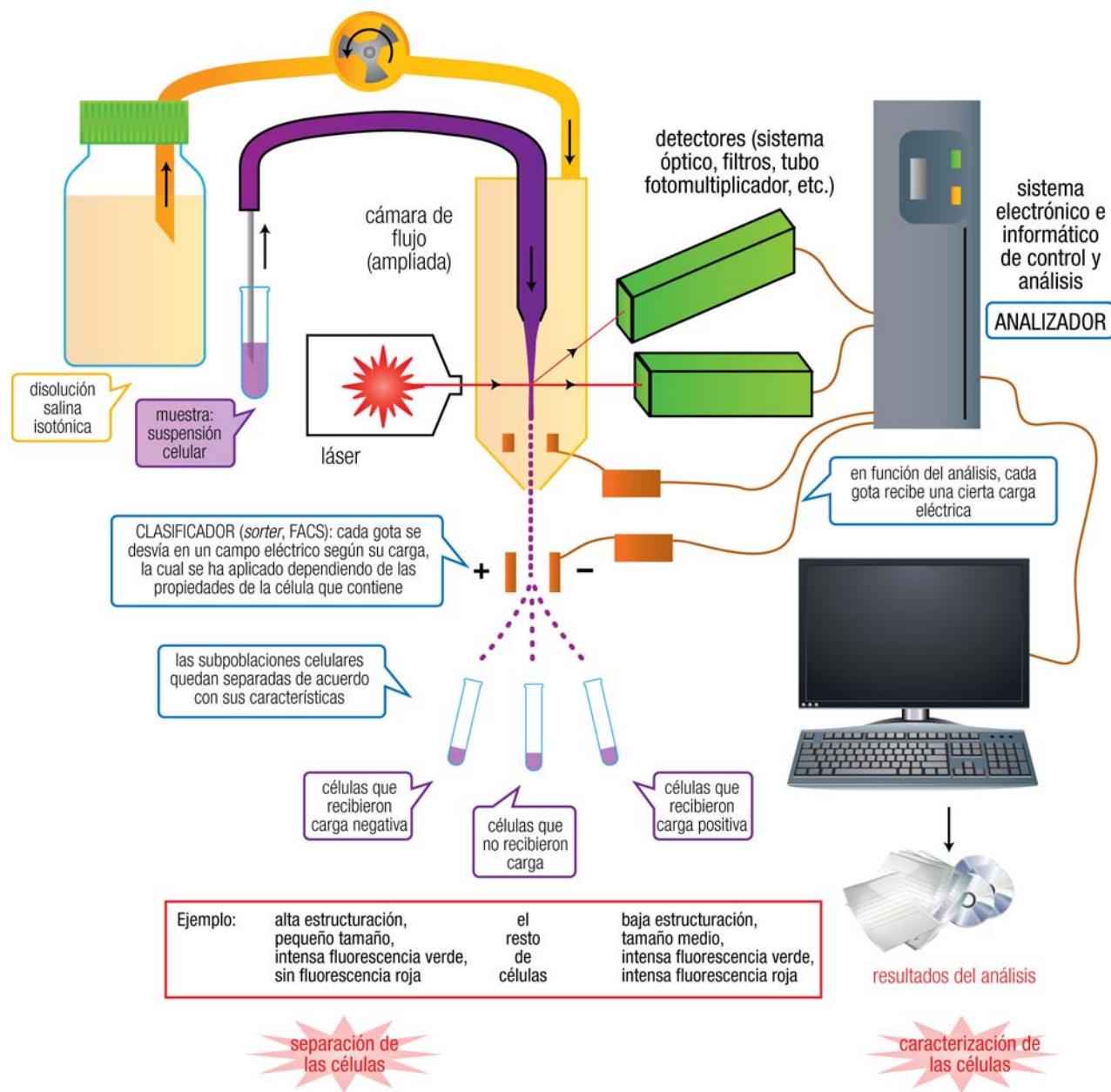
El procedimiento de citometría de flujo consta de dos etapas: caracterización y separación. De hecho, algunos instrumentos (analizadores) sólo realizan la primera, mientras que otros (clasificadores) son capaces, además, de llevar a cabo la segunda.

a) Caracterización de la población celular: analizador, citómetro de flujo o citofluorímetro

El sistema consigue que la suspensión celular llegue en condiciones de flujo laminar, formando un chorro muy fino que, como consecuencia de la dilución, contiene las células individuales en sucesión. Éstas pasan así de una en una a través de un haz luminoso láser, cuya longitud de onda permite la excitación de los marcadores fluorescentes incorporados previamente a la célula. En la luz emergente de cada célula se analizan la dispersión y la intensidad de la fluorescencia (emisión).

Parámetro medido	Propiedades celulares analizadas
Dispersión frontal de la luz (refracción) o FSC	Tamaño (diámetro y razón de volumen núcleo/citoplasma)
Dispersión perpendicular de la luz (reflexión) o SSC	Arquitectura interna (grado de estructuración subcelular, contenido de orgánulos, granularidad, rugosidad, etc.)
Intensidad de fluorescencia (a una o varias longitudes de onda)	Cantidad de fluorocromos marcadores previamente incorporados a la célula. Comúnmente se usan anticuerpos, marcados con fluorocromos, contra antígenos de la superficie celular, o bien colorantes o sondas que se fijan al DNA

Nota: FSC procede de *forward scatter*; SSC de *side scatter*.



b) Separación de células: clasificador de células activado por fluorescencia (FACS)

La corriente de líquido, tras pasar por el punto de detección, se transforma en una sucesión de gotas muy pequeñas (en general como consecuencia de una vibración aplicada por ultrasonidos a la cámara de flujo), que contienen como máximo una sola célula. Al mismo tiempo, en función de los resultados del análisis anterior de la célula, el clasificador aplica automáticamente a cada gota un pulso de voltaje (entre ± 50 y ± 150 V) que le proporciona una carga eléctrica.

El usuario define previamente el pulso que se ha de aplicar a cada gota (y, por tanto, el signo y cuantía de la carga que adquirirá), en función de criterios sobre los parámetros del análisis en el citómetro. El sistema informático que controla analizador y clasificador permite tomar dicha decisión en el breve intervalo transcurrido entre el paso de cada célula por el detector y la formación de la gota que la contiene. Al mismo tiempo, todos los datos de cada célula se almacenan para su estudio posterior.

A continuación, las gotas cargadas pasan por un campo eléctrico, donde se desvían de acuerdo con su carga para caer en distintos tubos de recolección de muestra; de esta manera se consigue la separación de subpoblaciones celulares de acuerdo con los criterios definidos (combinación de tamaño, estructuración y fluorescencia).

Como meros ejemplos de las innumerables aplicaciones de esta técnica se pueden citar: *a)* el estudio de la superficie celular con anticuerpos monoclonales, siendo posible seleccionar una célula, con un antígeno específico de superficie, de entre 1.000 que no lo tengan y clasificar miles de células por segundo; *b)* la valoración del contenido en DNA y RNA de una célula (por ejemplo, en estudios del ciclo celular), y *c)* la determinación de su forma y tamaño, por ejemplo en neoplasias.

10.3.3 Separación directa de tipos celulares

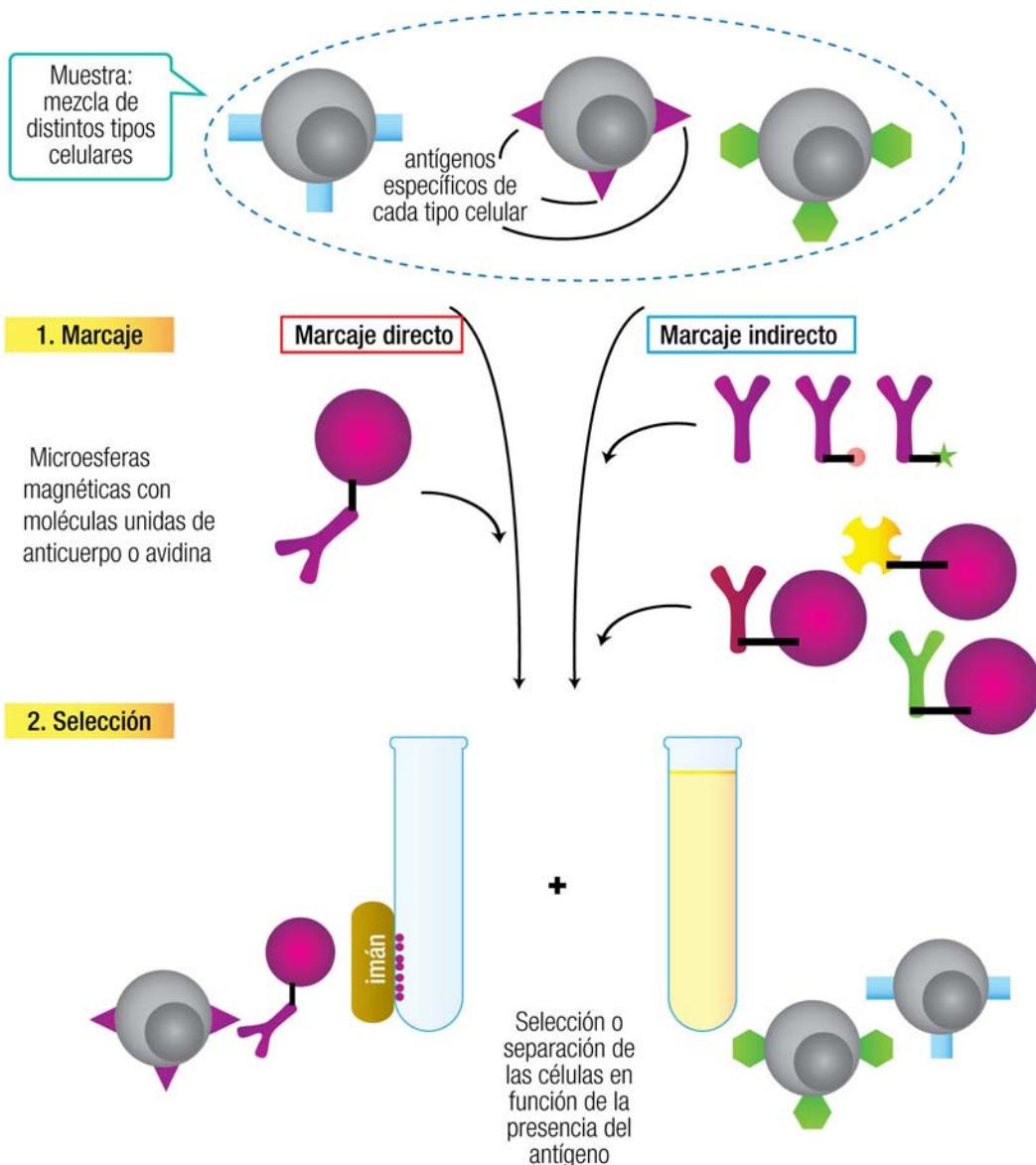
Aun siendo un caso particular, es de uso habitual y no puede dejar de mencionarse la posibilidad de aislar ciertos tipos celulares gracias a su tendencia a adherirse fuertemente a la superficie de un soporte de vidrio o plástico (**adhesión celular**), en general la propia placa de cultivo.

Con la llegada de los **anticuerpos monoclonales** este principio se pudo ampliar a prácticamente cualquier tipo de célula. La posibilidad de disponer de estos anticuerpos ha contribuido de forma notable a la separación de células en el laboratorio clínico y de investigación. Su exquisita selectividad en la unión a epítopos antigenicos de la superficie ha permitido diseñar métodos diversos para aislar y caracterizar de forma directa los tipos celulares presentes en una mezcla heterogénea. Estos métodos pueden aplicarse, por ejemplo, para definir las células que deben trasplantarse en caso de enfermedades graves (células precursoras de médula ósea, evitando células T maduras inmunocompetentes), para determinar el fenotipo de tejidos para trasplantes, para establecer los tipos celulares responsables de procesos en la respuesta inmune, para cuantificar subpoblaciones celulares, etc.

El uso de anticuerpos ha dado lugar a una gran diversidad de **métodos inmunoanalíticos** y de separación, que se pueden considerar parte de los **métodos de afinidad**. Por ejemplo, las células de un tipo pueden unirse, a través de sus antígenos, a anticuerpos monoclonales previamente fijados en superficies de plástico (tubos, placas de cultivo, varillas, microesferas, etc.), o matrices de colágeno, polisacárido, etc., lo que permite la separación de ese tipo celular. Como ejemplo singular, de entre la multitud de variantes desarrolladas, se comentan a continuación los métodos basados en el uso de microesferas magnéticas.

10.3.3.1 Separación por afinidad con partículas magnéticas

Estos procedimientos, de gran sencillez y por ello actualmente en expansión, hacen uso de **microesferas** disponibles comercialmente en varios tamaños (entre 0,5 y 10 μm , es decir, menores que las células) y con distintas posibilidades de aplicación. Están formadas por un material superparamagnético (esencialmente, un óxido de hierro) recubierto de una capa fina de un polímero plástico que permite la adsorción o la unión covalente de distintas moléculas. Con frecuencia se les une un anticuerpo, para formar *anticuerpos magnéticos* o *microesferas inmunomagnéticas*, o bien avidina o estreptavidina, proteínas ambas con gran afinidad por la biotina (*microesferas magnéticas de afinidad*).



10.4



Web 10.8. Aplicación de microesferas magnéticas para la separación de tipos celulares.

Estos métodos de selección de subpoblaciones celulares se pueden diseñar con dos planteamientos:

- **Selección positiva:** las microesferas magnéticas se unen a las células que interesa conservar para su estudio. Es, por su simplicidad, el abordaje recomendado. Sin embargo, requiere disponer de un método para romper la unión entre la célula y la microesfera, el anticuerpo o la avidina. Esto puede suponer una limitación dependiendo del uso posterior que se quiera dar a las células.
- **Selección negativa:** se marcan con las microesferas todos los tipos celulares excepto el de interés. Es la única solución cuando no se dispone de anticuerpos específicos o si existe interferencia de los anticuerpos o las microesferas con posteriores experimentos. Posee la ventaja de que se obtienen las células intactas, sin nada unido a ellas, y así se pueden utilizar de forma directa.

A menudo, el mejor resultado se obtiene combinando varias selecciones negativas previas, que enriquecen la muestra antes de proceder a la selección positiva final.

Más adelante se exponen aplicaciones similares de estas microesferas magnéticas para la extracción directa de ácidos nucleicos (pág. 137).

10.4 CARACTERIZACIÓN CELULAR Y MEDIDAS DE VIABILIDAD

Es importante proceder en esta fase preanalítica a la caracterización de las células por alguno de los procedimientos de los que se dispone en la biología moderna: **microscopía óptica**, como punto de partida más obvio; **microscopía de contraste de fase**, en paralelo a la anterior, para observar con más detalle la superficie celular; **microscopía electrónica de barrido** (SEM, de *scanning electron microscopy*), para definir la superficie celular y la naturaleza de las interacciones entre células, o **microscopía electrónica de transmisión** (TEM, de *transmission electron microscopy*), para observar cortes celulares y examinar los orgánulos o inclusiones. También se caracterizan las células mediante la **citometría de flujo** o el uso de **microesferas magnéticas**, técnicas ya comentadas.

Asimismo, es importante confirmar el estado vital de las muestras celulares durante una u otra etapa de la fase preanalítica. El concepto de **viabilidad** alude a la capacidad de una célula para realizar sus procesos fisiológicos y bioquímicos, en especial en lo que respecta a su metabolismo y a su capacidad de división. Sin embargo, en la práctica el término se utiliza con criterios diversos; por ejemplo, es común que se hable de viabilidad refiriéndose simplemente a la integridad de la célula, o bien a su actividad metabólica o, finalmente, a su capacidad de división (proliferación). Para todos ellos existen numerosos **ensayos de viabilidad**; los más empleados se describen a continuación.

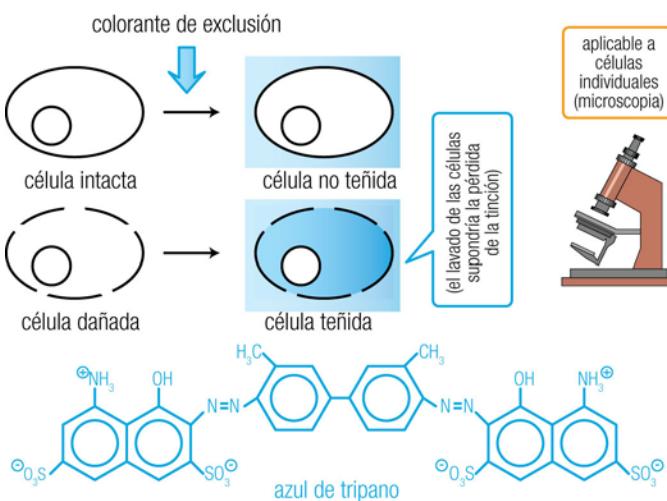
En el diseño de todos los ensayos de viabilidad hay dos parámetros fundamentales: por un lado, la especificidad y la detectabilidad de los reactivos y, por otro, la capacidad de éstos para penetrar en las células.

Respecto al primer punto, clásicamente se han empleado colorantes, que pueden detectarse en células individuales por microscopía y en suspensiones por absorciometría (habitualmente, ensayos simultáneos de múltiples muestras en placas multipicillo). Cada vez con más frecuencia estos ensayos se ven reemplazados por los que usan fluorocromos, cuya detectabilidad es muy superior, puede combinarse para medir a un tiempo varios fluorocromos distintos y permite, además, el uso de la citometría de flujo.

En cuanto a la permeabilidad a través de las membranas celulares, se trata del factor clave que permite que el ensayo diferencie células intactas de células muertas o dañadas. Básicamente la cuestión se reduce a la polaridad y carga eléctrica del reactivo o de su producto; los compuestos polares no atraviesan las membranas, a no ser que éstas se encuentren físicamente dañadas.

10.4.1 Ensayos de exclusión

Corresponden al recuento de células que no son capaces de incorporar un colorante (lo *excluyen*).



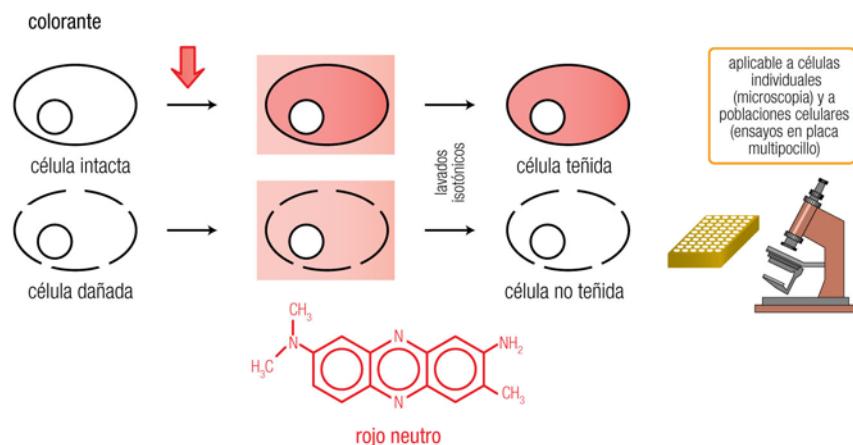
El ejemplo clásico más conocido es el **azul de tripano**, un colorante soluble en agua cuyos grupos cargados amino y sulfonato le impiden atravesar las membranas celulares intactas; por tanto, tiñe sólo las células muertas o dañadas. **Las células vivas no se tiñen**, pero debe resaltarse que el ensayo sólo valora la integridad de su membrana, no la verdadera viabilidad celular.

Un ensayo similar utiliza un **compuesto fluorescente** con carga (no atraviesa la membrana) y preparado para reaccionar con grupos amino. Las células intactas sólo se tiñen en superficie, débilmente, mientras que las

células dañadas se tiñen en todo su interior, de forma intensa. El uso de un fluorocromo –en lugar de un colorante– permite su aplicación en microscopía de fluorescencia y, en particular, en citometría de flujo.

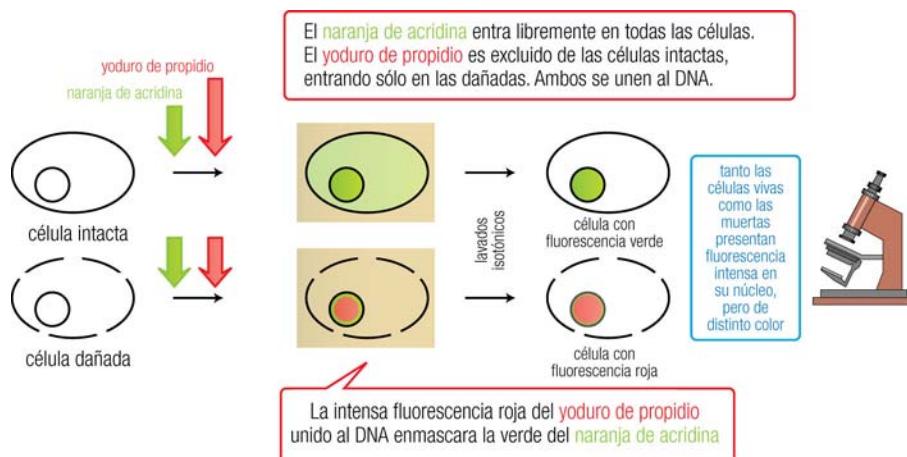
10.4.2 Recuento de células viables por incorporación de un marcador

Es un método alternativo, basado en la entrada de un compuesto a través de la membrana de las células intactas y su fijación en el interior.

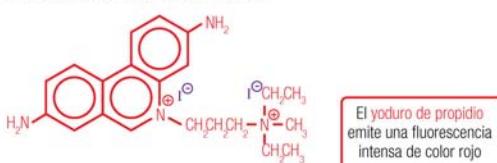


10.6

En algunos casos, el colorante (en realidad, casi siempre un fluorocromo) se une a los ácidos nucleicos, particularmente al DNA. Esto permite diseñar ensayos muy útiles en los que se tiñen, aunque con distinto color, tanto las células viables como las dañadas. El parámetro clave es siempre la lipofilia de los fluorocromos, que los hace permeantes o no a través de la membrana celular intacta.



Ambos compuestos son agentes intercalantes: su estructura aromática plana favorece su inserción entre los pares de bases apilados del DNA. Además, poseen propiedades fluorescentes: al ser excitados con luz UV emiten luz visible

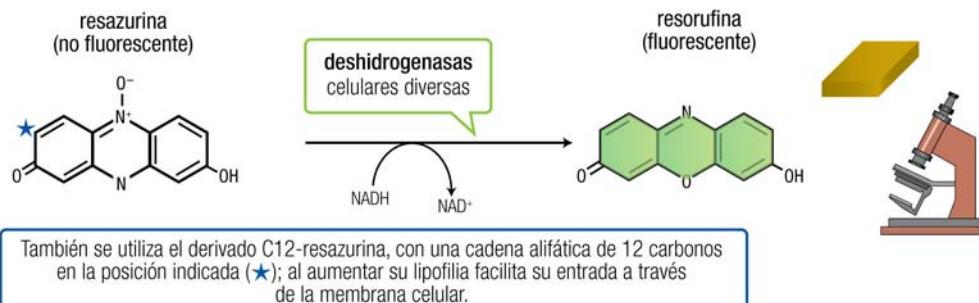
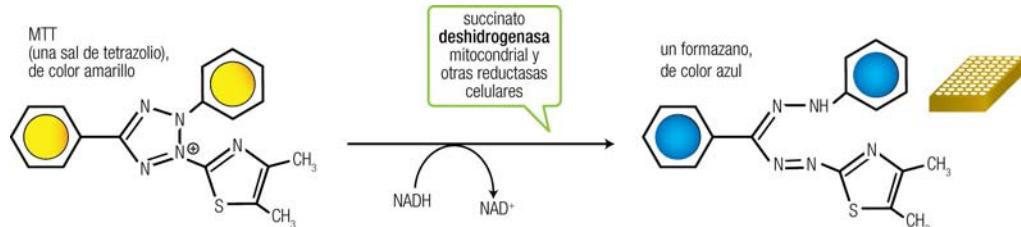
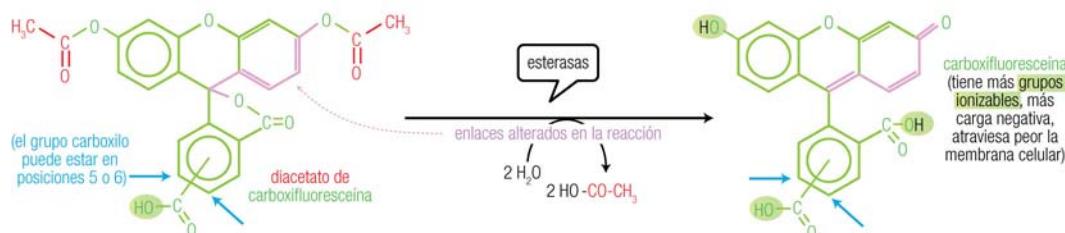
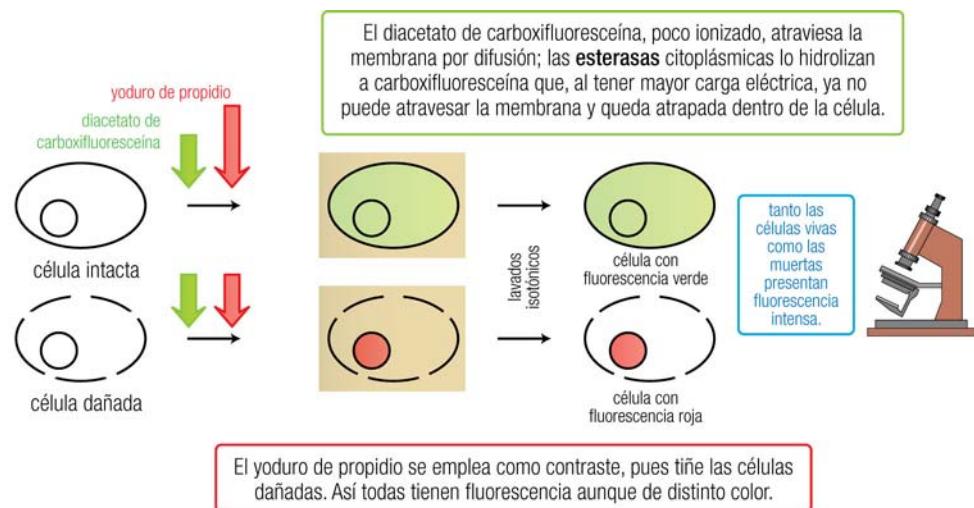


- El naranja de acridina emite una fluorescencia de distinto color según se encuentre unido a ácidos nucleicos bicatenarios o monocatenarios
- verde con DNA bicatenario
 - tonos intermedios (de amarillo a naranja) con DNA parcialmente desnaturalizado
 - roja con DNA monocatenario (desnaturalizado completamente) o con RNA

10.7

10.4.3 Ensayos por medida de la función metabólica

- Actividades enzimáticas:** la presencia en las células de actividades enzimáticas genéricas, tales como esterasas y deshidrogenasas, permite plantear ensayos de viabilidad basados en la generación de productos con grupos cromóforos o fluoróforos, procedentes de la transformación enzimática de sustratos sintéticos. Una alternativa, para tipos celulares concretos, es la determinación de una actividad enzimática específica, en cuyo caso es particularmente importante evitar que las condiciones experimentales afecten a dicha enzima.



10.9

- Síntesis proteica:** incorporación de radiactividad a las proteínas formadas durante el cultivo de las células en presencia de aminoácidos marcados (como [^{35}S]Met o bien cualquier otro marcado con ^{14}C o 3H).

- **Síntesis de DNA:** incorporación de radiactividad, u otro tipo de marca, al DNA formado durante el cultivo de las células en presencia de nucleótidos marcados. Refleja el concepto más exigente de viabilidad, la capacidad de crecimiento o, más estrictamente, proliferación celular. Es común el uso de timidina tritiada o de bromodesoxiuridina.



Web 10.9. Fundamento del uso de timidina tritiada y bromodesoxiuridina en ensayos de proliferación.

10.5 LISIS DE LAS CÉLULAS

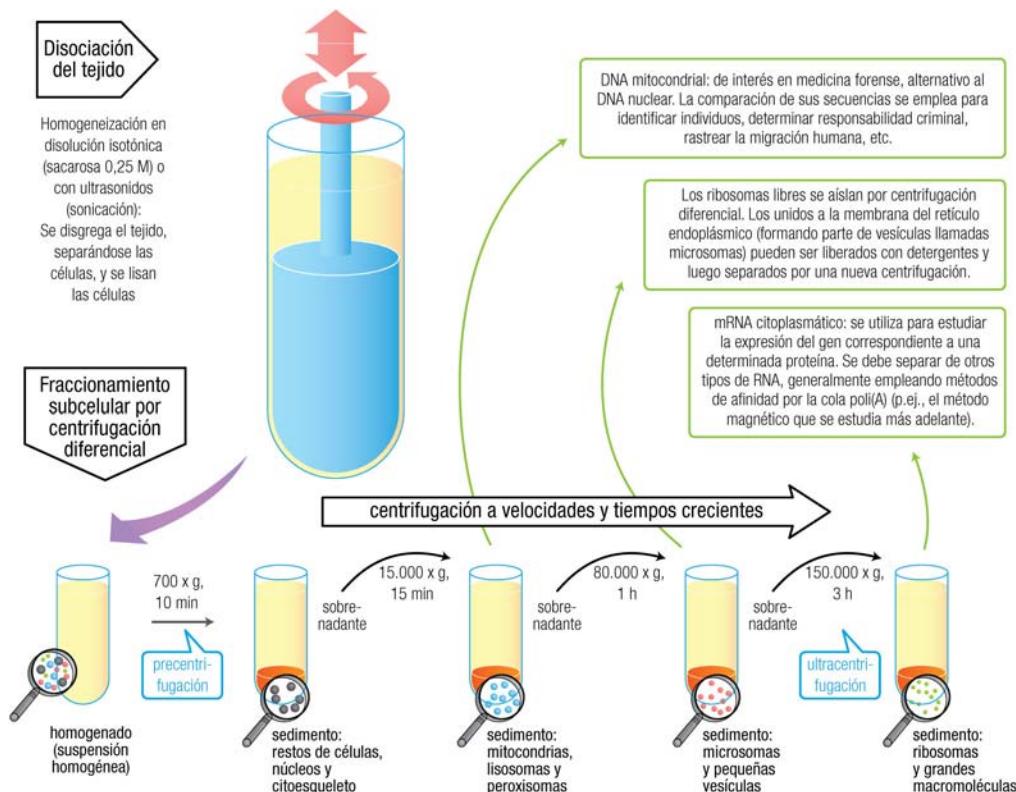
Esta etapa supone el punto de conexión entre la preparación de la muestra celular y la de **muestras de genoma o ácido nucleico**; comienza, por tanto, la fase analítica, que comprende el fraccionamiento subcelular, la extracción directa o el fraccionamiento de los ácidos nucleicos y la fragmentación del DNA. En algunos casos la lisis es simultánea con la preparación de la muestra (por ejemplo, la homogeneización).

Para la posterior purificación de ácidos nucleicos es imprescindible, obviamente, liberar primero el contenido celular, de modo que sea posible separar las distintas biomoléculas en disolución. La lisis celular requiere la ruptura de la membrana, que suele conseguirse empleando un medio **hipotónico**. Es también habitual añadir **detergentes** para solubilizar los lípidos. De cara a preservar la integridad de los ácidos nucleicos que luego se han de purificar es, asimismo, frecuente incluir en el medio de lisis algún **inhibidor de nucleasas**, pues éstas –particularmente las RNAsas– están a menudo presentes como componentes de las células.

La lisis se puede aplicar tanto a células en suspensión como a las adheridas a placas de cultivo. En la preparación de células sanguíneas es frecuente la aplicación de reactivos de lisis diferencial (es decir, que afectan sólo a determinados tipos celulares). Así, para el recuento y estudio de leucocitos se lisan los eritrocitos con un medio ácido (por ejemplo, líquido de Türk, con ácido acético y violeta de genciana) o con una disolución de cloruro amónico.

10.6 PREPARACIÓN DE FRACCIONES SUBCELULARES

El procedimiento clásico general parte del sobrenadante que resulta de la disociación mecánica de un tejido mediante homogeneización en un medio isotónico y posterior lisis celular. Mediante **centrifugación diferencial** se separan sucesivamente núcleos, orgánulos y otras estructuras subcelulares.



10.7 TRATAMIENTOS ADICIONALES O COMPLEMENTARIOS

Es esencial combinar las operaciones precedentes y posteriores con una serie de tratamientos adicionales o complementarios. Por ejemplo, el lisado celular se suele someter a tratamientos que eviten la destrucción de sus moléculas componentes o bien que permitan eliminar algunas de ellas. Para mantener la integridad de las muestras, especialmente en el caso de algunos tejidos (como páncreas y bazo), se debe evitar la acción de nucleasas endógenas (DNAsas y RNAsas); de ahí la conveniencia de usar muestras frescas o bien mantenerlas congeladas y particularmente de añadir **inhibidores de nucleasas**. Para la solubilización se acude, durante la homogeneización o resuspensión, al empleo de agentes que disocian y desnaturizan las proteínas, como **detergentes, fuerza iónica elevada, caótropos y reductores**.



Web 10.10. Detergentes, caótropos y otros reactivos auxiliares.

Para la inhibición de nucleasas en general suele utilizarse **EDTA** (sales del ácido etilendiaminetetraacético), compuesto quelante del Mg^{2+} , ya que este ion es un cofactor requerido por la mayor parte de las nucleasas. Por lo que respecta específicamente a las ribonucleasas, resultan inactivadas de modo covalente por el **dietilpirocarbonato** (DEPC), que reacciona con una histidina de su centro activo. Este reactivo se usa, por ello, en los procedimientos de extracción de RNA para tratar todo el material (tubos, etc.) y el agua con la que se preparan todos los reactivos.

Como precaución general, tanto en la obtención de RNA como de DNA, el material de vidrio y plástico debe estar completamente libre de ácidos nucleicos y nucleasas exógenos, y debe, evidentemente, evitarse la contaminación por el analista mediante el uso de guantes y precauciones como el manejo separado de las muestras de distinto origen.

10.8 EXTRACCIÓN Y PURIFICACIÓN DE ÁCIDOS NUCLEICOS

Es difícil describir la gran diversidad de métodos que se han desarrollado para la purificación de DNA y RNA, y más aún hacerlo racionalmente, bajo un planteamiento didáctico. Las variaciones son innumerables y están dictadas por criterios semiempíricos, aunque se basen en principios generales que sí son interpretables en función de las propiedades fisicoquímicas ya descritas. La dificultad es mayor, si cabe, con los métodos directos, de más reciente aparición, cuyos fundamentos están a menudo ocultos tras los intereses comerciales y las patentes. Se intenta por ello en este texto proporcionar una visión generalista de los principios y herramientas comunes, sin pretender tipificar los procedimientos descritos como los únicos aplicables.

El primer factor que debe considerarse, al igual que en cualquier proceso de purificación, es la necesidad de separar y eliminar las demás biomoléculas presentes en la muestra. Para ello, deben considerarse las propiedades fisicoquímicas de DNA, RNA, proteínas y resto de componentes celulares. De acuerdo con ellas, algunas de las herramientas clave en un proceso de purificación de ácidos nucleicos son las siguientes:

- El uso oportuno de enzimas que degradan específicamente un solo tipo de molécula: proteasas, DNAsas y RNAsas (pág. 212), o bien de inhibidores para evitar su acción.
- El uso de detergentes para la desnaturización de las proteínas, lo que facilita su posterior insolubilización y retirada, sin afectar a la estructura de los ácidos nucleicos. De modo similar, compuestos con actividad caotrópica, que desnaturizan las proteínas; destacan entre ellos, por su aplicación, urea, cloruro de guanidinio y tiocianato de guanidinio.
- La acción diferencial del medio alcalino sobre DNA y RNA (pág. 32), de tal modo que el primero sólo se desnaturiza (un proceso que, aparte de mantener su integridad, puede ser reversible, pág. 164), mientras que el segundo se hidroliza rindiendo oligonucleótidos y nucleótidos libres (cuyas propiedades son muy diferentes a las de sus macromoléculas originarias, los ácidos nucleicos y permiten, pues, separarlos con facilidad).
- El ajuste de la fuerza iónica (concentración de sales), que permite controlar la solubilidad diferencial de DNA, RNA y proteínas (en particular si éstas se han desnaturizado previamente).
- La escasa solubilidad de los ácidos nucleicos (y también de las proteínas) en alcoholos. Prácticamente todos los métodos de purificación incluyen, al menos, una precipitación alcohólica.

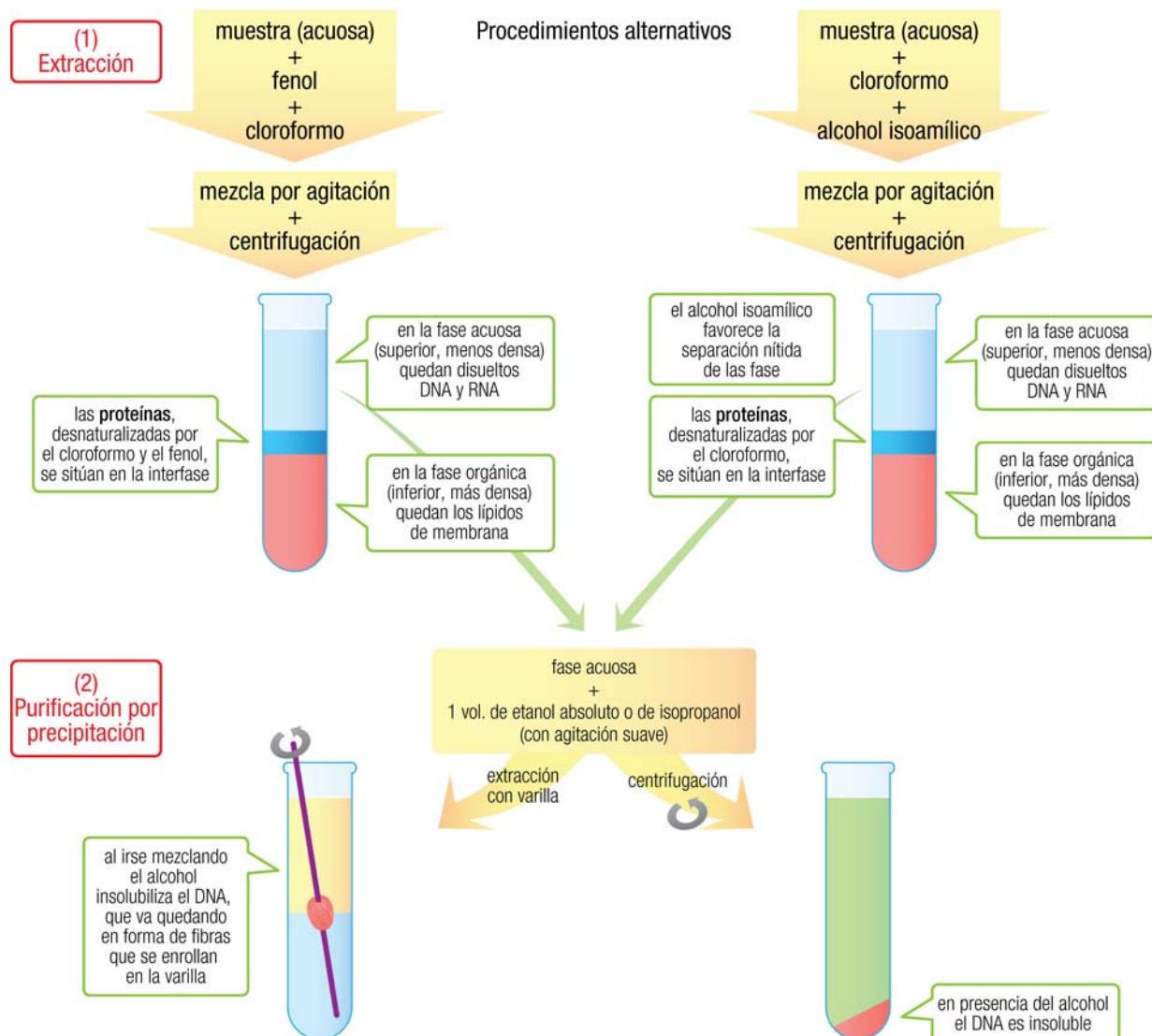
10.8.1 Extracción de ácidos nucleicos por solubilidad

10.8.1.1 Precipitación con alcoholes

Como se verá a continuación, los alcoholes se utilizan en prácticamente todos los métodos de purificación de ácidos nucleicos. Si la concentración de alcohol (habitualmente etanol o isopropanol) es suficientemente elevada, tanto DNA como RNA se hacen insolubles y pueden precipitarse mediante centrifugación. Esta precipitación requiere la presencia de sales. Los alcoholes retiran el agua que hidrata los ácidos nucleicos y, al ser menos polares que ella –dicho con más precisión, son disolventes con una menor constante dieléctrica– favorecen la unión de los cationes a los grupos fosfato, convirtiendo a los ácidos nucleicos en moléculas neutras que pueden agregarse y precipitar. La precipitación alcohólica es a menudo la etapa final, en la que se consigue completar la purificación y, además, concentrar el material.

10.8.1.2 Extracción del DNA mediante fases inmiscibles

El protocolo tradicional básico consiste en el tratamiento del homogenizado con cloroformo, cuyo papel es doble: disolver los lípidos y desnaturizar las proteínas. Muchos métodos emplean, además, fenol, que contribuye asimismo a la desnaturización de las proteínas. Para una extracción eficaz del DNA hacia la fase acuosa, ésta debe tener un pH neutro o ligeramente alcalino.



10.8.1.3 Extracción del RNA

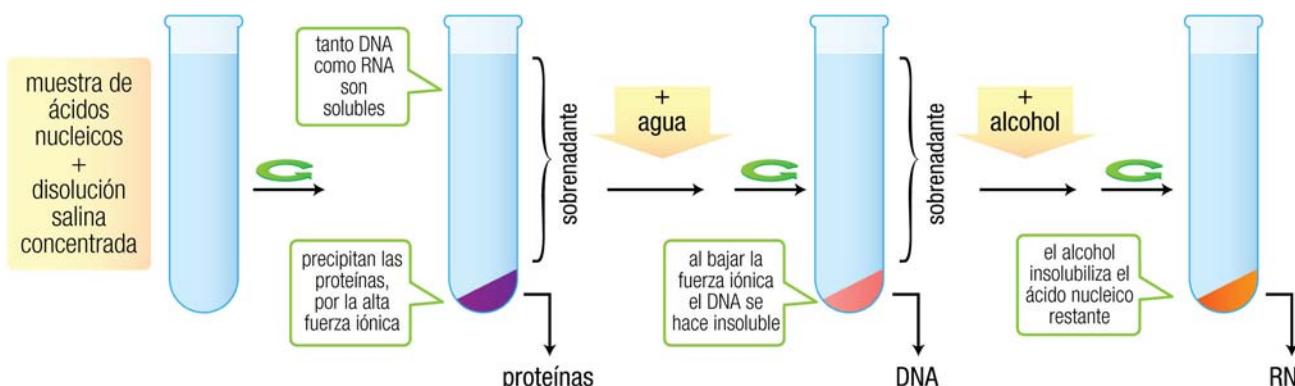
La extracción del RNA se realiza de manera similar, salvo dos modificaciones. En primer lugar, suele prepararse el homogenado inicial en presencia de una alta concentración de **tiocianato de guanidinio** (4 M). Éste es un fuerte caótropo que desnaturaliza las proteínas de modo que, además de facilitar después su retirada, asegura la inactivación de las RNAsas, esencial para poder obtener el RNA íntegro. Por otra parte, la extracción con cloroformo y fenol se hace a un pH ácido (entre 4 y 5), condiciones en las que el DNA ya no se mantiene en la fase acuosa, sino que se distribuye hacia la fase orgánica o la interfase, al igual que las proteínas. El RNA, por el contrario, queda disuelto en la fase acuosa y puede recuperarse mediante precipitación con etanol o isopropanol. En algunos métodos el cloroformo se sustituye por 1-bromo-3-cloropropano (BCP), menos tóxico y más eficaz en la retirada del DNA.

Otros métodos prescinden de la extracción con fenol y cloroformo, y utilizan cloruro de litio (entre 0,5 y 2,5 M), que precipita selectivamente el RNA por neutralización de su carga.

Finalmente, se toma la precaución ya mencionada de inactivar las posibles RNAsas tratando tanto el material como el agua de todos los reactivos con dietilpirocarnato (DEPC, inhibidor covalente de las ribonucleasas). Además, en ocasiones se incluye en las etapas finales una DNasa para asegurar la ausencia total de contaminación por DNA.

10.8.1.4 Purificación de ácidos nucleicos por precipitación salina diferencial

Se trata de un procedimiento alternativo a los anteriores, basado en las diferencias de solubilidad de DNA y RNA en función de la concentración de sales: tanto DNA como RNA son solubles en disolución salina concentrada, mientras que a baja concentración de sales sólo es soluble el RNA.



Para esta precipitación salina se suele utilizar NaCl, pero puede también emplearse acetato sódico saturado, en combinación con alcohol orgánicos, que conduce a resultados incluso mejores que con fenol y cloroformo, sobre todo para cantidades pequeñas de DNA.

10.8.2 Extracción directa de ácidos nucleicos

Los métodos de extracción descritos, basados en solubilidad y particularmente en disolventes inmiscibles, son los clásicos, suponen la norma de referencia y se acude a ellos cuando se desea obtener una gran cantidad de ácido nucleico o cuando es importante un grado elevado de pureza; por otra parte, resultan laboriosos (la extracción a menudo debe repetirse varias veces) y utilizan reactivos tóxicos.

A veces no es posible emplear los procedimientos descritos, dado el bajo contenido en DNA de algunas muestras, pero, sin embargo, puede estudiarse el DNA de la muestra sin necesidad de manipulación alguna, salvo la lisis de las células. Éste es el caso de la amplificación y detección del DNA mediante la técnica de PCR (v. Capítulo 14).

Por otra parte, en muchas ocasiones el tamaño de la muestra es pequeño, la pureza no es un requisito esencial, prima la rapidez en el procesado simultáneo de muestras numerosas, en resumen, otros métodos más simples y rápidos pueden ser una solución adecuada. Se han desarrollado más recientemente numerosos

métodos de extracción que cumplen estos requisitos. Se habla de *extracción directa* del DNA pues no se requiere la manipulación previa de la muestra y, en poco más de una etapa, permiten obtener su DNA o RNA.

El fundamento de estos métodos directos no es tan claro. La mayoría se basan en la adsorción selectiva del ácido nucleico, por ejemplo sobre soportes derivados de sílice o membranas porosas especialmente formuladas. Un procedimiento típico puede describirse así:

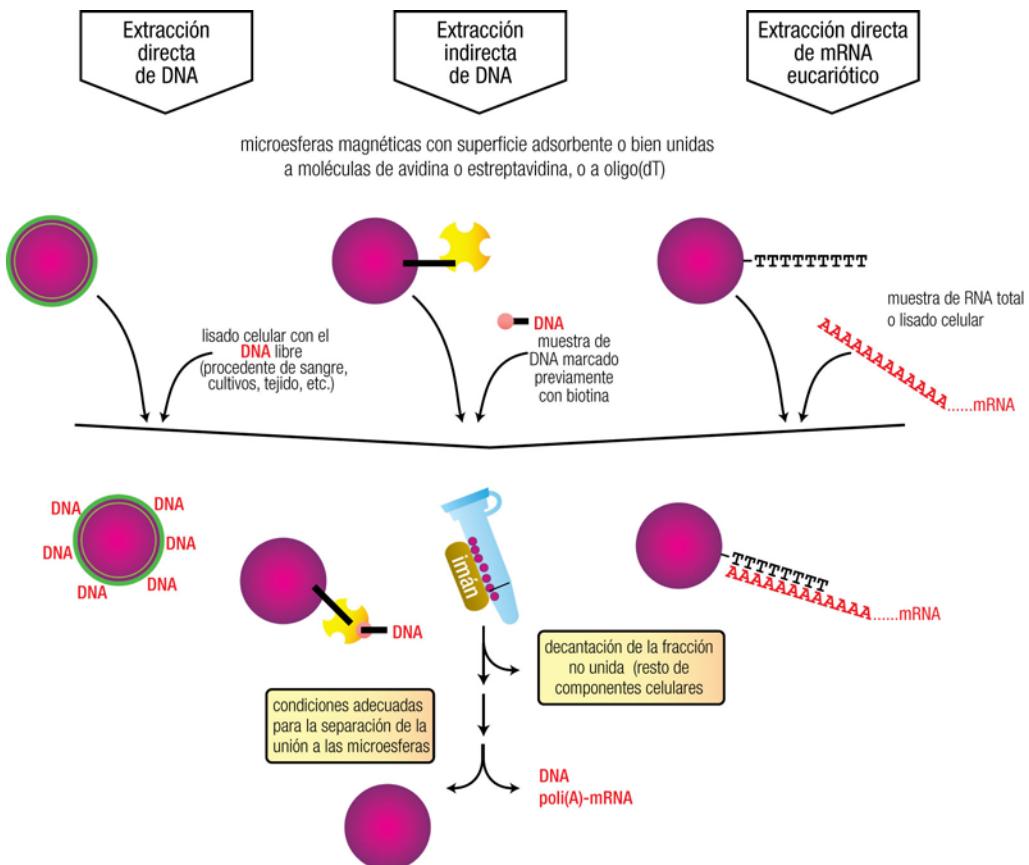
- *Lisis y pretratamiento*: la muestra (sangre, homogenado de tejido, células cultivadas, etc.) se somete a la acción de un tampón de lisis. Éste puede incluir detergentes, proteasas o agentes desnaturizantes de proteínas.
- Se aplica el lisado sobre una pequeña columna que contiene la membrana o material adsorbente.
- *Etapa de adsorción*: se fuerza el paso del líquido a través de la columna, bien mediante presión o bien por centrifugación. De ese modo, irán saliendo de la columna los componentes que no se han adsorbido; habitualmente esa fracción comprende todo excepto el DNA o el RNA (dependiendo del diseño del material adsorbente y de los medios líquidos empleados).
- *Etapa de lavado*: se añade más tampón para desplazar por completo todos los componentes no fijados.
- *Etapa de elución*: finalmente, se añade un tampón o disolvente diferente que desestabilice la unión del ácido nucleico al soporte.



Web 10.11. Ejemplo de extracción directa del DNA.

Como una aplicación particular, se ha extendido el uso para pruebas de identidad de las *tarjetas FTA®*, papel especialmente tratado que permite absorber una muestra (gota de sangre), lisar sus células, desnaturizar sus proteínas y proteger el DNA, que queda intacto fijado en el papel. Estas tarjetas pueden conservarse sin especiales precauciones durante varios años, a temperatura ambiente, enviarse por correo postal, etc.

También se emplean con frecuencia las **microesferas magnéticas**, análogas a las utilizadas para separación de células (pág. 130), sobre cuya superficie se fijan los ácidos nucleicos, bien por adsorción directa o bien a través de otras moléculas con afinidad por ellos. Se consiguen así métodos de extracción simples que se pueden aplicar a una amplia variedad de muestras.



www**Web 10.12. Aplicación de microesferas magnéticas para la purificación de ácidos nucleicos.**

El RNA mensajero eucariótico, gracias a la singularidad de su *cola de poli(A)* (pág. 302), puede purificarse fácilmente sobre soportes de afinidad. Por ejemplo, microesferas magnéticas con cadenas oligo(dT), columnas de oligo(dT)-celulosa o de poli(U)-*Sepharosa*; todas ellas retienen el RNA mensajero por complementariedad de bases con la cola de poli(A).

10.9 RECOGIDA Y CONSERVACIÓN DE LAS MUESTRAS

Una vez separados y extraídos, los ácidos nucleicos deben conservarse de forma adecuada. El precipitado seco de ácidos nucleicos o el “ovillo” de fibras de DNA obtenido con la varilla pueden disolverse en disolución amortiguadora y conservarse a 4 °C, o disolverse en agua y mantenerse congelados para su utilización posterior. El DNA genómico en disolución no se debe congelar, para evitar la rotura de las moléculas por las fuerzas de cizalla provocadas al solidificarse el hielo y al descongelarse. Por el contrario, el RNA sí puede conservarse congelado a -70 °C en disolución acuosa, cuidando de añadir un inhibidor de ribonucleasas.

10.10 FRACTIONAMIENTO DE ÁCIDOS NUCLEICOS

10.10.1 Ultracentrifugación

Este método se basa en el aumento de la **velocidad de sedimentación** mediante campos centrífugos lo suficientemente elevados (superior a $100.000 \times g$) como para superar las fuerzas de difusión. Se emplea en especial para el DNA, tanto con fines de preparación (gran cantidad de material) como analíticos (pequeña cantidad de material).

10.10.1.1 Centrifugación zonal

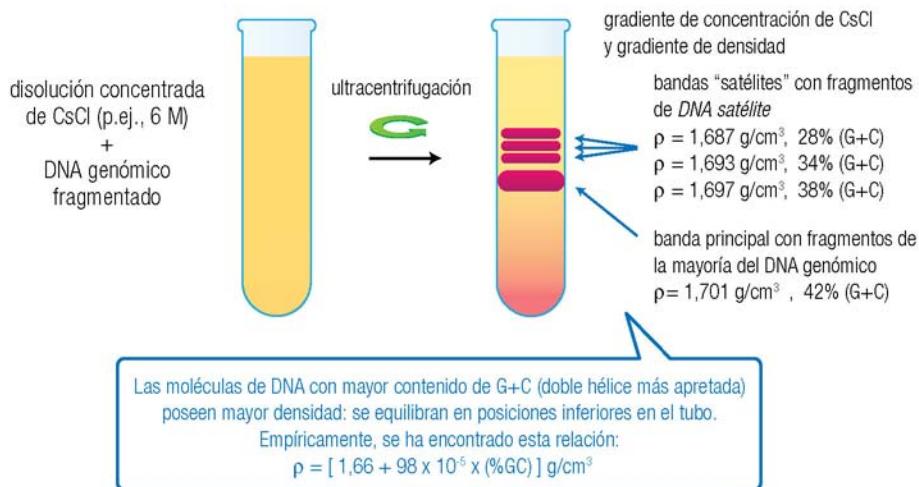
De forma similar a lo estudiado para la separación de tipos celulares (pág. 126), se puede emplear esta modalidad de centrifugación sobre gradientes de densidad para separar moléculas de ácido nucleico, en función de su velocidad de sedimentación (determinada por su coeficiente de sedimentación, en unidades svedberg, S) (v. web 10.2). Se consigue que las moléculas de ácido nucleico sedimenten a velocidad constante, con lo que se sitúan en *bandas*, cuya posición depende del tamaño y la forma de la molécula. Por ejemplo, pueden separarse así los cuatro tipos de RNA ribosómico eucariótico (pág. 63). El material de las bandas puede recuperarse después por separado (por ejemplo, mediante punción del fondo del tubo) y cuantificarse midiendo su absorción a 260 nm.

10.10.1.2 Centrifugación isopícnica

El método isopícnico o de equilibrio de sedimentación es quizás el más utilizado para la separación de distintos tipos de ácido nucleico, en especial DNA. Debe recordarse que, en este caso, las macromoléculas se separan en función de su densidad, en general usando un gradiente que se va formando durante la ultracentrifugación a la vez que se separan los componentes de la muestra (gradiente *autoformado*, pág. 125). El medio habitual para preparar el gradiente es una disolución de cloruro de cesio que, gracias a su alta solubilidad y baja viscosidad, permite alcanzar densidades lo suficientemente elevadas y formar gradientes con un intervalo estrecho de densidades. La muestra se mezcla con dicha disolución de CsCl y se procede a la ultracentrifugación durante un tiempo largo (decenas de horas), hasta alcanzarse el equilibrio de sedimentación. Cada tipo de ácido nucleico se sitúa, formando una *banda*, en aquella posición del gradiente donde la densidad del medio iguala a la de la molécula.

Ejemplo de aplicación de la centrifugación isopícnica:

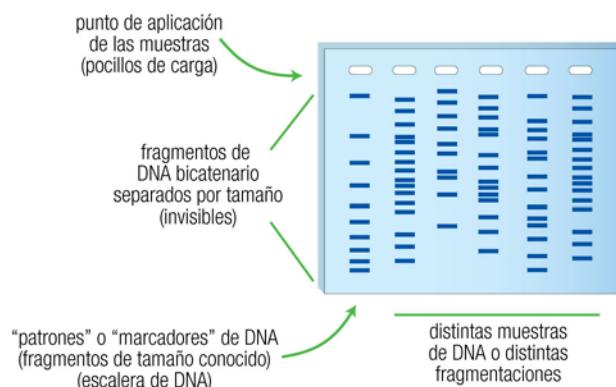
Algunas regiones del genoma están formadas por la repetición de secuencias cortas. Debido a ello, su composición de bases nitrogenadas es diferente de la del resto del genoma, que es más promediada. Como consecuencia, cuando se fragmenta el DNA total y se analiza en gradientes de densidad de CsCl, los fragmentos procedentes de ese DNA altamente repetitivo se separan de la banda principal, formando *bandas satélites*; de ahí vino el nombre de DNA *satélite* (descrito en el Capítulo 9).



10.14

10.10.2 Electroforesis

Como es bien sabido, esta técnica se basa en la capacidad de las macromoléculas cargadas para desplazarse en un campo eléctrico, con velocidad proporcional a su carga e inversamente proporcional a su tamaño. Las moléculas de DNA poseen, a pH neutro o alcalino, una carga negativa uniforme por unidad de masa, lo que hace que su movilidad hacia el ánodo (+) sólo venga determinada por el tamaño de la molécula. Dicho tamaño, cuando la molécula es lineal, depende únicamente de su número de pares de bases. Cuando la molécula es circular (como es el caso de los plásmidos y el DNA cromosómico bacteriano) el tamaño depende, además, de la conformación, lo que incluye el estado de superenrollamiento (pág. 79). Debe destacarse, además, que los cromosomas eucarióticos son demasiado grandes para ser identificados con este tipo de análisis, por lo que lo más común es que se aplique a DNA fragmentado con nucleasas.



10.15



Web 10.13. Separación del DNA mediante electroforesis.

10.10.2.1 Soporte para la electroforesis

La separación de los fragmentos de DNA bicatenario se realiza mediante electroforesis en geles de agarosa o de poliacrilamida (que suelen desarrollarse, respectivamente, en horizontal y en vertical). En ambos casos, el gel se prepara con una concentración adecuada para el tamaño de los fragmentos de DNA que se quieren separar. Los fragmentos más pequeños se pueden separar correctamente en geles de poliacrilamida; para tamaños algo mayores se usan geles de agarosa, más porosos. Modificando la concentración de acrilamida o agarosa se consigue variar el intervalo de tamaños que es posible separar (intervalo de resolución). En ocasiones se utilizan geles que combinan agarosa y poliacrilamida.

Composición del gel:	Intervalo de resolución conseguido: tamaño de DNA (kb)	
% de acrilamida	% de agarosa	
20		0,006 - 0,1
15		0,025 - 0,15
12		0,04 - 0,2
8		0,06 - 0,4
5		0,08 - 0,5
3,5		1 - 2
	2,0	0,1 - 2
	1,5	0,2 - 3
	1,2	0,4 - 6
	0,9	0,5 - 7
	0,7	0,8 - 10
	0,6	1 - 20
	0,3	5 - 60

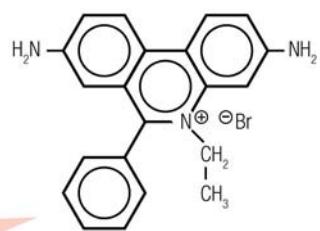
10.16

10.10.2.2 Detección del DNA tras su separación ("revelado")

Una vez separadas, las moléculas de DNA se visualizan habitualmente a través de la fluorescencia de distintos compuestos que se unen a ellas (*intercalantes*, casi siempre). Aunque el más utilizado ha sido el bromuro de etidio, va siendo sustituido por otros fluorocromos con mayor poder de detección y menor toxicidad.

Bromuro de etidio:

Agente intercalante, se introduce entre las bases apiladas del DNA, abre un poco la doble hélice, provoca un desenrollamiento local. Su uorescencia aumenta mucho cuando se une al DNA (color naranja al iluminar con UV): método de tinción o revelado del DNA, especialmente usado en electroforesis y en centrifugación.



El bromuro de etidio debe manejarse con precaución porque posee carácter mutágeno

10.17



Web 10.14. Estructura y acción de compuestos intercalantes.

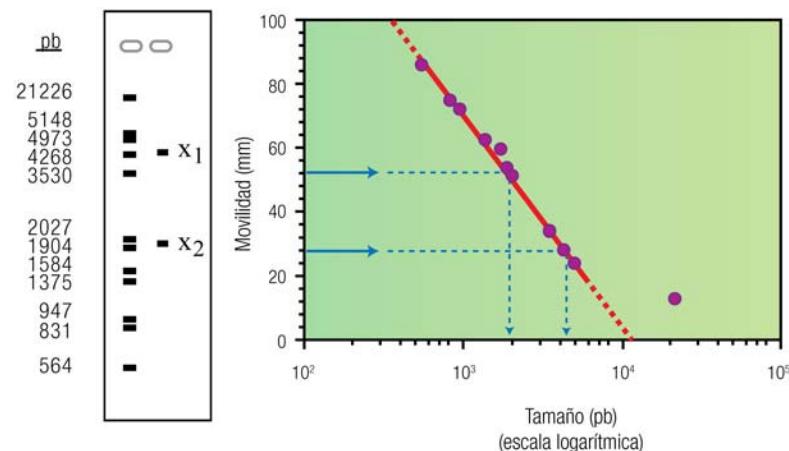
10.10.2.3 Calibración de la electroforesis para el cálculo del tamaño de los fragmentos de DNA

Las moléculas lineales de DNA bicatenario migran en la electroforesis a una velocidad inversamente proporcional a su tamaño (dentro del intervalo de resolución del gel empleado). Esto permite estimar la longitud de los fragmentos estudiados.

Se aplica en uno de los pocillos del gel una mezcla de fragmentos de DNA de tamaño conocido, denominados "patrones", "marcadores de masa molecular" o "escalera de DNA" (*DNA ladder*). Tras su separación en la electroforesis, se revelan y se representa para todas las bandas la movilidad (distancia de avance) frente al logaritmo del tamaño (longitud en pb). Sobre la recta ajustada se interpolan las distancias de avance de las muestras problema, calculando así su tamaño en pb.

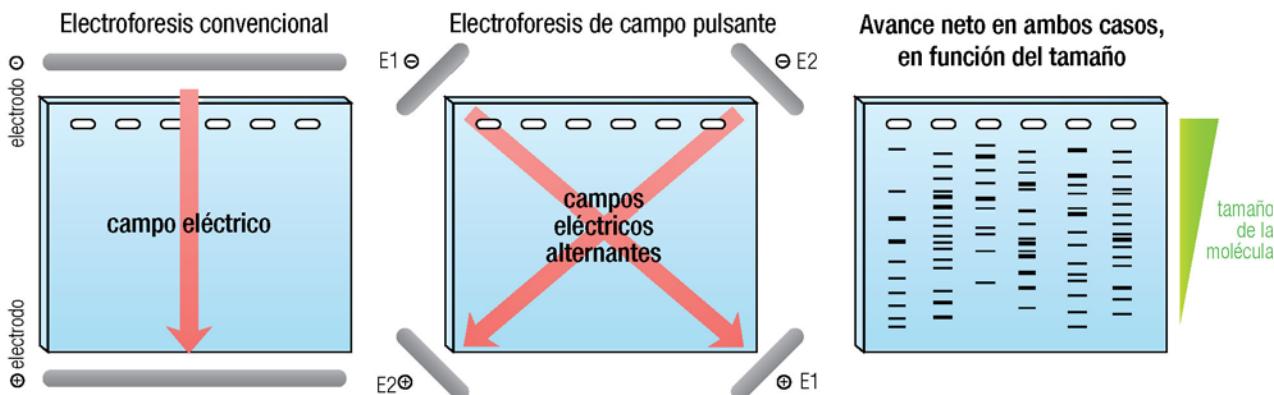
En este ejemplo, los patrones empleados son fragmentos bien conocidos, aquellos obtenidos a partir del DNA del virus λ por la acción de las endonucleasas de restricción *Eco RI* e *Hind III*.

10.18



10.10.2.4 Electroforesis de campo pulsante

Las moléculas o fragmentos de DNA de longitud superior a 50 kb no se pueden resolver (separar) empleando la electroforesis convencional en gel de agarosa. Esto se debe, en primer lugar, a la dificultad de manejo de los geles necesarios: con menos del 0,4% de agarosa la consistencia del gel es muy baja. Además, las moléculas de DNA poseen un tamaño excesivamente grande para atravesar los poros del gel y ya no se separan en función de su tamaño. Para solventar este problema se ha ideado una modificación de la técnica, conocida como **electroforesis de campo pulsante** (o pulsado, *pulsed-field gel electrophoresis*, PFGE). Se emplean geles de agarosa al 1%, pero se altera periódicamente la orientación del campo eléctrico aplicado, activando de forma alternativa dos parejas de electrodos. El frecuente cambio de dirección del campo eléctrico consigue que las moléculas se desplieguen y avancen a través de los poros en conformación extendida. Cuanto mayor es su tamaño con más dificultad se reorientan, por lo que avanzan más despacio. Así se consigue separar fragmentos de DNA de hasta varias megabases, lo que supone un progreso significativo, aunque aún no sirve para estudiar los cromosomas humanos enteros (de 50 a 250 Mb cada uno).



10.19

Las muestras han de contener el DNA sin fragmentar, por lo que no sirven las de DNA purificado, que, en general, ya se ha fragmentado por debajo de 100 kb como consecuencia de las fuerzas de cizalla experimentadas durante los procesos habituales de aislamiento. Para obtener preparaciones con los cromosomas intactos se mezclan las células con agarosa caliente (37 °C) y se vierte en pequeños moldes de unos milímetros, de forma que al solidificarse la agarosa (4 °C) forma bloques que incluyen las células intactas. En ellos se realizan la lisis celular y la digestión de las proteínas (por ejemplo, con EDTA, SDS y proteinasa K, que difunden a través de los poros del gel) sin alterar las grandes moléculas de DNA. Tras una diátesis para eliminar los restos de la digestión, se dispone de un pequeño bloque de agarosa con el DNA intacto en su interior, y dicho bloque se deposita en un pocillo del gel de electroforesis. Al aplicar la corriente, el DNA sale del bloque de preparación y entra en el gel, donde comienza su separación.

Sección II

Transmisión de la información genética y tecnologías relacionadas



Replicación del DNA

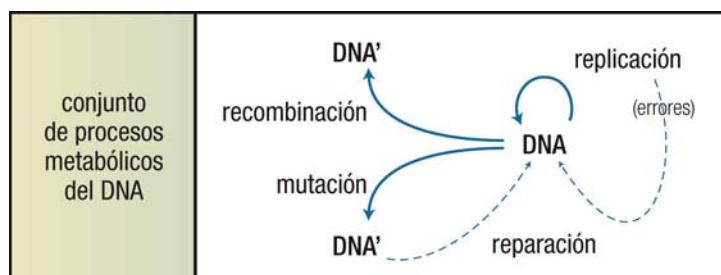


11.1 CARACTERÍSTICAS GENERALES DE LA REPLICACIÓN	145	
11.1.1 Carácter semiconservador	146	
11.1.1.1 Demostración experimental de la replicación semiconservadora	147	
11.1.2 Síntesis simultánea, secuencial y bidireccional	148	
11.1.3 Inicio monofocal o multifocal	148	
11.2 ENZIMOLOGÍA DE LA REPLICACIÓN	148	
11.2.1 Requerimientos de la reacción de síntesis de DNA	148	
11.2.2 Mecanismo de la reacción	149	
11.2.3 Dirección de la síntesis	149	
11.2.4 Tipos de polimerasas	149	
11.2.4.1 DNA polimerasas procarióticas: actividades polimerasa y exonucleasas.	149	
11.2.4.2 DNA polimerasas eucarióticas: actividades polimerasa y exonucleasa	151	
11.2.5 Velocidad de replicación	152	
11.3 ETAPAS EN EL PROCESO DE REPLICACIÓN	153	
11.3.1 Iniciación	153	
11.3.2 Elongación	154	
11.3.2.1 Conexión entre la iniciación y la elongación: proteínas necesarias	154	
a) Helicinas	154	
b) Proteínas ligantes de DNA monocatenario	155	
c) Topoisomerasas	155	
d) Primasa	156	
11.3.2.2 Asimetría de la replicación en ambas hebras	156	
11.3.2.3 Mecanismo de la elongación	157	
11.3.2.4 Maduración de los fragmentos de Okazaki	158	
11.3.3 Terminación	158	
11.3.3.1 Final de la elongación	158	
11.3.3.2 Replicación de los telómeros	158	
11.3.3.3 Telomerasa: componentes y características	160	
11.3.3.4 Mecanismo de acción de la telomerasa.	160	
11.3.3.5 Funcionalidad de la actividad telomerasa en las células	160	
11.4 REPLICACIÓN MITOCONDRIAL	161	

11.1 CARACTERÍSTICAS GENERALES DE LA REPLICACIÓN

La replicación es el proceso, conceptualmente sencillo pero molecularmente complejo, mediante el cual a partir de una molécula de DNA *progenitora* o *parental* se sintetiza una nueva, originándose así dos moléculas de DNA *hijas*, de secuencia idéntica a la del DNA original. Esto permite el paso de la información genética a la descendencia (tanto la de una célula como la del individuo).

Aunque la replicación constituye el aspecto esencial del **metabolismo del DNA**, también se deben considerar como parte del mismo otros procesos relacionados directa o indirectamente: la **recombinación** o reordenamiento de la información genética, la **mutación** o alteración de la secuencia, y la **reparación** de las alteraciones o daños en el DNA. Cabe también añadir la degradación o catabolismo del DNA hacia nucleótidos, mediado por las nucleasas.



11.1

La replicación se produce de forma coordinada con la división celular (Capítulo 8), concretamente en la fase S, durante las interfases previas tanto a la mitosis como a la meiosis I, de forma que, en el caso de la mitosis, las dos células hijas reciban la misma dotación genética que tenía la célula progenitora. En cuanto a la meiosis,

puesto que da lugar a cuatro células hijas, la replicación previa permite que éstas reciban la mitad de dicha dotación genética, y no una cuarta parte (pág. 104). La relación entre el estado de replicación y la progresión por el ciclo celular se puede resumir en los siguientes principios generales:

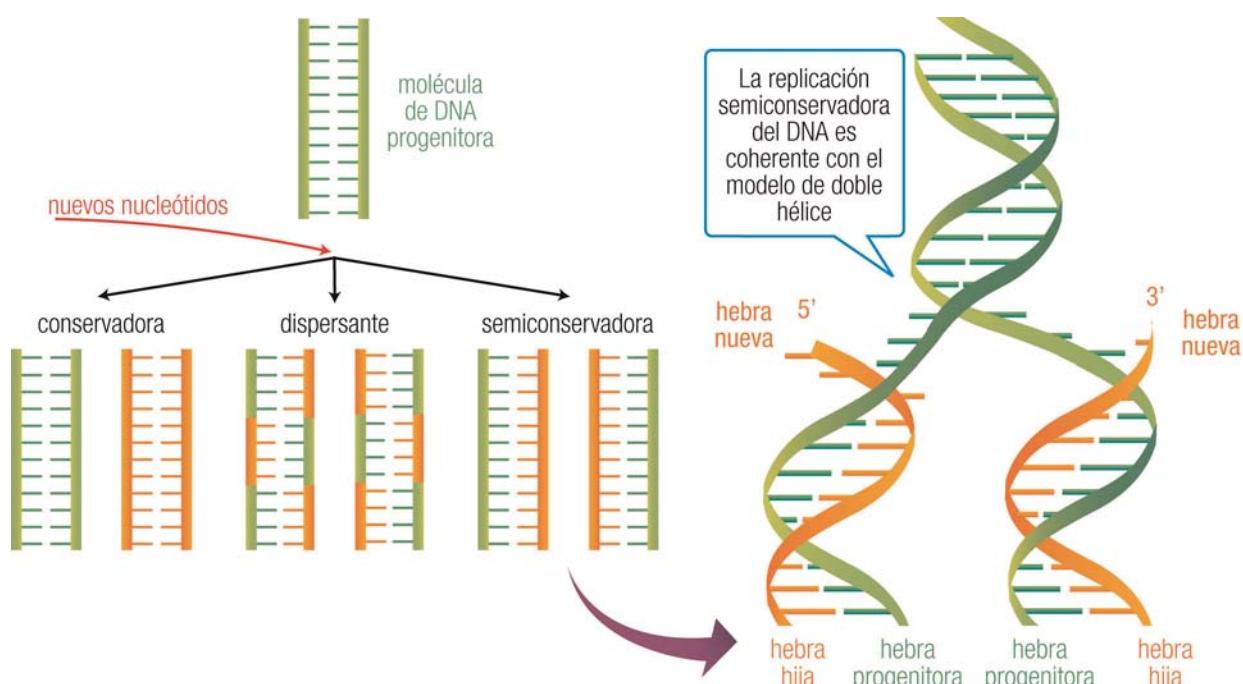
- El inicio de la replicación del DNA obliga a la célula a emprender una división.
- Una vez iniciada la replicación, no puede tener lugar la consiguiente división hasta que se haya completado dicha replicación. En realidad, el final de la replicación constituye posiblemente una señal para la división.
- El proceso de replicación se lleva a cabo sin necesidad de una separación completa de la cadena progenitora: la doble hélice se desenrolla gradualmente y sus dos hebras se van separando a la par que se produce su replicación.

Antes de estudiar los detalles moleculares de la síntesis de un nuevo DNA, procede analizar las características principales del proceso, atendiendo a un criterio eminentemente didáctico. Éstas son: carácter **semiconservador**, realización **simultánea** en ambas hebras, de forma secuencial y con carácter **bidireccional**, y origen monofocal (procariotas) o multifocal (eucariotas). Otra característica destacada, el carácter **semidiscontinuo**, se explicará una vez que se conozca el mecanismo de la reacción.

11.1.1 Carácter semiconservador

Aunque hipotéticamente se pueden plantear tres mecanismos para la replicación, pronto se demostró que sólo el tercero tiene lugar en las células:

Replicación conservadora	Replicación dispersante	Replicación semiconservadora
Un complejo enzimático recorrería el DNA, reconociendo la secuencia en ambas hebras para copiarlas simultáneamente, de forma que el DNA progenitor bicitenario se conservaría intacto mientras que la molécula hija de DNA contendría dos hebras totalmente nuevas	La síntesis se produciría por fragmentos, que se combinarían con partes del DNA progenitor, dando lugar a las dos moléculas hijas de DNA, ambas con fragmentos tanto nuevos como de las dos hebras del DNA progenitor	Es el mecanismo real. Las dos hebras del DNA <i>progenitor</i> sirven de molde para la síntesis de sus respectivas hebras complementarias. Cada molécula hija de DNA está formada por una hebra <i>progenitora</i> y otra <i>hija</i> (recién sintetizada). El mecanismo se repite en sucesivas generaciones de células. El nombre alude, precisamente, a que siempre se conserva la mitad de la molécula original



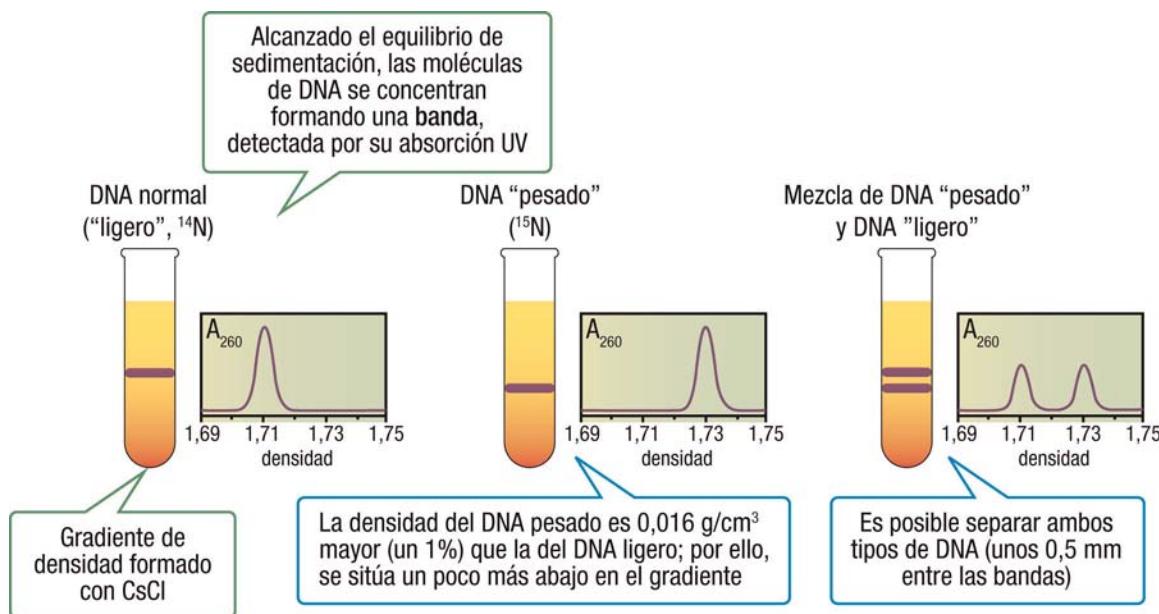
11.1.1.1 Demostración experimental de la replicación semiconservadora

El mecanismo real de la replicación lo demostraron por primera vez Mathew Meselson y Franklin W. Stahl en 1958, mediante experimentos muy completos, sencillos y representativos de muchos otros realizados en diversas áreas de la biología molecular. Aunque los aspectos moleculares de la replicación se establecieron posteriormente con el estudio de la DNA polimerasa, el experimento de Meselson y Stahl tiene interés, aparte de por su propio diseño, por haber sido la primera vez que se aplicaron tanto la técnica de sedimentación en gradiente de densidad como la introducción de isótopos en las moléculas de DNA (^{15}N , infrecuente, y ^{14}N , isótopo habitual).

Para comprender la novedad y trascendencia de estos experimentos en su momento y contexto, se van a considerar los antecedentes y el planteamiento antes de describir el resultado. Se aprovecha así este ejemplo para ilustrar el método científico, en especial cómo un diseño adecuado del experimento, que parece simple, permite obtener una información notable.

a) Antecedentes

El primer lugar, se examina la posibilidad de aplicar la técnica de equilibrio de sedimentación en gradiente de densidad, o centrifugación isopícnica (pág. 126), al análisis de muestras de DNA cuya densidad es diferente por estar marcado o no con el isótopo ^{15}N .



b) Planteamiento y diseño del experimento

Aplicando esta técnica de marcaje del DNA con ^{15}N y su análisis mediante centrifugación isopícnica, se diseñó el experimento de forma que los resultados obtenidos permitieran diferenciar cuál de las hipótesis era la correcta.



Web 11.1. Experimento de Meselson y Stahl.

c) Resultados obtenidos y conclusión

El resultado observado al realizar el experimento fue de una banda de DNA con densidad intermedia en la primera generación de células y dos bandas en la segunda generación, con densidades baja e intermedia. Estos datos experimentales coinciden con los resultados previsibles a partir de la **hipótesis semiconservadora**, de modo que ésta debe ser la correcta; es decir, en la replicación del DNA cada una de las moléculas hijas conserva una hebra del DNA progenitor y tiene otra hebra de nueva síntesis.

11.1.2 Síntesis simultánea, secuencial y bidireccional

De cara a la descripción del mecanismo de la replicación es importante destacar algunas características, observadas por primera vez en *E. coli* pero igualmente válidas para eucariotas. La principal estrategia experimental para obtener estas conclusiones ha sido añadir al medio de cultivo timidina tritiada, con lo que se marca el DNA de nueva síntesis y luego puede detectarse por métodos autorradiográficos.

- La síntesis no comienza al azar, sino en puntos concretos de la molécula, denominados **orígenes** de replicación. En cada uno de ellos se produce la apertura local de la doble hélice y comienza la síntesis **simultánea** de las dos nuevas hebras (no debe confundirse simultáneo con el carácter continuo de la síntesis de una sola de las hebras, que se estudiará más adelante).
- Como consecuencia, las dos hebras nuevas se van alargando de forma progresiva, por adición **secuencial** de nucleótidos.
- La separación de las hebras progenitoras, que comienza en cada origen, progresiva en ambas direcciones, por lo que se afirma que la replicación es **bidireccional**. Los puntos de transición entre doble hebra y hebras sencillas se denominan horquillas de replicación, y van alejándose entre sí.
- Finalmente, no debe pasarse por alto el hecho de que, dado que las dos hebras del DNA son antiparalelas, su separación con el avance de la horquilla progresiva en sentido opuesto, para una hebra de 3' hacia 5' y para la otra en sentido 5' → 3' (y a la inversa en la otra horquilla).



Web 11.2. Formación y avance de las horquillas de replicación.

11.1.3 Inicio monofocal o multifocal

En el cromosoma circular único de procariotas la replicación comienza siempre en un punto determinado, denominado origen (y etiquetado como *oriC*); se afirma que la replicación es **monofocal**. En consecuencia, progresiva formando dos horquillas de replicación. Por el contrario, en eucariotas la replicación es **multifocal**, pues en cada uno de los cromosomas existen múltiples orígenes de replicación (centenares o miles), que dan lugar a dos horquillas cada uno. Gracias a ello se hace posible completar en un tiempo razonable la replicación de los cromosomas, mucho más largos que en procariotas.

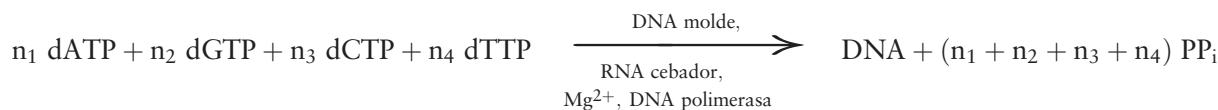
11.2 ENZIMOLOGÍA DE LA REPLICACIÓN

11.2.1 Requerimientos de la reacción de síntesis de DNA

- **Sustratos:** se utilizan como sustratos el conjunto de los cuatro desoxinucleósidos-trifosfato: dATP, dGTP, dCTP y dTTP. De cada uno de ellos queda incorporado en el nuevo DNA la parte dNMP de la molécula. Tanto dNMP como dNDP no sirven como sustratos.
- **Cofactores:** para una actividad óptima se requiere un ion metálico divalente como cofactor, asociado a los dNTP y a la polimerasa. Aunque *in vitro* este papel pueden desempeñarlo tanto Mn^{2+} como Mg^{2+} , es este último el que actúa *in vivo*.
- **Molde o plantilla:** ésta es, quizás, la principal característica de la reacción. El orden correcto de incorporación de los nucleótidos viene determinado por su complementariedad de bases con la secuencia de cada hebra de DNA, que actúa así como *molde o plantilla*.
- **Cebador:** la síntesis de una nueva hebra de DNA no puede comenzarse a partir de dos nucleótidos, sino que requiere un fragmento monocatenario iniciador, denominado *cebador*, que aporte un grupo 3'-OH libre. Se trata de un oligonucleótido, por lo común RNA, que debe estar emparejado con la hebra progenitora de DNA, de forma complementaria y antiparalela. Este requisito puede expresarse afirmando que la replicación no es autoiniciadora (a diferencia de la transcripción, pág. 268), sino que sólo **elonga** o alarga moléculas preexistentes.

II.2.2 Mecanismo de la reacción

Se inicia la reacción con el emparejamiento de un dNTP complementario al nucleótido de la hebra molde (DNA) en la posición vecinal al extremo 3' de la hebra en crecimiento (initialmente, el RNA cebador; luego, la hebra que se está sintetizando). La reacción propiamente dicha consiste en la unión del dNTP seleccionado, que queda como dNMP, liberándose PP_i. Ello implica la formación de un enlace fosfoéster entre el fosfato α del dNTP que se incorpora y el 3'-OH libre de la cadena en crecimiento (initialmente, del cebador). La ruptura del enlace fosfoanhídrido (α - β) del dNTP proporciona la energía necesaria para impulsar la reacción, ayudada por la subsiguiente hidrólisis del PP_i a 2 P_i catalizada por la pirofosfatasa. Tras muchas reacciones similares sucesivas, el producto es la nueva hebra de DNA:



Web 11.3. Mecanismo de la reacción de síntesis de DNA.

II.2.3 Dirección de la síntesis

Puesto que la adición de nucleótidos se realiza siempre mediante 5'-trifosfatos y sobre un extremo 3'-OH libre, la hebra crece por su extremo 3', y se afirma que la síntesis transcurre **en dirección 5' → 3'**. Éste es un mecanismo universal: todas las DNA polimerasas, tanto de procariotas como de eucariotas, así como todas las RNA polimerasas, sintetizan exclusivamente en dicho sentido.

II.2.4 Tipos de polimerasas

Aunque la reacción de síntesis de DNA es universal, es decir, tiene lugar de idéntica forma en todos los organismos y ubicaciones subcelulares, su catálisis la efectúan en cada caso distintas enzimas, bajo la denominación común de **DNA polimerasas** (abreviadas como DNAPol). Estrictamente su nombre es más completo, *polimerasas de DNA dirigidas por DNA* (o *dependientes de DNA*), para indicar tanto la molécula que sintetizan como la que emplean como molde. Más adelante se estudiarán otras polimerasas de DNA o de RNA que utilizan distintos moldes; para ofrecer una visión unificada se resumen a continuación:

Molécula sintetizada	Molécula molde	Nombre completo	Ejemplos (nombre común)
DNA	DNA	Polimerasa de DNA dirigida por DNA	<ul style="list-style-type: none"> • DNA polimerasas
RNA	DNA	Polimerasa de RNA dirigida por DNA	<ul style="list-style-type: none"> • RNA polimerasas o transcriptasas • Primasas
DNA	RNA	Polimerasa de DNA dirigida por RNA	<ul style="list-style-type: none"> • Transcriptasas inversas • Telomerasas
RNA	RNA	Polimerasa de RNA dirigida por RNA	<ul style="list-style-type: none"> • RNA replicasas

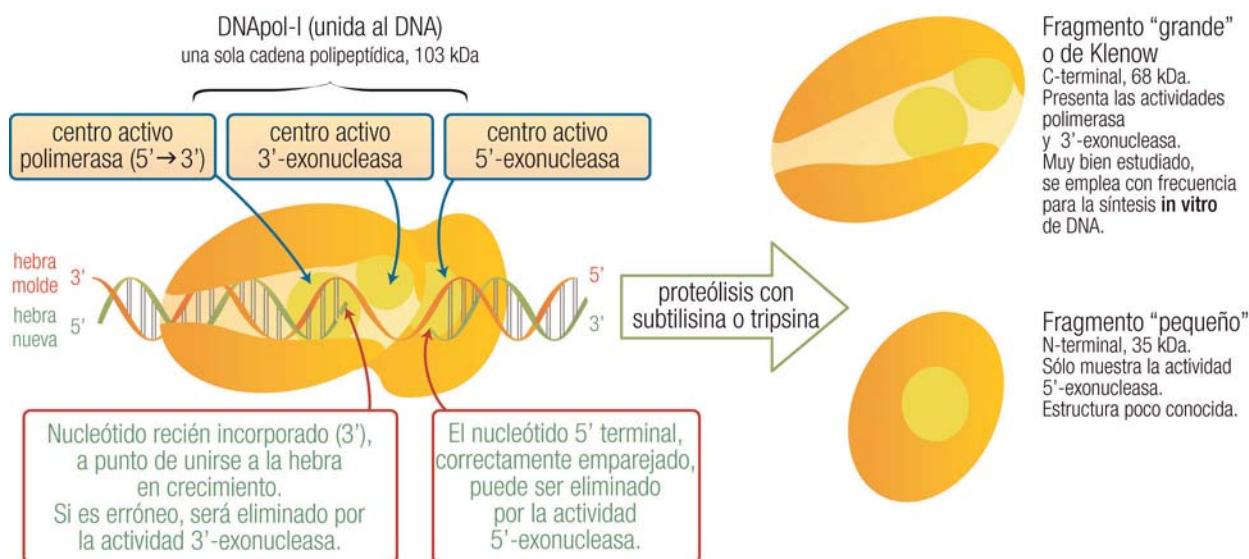
II.2.4.1 DNA polimerasas procarióticas: actividades polimerasa y exonucleasas

La primera DNA polimerasa descubierta (Arthur Kornberg, 1958) y la mejor estudiada hasta la fecha es la DNAPol-I de *Escherichia coli*. La comprobación de que su actividad no es suficiente para justificar la velocidad de replicación del cromosoma bacteriano *in vivo* condujo a la búsqueda y descubrimiento de otras enzimas: DNAPol-II y DNAPol-III y, más tarde, DNAPol-IV y DNAPol-V. Todas ellas coexisten en *E. coli*, con funciones especializadas: pol-III realiza la mayor parte de la tarea de replicación, mientras que las otras se

encargan de distintas funciones auxiliares (sus características particulares se exponen en la tabla de pág. 151, junto con las de las polimerasas eucarióticas).

Varias de las DNA polimerasas presentan, además de su actividad enzimática como polimerasa, una o dos actividades exonucleasa. La actividad 3'-exonucleasa (también denominada exonucleasa 3' → 5') permite hidrolizar el nucleótido situado en el extremo 3' de la hebra de DNA, sólo un nucleótido cada vez. Esta eliminación tiene lugar de forma inmediata, cuando el nucleótido recién incorporado por la actividad polimerasa a la cadena en crecimiento es erróneo, es decir, no empareja correctamente con el de la hebra molde de DNA. La polimerasa es así capaz de reconocer el error, detener la adición de más nucleótidos y liberar ese último dNMP mal incorporado. Los centros activos polimerasa y exonucleasa son diferentes. En resumen, las DNA polimerasas corrigen sus propios errores, en sentido (3' → 5') contrario al de la polimerización (5' → 3'); por ello, esta actividad 3'-exonucleasa se denomina también actividad correctora de pruebas (o de galeras, por analogía al proceso en imprenta) o, alternativamente, exonucleasa revisora. De este modo, mejora la fidelidad de la replicación, aumentándola entre 10^2 y 10^3 veces sobre la que aporta la incorporación de nucleótidos complementarios por la polimerasa (que tiene una tasa de error de un nucleótido por cada 10^4 o 10^5).

CARACTERÍSTICAS DE LA DNA POLIMERASA I DE *E. coli*



11.4



Web 11.4. Estructura de la DNA polimerasa, centros activos y unión al DNA.

La DNApol-I de procariotas es única entre las polimerasas por poseer, además, una segunda actividad exonucleasa, en este caso 5'-exonucleasa (o exonucleasa 5' → 3'), que le permite hidrolizar secuencialmente DNA o RNA a partir del extremo 5' de la hebra. De nuevo, el centro activo es independiente de los anteriores; en este caso está situado en el fragmento pequeño. Esta actividad desempeña un papel fundamental en la replicación del DNA procariótico, al eliminar el RNA cebador y sustituirlo por DNA durante la síntesis discontinua de la hebra retardada (pág. 158). También complementa la actividad 3'-exonucleasa corrigiendo errores de distinto tipo, y en ocasiones se denomina actividad de reparación; por ejemplo, interviene en la escisión de los dímeros de pirimidina que se forman por exposición del DNA a luz UV (pág. 398). Experimentalmente, se aprovecha la actividad 5'-exonucleasa, por ejemplo, para la preparación de sondas marcadas (pág. 187).

Algunas características de las DNA polimerasas, desde el punto de vista bioquímico y funcional

Polimerasa:	Procariontas			Eucariotas							
	I	II	III	α	β	γ	δ	ε			
Ubicación subcelular				Núcleo	Núcleo	Mitocondria	Núcleo	Núcleo			
Actividades enzimáticas:											
Primasa (inicio)	No	No	No	Sí	No	No	No	No			
Polimerasa 5' → 3' (elongación)	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí			
3'-Exonucleasa (o 3' → 5') (corrección de pruebas)	Sí	Sí	Sí	No	No	Sí	Sí	Sí			
5'-Exonucleasa (o 5' → 3') (eliminación de cebadores)	Sí	No	No	No	No	No	No	No			
Función en la célula:											
Replicación (síntesis continuada de hebras nuevas, adelantada y retardada)	No	No	Sí	Sólo inicialmente	No	Sí	Sí	Sí			
Empalme de fragmentos de Okazaki	Sí	No	No	No	No	Sí (con nucleasas)					
Velocidad de replicación:											
Velocidad de polimerización (nucleótidos/segundo)	16-20	2-7	250-1.000								
Procesividad (nucleótidos añadidos antes de su disociación)	Baja 3-200	Media ≥ 10.000	Muy alta $\geq 5 \times 10^5$	Baja 100-200	Baja	Elevada	Elevada con PCNA ≥ 5.000	Elevada con/sin PCNA			

II.2.4.2 DNA polimerasas eucarióticas: actividades polimerasa y exonucleasa

En **eucariotas**, las DNA polimerasas son diferentes a las de procariontas, aunque presentan cierta similitud de secuencia y estructura. Se han descrito más de **cinco enzimas**. La **DNApol α** posee dos actividades enzimáticas independientes: polimerasa (de DNA) y primasa (polimerasa de RNA). Gracias a su actividad primasa se encarga de formar un RNA cebador –necesario para iniciar la replicación– y, por su actividad polimerasa, de la elongación inicial de ese cebador. Esta tarea se cede luego a las **DNApol δ** y **ε**, responsables de la mayor parte de la elongación de ambas hebras; de hecho, ambas tienen actividades similares, pero aún no está claro si desempeñan indistintamente la misma función o si están especializadas. La **DNApol β** no interviene en la replicación, sino que está implicada en la reparación de errores o daños en el DNA (procesos en los que también participan otras polimerasas, incluyendo **δ** y **ε**). Mientras estas cuatro enzimas son responsables de la replicación, reparación y recombinación del DNA nuclear, el DNA mitocondrial lo sintetiza una polimerasa diferente, la **DNApol γ**. Por último, se han descubierto otras polimerasas nucleares, peor conocidas (θ [theta], ζ [zeta], λ [lambda], μ [mu], κ [kappa], η [eta], ι [iota]), que están implicadas sobre todo en reparación y recombinación.

En cuanto a la actividad exonucleasa de las DNA polimerasas eucarióticas, sólo las 3 que se encargan de la elongación (**δ**, **γ** y **ε**, v. tabla) poseen actividad **3'-exonucleasa**, la responsable de la capacidad de corrección de errores cometidos al incorporar los nucleótidos a la hebra que se está sintetizando. Ninguna de las polimerasas eucarióticas parece tener actividad **5'-exonucleasa**; la función de eliminación de cebadores la desempeñan dos nucleasas independientes, aunque las polimerasas **δ** y **ε** sí se encargan de llenar el hueco con DNA; todos estos detalles se estudian posteriormente (pág. 158).

La función específica de cada polimerasa sólo se puede comprender en el contexto de los mecanismos de iniciación y elongación de las cadenas de DNA, que se estudian con detalle dentro de la etapa de elongación (pág. 157).

Bajo el alcance de este libro no parece adecuado describir las propiedades estructurales de las distintas polimerasas. En la tabla siguiente se resumen los datos relativos a tamaño y composición en subunidades, pero debe tenerse en cuenta que la variabilidad de cifras es elevada, en especial para las enzimas eucarióticas, e incluso la interpretación de éstas en los distintos organismos estudiados, como para su generalización a todos los eucariotas.

Características estructurales de las DNA polimerasas

Polimerasa:	Prokariotas ¹			Eucariotas ²				
	I	II	III	α	β	γ	δ	ε
Masa total (kDa) ⁽³⁾	103	90	≈ 900	356	39	160	173	313
N.º de subunidades	1	≥ 4	22	4	1	3	2 o 3	4
Masa de subunidades			⁽⁴⁾	⁽⁵⁾		125	125	258
						35	48	55

¹Al menos en *E. coli* se han identificado otras dos polimerasas, IV y V.

²Los datos corresponden a humanos. Existen otras DNA polimerasas minoritarias o no esenciales.

³Masa molecular de la holoenzima, o enzima funcional con todas sus subunidades.

⁴La DNAPol-III de *E. coli* está formada por 22 subunidades de 10 tipos, con una composición global $(\alpha, \epsilon, \theta)_2\tau_2(\gamma, \delta, \delta', \theta, \psi)_2\beta_4$. De ellas, la agrupación $(\alpha, \epsilon, \theta)$ se conoce como “núcleo” o “centro” de la polimerasa. El centro activo polimerasa está en la subunidad α (130 kDa), mientras que el centro 3'-exonucleasa o actividad correctora de pruebas está en la ϵ .

⁵La DNAPol α humana está formada por 4 subunidades diferentes: 180 kDa, con el centro activo polimerasa; 48 kDa, con el centro activo primasa; 58 kDa, cofactor de la primasa; 70 kDa, reguladora de la polimerasa.

11.2.5 Velocidad de replicación

Dado el gran tamaño del genoma de una célula, para que su replicación completa se lleve a cabo en un tiempo razonable el proceso de síntesis debe ser muy rápido. En prokariotas, la velocidad alcanza 1.000 nucleótidos por segundo. En eucariotas, el mayor tamaño del genoma hace pensar en la necesidad de una gran velocidad de replicación que, sin embargo, no precisa ser tan elevada gracias a la existencia de varios orígenes de replicación en cada cromosoma, lo que reduce la longitud de DNA que debe sintetizar cada molécula de polimerasa.

La consecución de una alta velocidad de síntesis sólo es posible gracias a un mecanismo peculiar, según el cual la molécula de DNA polimerasa se mantiene unida a la cadena de DNA molde, a la vez que asociada por su centro activo al extremo 3'-OH de la hebra creciente. De esta forma puede incorporar numerosos nucleótidos sucesivamente, con una reducción del tiempo necesario para la adición de cada uno: por ejemplo, la polimerasa puede tardar un minuto en soltarse y volverse a unir al DNA (por mera difusión de las moléculas), mientras que para añadir un nucleótido sin haberse soltado le bastan unos milisegundos. Esta capacidad de las polimerasas, variable entre ellas, se recoge bajo el término *procesividad*, que se define como el número de nucleótidos que son incorporados a la hebra antes de que la enzima se separe del molde (coloquialmente se diría añadidos “de un tirón”). Un valor elevado de procesividad en una polimerasa es mucho más importante para la replicación que una velocidad alta de catálisis.

Las polimerasas de mayor procesividad son las que intervienen en la elongación de las cadenas de DNA (v. tabla en pág. 151), mientras que las de baja procesividad se encargan de tareas que podrían calificarse de auxiliares, tales como la síntesis y eliminación de cebadores, su sustitución por DNA y la reparación. En general, la alta procesividad se consigue gracias a una molécula que “abraza” o rodea la cadena de DNA, evitando así que la enzima se separe; este mecanismo se denomina de “abrazadera deslizante”, pues permite el deslizamiento continuado de la polimerasa a lo largo de la cadena de DNA, sin separarse de ella.

II.3 ETAPAS EN EL PROCESO DE REPLICACIÓN

Mientras que en procariotas es probable que la replicación tenga lugar en conexión con la membrana celular, pues el único cromosoma está anclado a ella en la zona nucleoide, en eucariotas el proceso ocurre, lógicamente, en la matriz nuclear. El DNA se replica una vez por ciclo celular, durante la fase S.

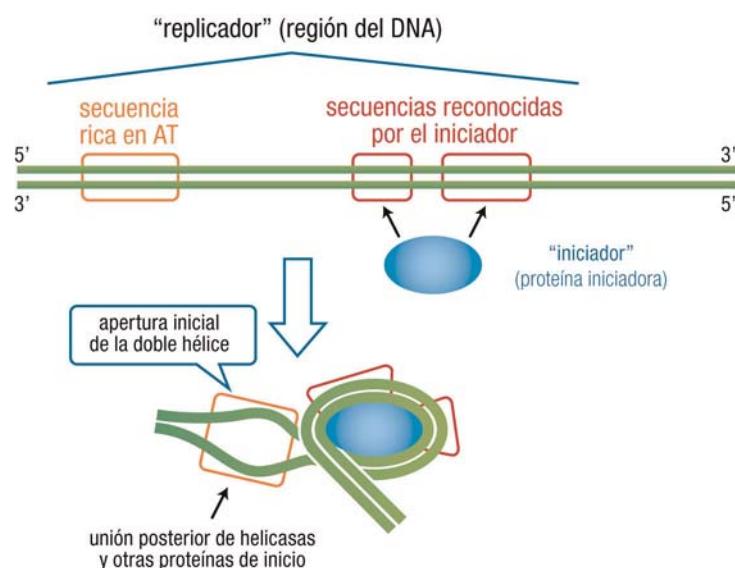
El mecanismo del proceso de replicación permite definir una unidad funcional denominada **replicón**, como aquella región de DNA que se replica a partir de cada origen (v. web 11.2). Cada replicón contiene, pues, un origen de replicación y se replica mediante dos horquillas. Mientras en procariotas el cromosoma completo constituye un solo replicón, en eucariotas son numerosos y sufren la replicación de forma casi simultánea, lo que supone la existencia de un complejo mecanismo de control que coordine el inicio, elongación y terminación en todos los replicones de cada cromosoma. El conocimiento preciso de estos aspectos está pendiente, al igual que otros relacionados, por ejemplo, con la replicación de los centrómeros y de otras regiones heterocromáticas altamente condensadas.

Centrándonos en el inicio y progreso de la replicación en cada replicón, deben plantearse cuestiones relativas a cómo son las interacciones específicas entre DNA y proteínas que desencadenan la iniciación, a la consiguiente actuación de estas proteínas y al control del proceso. Por delante de la horquilla de replicación debe producirse la descondensación de los cromosomas, la disociación de los nucleosomas y la separación de las histonas y las proteínas no histónicas (Capítulo 7). Por detrás de la horquilla, deben reutilizarse las histonas y proteínas no histónicas para la reasociación de los nucleosomas, volverse a condensar la cromatina, etc.

II.3.1 Iniciación

La replicación siempre comienza en puntos definidos de la molécula de DNA, con secuencias características. Como se ha comentado (web 11.2), en **procariotas** el cromosoma tiene un origen de replicación único (llamado *oriC*) mientras que en **eucariotas** la enorme longitud de los cromosomas requiere la existencia de numerosos orígenes de replicación que, aunque se conocen con menos precisión, también poseen secuencias características.

La definición de los puntos de inicio depende de dos regiones en el DNA: el **replicador**, que es la zona que determina la apertura inicial de la doble hebra, y el **origen** propiamente dicho, el punto donde comienza la síntesis sobre el molde de la molécula abierta. El replicador incluye zonas donde la doble hélice puede desenrollarse con facilidad (secuencias ricas en pares AT) y otras que interaccionan específicamente con **proteínas iniciadoras**. La unión de estas proteínas provoca la flexión de la cadena de DNA, generando una tensión superhelicoidal negativa que, sumada a la interacción más débil entre las bases A y T, contribuye a la separación de las hebras. En eucariotas, en particular, se han identificado también proteínas iniciadoras que, en lugar de flexionar el DNA, se unen a regiones palindrómicas que pueden formar estructuras cruciformes (pág. 55); el resultado es análogo, pues las cruciformes reducen la tensión de superenrollamiento y así facilitan la apertura de la doble hélice.



Las proteínas iniciadoras también actúan uniendo otras proteínas, que son así reclutadas al sitio de inicio. Una vez abierta la doble hebra (lo que supone una desnaturización local, v. pág. 164), tanto la helicasa como el resto de proteínas responsables del inicio –incluidas las propias DNA polimerasas– ya pueden asociarse al DNA y comenzar el proceso de síntesis.



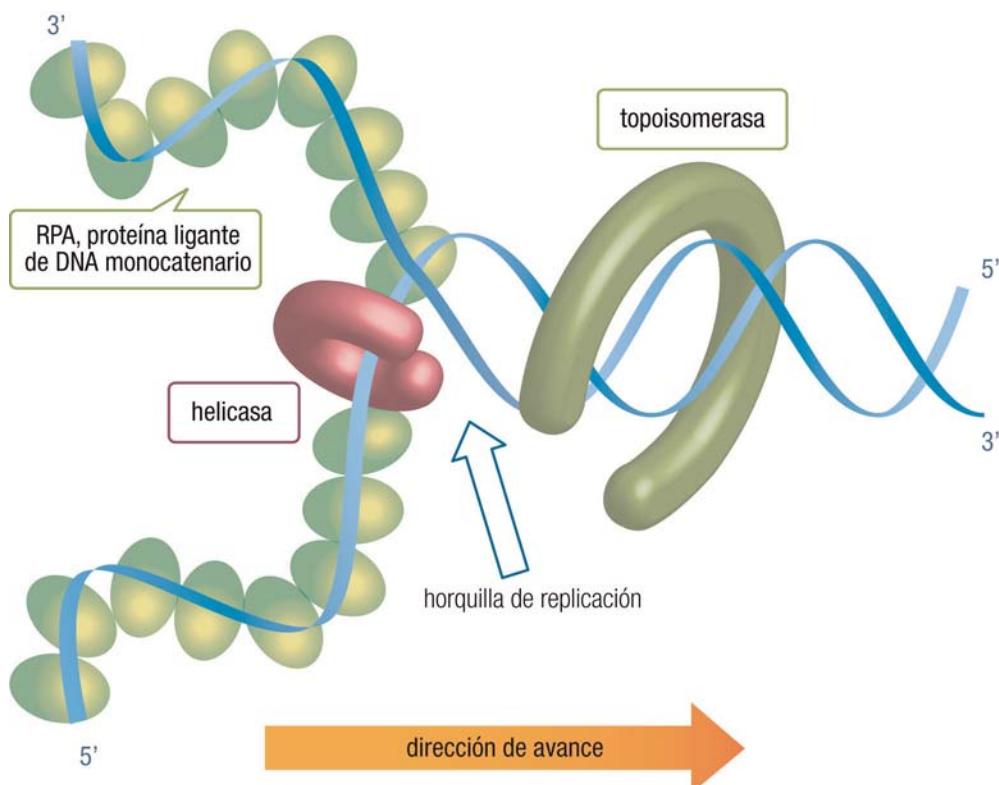
Web 11.6. Inicio de la replicación.

11.3.2 Elongación

11.3.2.1 Conexión entre la iniciación y la elongación: proteínas necesarias

Una vez formado el complejo de iniciación, la replicación precisa el desenrollamiento de la doble hélice, a fin de hacer accesibles las bases como molde para formar las nuevas hebras. El desenrollamiento inicial en el tramo correspondiente al origen de replicación, por efecto de las proteínas iniciadoras, debe continuar avanzando por delante de la polimerasa, encabezando la horquilla de replicación. Además, a medida que se van sintetizando las hebras nuevas se debe ir recuperando el enrollamiento en las dos dobles hélices recién formadas. En estas operaciones de avance, desenrollamiento o relajación y enrollamiento o compactación intervienen diversas proteínas especializadas. El complejo multiproteico formado por éstas y por las DNA polimerasas, que viaja asociado a cada horquilla realizando la replicación, se conoce como **replisoma**.

Aunque la identidad molecular y el nombre de las proteínas implicadas varían dependiendo del tipo de organismo (bacterias, levaduras, mamíferos, etc.) básicamente todos utilizan el mismo conjunto de actividades enzimáticas y siguen mecanismos análogos.



11.6

a) Helicasas

Estas proteínas se unen al DNA una vez que éste se ha abierto en el origen y actúan propagando la separación de sus dos hebras, en un proceso que requiere la hidrólisis simultánea de ATP. En general son hexámeros en forma de anillo que rodea a una de las hebras del DNA. Una vez creadas las dos horquillas de replicación,

comienzan a avanzar en sentidos opuestos, cada una liderada por una helicasa seguida inmediatamente por el resto de componentes del replisoma o complejo de elongación asociado a la horquilla.

b) Proteínas ligantes de DNA monocatenario

Una vez separadas las hebras por la helicasa, intervienen las “proteínas ligantes de DNA monocatenario”, que evitan el reemparejamiento de las hebras. De esta forma, las hebras molde pueden acoger a los nucleótidos que se incorporan en la replicación. Se las denomina **proteínas SSB** (*single strand binding proteins*), también de forma genérica pero de modo específico en procariotas. En eucariotas desempeña esta función la **proteína de replicación A (RPA)**.

c) Topoisomerasas

La apertura de la doble hélice ocasionada por la progresión de la helicasa a lo largo del DNA bicatenario conlleva la aparición de superenrollamientos positivos por delante de la horquilla. Si se tratase de un DNA lineal y corto, el superenrollamiento se podría eliminar por simple rotación de la molécula, en torno a su eje, en sentido contrario. Sin embargo, esto no es posible en los DNA circulares (procarióticos) ni en los lineales de gran tamaño (eucarióticos), que están restringidos topológicamente (Capítulo 7) debido a su naturaleza circular y a su unión a proteínas, respectivamente. Para aliviar esa tensión de torsión se requiere, pues, algún tipo de mecanismo giratorio; de lo contrario, el superenrollamiento acumulado por delante de la horquilla llegaría a impedir el avance de la replicación. Este problema se resuelve gracias a la acción de las topoisomerasas, enzimas que alteran el estado de superenrollamiento del DNA sin modificar su estructura en otros aspectos. Es decir, interconvierten topoisómeros modificando el *número de enlace* del DNA (número de vueltas que da una hebra respecto a la otra, pág. 82). Las topoisomerasas también intervienen en los procesos de recombinación y, posiblemente, en la transposición, constituyendo un mecanismo básico general de modificación topológica de la molécula de DNA.



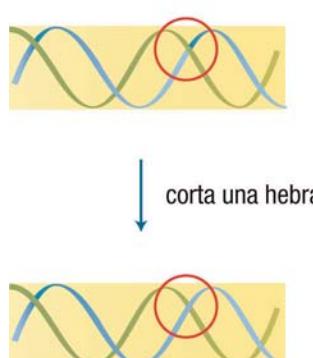
Web 11.7. Superenrollamiento asociado al avance de la replicación.

Atendiendo a su modo de acción, se distinguen dos familias de topoisomerasas:

Las **topoisomerasas de tipo I** escinden de manera transitoria una hebra del DNA (un enlace fosfodiéster), dejan pasar la otra por la brecha abierta y vuelven a ligar la primera hebra. Como consecuencia, el número de enlace aumenta o disminuye en una unidad. La reacción transcurre a través de un intermediario covalente entre uno de los extremos del corte y un residuo de tirosina de la enzima. Al permanecer la enzima unida así al DNA abierto durante la acción catalítica, se controla en todo momento el grado de tensión liberado.

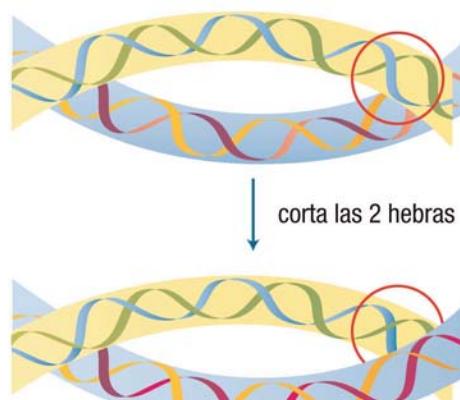
Acción de las topoisomerasas

de tipo I:



corta una hebra

de tipo II:



corta las 2 hebras

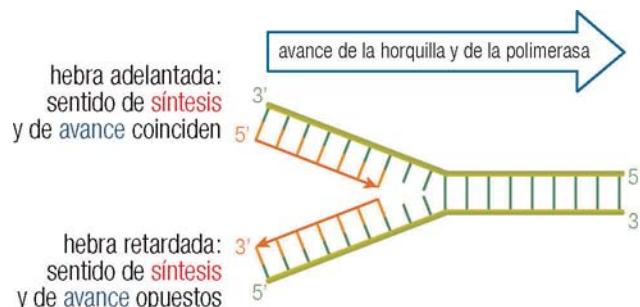
Las topoisomerasas de tipo II, cuyo ejemplo clásico es la **girasa** procariótica, cortan de manera transitoria las dos hebras del DNA, permiten el paso de otra doble hebra, intacta, por la brecha abierta y finalizan resellando los cortes. Por consiguiente, el número de enlace aumenta o disminuye en dos unidades. La misión de estas enzimas es generalmente inducir superenrollamientos negativos que facilitan el posterior desemparejamiento local. El consumo energético del proceso es importante: se hidroliza una molécula de ATP por cada corte de la cadena.

d) Primasa

Como ya se ha indicado, las DNA polimerasas no son capaces de iniciar la síntesis de una molécula de DNA, es decir, unir los dos primeros nucleótidos, sino sólo de elongar una hebra preexistente; por ello, el inicio de las hebras nuevas requiere un cebador. Éste lo sintetiza una polimerasa de RNA dirigida por DNA (pág. 149), denominada **primasa** (de *primer*, cebador en inglés), que produce cadenas cortas de RNA, complementarias de cada una de las hebras desemparejadas en el origen de replicación. En procariotas, la primasa es una proteína independiente que forma parte de un complejo proteico denominado *primosoma*, asociado a la horquilla. En eucariotas, la actividad primasa reside en una de las subunidades de la DNA polimerasa α , distinta de la subunidad que posee el centro activo polimerasa (v. tabla en pág. 152); por ello, esta proteína con dos actividades se llama **DNApol α /prim**.

En resumen, la replicación no sólo depende de la DNA polimerasa, sino de numerosas enzimas y factores proteicos (más de 20 en *E. coli*). Este conjunto de proteínas que, como se ha indicado, recibe el nombre de *replisoma* o *complejo de replicación*, se asocia en torno a la horquilla de replicación, participando de forma directa o indirecta en la síntesis simultánea de ambas hebras.

11.3.2.2 Asimetría de la replicación en ambas hebras



11.8

Conceptualmente, el principal problema de la replicación surge al constatar que ambas hebras se copian de manera simultánea, a medida que avanza la horquilla, mediante enzimas (DNA polimerasas) que sólo son capaces de sintetizar en dirección $5' \rightarrow 3'$. Al abrirse la horquilla, una de las hebras progenitoras se va exponiendo en el sentido correcto para actuar de molde, $3' \rightarrow 5'$, por lo que la síntesis de su hebra complementaria puede transcurrir por simple avance de la polimerasa a lo largo del molde a la par que avanza la horquilla. Esta hebra nueva recibe el nombre de **hebra adelantada** (*leading strand*; a veces, hebra líder o hebra guía); en ocasiones, su hebra molde recibe el mismo nombre, pero para evitar confusiones la llamaremos “molde de la hebra adelantada”. Por el contrario, la otra hebra progenitora se expone en sentido $5' \rightarrow 3'$, por lo que no puede actuar como molde a medida que avanza la horquilla. Como se verá en seguida, la síntesis de su complementaria requiere un mecanismo particular que, entre otras cosas, supone un desfase con la síntesis de la hebra adelantada, por lo que se denomina **hebra retardada o retrasada** (*lagging strand*).

La solución al problema de síntesis de la hebra retardada se desveló al observar que durante la replicación era posible aislar pequeñas moléculas de DNA recién sintetizado (más cortas en eucariotas que en procariotas), conocidas como **fragmentos de Okazaki** en honor a Reiji y Tuneko Okazaki, quienes fueron sus descubridores. Se pudo entonces proponer un mecanismo que explicase la síntesis de la hebra retardada en paralelo al desplazamiento de la horquilla de replicación. Dicho mecanismo consiste en la síntesis discontinua de la hebra retardada: cada fragmento de Okazaki corresponde a una porción de dicha hebra, cuya síntesis sólo comienza

(de ahí el nombre) cuando el avance de la horquilla ha liberado suficiente longitud de DNA monocatenario como para que la polimerasa lo utilice como molde, en sentido $3' \rightarrow 5'$. Se propone para ello que la hebra sencilla forma un bucle para permitir la síntesis en el sentido $5' \rightarrow 3'$ propio de la polimerasa. La síntesis conjunta de ambas hebras (la adelantada, de forma continua, y la retardada, de forma discontinua), por sendas moléculas de DNA polimerasa se dice que es **semidiscontinua**.

Lógicamente, la síntesis de cada fragmento de Okazaki requiere un RNA cebador, sintetizado por una actividad primasa (parte del *primosoma* en procariotas, DNAPol α /prim en eucariotas) que viaja con la horquilla y que también sintetizó en el origen de replicación el único cebador requerido por la hebra adelantada.

11.3.2.3 Mecanismo de la elongación

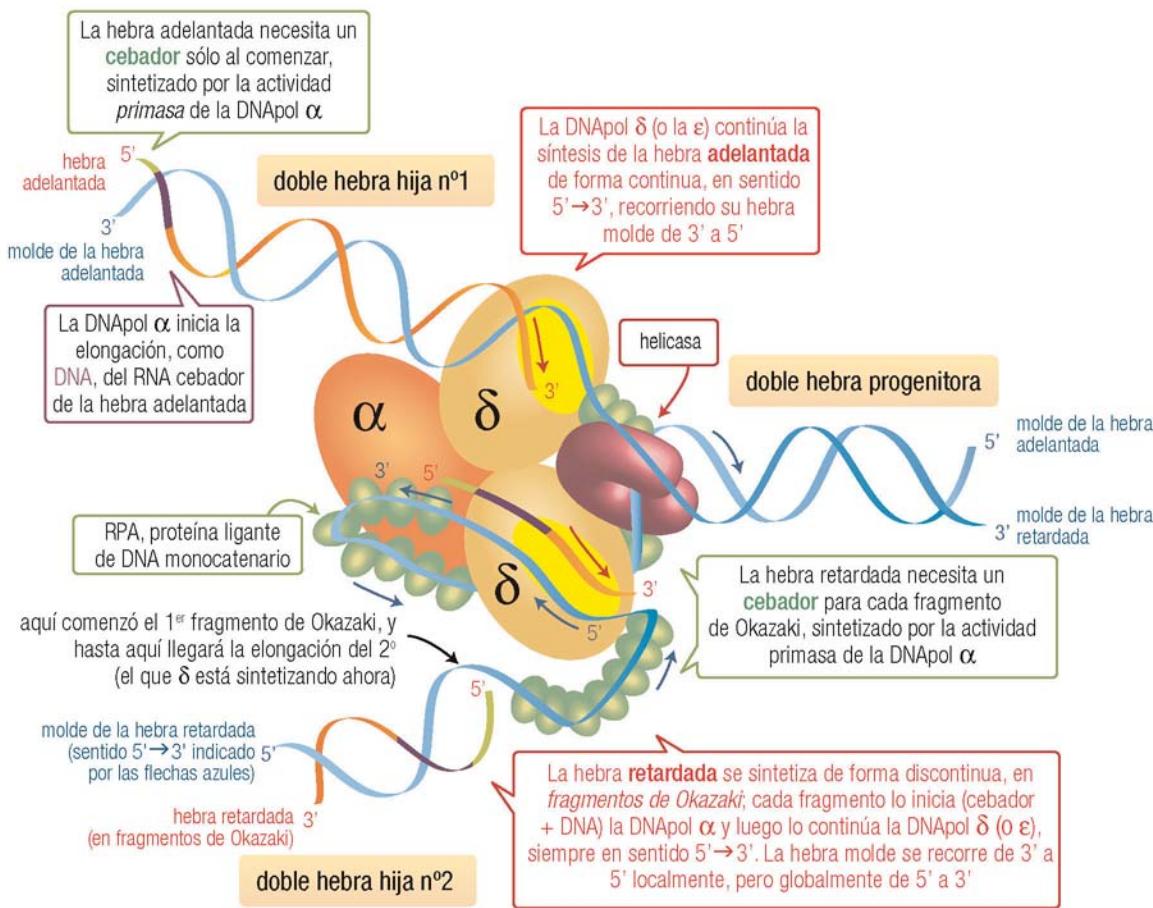


Web 11.8. Mecanismo de la elongación en replicación y organización del replisoma.

Representación global aproximada de la organización de las proteínas responsables de la replicación (componentes del replisoma) en torno a la horquilla de replicación.

La situación mostrada corresponde a un momento intermedio de la elongación, con un fragmento de Okazaki sintetizado a medias (p.ej., el nº 2 en las figuras anteriores, etapa B).

Por claridad, no se muestran otros factores proteicos, como el PCNA que va asociado a las DNAPol δ , el sistema de nucleasas FEN1/RNasa H1 que madura los fragmentos de Okazaki, la DNA ligasa I, y las topoisomerasas que regulan el superenrollamiento en las dobles hebras progenitora e hijas.



11.3.2.4 Maduración de los fragmentos de Okazaki

Para completar la síntesis de la hebra retardada deben unirse los fragmentos de Okazaki. Este proceso, denominado globalmente maduración, requiere la eliminación de los cebadores, la elongación del fragmento adyacente para llenar con DNA el hueco dejado por cada RNA cebador y la unión o empalme de los extremos resultantes para dar una hebra continua. Todo ello tiene lugar posiblemente en las proximidades de la propia horquilla, o bien posteriormente, lejos de ella.



Web 11.9. Maduración de los fragmentos de Okazaki: eliminación de cebadores y reemplazo con DNA.

La última etapa en esta maduración de la hebra retardada, la unión de los fragmentos, la catalizan DNA ligasas (o simplemente ligasas). Estas enzimas forman el enlace fosfoéster que falta entre el grupo OH del extremo 3' de un fragmento de Okazaki y el fosfato del extremo 5' del siguiente, asegurando así que la hebra retardada sea una cadena continua. Además de este papel fisiológico, las ligasas son enzimas de gran utilidad en el laboratorio, junto con las nucleasas de restricción, especialmente para la preparación de moléculas de DNA recombinante (pág. 218).

11.3.3 Terminación

Una vez descritos los detalles particulares del inicio y la elongación de la replicación corresponde estudiar los mecanismos de la terminación. Debe indicarse que este proceso no se conoce tan bien como los anteriores, entre otras razones por la propia amplitud y complejidad del concepto de terminación. Se tratará aquí de exponer los problemas que plantea la terminación desde un punto de vista lógico y asumiendo su limitada confirmación experimental.

11.3.3.1 Final de la elongación

Se puede considerar, conceptualmente, que la terminación comprende varios aspectos. El primero de ellos es la finalización propiamente dicha de la elongación por la DNA polimerasa, que se verifica en las dos horquillas de replicación de cada replicón, previsiblemente cuando en su avance alcancen a las horquillas respectivas de replicones adyacentes. Éste es un proceso prácticamente desconocido en términos moleculares, pero cuya resolución no plantea dificultad conceptual, pues el relleno de huecos y la unión de los dos tramos de nueva hebra pueden resolverse de la misma forma que la maduración y unión de los fragmentos de Okazaki.

Por otra parte, como segundo problema, falta por aclarar cuál es el número de replicones implicados de forma simultánea en la replicación de cada cromosoma, así como los mecanismos de control que coordinan el proceso en todos ellos. A diferencia del caso anterior, no existe aún una hipótesis previsible.



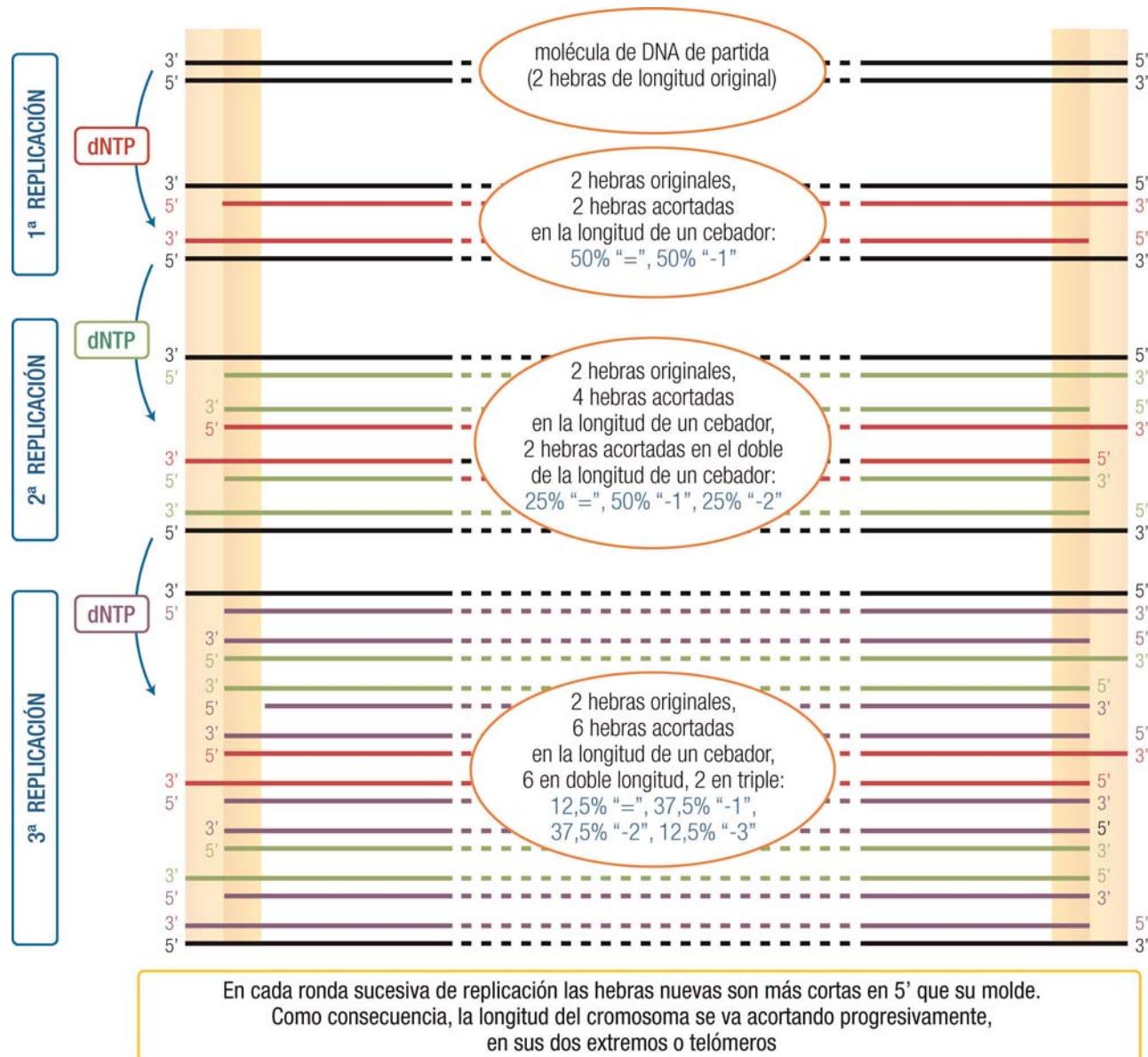
Web 11.10. Sentido de la replicación en los diferentes replicones y replicación incompleta de los extremos.

11.3.3.2 Replicación de los telómeros

La replicación de los extremos del cromosoma eucariótico supone el tercer problema de la terminación, pues, como se verá en seguida, no puede llevarse a cabo mediante la elongación por la DNA polimerasa. Este proceso es, por ahora, el mejor conocido de la terminación, posiblemente por su mayor trascendencia fisiológica y patológica.

Como ya se ha estudiado, los telómeros son regiones situadas en los extremos de los cromosomas lineales de la mayoría de eucariotas (pág. 94), a las que se unen ciertas proteínas. Están constituidos por una de las familias de DNA repetitivo no codificador (pág. 116), formando bloques de 1 a 12 kb de longitud, con unidades de 6 a 8 pb repetidas en tandem centenares o miles de veces. Para entender la implicación de los telómeros en la replicación deben recordarse sus dos características principales: la abundancia de repeticiones en tandem y el carácter *cohesivo* de los extremos (que recuerda, aunque sin relación funcional alguna, a los extremos resultantes de la acción de ciertas enzimas de restricción sobre un DNA, pág. 215).

El carácter cohesivo, en forma de salientes 3' en una de las hebras de cada extremo de los cromosomas, tanto progenitores como filiales, se explica por el mecanismo de replicación discontinua de la hebra retardada (v. web 11.10). Como consecuencia adicional de dichos salientes 3', se produce un acortamiento progresivo de las hebras de DNA en sucesivas rondas de replicación.



11.10



Web 11.11. Acortamiento de los telómeros en la replicación.

Este acortamiento de las regiones teloméricas (calculado entre 50 y 200 pb por cada división celular) supone la disminución progresiva de la longitud del cromosoma a lo largo de sucesivas divisiones celulares, con el consiguiente problema para la perpetuación del mensaje genético. Este problema se resuelve en la célula gracias a la actuación, también durante la fase S del ciclo celular (Capítulo 8), de un nuevo tipo de enzima, la **telomerasa**, que alarga los telómeros, compensando así la pérdida anterior.

11.3.3 Telomerasa: componentes y características

Se ha propuesto que la telomerasa es un resto evolutivo de una ribozima (pág. 71) que pudo servir anteriormente para catalizar la síntesis de DNA; también, que pudo suponer un intermediario en la transición evolutiva desde las RNA replicasas a las transcriptasas inversas proteicas. Se trata de una enzima de tipo DNA polimerasa, concretamente una *polimerasa de DNA dirigida por RNA* (pág. 149) y, por tanto, es una transcriptasa inversa (págs. 207 y 223), pero con características especiales. En efecto, se trata de una ribonucleoproteína que sólo es capaz de sintetizar una secuencia determinada de DNA, y que está formada por dos componentes:

1. **Componente ribonucleico o “RNA de la telomerasa”** (*TR* o *TER*, de *telomerase RNA*), de tamaño variable según especies (desde 146 nucleótidos hasta 1.544), que forma parte integral de la enzima. Incluye una secuencia, característica para cada especie, de 9 a 28 nt y complementaria a la secuencia telomérica respectiva. Por ejemplo, el RNA componente de la telomerasa humana (*hTR*) tiene 450 nt entre los que hay una secuencia rica en C, (5')CCCUAA(3'), complementaria y antiparalela a la hebra rica en G de la repetición telomérica (5')TTAGGG(3'). Gracias a ella, como se verá luego, el TR sirve de molde para la síntesis de nuevas secuencias teloméricas.
2. **Componente proteico**, que contiene la actividad enzimática, o “transcriptasa inversa de la telomerasa” (*TRT* o *TERT*, de *telomerase reverse transcriptase*). Ésta es la actividad telomerasa propiamente dicha. A diferencia de la transcriptasa inversa vírica y del resto de DNA polimerasas, para sintetizar el DNA telomérico no utiliza un cebador de RNA sintetizado a tal fin, sino que elonga el extremo telomérico 3' precedente. La segunda diferencia importante es que el molde necesario para su acción es la parte RNA de la propia enzima (el TR). Generalmente el TRT suele ir acompañado de otras subunidades no catalíticas, también proteicas.

11.3.4 Mecanismo de acción de la telomerasa



Web 11.12. Mecanismo de acción de la telomerasa.

11.3.5 Funcionalidad de la actividad telomerasa en las células

En las células somáticas de la mayoría de tejidos adultos apenas existe actividad telomerasa, mientras que sí está presente en células de la línea germinal, en tejidos fetales y en ciertas células troncales (principalmente, hematopoyéticas), todas ellas células poco diferenciadas y en rápida proliferación. Esto se debe a la regulación de la expresión génica, al igual que para cualquier otra proteína (Capítulo 18), pero aquí no se necesita entrar en los mecanismos por los que se evita la expresión de la telomerasa, sino que interesa destacar sus consecuencias.

En ausencia de telomerasa, a medida que las células se van dividiendo –y envejeciendo– sufren un proceso constante de acortamiento de sus telómeros (pág. 159). Dado el carácter repetitivo y no codificante de éstos, inicialmente esa pérdida de los extremos de cada cromosoma no tiene consecuencias, pero llega un punto, agotadas todas las repeticiones teloméricas, en que afecta a regiones codificantes, y entonces la función celular se ve perjudicada y la célula termina por morir. Se ha considerado este acortamiento como parte de un *reloj biológico o celular* que mide el proceso de envejecimiento, “llevando la cuenta” del número de divisiones hasta que la célula debe morir, para controlar así su proliferación correcta en los tejidos.

Puesto que en la mayor parte de las células de un organismo no hay actividad telomerasa, este acortamiento es fisiológico y se produce continuamente durante toda la vida de un individuo. En efecto, se ha demostrado que la pérdida de repeticiones teloméricas aumenta proporcionalmente al número de divisiones mitóticas experimentadas o a la edad del individuo al que pertenecen las células. En consonancia con este hecho, se ha observado que la introducción artificial de telomerasa en células somáticas normales repone la pérdida de telómeros. Se concluye, por tanto, que el acortamiento telomérico está relacionado de forma inversa con el grado de inmortalidad celular y, en definitiva, que el envejecimiento es una consecuencia de la pérdida de telómeros con la edad. Ello está también de acuerdo con la pérdida acelerada de telómeros observada en pacientes que sufren varios síndromes de envejecimiento prematuro.

La telomerasa está presente, por el contrario, en casi todas las células que proliferan, en especial las células tumorales, y contribuye, junto a otros factores, a la proliferación indefinida de los tumores. Así, el aumento de actividad telomerasa observado en fases tardías del progreso tumoral puede ser necesario para el crecimiento del tumor, al impedir la muerte celular.

Por consiguiente, se ha propuesto la inhibición de la actividad telomerasa como un principio de terapia antitumoral y la telomerasa se considera una diana terapéutica. Se buscan inhibidores específicos del componente catalítico de la telomerasa humana, fármacos y factores que actúen modulando la telomerasa o cualquier otro componente asociado a los telómeros. Estos aspectos tienen un gran interés aplicado en relación con los problemas de envejecimiento y cáncer.

11.4 REPLICACIÓN MITOCONDRIAL

La replicación del cromosoma mitocondrial humano se realiza por un mecanismo peculiar, diferente al estudiado para los casos de procariotas y genoma nuclear de eucariotas. Previamente, se deben introducir algunos conceptos relativos a esta molécula de DNA.

En el cromosoma mitocondrial, circular, una de las hebras posee un mayor contenido de bases purínicas, por lo que se conoce como **hebra pesada** o H (de *heavy*, en inglés), mientras que la otra es la **hebra ligera** o L (*light*). Por otra parte, una zona particular de la molécula recibe el nombre de **asa de desdoblamiento** o de **desplazamiento**, o bucle D (*D-loop*); se trata de la única región no codificante del DNA mitocondrial, que desempeña el papel de región de control, pues contiene el origen de replicación de la hebra H y los orígenes de transcripción de ambas hebras.

La replicación requiere un cebador, sintetizado por la RNA polimerasa mitocondrial usando la hebra L como molde (a partir del promotor P_L, situado también en la región del bucle D, pág. 276). Entonces la DNA polimerasa γ comienza la replicación propiamente dicha a partir del origen O_H, elongando el cebador para sintetizar la hebra H hija. El progreso de la elongación requiere la participación de proteínas ligantes de DNA monocatenario, mtSSB. Una vez que la horquilla de replicación ha sobrepasado la posición del origen O_L se comienza sobre él la replicación de la otra hebra, empleando como molde la hebra H y sintetizando la nueva hebra L.



Web 11.13. Replicación del cromosoma mitocondrial.

La primera característica singular de la replicación del DNA mitocondrial es que no tiene lugar de manera específica en la fase S, sino a lo largo de todo el ciclo celular. Esto concuerda con la división de las mitocondrias durante la interfase, en desfase con la división de otros orgánulos y de la célula. Además, no todas las moléculas de DNA se replican una vez por ciclo, sino que lo hacen al azar.

La segunda peculiaridad que diferencia esta replicación de la del genoma nuclear es su avance **unidireccional**, mediante dos horquillas que parten de orígenes distintos y avanzan en sentido opuesto sintetizando una sola hebra cada una. Como consecuencia de este mecanismo, la replicación también se diferencia de la del DNA nuclear por ser **no simultánea, bifocal y continua** (no se forman fragmentos de Okazaki, las dos hebras se sintetizan como adelantadas).

Carácter	Replicación del genoma nuclear de eucariotas	Replicación del genoma mitocondrial de eucariotas	Replicación del genoma procariótico
Semiconservadora	Sí	Sí	Sí
Simultánea en ambas hebras	Sí	No	Sí
Secuencial	Sí	Sí	Sí
Bidireccional	Sí	No	Sí
Origen	Múltiples (multifocal)	Dos, uno para cada hebra (bifocal)	Único (monofocal)
Semidiscontinua	Sí	No (continua en ambas hebras)	Sí

Hibridación de ácidos nucleicos



12.1 INTRODUCCIÓN: DESNATURALIZACIÓN Y RENATURALIZACIÓN DEL DNA	164		
12.1.1 Agentes desnaturalizantes	164		
12.1.1.1 Desnaturalización y renaturalización por efecto de la temperatura	164		
12.1.1.2 Desnaturalización por ácidos y bases	165		
12.1.1.3 Desnaturalización por agentes químicos	165		
12.1.2 Influencia de la desnaturalización sobre las propiedades del DNA	165		
12.1.2.1 Influencia sobre la viscosidad	165		
12.1.2.2 Influencia sobre la absorción ultravioleta	166		
12.1.2.3 Significado práctico de las curvas de desnaturalización	167		
12.2 ANÁLISIS MOLECULAR DE LA HIBRIDACIÓN DE ÁCIDOS NUCLEICOS	168		
12.2.1 Principio básico del ensayo de hibridación	168		
12.2.2 Influencia de algunos factores sobre la hibridación	168		
12.2.2.1 Rigor de la hibridación	168		
12.3 MÉTODOS DE ENSAYO CON HIBRIDACIÓN	169		
12.3.1 Hibridación en fase líquida	169		
12.3.2 Hibridación en soporte sólido	170		
12.3.2.1 Hibridación "dot-blot" y "slot-blot"	171		
12.3.2.2 Hibridación de Southern	171		
12.3.2.3 Hibridación Northern	172		
12.3.3 Matrices de DNA	172		
12.3.3.1 Matrices de oligonucleótidos:	173		
chips de DNA	173		
a Matrices de genotipado	174		
b Matrices de expresión	174		
12.3.3.2 Matrices de ácidos nucleicos	175		
12.3.4 Hibridación <i>in situ</i>	175		
12.3.4.1 Hibridación <i>in situ</i> tisular	176		
12.3.4.2 Hibridación <i>in situ</i> cromosómica	177		
12.3.5 Técnicas relacionadas	178		
12.3.5.1 Transferencia Western	178		
12.4 PREPARACIÓN DE SONDAS	179		
12.4.1 Clasificación de las sondas, en función del método de preparación	180		
12.4.2 Sondas convencionales	180		
12.4.2.1 Sondas de DNA obtenidas por clonación celular (tecnología del DNA recombinante)	180		
12.4.2.2 Sondas de DNA obtenidas por clonación acelular (amplificación por PCR)	181		
12.4.2.3 Sondas de RNA (ribosondas)	181		
12.4.3 Sondas sintéticas	182		
12.4.3.1 Síntesis química de oligonucleótidos	182		
12.4.3.2 Diseño de una sonda sintética	184		
12.5 MARCAJE DE LAS SONDAS	185		
12.5.1 Marcaje directo en la molécula de sonda	186		
12.5.1.1 Marcaje con DNA polimerasa	187		
a Método de traslado de la mella	187		
b Método de iniciación o cebado aleatorio	187		
12.5.1.2 Marcaje terminal	187		
a Marcaje con polinucleótido quinasa	187		
b Marcaje terminal por relleno	187		
12.5.2 Marcaje externo a la sonda	188		
12.5.3 Marcaje indirecto	188		
12.5.4 Tipos de marcadores, etiquetas o indicadores y su detección	189		
12.5.4.1 Marcadores radiactivos	189		
12.5.4.2 Marcadores no radiactivos	189		
12.5.4.3 Etiquetas o indicadores	190		

La hibridación es un fenómeno de importante aplicación práctica en todas las áreas de la biología molecular e ingeniería genética, y en especial en la patología molecular. Para comprender estas aplicaciones es preciso estudiar previamente la base molecular que fundamenta este proceso y los métodos de ensayo desarrollados para su aplicación, así como las técnicas de preparación de sondas.

La hibridación está basada en el proceso de renaturalización de dos cadenas sencillas de un DNA, que se observa, por ejemplo, al enfriar de forma lenta una disolución del DNA bicatenario previamente desnaturalizado por calor. La especificidad del emparejamiento mediante puentes de hidrógeno entre bases complementarias (A-T, A-U o G-C) de moléculas monocatenarias contiguas da lugar a estructuras bicatenarias, de gran estabilidad. Cuando ambas moléculas corresponden a las hebras originales de un mismo ácido nucleico se habla de renaturalización, mientras que si las dos hebras proceden de moléculas diferentes adquiere sentido el término hibridación. El resultado pueden ser híbridos DNA-DNA, híbridos RNA-RNA o híbridos DNA-RNA.

12.1 INTRODUCCIÓN: DESNATURALIZACIÓN Y RENATURALIZACIÓN DEL DNA

Puesto que la idea de la hibridación surge de la observación del proceso de renaturalización de un DNA previamente desnaturalizado, se estudian como introducción los procesos de desnaturalización y renaturalización.

De forma similar a lo que ocurre en las proteínas, los ácidos nucleicos se **desnaturalizan** por efecto de agentes químicos o físicos, perdiendo su conformación tridimensional, y en especial la naturaleza bicatenaria del DNA (doble hélice). La desnaturalización se debe a la desestabilización tanto de los puentes de hidrógeno entre bases como de las interacciones de apilamiento entre bases y a la pérdida del efecto hidrófobo. Al desnaturalizarse, las dos hebras del DNA se separan y pasan a una conformación al azar (llamada a menudo *ovillo aleatorio*), sin que se altere la estructura primaria, pues no hay ruptura de enlaces covalentes. En el caso de los RNA, con muy distinta conformación tridimensional a pesar de su analogía en estructura primaria, la desnaturalización afecta fundamentalmente a los tRNA y rRNA, ya que son los que presentan más regiones bicatenarias locales. Al no dar lugar a aplicaciones experimentales, la desnaturalización del RNA no se considera de forma particular en este capítulo.

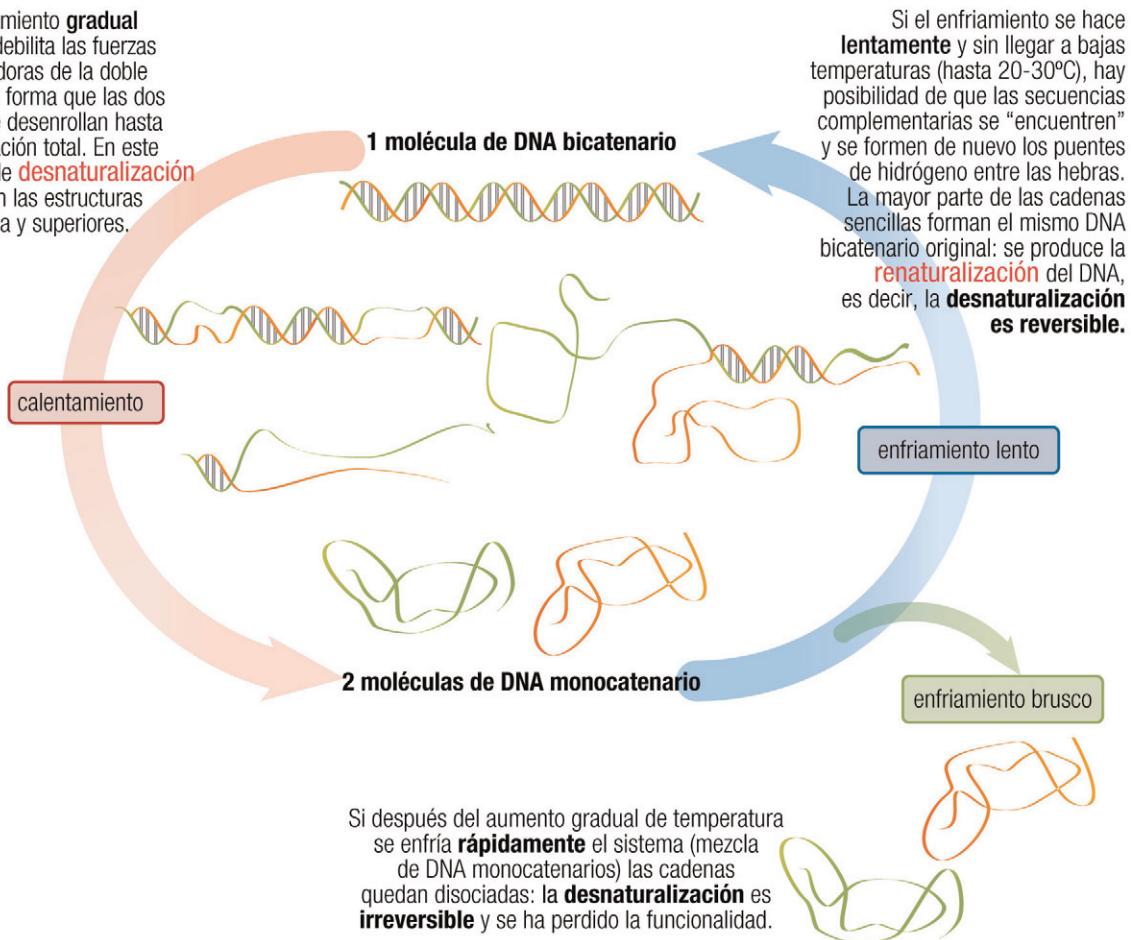
12.1.1 Agentes desnaturalizantes

Aunque pueden considerarse desnaturalizantes los mismos factores que actúan sobre las proteínas, sólo se citarán aquí, de forma breve, los ácidos y bases y ciertos agentes químicos, para centrar la atención en el efecto de la temperatura, como factor representativo. Aunque cada uno de ellos posee un mecanismo particular de actuación, todos debilitan las interacciones por puentes de hidrógeno o por apilamiento de bases, conduciendo a un mismo proceso de desnaturalización y de separación de las hebras.

12.1.1.1 Desnaturalización y renaturalización por efecto de la temperatura

La temperatura, el agente desnaturalizante más representativo, actúa de la forma siguiente:

El calentamiento **gradual** del DNA debilita las fuerzas estabilizadoras de la doble hélice, de forma que las dos hebras se desenrollan hasta su separación total. En este proceso de **desnaturalización** se pierden las estructuras secundaria y superiores.



En general, se emplea el proceso de desnaturalización para el estudio de la estructura del DNA, mientras que la renaturalización es útil como base de las técnicas de hibridación, de enorme trascendencia en biología molecular e ingeniería genética. La combinación de desnaturalización y renaturalización es especialmente crítica en la técnica de PCR (Capítulo 14).

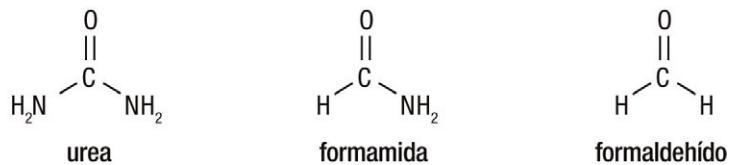
12.1.1.2 Desnaturalización por ácidos y bases

Como ya se ha estudiado (pág. 41), el emparejamiento de las bases disminuye o se debilita por valores de pH alejados de la neutralidad, con la consiguiente desnaturalización del DNA. Sin embargo, un medio ácido puede conducir también a la ruptura de la estructura primaria de los ácidos nucleicos (recuérdese que el medio ácido fuerte rompe los enlaces fosfodiéster y N-glicosídicos, mientras que en condiciones más suaves –por ejemplo, a pH alrededor de 3– se separan las purinas; pág. 31). Por ello, para la desnaturalización del DNA es más habitual emplear condiciones alcalinas (en las cuales el DNA, a diferencia del RNA, es estable). La neutralización posterior del medio puede permitir la recuperación de la estructura de partida, es decir, la renaturalización del DNA (con consideraciones similares a las de la desnaturalización por calor).

12.1.1.3 Desnaturalización por agentes químicos

Algunos de los compuestos empleados con frecuencia para desnaturalizar el DNA son moléculas sencillas, altamente polares, con grupos amino y carbonilo. Éstos les permiten competir en la formación de enlaces de hidrógeno con los grupos amino y carbonilo propios de las bases, lo que facilita la ruptura de sus emparejamientos y la consiguiente separación de las hebras.

Agentes químicos
desnaturalizantes
de uso más frecuente:



12.2

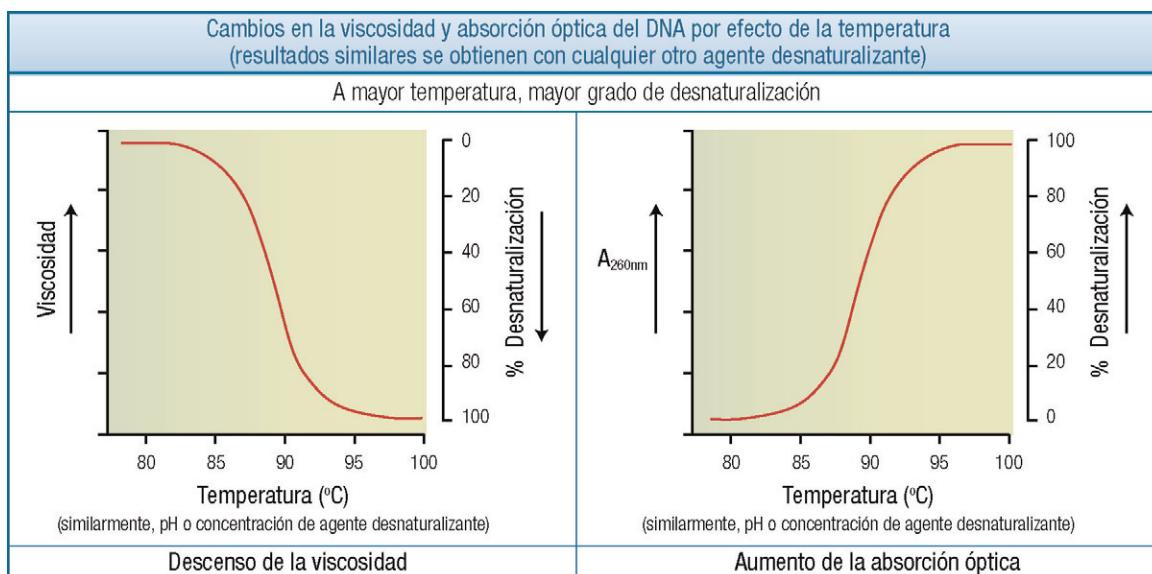
Dependiendo del tipo de acción del agente desnaturalizante y de la forma como ese agente se elimine, se puede conseguir la renaturalización del DNA, de modo análogo a los casos anteriores.

12.1.2 Influencia de la desnaturalización sobre las propiedades del DNA

En la práctica, la desnaturalización del DNA se puede seguir fácilmente midiendo aquellas propiedades físicas que permiten diferenciar el DNA bicatenario y el monocatenario. Varias de ellas experimentan profundos cambios durante el proceso: aumento de la absorción óptica, cambio en la rotación óptica, incremento de la densidad de flotación y descenso de la viscosidad. Las más utilizadas son la absorción en el ultravioleta (a 260 nm) y la viscosidad. A partir de estos cambios también puede obtenerse información de tipo estructural acerca del propio ácido nucleico (aunque este aspecto no se considera aquí).

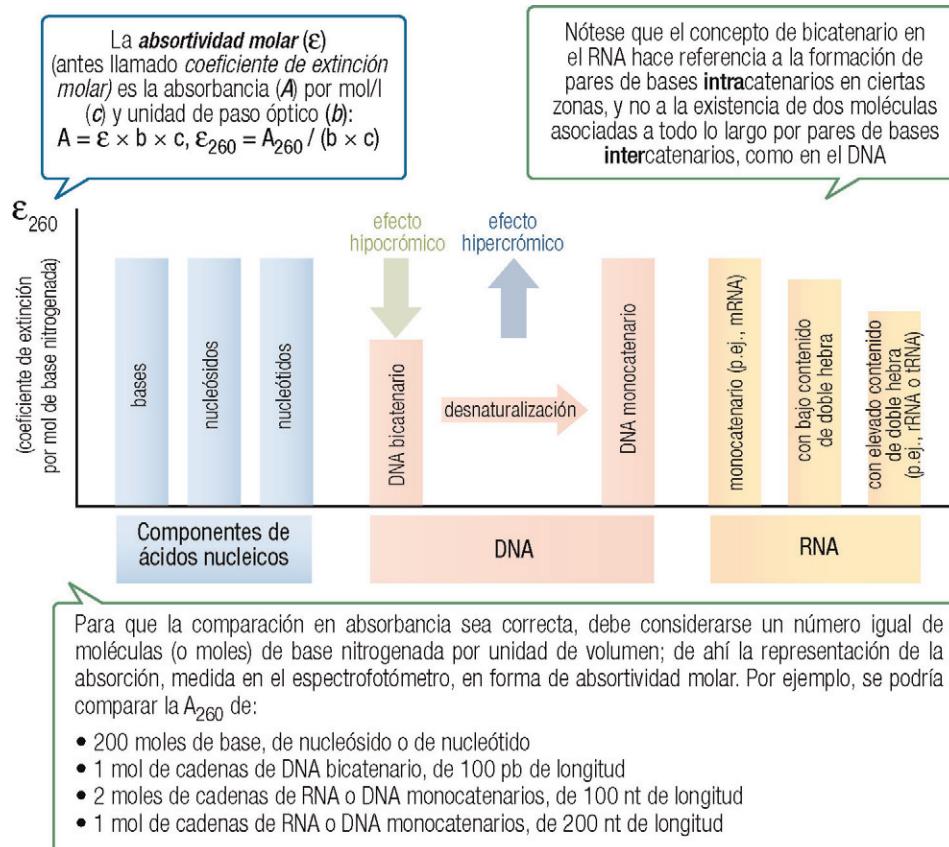
12.1.2.1 Influencia sobre la viscosidad

La relativa rigidez de la molécula en doble hélice y su inmensa longitud en relación con su diámetro hacen que las disoluciones de DNA bicatenario sean muy viscosas. Por el contrario, en forma monocatenaria el plegamiento al azar de la cadena conduce a una viscosidad menor. De acuerdo con ello, al aumentar la temperatura, el pH o la concentración del agente químico y producirse la desnaturalización, se observa una disminución en la viscosidad de la disolución de DNA.



12.1.2.2 Influencia sobre la absorción ultravioleta

Debe recordarse que las bases nitrogenadas, los nucleósidos y los nucleótidos poseen valores máximos de absorción de la luz cerca de la longitud de onda de 260 nm, debido a la presencia en su molécula de sistemas conjugados de dobles enlaces (págs. 16 y 24). Esta propiedad se mantiene en los ácidos nucleicos (v. web 3.1), pero la magnitud de la absorción no es la misma, ya que el apilamiento y los puentes de hidrógeno la reducen (esto recibe el nombre de **efecto hipocrómico**). Por ello, el DNA monocatenario absorbe en magnitud similar a los nucleótidos libres, mientras que el DNA bicatenario absorbe menos. Como consecuencia, la desnaturalización del DNA va acompañada de un aumento en la A₂₆₀ (denominado **efecto hiperócrómico**).



Obsérvese comparativamente la relación existente entre disminución de la viscosidad o aumento de A_{260} y el proceso de desnaturación del DNA.

El RNA, al presentar en su molécula un grado variable de regiones bicatenarias, presentará una absorción intermedia entre la de un DNA bicatenario de tamaño equivalente y uno monocatenario (desnaturalizado); la magnitud depende de la complejidad de su estructura secundaria y, por tanto, también del tipo de RNA (mensajero, ribosómico o transferente).

12.1.2.3 Significado práctico de las curvas de desnaturación

La representación gráfica de la variación de viscosidad o de A_{260} frente a la temperatura (o al pH, o a la concentración de agente desnaturizante) permite obtener las **curvas de desnaturación**, que son características para cada DNA en disolución. El carácter sigmoidal de estas curvas indica que la desnaturación se produce de forma brusca, en un margen muy estrecho de temperatura, de pH o de concentración de agente desnaturizante. Esto contrasta con la transición más gradual observada en la desnaturación térmica de las proteínas, y sugiere que las interacciones que mantienen la doble hélice (apilamiento y puentes de hidrógeno) son cooperativas, es decir, la separación de las hebras en un punto facilita en gran medida que se sigan separando en el resto de la molécula.

Al ser la transición tan brusca, la desnaturación térmica del DNA es similar en cierto modo a la fusión de un sólido, por lo que se habla también de **curvas de fusión**, y la temperatura central del intervalo de transición (50% de desnaturación) se llama **temperatura de fusión** (T_f o T_m , del inglés *melting*). Bajo condiciones definidas, cada DNA muestra una curva de fusión con valores característicos de T_m y de magnitud del efecto hipocrómico (diferencia entre A_{260} máxima y mínima, o cambio de viscosidad). Estos datos son una medida de la estabilidad estructural de la molécula de ácido nucleico. Su valor depende, además, del pH y la fuerza iónica del medio, y de la composición en bases del ácido nucleico.

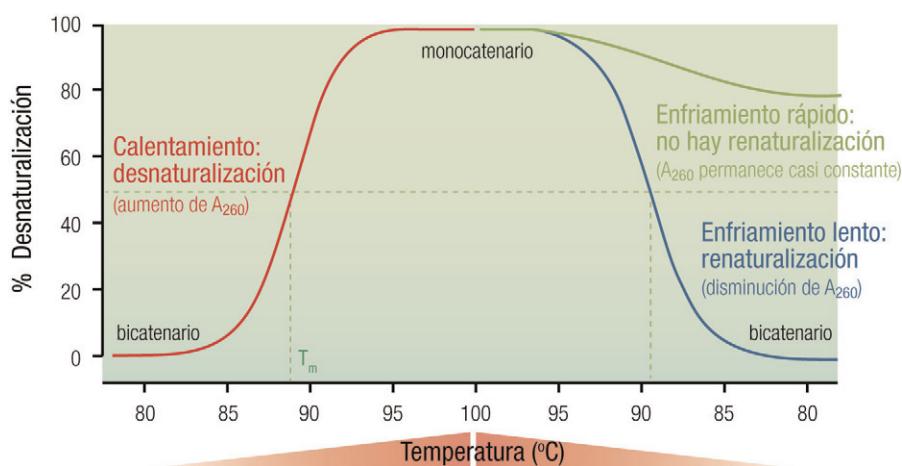


Web 12.1. Variación de la temperatura de fusión del DNA dependiendo de su composición y de otros factores.

Es importante resaltar la relación entre T_m y el contenido en pares de bases G≡C o A=T del DNA: cuanto más abundantes son G≡C, más desplazada hacia la derecha aparece la curva, es decir, se necesita mayor temperatura para la desnaturación, T_m es mayor. Y viceversa, cuantos más pares A = T, mayor desplazamiento a la izquierda y menor T_m . Esto se debe a que, al ser más fuerte la interacción en los pares G≡C (3 puentes de hidrógeno), necesitan mayor energía para disociarse.

En efecto, al determinar la temperatura de fusión para varios DNA de orígenes diversos, se ha encontrado una correlación lineal con su contenido de G + C (que, de acuerdo con una de las reglas de Chargaff, ya estudiadas, varía entre especies y puede emplearse como índice en estudios evolutivos, pág. 38). Ello permite utilizar la T_m de una muestra de DNA para estimar su composición en bases. Por ejemplo, se puede emplear para determinar la presencia de regiones ricas en G + C o de “islotes CpG”, asociados con frecuencia en el genoma eucariótico a regiones promotoras, reguladoras de la transcripción y comúnmente metiladas (pág. 282).

Finalmente, la representación gráfica de las curvas de desnaturación y renaturalización de un DNA permite observar, como se indicó al principio, que la similitud entre ambos procesos depende de las condiciones en las que se intente la renaturalización. Obviamente, ambas curvas discurren en sentidos opuestos, pero sólo si el enfriamiento es lento tiene lugar una renaturalización, mientras que un enfriamiento rápido impide la reasociación correcta de las hebras:



12.2 ANÁLISIS MOLECULAR DE LA HIBRIDACIÓN DE ÁCIDOS NUCLEICOS

Debe tenerse en cuenta que la renaturalización depende únicamente de la complementariedad de secuencia de las hebras. Por ello, aunque lo más probable es que éstas se reasocien formando las mismas parejas o moléculas bicatenarias originales, son posibles otras reasociaciones. En particular, si se parte de una mezcla desnaturizada de moléculas distintas de DNA, pero con secuencias parecidas en parte (por ejemplo, DNA que codifica la misma proteína en especies distintas), al renaturalizar se puede producir la asociación parcial de dos hebras procedentes de moléculas bicatenarias distintas. Se forman así **DNA híbridos**, con unas regiones emparejadas y otras no. Esta formación de moléculas bicatenarias de DNA cuyas hebras tienen distinto origen (distinta especie o, en general, distinto ácido nucleico) recibe el nombre de **hibridación**. Es característico el carácter artificial de su consecución, de ahí que también se considere hibridación la formación de híbridos entre DNA y RNA, o entre un ácido nucleico y un oligonucleótido natural o sintético, siempre en función de la existencia de secuencias total o parcialmente complementarias.

La estabilidad del híbrido es tanto mayor cuanto más elevada sea la proporción de bases complementarias y mayor la longitud de las secuencias emparejadas. El hecho de que cadenas diferentes de ácido nucleico hibriden entre sí indica que los organismos de donde proceden comparten un cierto grado de herencia evolutiva común, y que sus RNA y proteínas tienen estructuras y funciones similares. Se dice que los ácidos nucleicos que codifican estas proteínas y RNA tienen **secuencias similares, homólogas**. A mayor grado de hibridación, mayor conexión evolutiva entre especies, y viceversa; por ejemplo, el DNA genómico humano hibrida en mayor proporción con el DNA de ratón que con el de levadura. Sin embargo, no es ésta la única utilidad de la hibridación, sino que da lugar a numerosas aplicaciones, muchas entre ellas en el diagnóstico clínico.

12.2.1 Principio básico del ensayo de hibridación

Todos los ensayos de hibridación se basan en la gran especificidad de la interacción entre bases complementarias. Para que la hibridación permita identificar en un genoma, u otra muestra cualquiera de ácido nucleico, una secuencia particular se requieren dos **elementos básicos**:

- La presencia de la **secuencia diana** en uno de los fragmentos de la **muestra** de ácido nucleico, generalmente obtenidos usando enzimas de restricción.
- El empleo de una **sonda** (sonda molecular, sonda de hibridación), que es un oligonucleótido (fragmento corto de ácido nucleico) de secuencia conocida, marcado de forma que permita su detección. La secuencia de la sonda ha de ser complementaria a la secuencia diana que se pretenda detectar, y su marcaje puede ser con un isótopo radiactivo, un cromóforo, un fluoróforo, una enzima, etc. (pág. 185).



Web 12.2. Fundamento y etapas de los ensayos de hibridación.

12.2.2 Influencia de algunos factores sobre la hibridación

Además de la composición de bases de las cadenas, existe una serie de parámetros que afectan a la posibilidad de renaturalización de dos moléculas monocatenarias y, por tanto, condicionan experimentalmente la realización de un ensayo de hibridación. Estos factores, cuya descripción detallada corresponde a un tratado específico de metodología en biología molecular, están esencialmente relacionados con la concentración de DNA o RNA diana (no sólo la concentración de ácido nucleico en la muestra, sino también el número de copias de la secuencia diana que contiene dicha muestra), la concentración y tamaño de la sonda, la temperatura y tiempo de hibridación, y las características del medio de reacción, como fuerza iónica, pH, concentración de agentes desnaturizantes, adición de ciertos polímeros, etc. A continuación se comenta sólo un aspecto, de interés general en los ensayos de hibridación, determinado por algunos de dichos factores.

12.2.2.1 Rigor de la hibridación

Se define el **rigor** (en inglés, *stringency*) como el grado de especificidad en el emparejamiento conseguido en los híbridos; supone, pues, una medida de la capacidad de reconocimiento mutuo de dos secuencias. Al

realizar la hibridación, habitualmente se busca un emparejamiento completo entre sonda y diana, pero en ocasiones interesa tolerar un cierto grado de desemparejamiento, con el fin de detectar regiones del DNA cuya secuencia no es exactamente igual a la diana, pero presenta una elevada homología. Por otro lado, el control del rigor en cada aplicación y caso concreto es esencial: ni debe ser tan alto que impida la unión estable de la sonda a su diana, ni tan bajo que permita la hibridación inespecífica de la sonda con secuencias poco relacionadas, dando lugar a una señal de fondo elevada o a falsos positivos.

El rigor depende de diversos factores experimentales: temperatura y tiempo de renaturalización (o hibridación), longitud y concentración de las moléculas, composición de bases y ambiente químico (presencia de cationes, formamida, urea, pH, etc.). Cuanto más elevada es la temperatura durante la fase de hibridación, más específica debe ser la interacción entre las hebras para que se mantengan unidas. Los cationes monovalentes (Na^+) estabilizan la estructura bicatenaria, pues apantanallan la repulsión electrostática entre los grupos fosfato, facilitando la aproximación de ambas hebras. Por último, la formamida y la urea afectan a los enlaces de hidrógeno (pág. 165) y son por ello agentes que desestabilizan la doble hélice, de modo que sólo se mantendrán los emparejamientos más perfectos.

En consecuencia, los principales factores que se suelen manejar en la práctica para maximizar el rigor de la hibridación, y permitir sólo los híbridos que estén emparejados completamente, son la elevación de la temperatura de hibridación, la disminución de la concentración de Na^+ y el aumento de la de formamida. Viceversa, para reducir el rigor de la hibridación, es decir, para permitir la formación de híbridos parcialmente desemparejados, se emplean temperaturas de hibridación relativamente bajas, concentraciones altas de Na^+ y bajas de formamida.

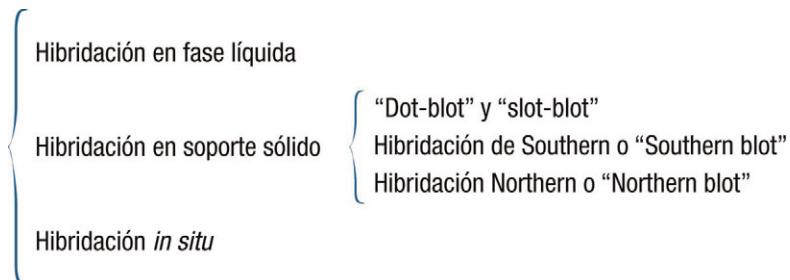


Web 12.3. Rigor en la hibridación.

12.3 MÉTODOS DE ENSAYO CON HIBRIDACIÓN

Todos los ensayos de hibridación se basan en la mezcla de moléculas monocatenarias de ácido nucleico muestra o diana, no marcado, con una sonda de secuencia conocida, marcada, bajo condiciones experimentales que permitan el emparejamiento entre bases complementarias. Si la sonda o el DNA diana son bicatenarios, deben desnaturalizarse previamente, en general por calentamiento o por tratamiento alcalino. Una vez mezcladas, se permite la renaturalización en la cual, como ya se ha comentado, se pueden formar híbridos sonda:diana. Se ha de disponer, además, de un método para detectar los híbridos formados, que viene determinado por el tipo de marcaje empleado en la sonda.

Los ensayos de hibridación pueden clasificarse en tres grupos:

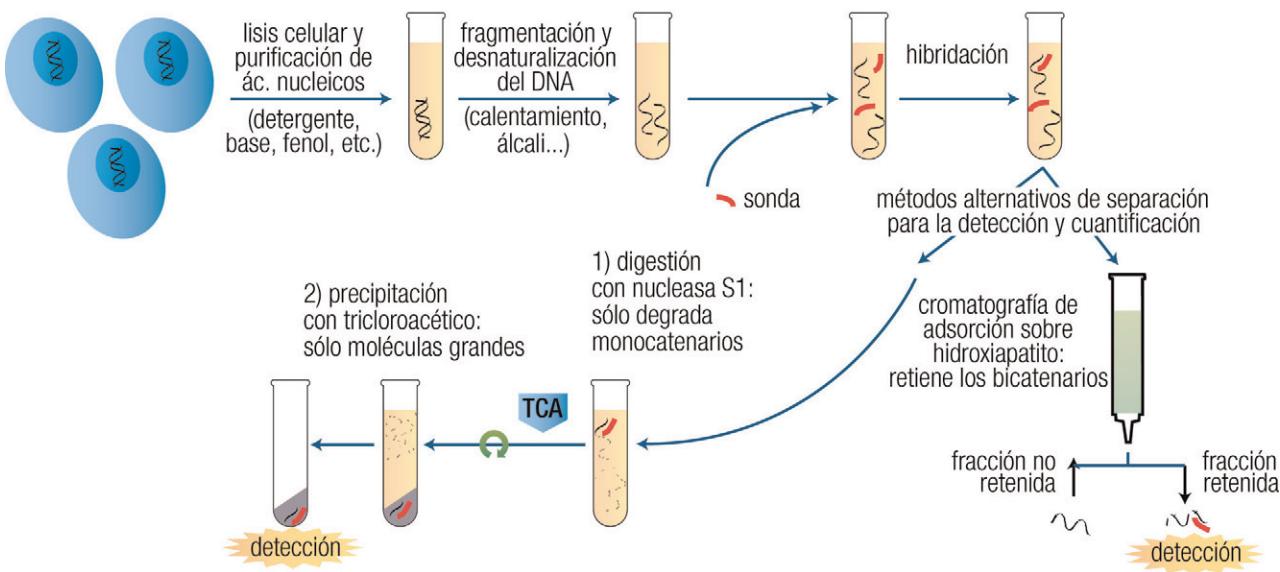


12.6

12.3.1 Hibridación en fase líquida

Éste fue el primer formato de hibridación diseñado. Utilizando una muestra de DNA o RNA en disolución, se realiza su hibridación con la sonda; al estar ambas disueltas, la cinética es más rápida. Además, todas las moléculas diana son igualmente accesibles, por lo que la técnica es más cuantitativa que la hibridación en soporte sólido. Sin embargo, es el método que menos se utiliza actualmente, y ha sido desplazado por otros más convenientes. Se aplica, por ejemplo, en el estudio de la estructura de los mRNA empleando nucleasa S1.

células



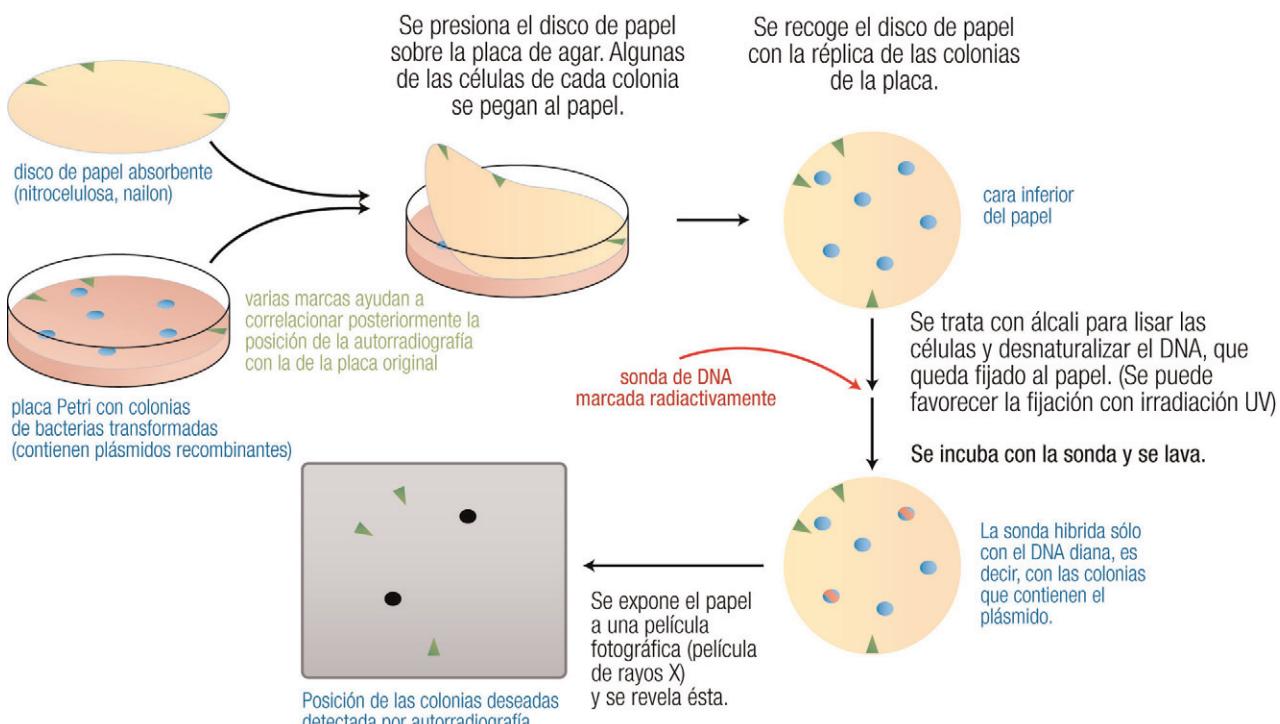
12.7

12.3.2 Hibridación en soporte sólido

Se trata de un tipo de ensayo más simple, que permite procesar varias muestras de forma simultánea y facilita el control de la hibridación. Aunque la cinética de la hibridación es más lenta y el proceso menos eficiente, se utiliza mucho más que la hibridación en disolución, porque es más sencillo de realizar y más versátil.

Estos métodos han adquirido hoy día un interés particular como herramientas muy válidas para el análisis de ácidos nucleicos y el diagnóstico de enfermedades moleculares con base genética. Como antecedentes, pueden citarse la inmovilización del DNA en nitrocelulosa, su detección por primera vez por hibridación y el descubrimiento de un método para la lisis *in situ* de colonias bacterianas en filtros de celulosa.

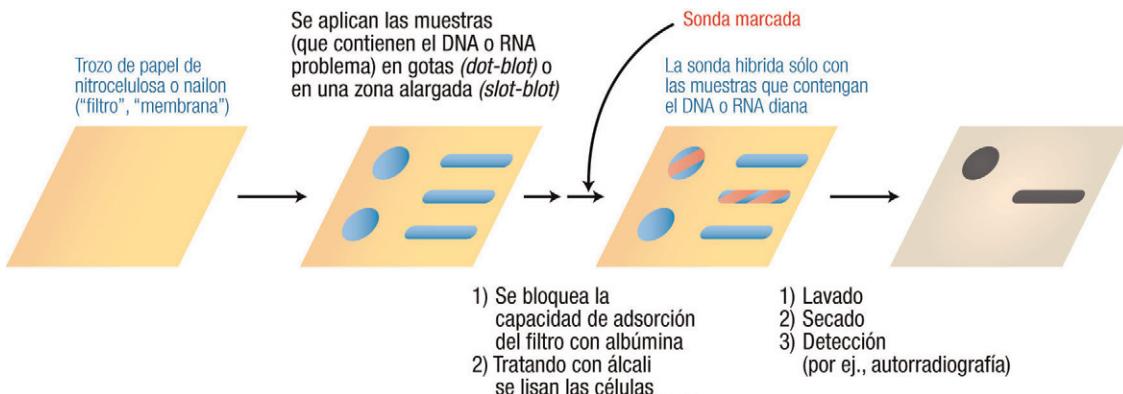
El esquema general de este tipo de ensayos se puede ilustrar con el método comúnmente empleado para identificar, de entre varias colonias bacterianas en placas de cultivo con agar, las que contienen un gen o fragmento de DNA de interés bajo la forma de plásmido recombinante (Capítulo 15).



12.8

12.3.2.1 Hibridación "dot-blot" y "slot-blot"

Éste es el método más simple y común, gracias a la posibilidad de emplearlo con muestras de ácido nucleico sin purificar. La muestra se aplica, bien gota a gota (*dot-blot*) o en manchas alargadas (*slot-blot*), sobre un “filtro” de nailon o nitrocelulosa, cuyas propiedades de adsorción fijan las moléculas de ácido nucleico y proteína.

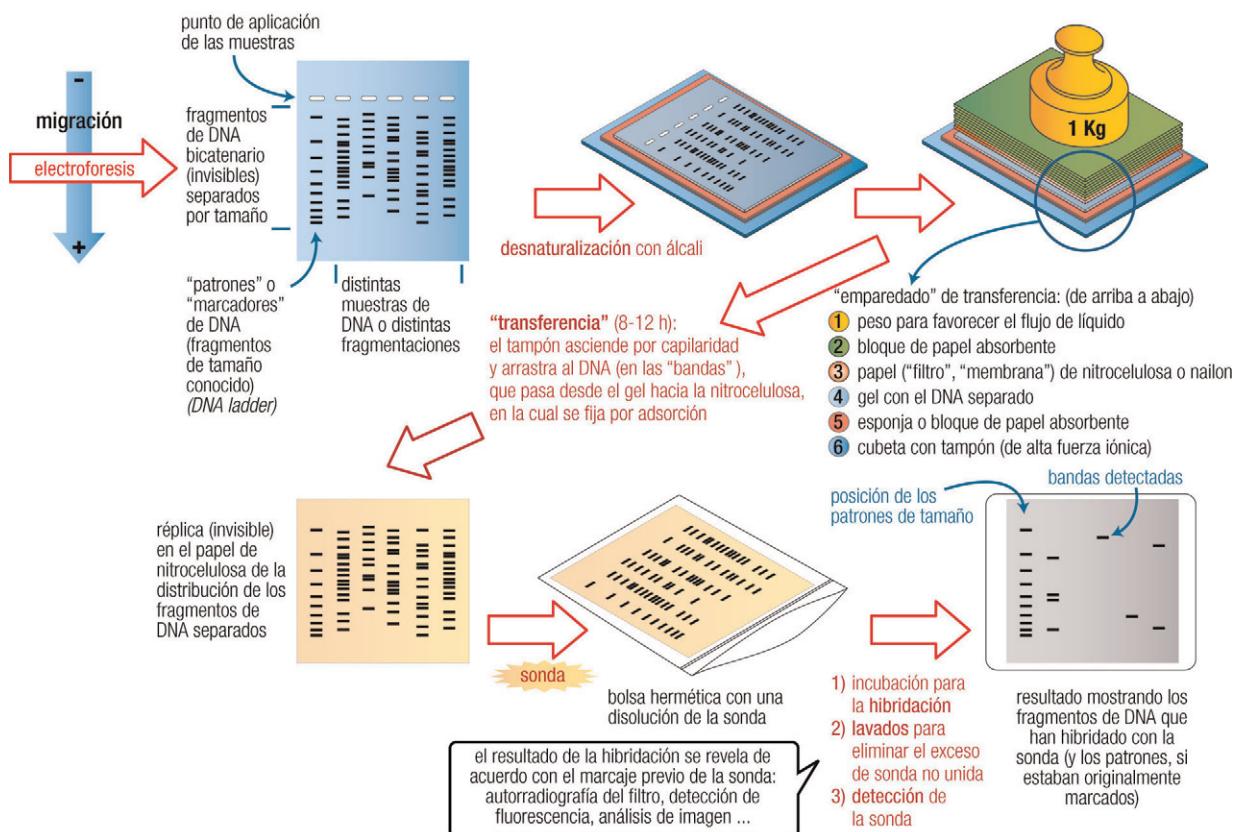


12.9

12.3.2.2 Hibridación de Southern

Fue desarrollada por Edwin M. Southern, al que debe su nombre (*Southern blot*). Es una técnica simple y fácil de realizar que detecta fragmentos de DNA, separados por tamaño mediante electroforesis en gel (pág. 140). Combina la separación electroforética del DNA con su transferencia a un soporte sólido o “filtro” (nailon o nitrocelulosa) para su hibridación. Tanto ésta como la posterior detección se realizan de forma más eficaz en el filtro de lo que se haría en el gel. Esta técnica ha contribuido de forma notable a aumentar el potencial de la hibridación como herramienta analítica, constituyendo hoy día en bioquímica clínica un método habitual de análisis del DNA extraído de muestras de sangre, esputos, tejido y células (linfocitos, amniocitos, etc.).

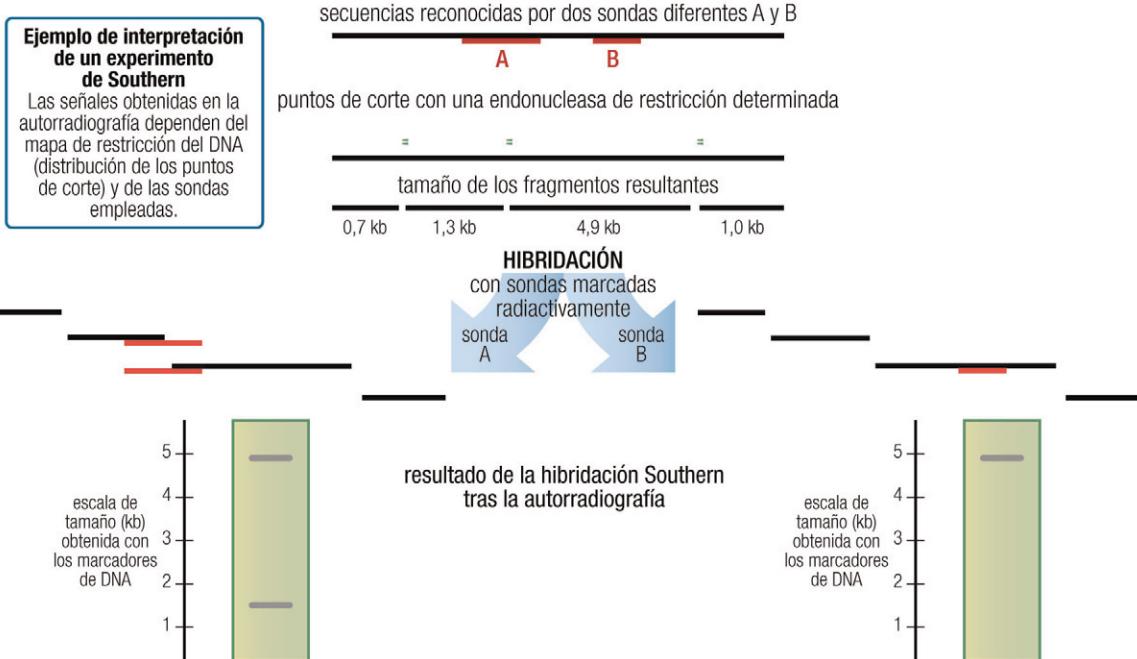
Metodológicamente, implica cinco etapas bien diferenciadas: separación electroforética, desnaturización, transferencia, hibridación y detección:



12.10



Web 12.4. Hibridación de Southern: procedimiento.



12.3.2.3 Hibridación northern

A diferencia del caso anterior, su nombre (*northern blot*) no corresponde al de ningún investigador, sino a un simple juego de palabras (*southern* = del sur, *northern* = del norte). El procedimiento es idéntico, pero el ácido nucleico analizado es RNA. Por su menor tamaño, puede no ser necesaria su fragmentación con endonucleasas, pero debe tenerse especial cuidado para evitar la acción de RNAsas, tanto endógenas como exógenas (contaminación muy común, incluso por las manos del analista, y difíciles de inactivar). El uso de agentes desnaturalizantes sigue siendo necesario para evitar emparejamientos intracatenarios, que alterarían el patrón de separación por tamaño.

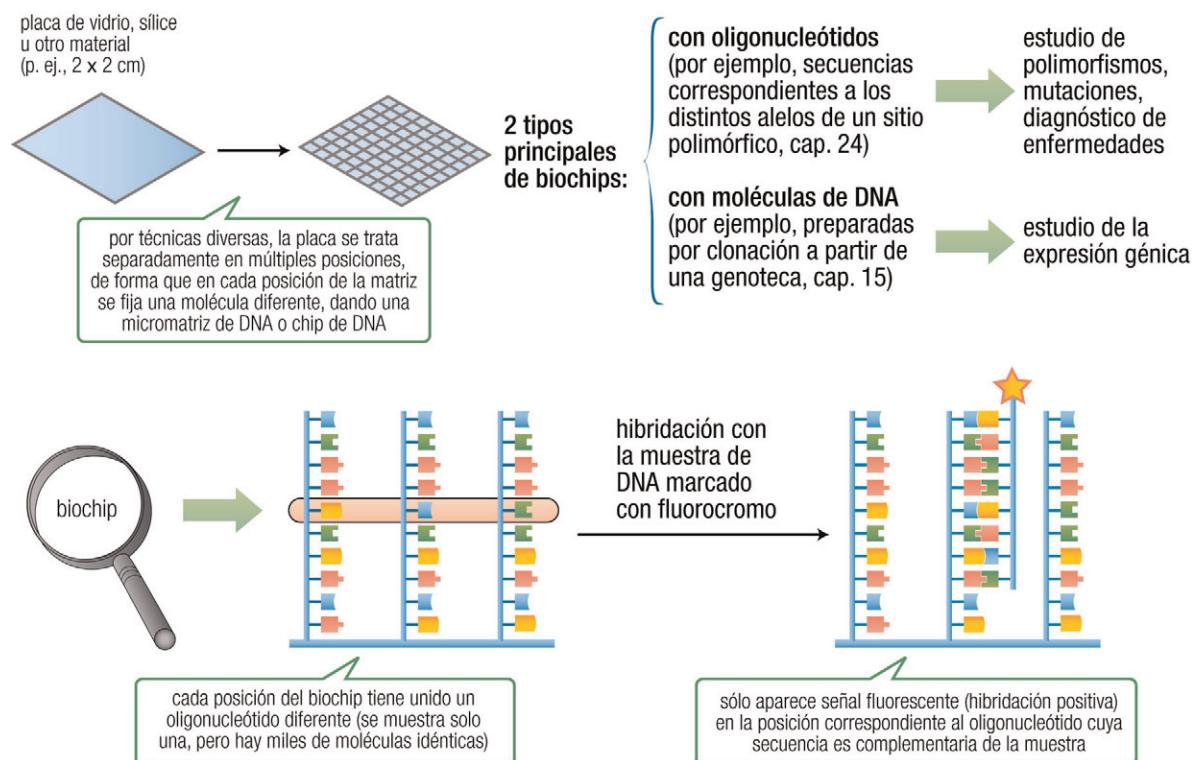
La técnica de hibridación northern se emplea principalmente para obtener información sobre el tamaño de RNA y sobre el modelo de expresión de genes específicos. Éstos, una vez clonados, pueden utilizarse como sondas para hibridar con muestras de RNA aisladas de una variedad de tejidos diferentes, y así conocer los tipos celulares en los que se expresa el gen, la abundancia relativa y tamaño de los transcritos, y diversos detalles de su procesamiento posterior.

12.3.3 Matrices de DNA

Los llamados **biochips**, **genochips**, **genosensores** o **chips de DNA**, fruto de una confluencia de la biología molecular con la automatización y la informática, están revolucionando el diagnóstico molecular, al discriminar de una manera sencilla, rápida y automatizada el componente genético de las enfermedades. Se llaman también **micromatrices de DNA** (o, dependiendo del país, microarreglos de DNA; en inglés, *DNA micro-arrays*), debido a la disposición de las muestras en una retícula ortogonal.

Se trata de placas de silicio, vidrio u otro material, de unos centímetros de tamaño, a cuya superficie se han adherido según una disposición ordenada, conocida, diversas moléculas cortas de DNA. La tecnología empleada permite fijar en cada biochip entre cientos y cientos de miles de moléculas diferentes, por lo que se puede hacer un número de ensayos muy elevado de forma simultánea sobre una muestra muy pequeña (por ejemplo, una gota de sangre o menos de un microgramo de mRNA total purificado de un tipo celular concreto).

Esencialmente, se trata de un ensayo de hibridación en soporte sólido, en el que el papel de sonda lo desempeñan las moléculas unidas al biochip y en este caso se marca el DNA o RNA de la muestra. La muestra, desnaturalizada, se vierte sobre el biochip para que hibride con las secuencias de éste, se lava y se detecta en qué posiciones aparece el marcaje.



12.12

Se utilizan principalmente dos tipos de matrices de DNA:

- **Matrices preparadas con fragmentos de ácidos nucleicos**, obtenidos a partir de muestras biológicas o mediante clonación. Habitualmente, las muestras de DNA se aplican (por ejemplo, sobre un portaobjetos) bien por contacto, en forma de micropuntas, o bien sin contacto físico con el soporte, empleando tecnologías piezoelectrónica, térmica o de electroaerosol, similares a las usadas en las impresoras de inyección de tinta.
- **Matrices de oligonucleótidos**, que por lo común se sintetizan químicamente con las secuencias deseadas (pág. 182) mediante reacciones realizadas de forma directa sobre el chip, de modo que los oligos quedan anclados a él de modo covalente. En comparación con las anteriores, las matrices de este tipo suelen ser de menor tamaño (típicamente, sobre un cuadrado de vidrio o sílice de unos 2 cm de lado) y con un número de muestras muy superior (alcanzando hasta un millón).

En ambos casos, la idea es preparar una colección de moléculas de DNA monocatenario con secuencias conocidas o, al menos, con origen biológico conocido y relevancia para el estudio abordado. La normalización del formato y la miniaturización van ligadas a la automatización del ensayo, que emplea técnicas robóticas e informáticas para procesar un gran número de muestras de muy pequeño volumen, recolectar y analizar los resultados.

A continuación se comentan algunas características y aplicaciones frecuentes –a título de ejemplo– de ambos tipos de matriz de DNA.

12.3.3.1 Matrices de oligonucleótidos: chips de DNA

Como se ha mencionado, en este planteamiento la matriz se forma con colecciones de oligonucleótidos generados *in situ*, es decir, mediante síntesis química y anclados de forma covalente al soporte tanto desde el comienzo de su síntesis como durante todo el ensayo posterior. Las secuencias de dichos oligonucleótidos (entre 25 y 60 nt) se diseñan para corresponder con las que interesa detectar en las muestras de DNA o RNA. Éste es el caso en el que el término chip de DNA tiene su mayor implantación, pues la tecnología necesaria para fabricarlo procede de la empleada en circuitos integrados (fotolitografía, véase web 12.6) y el soporte a menudo es un chip de silicio.



Imagen procedente de <http://images.nih.gov/detail.cfm?id=140> – Licencia: dominio público.

12.13

a) Matrices de genotipado

Una de las principales aplicaciones de estas matrices es la cartografía de mutaciones y el análisis de polimorfismos. En este caso las matrices se emplean para analizar muestras de DNA genómico, con el fin de diagnosticar una enfermedad o, más comúnmente, para evaluar la predisposición a una enfermedad o su pronóstico y, asimismo, para la tipificación o clasificación molecular de subtipos de cáncer y de otras enfermedades.

Veamos un ejemplo que lo ilustra. Estudios previos han llevado a conocer unas centenas de loci con polimorfismos de un solo nucleótido (SNP, pág. 418) relacionados con la hipercolesterolemia familiar. Se conocen, asimismo, las distintas variantes alélicas presentes en la población en cada uno de esos loci y la incidencia de cada uno de estos alelos sobre la evolución clínica de la hipercolesterolemia. Se procede a diseñar secuencias (de unos 25 nt) que sean complementarias a los distintos alelos para cada SNP. Se sintetizan entonces todos esos oligonucleótidos *in situ* sobre las distintas posiciones de una matriz, o bien se sintetizan por separado y se aplican al soporte con alguna de las técnicas de impresión. Cada posición de la matriz contiene, obviamente, múltiples moléculas idénticas de oligonucleótido con una de las secuencias en concreto.

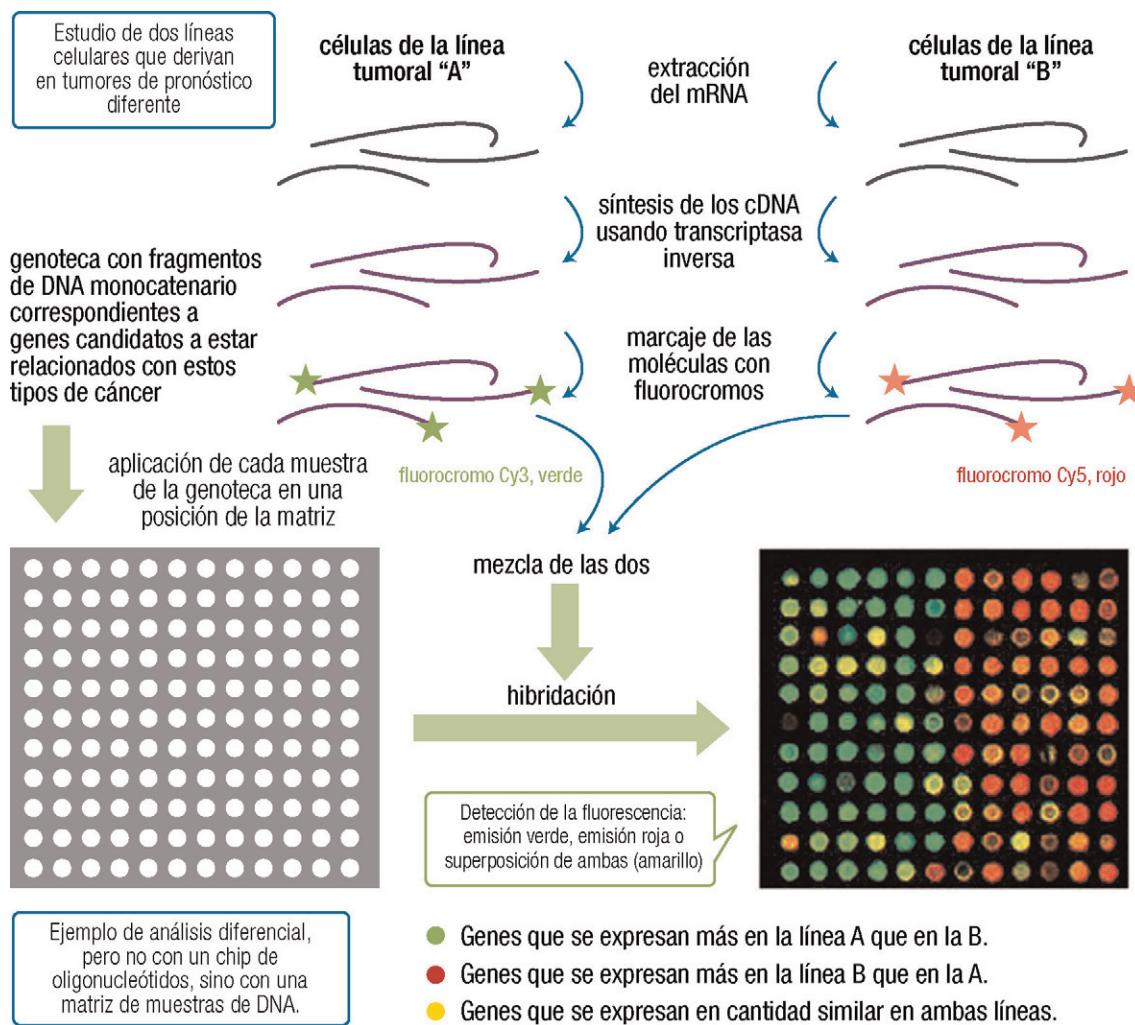
A partir de sangre o frotis bucal de los pacientes se extrae el DNA genómico, que se amplifica por PCR (Capítulo 14) con cebadores diseñados para flanquear cada SNP. El producto de esta PCR múltiple se marca con un nucleótido biotinilado, bien durante la PCR o tras finalizarla –por ejemplo, empleando polinucleótido quinasa o transferasa terminal (págs. 187 y 223)–. A continuación se realiza la hibridación de las muestras de DNA amplificado con la matriz de oligonucleótidos, y se revela empleando estreptavidina conjugada con un fluoróforo (pág. 189). El análisis de la fluorescencia obtenida en cada punto de la matriz permite interpretar cuál de los alelos de cada SNP está presente en el genoma del paciente. El ensayo permite así no sólo el diagnóstico, sino especialmente la prevención, la evaluación de riesgos y la elección de la terapia más adecuada a cada caso. Asimismo, el conjunto de datos que se va obteniendo permite incrementar el corpus de conocimiento sobre la relación entre cada una de las variantes alélicas y la evolución de la hipercolesterolemia.

b) Matrices de expresión

Otro gran campo de aplicación lo constituyen los ensayos de expresión génica. En este caso, las muestras analizadas corresponden al RNA que se expresa en el tejido de interés. La cuantificación de la señal en la matriz ofrece información sobre la actividad relativa de los genes para los que se han diseñado los oligonucleótidos componentes de la matriz.

En una de las aplicaciones más comunes, el denominado **análisis de expresión diferencial**, se preparan muestras de mRNA. Por ejemplo, pueden compararse muestras de individuos sanos con las de quienes presentan una enfermedad, o bien muestras obtenidas de enfermos que padecen dos tipos de cáncer diferentes. El objetivo es identificar genes relacionados con la aparición de la enfermedad o con uno de sus subtipos. De cada una de las muestras se purifica el mRNA total y se realiza *in vitro* una transcripción inversa para obtener el conjunto de moléculas de DNA complementario (pág. 223). Bien durante la propia reacción de retrotranscripción o bien tras terminarla se marcan los cDNA con un fluoróforo, diferente para cada una de las dos muestras que se quieren comparar. La matriz de oligonucleótidos se construye con secuencias correspondientes a los genes de interés, cuya expresión se pretende analizar, o bien con una biblioteca de secuencias aleatorias (pero conocidas), si es que se persigue la búsqueda de genes cuya implicación en la enfermedad es aún desconocida. Ambas muestras de cDNA marcado se mezclan y se aplican sobre la matriz en un ensayo de hibridación. La detección de la intensidad relativa de la fluorescencia en ambos colores para cada punto de la matriz proporciona la información de qué genes están sobreexpresados en una de las dos

muestras, cuáles están reprimidos o expresados más débilmente y cuáles no ven modificada su expresión y, en consecuencia, no tienen implicación en la enfermedad estudiada o en los subtipos comparados. Así se puede deducir cuáles son los genes cuya expresión está relacionada con cada tipo de situación (por ejemplo, con la evolución de un tipo de cáncer).



La imagen inferior derecha se ha adaptado de:

<http://breast-cancer-research.com/content/6/3/R157>. Jessen et al. Breast Cancer Res 2004;6:R157. Licencia: acceso abierto.

12.14

12.3.3.2 Matrices de ácidos nucleicos

Otra aproximación con abundantes ejemplos de aplicación consiste en preparar las matrices con fragmentos de moléculas de ácido nucleico (por ejemplo, de 300 a 500 nt) depositados sobre el soporte. Dichos fragmentos suelen proceder de muestras biológicas o de técnicas de clonación (Capítulos 13 y 14). En algunos casos se eligen por ser representativos de secuencias que interesa analizar, por ejemplo para buscar las diferencias genéticas entre pacientes enfermos e individuos sanos. En otros casos, lo que se fija en la matriz es un conjunto de muestras distintas sobre las que se desea localizar una secuencia o ensayar una actividad bioquímica, como puede ser el reconocimiento por una proteína. Como ejemplo, esto permite la búsqueda de factores de transcripción para una secuencia promotora (Capítulo 18). Dependiendo de la aplicación perseguida, las moléculas de DNA candidatas, inmovilizadas en la matriz, se hacen interactuar con DNA o RNA (ensayo de hibridación), o bien con un extracto proteico (ensayo de unión o reconocimiento proteína-DNA).

12.3.4 Hibridación *in situ*

Es un tipo especializado de hibridación en soporte sólido, en la que la sonda se aplica sobre muestras en su ubicación natural (célula o tejido), generalmente las preparadas para microscopía. Supone, por tanto, el

empleo de la hibridación bajo un abordaje citológico e histológico y, en cuanto a la metodología, es análoga a las técnicas immunocitoquímicas e immunohistoquímicas. De forma muy global, cabe distinguir entre sus aplicaciones tisular y cromosómica.

12.3.4.1 Hibridación *in situ* tisular

Se realiza sobre cortes de tejido, células intactas, núcleos aislados, etc. En muchos casos, se lleva a cabo directamente sobre tejidos fijados con formalina, incluidos en parafina, congelados, etc. Es una modalidad de hibridación particularmente relevante para detectar virus patógenos, para diagnosticar células cancerosas como consecuencia de cambios genéticos o para estudiar cambios en la expresión génica, detectando secuencias en el RNA celular. En este caso, tras una fijación suave de la preparación, se añaden sondas de DNA complementario (cDNA), monocatenario. La detección se realiza por autorradiografía y examen de la película fotográfica bajo el microscopio, o por microscopia de fluorescencia, si se empleó este marcaje para la sonda (Capítulo 12).

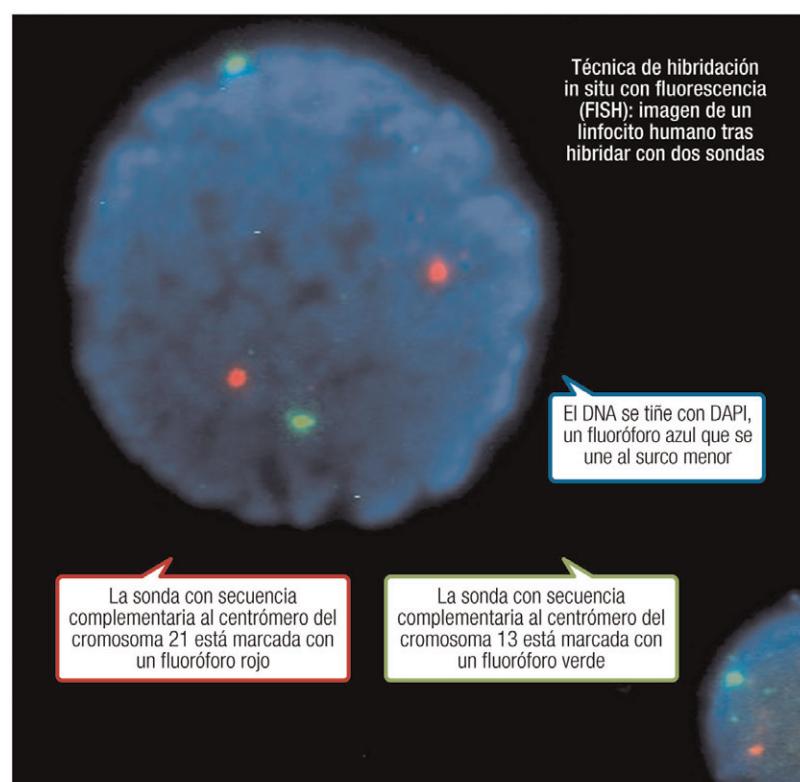
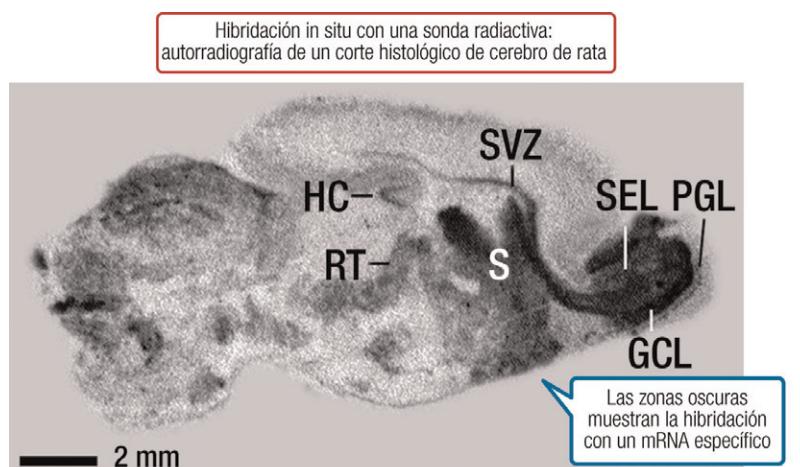


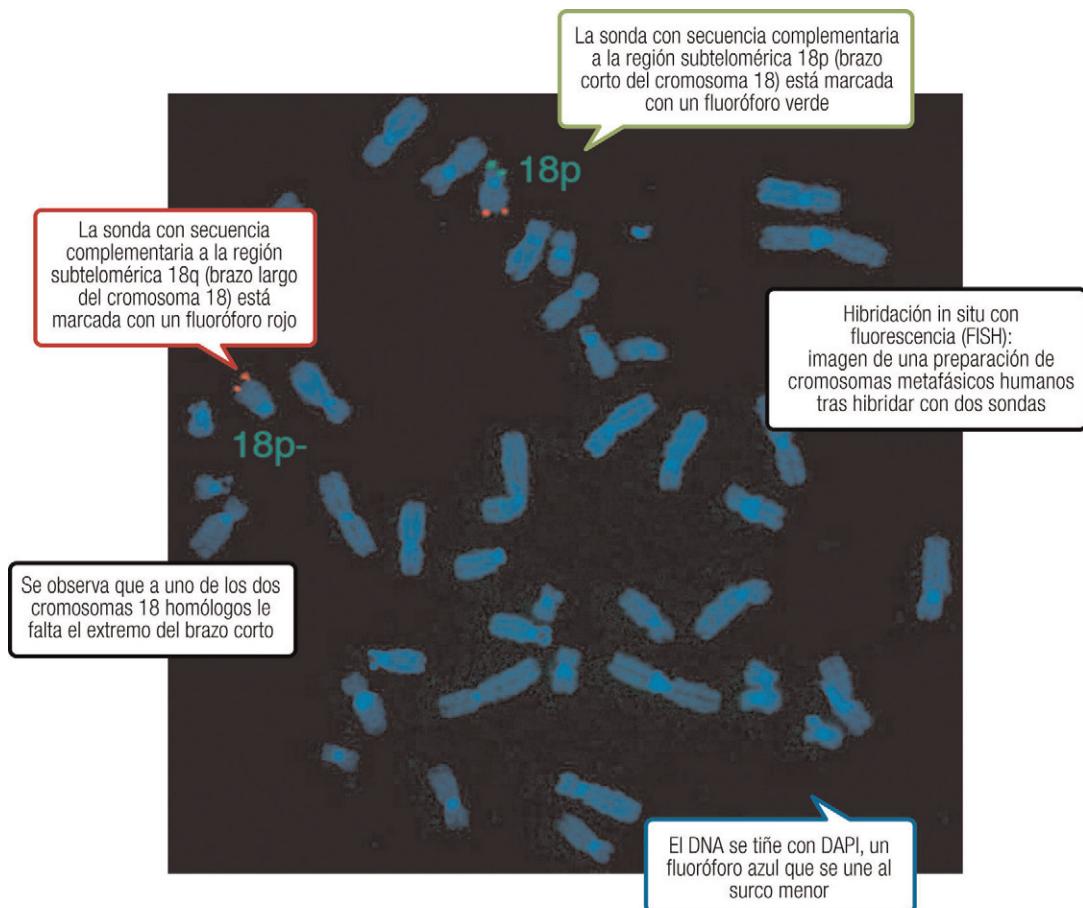
Imagen superior procedente de:
http://commons.wikimedia.org/wiki/File:Distribution_of_GAD67_splice_variants_in_rat_brain_at_P1_%28left_panels%29_and_P90_%28right_panels%29.png. Autores: Anke Popp, Anja Urbach, Otto W. Witte, Christiane Frahm. Licencia Creative Commons Genérica de Atribución 2.5.

Imagen inferior procedente de:
http://commons.wikimedia.org/wiki/File:FISH_13_21.jpg. Autor: Gregor1976. Licencia: dominio público

12.3.4.2 Hibridación *in situ* cromosómica

Se realiza sobre preparaciones en porta de cromosomas metafásicos o prometafásicos (pág. 91), tratadas con reactivos fijadores para preservar las características morfológicas del cromosoma. Para hacer accesible la secuencia de DNA, se trata la muestra con enzimas proteolíticas y se desnaturaliza el DNA mediante calor, álcali o formamida para permitir su hibridación con la sonda. Inicialmente se emplearon sondas radiactivas, pero problemas técnicos en la detección de la señal limitan su utilidad. Se ha conseguido aumentar la sensibilidad y la resolución mediante el uso de sondas marcadas con fluoróforos. En esta técnica, llamada **hibridación *in situ* con fluorescencia** o **FISH** (*fluorescence in situ hybridization*), la detección se hace bajo un microscopio de fluorescencia, bien de forma directa o por un método indirecto (págs. 188-189).

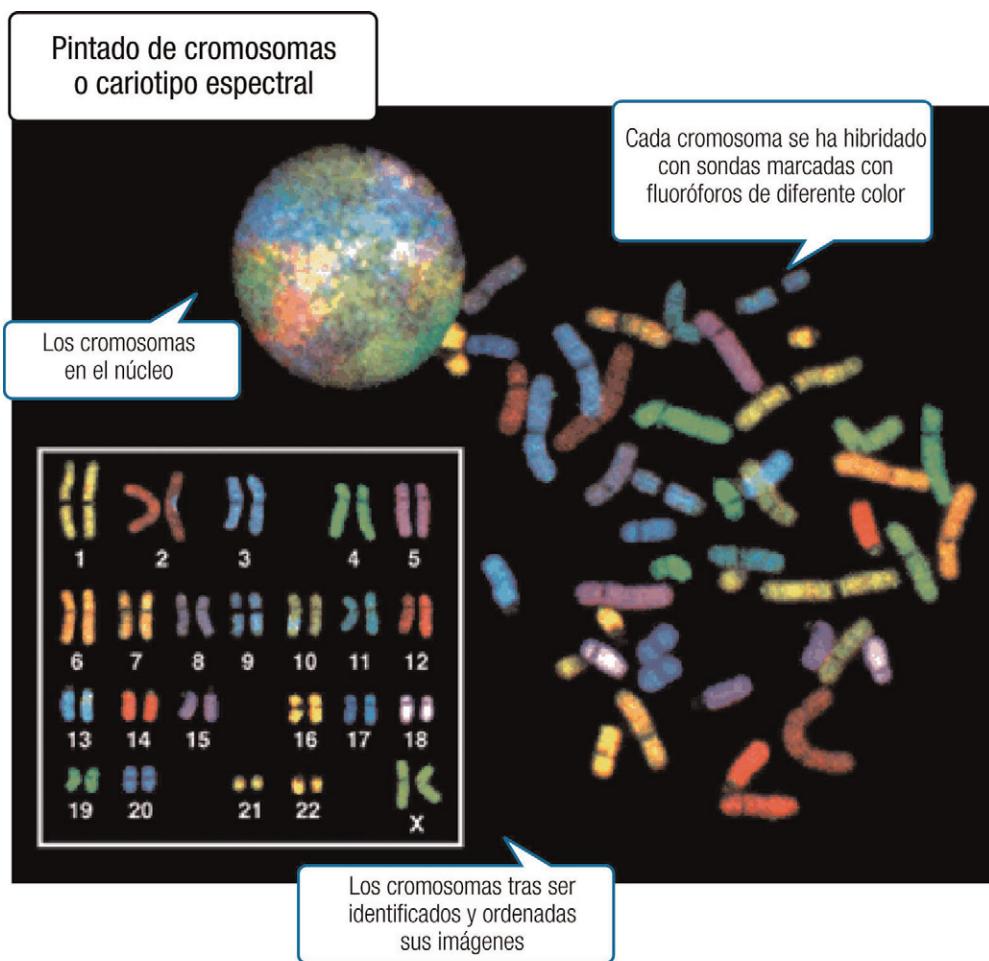
En cuanto a sus aplicaciones, dada la gran detectabilidad de los fluoróforos actuales se puede emplear la FISH con fines muy diversos y se ha convertido en una técnica con elevado potencial diagnóstico en cromosomopatías (Capítulo 25), genes tumorales, virología, trasplantes, etc.



12.16

Imagen procedente de <http://commons.wikimedia.org/wiki/File:FISH18.jpg>. Autores: B. Maranda, N. Lemieux, E. Lemyre. Licencia Creative Commons Genérica de Atribución 2.0

Uno de los métodos surgidos de la hibridación *in situ* cromosómica es el denominado **pintado de cromosomas** (*chromosome painting*), en el que se usa una mezcla de sondas, fragmentos de DNA procedentes de distintas partes de un mismo cromosoma, con lo cual se consigue fluorescencia a lo largo del cromosoma completo. Se facilita de esta forma la identificación visual de cada cromosoma (o de varios simultáneamente, si se emplean sondas con distintos fluorocromos), en lo que se ha denominado un “cariotipo molecular”. Esto es útil para estudiar grandes reorganizaciones de los cromosomas (págs. 447-448), que tienen lugar, por ejemplo, en procesos cancerosos.



12.17

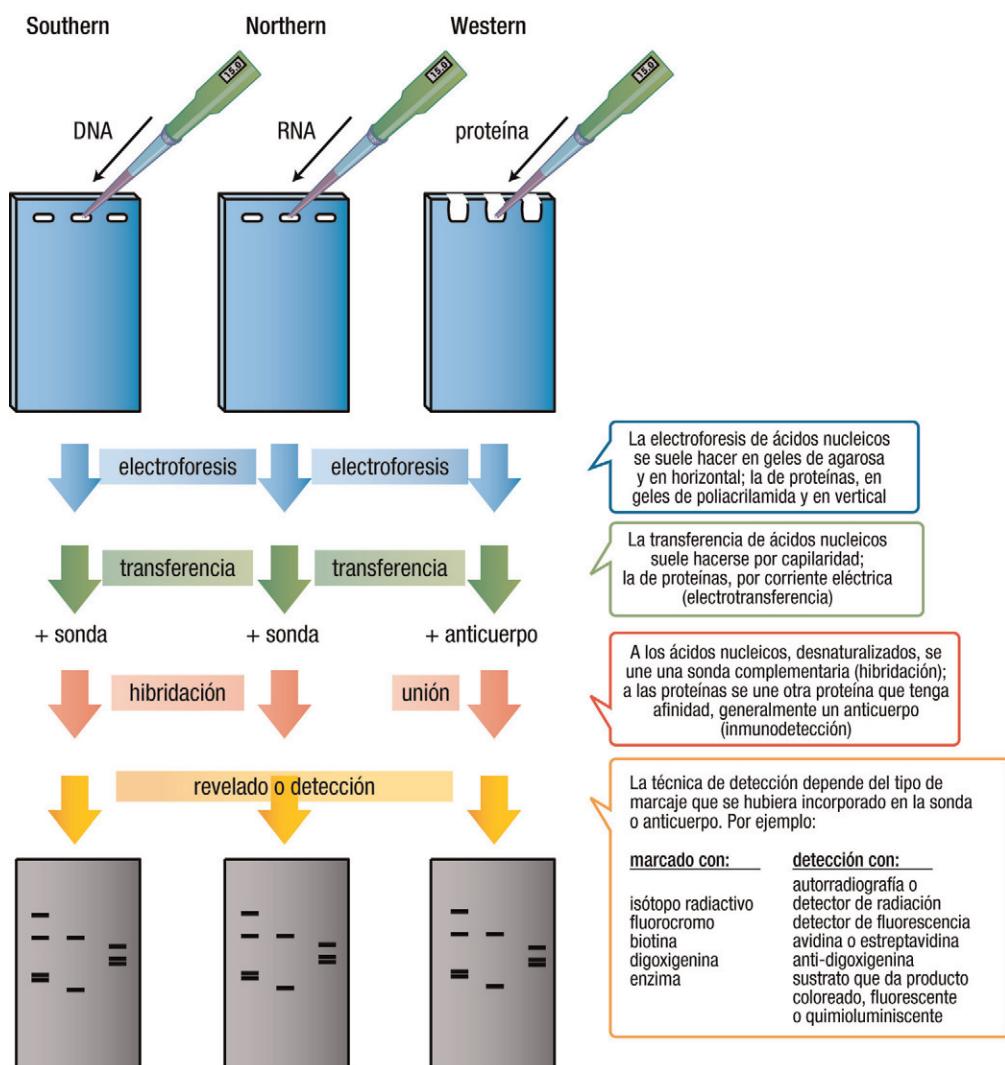
Imagen procedente de http://commons.wikimedia.org/wiki/File:Sky_spectral_karyotype.gif.
Licencia: dominio público

12.3.5 Técnicas relacionadas

Se comenta finalmente una técnica de análisis que, aun no estando basada en los principios de hibridación, posee tal similitud en los planteamientos que hace interesante su exposición aquí.

12.3.5.1 Transferencia Western

A pesar de que en este caso no existe hibridación entre ácidos nucleicos, se describe este método por su similitud técnica y mecanística con las hibridaciones Southern y northern. El ensayo Western (*western* = del oeste) se emplea para detectar las **proteínas** presentes en una muestra. Para ello, se las somete a electroforesis en gel de poliacrilamida (separándose en función de una combinación de carga, tamaño y forma, salvo en el caso frecuente de la electroforesis en presencia de SDS, que separa sólo en función del tamaño o masa molecular) y se transfieren a una membrana (*western blot*). La detección de las bandas proteicas en la membrana o filtro se realiza gracias a la unión específica de un anticuerpo dirigido contra la proteína o región proteica que se quiere estudiar. De forma similar al marcaje y detección de sondas de ácido nucleico, se puede marcar directamente el anticuerpo o detectarlo mediante la unión de una proteína ligante marcada a su vez, tanto empleando isótopos como técnicas no radiactivas. El paralelismo (y las diferencias) entre los 3 métodos se resumen en la figura.



12.18

De forma análoga, también se realizan ensayos “dot-blot” y “slot-blot” para proteínas, con un planteamiento similar al expuesto para ácidos nucleicos, salvo por el tipo de muestra, y con procedimientos de unión y detección completamente análogos al del *western* (en este caso, obviamente, al no realizar electroforesis no es precisa la transferencia).

Han surgido también otros numerosos métodos de análisis de ácidos nucleicos, variantes de los aquí expuestos, que combinan el reconocimiento de los ácidos nucleicos mediante sondas con la interacción específica entre aquéllos y proteínas, como los ensayos de retardo en gel o los métodos *southwestern* y *far-western*.



Web 12.5. Técnicas para analizar la interacción de ácidos nucleicos con proteínas.

12.4 PREPARACIÓN DE SONDAS

Como ya se ha indicado, la detección de ácidos nucleicos mediante ensayos de hibridación depende de la presencia de la secuencia diana que debe detectarse en la muestra y de la disponibilidad de una sonda u oligonucleótido de secuencia complementaria a la diana. Por tanto, la utilidad de la hibridación, el punto clave, radica en la sonda: de su naturaleza depende que el ensayo sea fructífero. El papel de la sonda equivale al del anticuerpo en un inmunoensayo y, al estar marcada, permite localizar y cuantificar cualquier secuencia diana con la cual se pueda hibridar. Gracias al empleo de sondas, la sensibilidad de la hibridación es muy elevada, hasta un millón de veces superior a la detección directa por otros métodos, por ejemplo la tinción tras electroforesis. La sonda se convierte así en un elemento esencial para cualquier estudio molecular, en especial para analizar la variabilidad o polimorfismo del genoma. De ahí la conveniencia de estudiar los métodos de preparación y marcaje de sondas, como aspectos esenciales para su empleo en ensayos de hibridación.

Se utiliza una gran variedad de métodos para obtener sondas, tanto de DNA (monocatenario y bicatenario) como de RNA. Obviamente, para su empleo en el ensayo de hibridación la sonda siempre ha de ser monocatenaria, para lo cual simplemente es preciso desnaturizar las sondas bicatenarias.

12.4.1 Clasificación de las sondas, en función del método de preparación

Las sondas se pueden clasificar según varios criterios, como su tamaño, longitud o tamaño de la secuencia diana, método de marcaje y detección, método de preparación, etc. Este último criterio es el más utilizado, distinguiéndose entre sondas preparadas por métodos convencionales (por clonación celular o acelular, Capítulos 14 y 15) y por métodos de síntesis química.

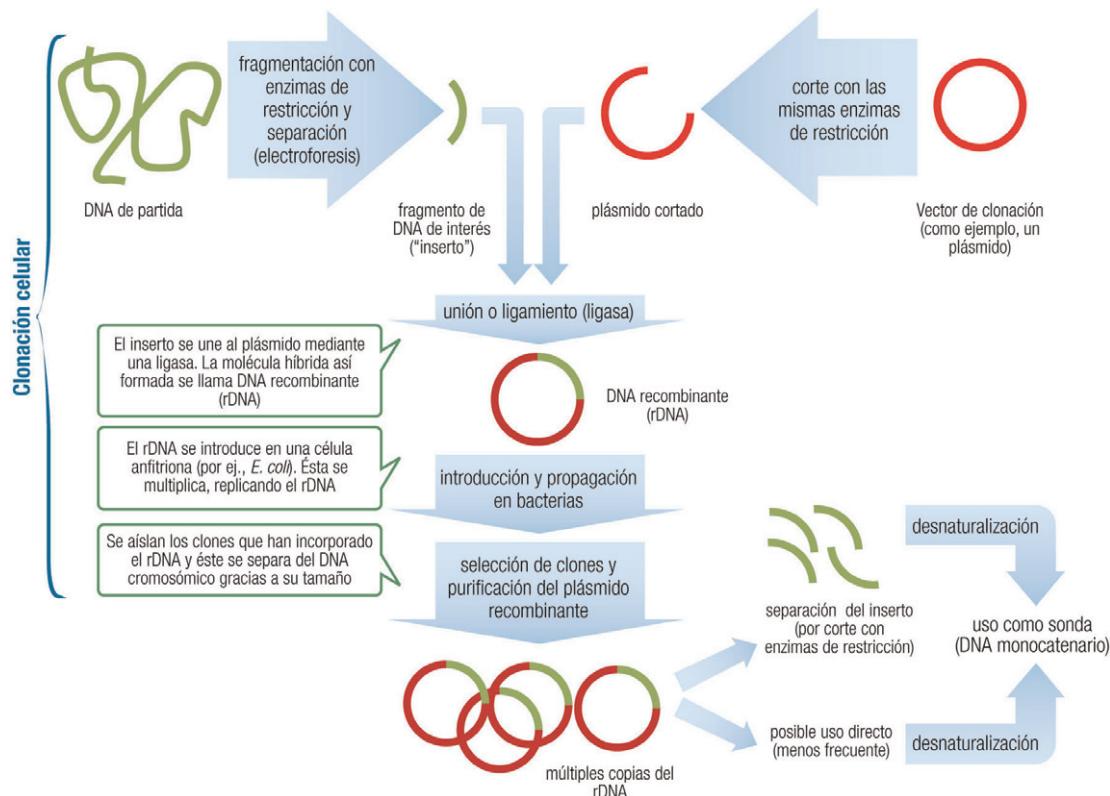
Tipo de sonda:	Convencionales			Sintéticas	
Método de obtención:	Clonación celular clásica	Clonación acelular (PCR)	PCR con transcriptasa inversa (RT-PCR)	Clonación celular en vector especial seguida de transcripción	Síntesis química
Material de partida:	DNA bicatenario		RNA monocatenario	DNA bicatenario	Nucleótidos
Naturaleza de la sonda formada:	Fragmento de DNA (oligodesoxirribonucleótido)			“Ribosonda”, o fragmento de RNA (oligorribonucleótido)	Oligonucleótido (“oligo”)

12.4.2 Sondas convencionales

También llamadas **sondas genómicas, clonadas o biológicas**, fueron las primeras que se utilizaron y han contribuido especialmente al progreso de la biología molecular.

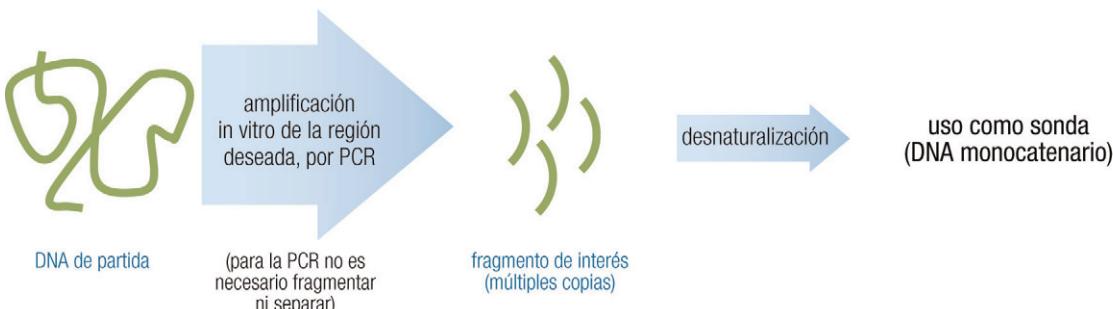
12.4.2.1 Sondas de DNA obtenidas por clonación celular (tecnología del DNA recombinante)

Poseen tamaño (longitud) variable, entre 100 bases y cientos de kb. Aunque los detalles experimentales de la clonación celular se estudian posteriormente (Capítulo 15), se adelantan aquí de forma breve las etapas principales del proceso.



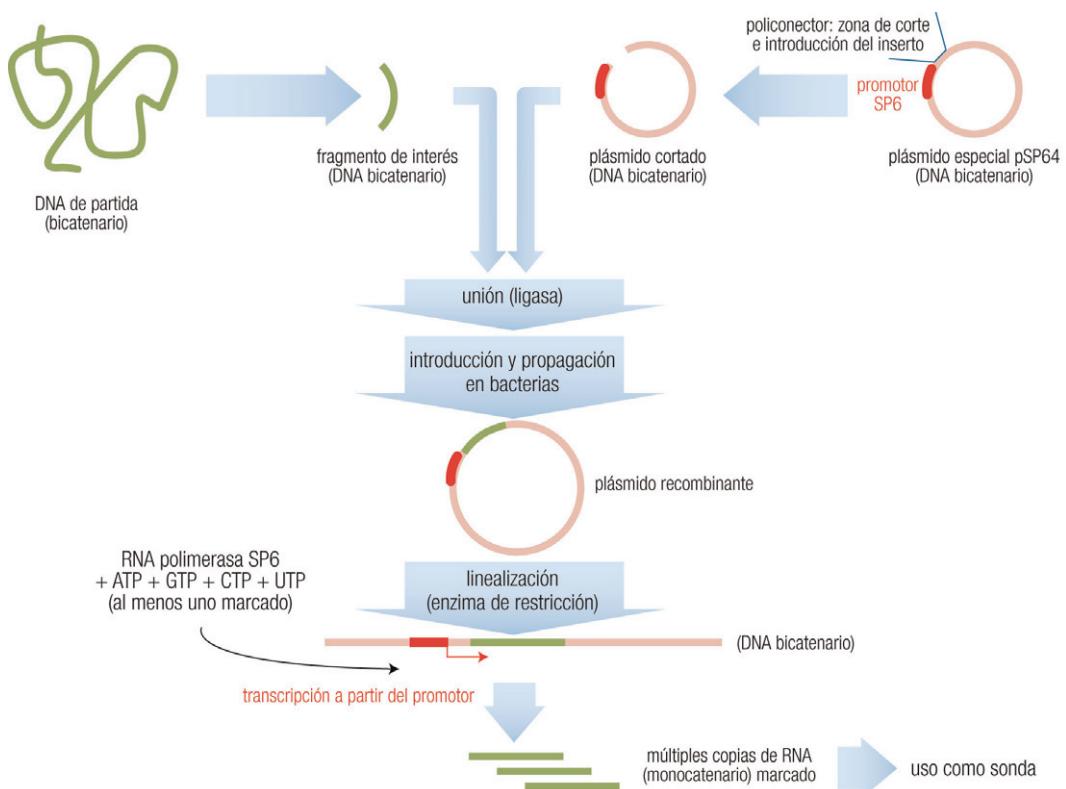
12.4.2.2 Sondas de DNA obtenidas por clonación acelular (amplificación por PCR)

Estas sondas suelen tener menor tamaño que las obtenidas por clonación celular, entre 0,1 y 20 kb. El método de preparación (PCR, Capítulo 14) consta de un menor número de etapas, por lo que las sondas se obtienen de forma más rápida y sencilla. En general, partiendo de DNA bicatenario se preparan sondas de la misma naturaleza, salvo en la PCR asimétrica, que conduce a DNA monocatenario, y en la RT-PCR, que permite obtener sondas de DNA bicatenario partiendo de un RNA (pág. 207).



12.4.2.3 Sondas de RNA (ribosondas)

La clonación celular se puede emplear también para preparar sondas de RNA de hasta unos miles de nucleótidos. El proceso, relativamente complejo, se basa en la transcripción de un DNA previamente clonado. Se usa un plásmido especial (como el *pSP64*), que se caracteriza por incluir el promotor de una RNA polimerasa de fago (*promotor SP6*). El DNA recombinante que resulta, una vez purificado y linealizado, se utiliza como molde para la transcripción por la RNA polimerasa *SP6*, que transcribe específicamente a partir de la secuencia promotora *SP6*, sintetizando así numerosas copias (monocatenarias) del inserto RNA. Si alguno de los NTP, sustratos de la RNA polimerasa, está marcado (pág. 186), el RNA se puede utilizar directamente como sonda; si no, deberá marcarse el RNA posteriormente a su obtención.



Ventajas de las ribosondas

Al no poderse formar el homodúplex sonda : sonda, se favorece la eficacia de la interacción de la sonda con la secuencia diana. Las sondas no unidas específicamente pueden ser eliminadas por digestión con una RNasa, que degrada moléculas monocatenarias pero respeta los híbridos RNA sonda : DNA diana. Se puede emplear como control negativo la sonda complementaria (secuencia en dirección *antisentido*, pág. 267, conseguida al incorporar el inserto junto al promotor en orientación contraria)

Desventajas de las ribosondas

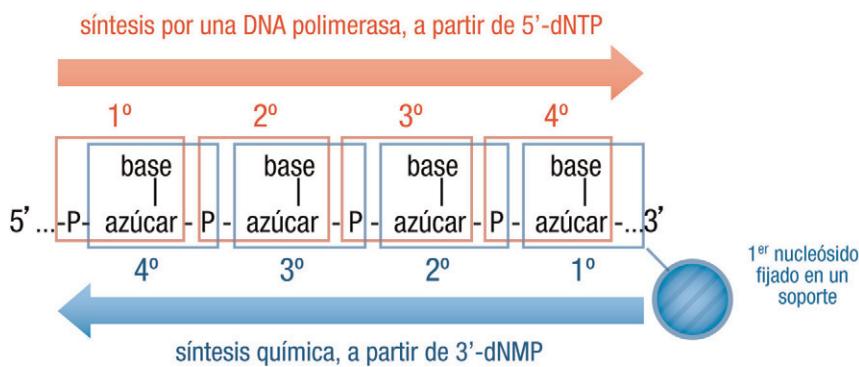
Es necesario usar un vector de clonación particular, con un promotor de transcripción. Se aumenta la complejidad al precisarse la etapa de reacción de la RNA polimerasa. Puesto que el RNA es mucho más lábil que el DNA, y las RNAsas son una contaminación muy frecuente, se requiere un manejo cuidadoso de las sondas de RNA para evitar su degradación y el uso de técnicas y material estériles y específicamente libres de RNAsas.

12.4.3 Sondas sintéticas

La capacidad de síntesis química es esencial en ingeniería genética no sólo para preparar sondas, sino también para modificar un producto génico de forma parcial o notable, o –idealmente– para sintetizar genes completos. En general, se obtienen sondas de DNA más cortas que por clonación. El método más popular es la síntesis de sondas cortas de DNA (de 15 a 45 nucleótidos), monocatenarias, denominadas oligonucleótidos o simplemente oligos. Éstos pueden ser también de RNA, aunque se sintetizan con mayor dificultad.

12.4.3.1 Síntesis química de oligonucleótidos

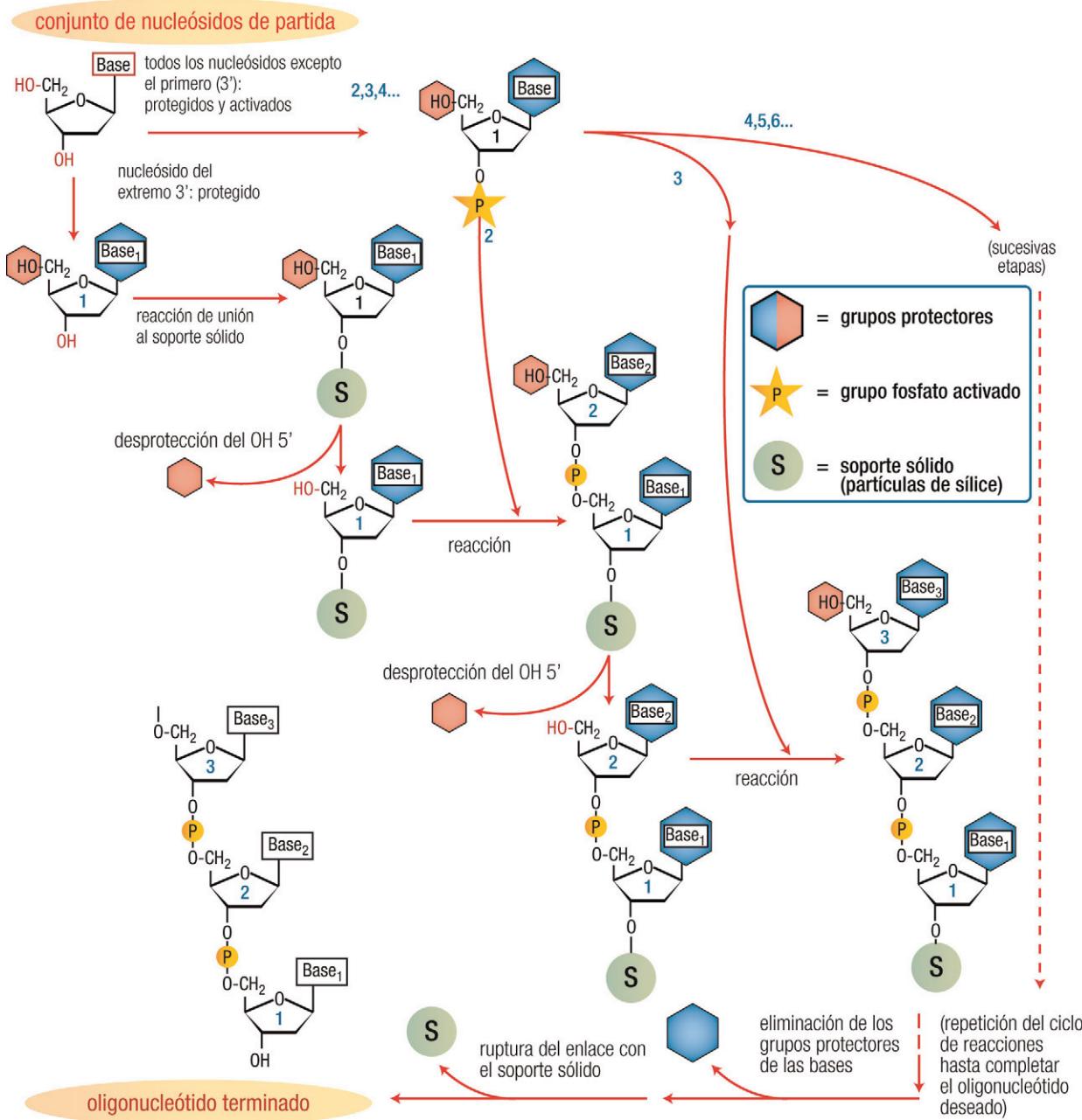
El aspecto básico de la síntesis es la formación repetitiva de un enlace éster entre el grupo 3'-fosfato de un nucleótido y el grupo 5'-OH de otro, para dar un enlace fosfodiéster. Es decir, se parte de los 3'-NMP y se avanza en sentido 3' → 5', a diferencia de la síntesis de ácidos nucleicos en las células (tanto replicación como transcripción emplean los 5'-NTP y avanzan de 5' a 3'). Los NMP se añaden uno a uno partiendo del primer nucleósido, que está unido por su posición 3' a un soporte sólido y constituirá el extremo 3' del ácido nucleico una vez sintetizado.



12.22

El principal problema es la gran reactividad de los dNMP, con un grupo -OH, un grupo fosfato y grupos amino o carbonilo en la base nitrogenada, por lo que se requieren métodos de bloqueo y desbloqueo que no afecten de manera irreversible a ninguno de los componentes de la molécula de nucleótido. Por otro lado, el grupo fosfato (en 3') ha de activarse para que reaccione con el -OH (en 5' del nucleótido anterior).

Método desarrollado en los años 70 por H. Gobind Khorana para la síntesis química de oligonucleótidos:
(se han omitido los detalles químicos para enfatizar la estrategia del método)



Web 12.6. Detalles de la síntesis química de oligos.

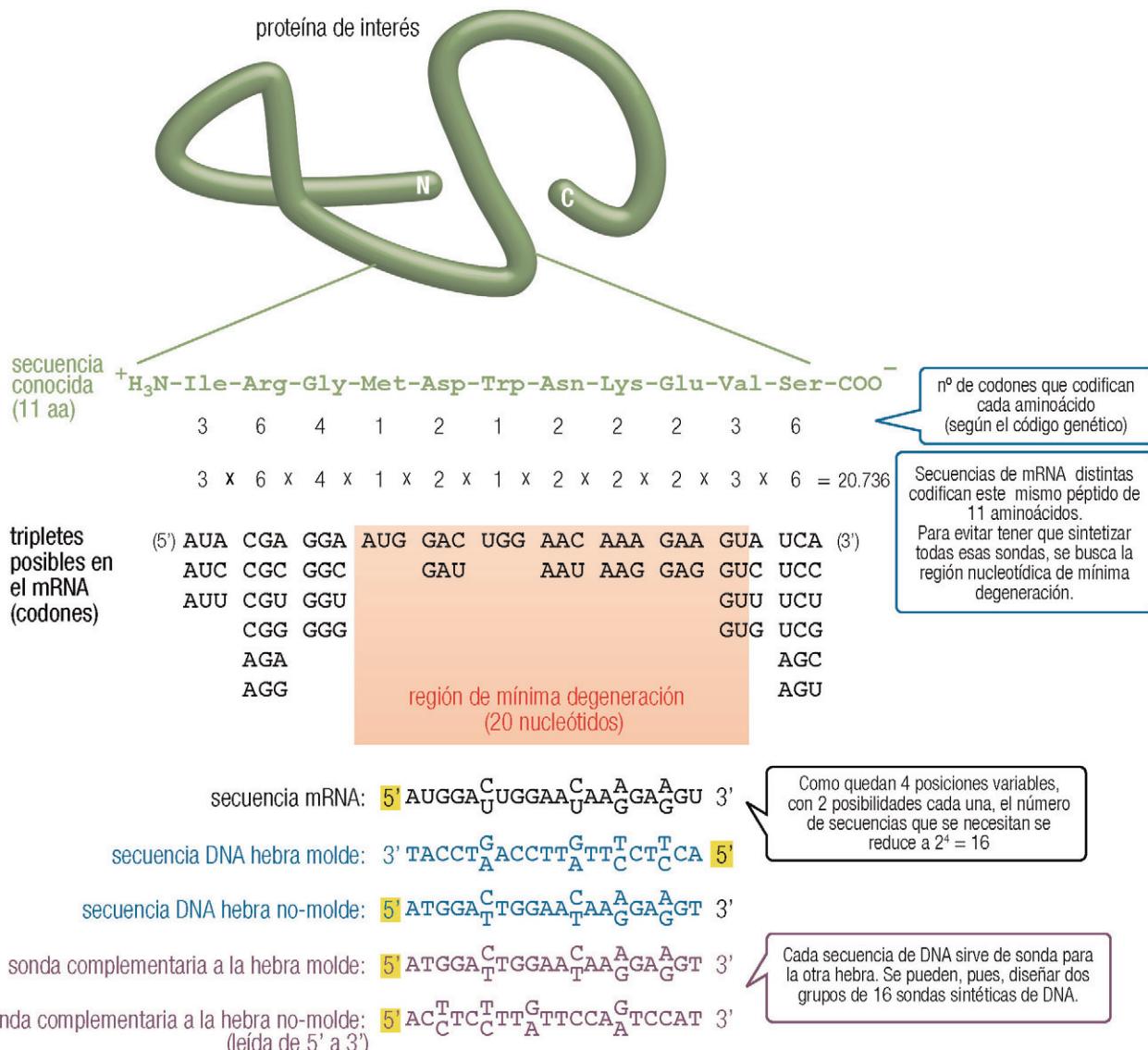
Estos procesos de síntesis están completamente automatizados, son eficientes y de bajo coste. Debido a que la longitud de estos oligos se ve limitada por el rendimiento de las reacciones de síntesis, la sensibilidad alcanzada con ellos es, en general, menor que con las sondas obtenidas por clonación. Sin embargo, la continua mejora en los métodos de síntesis química permite la obtención de oligos cada vez mayores.

Respecto a la síntesis química de oligorribonucleótidos, se utiliza mucho menos debido principalmente a la dificultad de proteger el grupo OH 2' sin afectar al 3', y habida cuenta de que el resultado de la hibridación y sus aplicaciones son los mismos que con los oligodesoxirribonucleótidos.

12.4.3.2 Diseño de una sonda sintética

Las sondas sintéticas se diseñan de forma que contengan la secuencia complementaria a una zona del DNA diana; por ejemplo, de una parte del gen que interesa detectar. El objetivo es preparar un oligonucleótido que pueda hibridarse con dicho DNA. Por ejemplo, con frecuencia se pretende localizar el gen que codifica una proteína determinada. Para diseñar la sonda, es necesario:

1. Conocer la secuencia completa de la proteína, o al menos de un fragmento de ella, al que corresponderá la secuencia diana de DNA.
2. A partir de esa secuencia de aminoácidos se establece (empleando el código genético) la secuencia de nucleótidos que habrá de tener el DNA. Sin embargo, debido a la degeneración del código genético (distintos trinucleótidos codifican un mismo aminoácido, pág. 318), hay un número elevado de secuencias de DNA que codifican el mismo polipéptido y, obviamente, no se puede conocer *a priori* cuál es la que está presente en el genoma de la célula, por lo que es preciso sintetizar sondas que cubran todas las posibilidades.
3. Para reducir el número de sondas necesarias, se selecciona el segmento de secuencia que posea la menor degeneración posible.
4. Una vez sintetizadas todas las sondas, se ensaya su hibridación con el DNA de partida; una de ellas será la correcta.



En este tipo de experimentos, los métodos de síntesis y secuenciación de proteínas y de DNA convergen y se complementan. Como se acaba de ver, la secuencia de una proteína aislada puede usarse para localizar e incluso aislar el gen correspondiente; de una forma similar, aunque inversa, un gen clonado (y, por tanto, secuenciado) puede usarse para aislar la proteína producto: la secuencia del gen permite diseñar y sintetizar químicamente un péptido corto que representa una parte del producto proteico del gen. Entonces, una de las posibilidades es, por ejemplo, preparar anticuerpos contra este péptido sintético y emplearlos para identificar y purificar la proteína para su estudio posterior.

12.5 MARCAJE DE LAS SONDAS

El empleo de sondas de DNA o RNA como herramientas analíticas para localizar secuencias diana mediante ensayos de hibridación requiere obligatoriamente un proceso de marcaje de la sonda, que permita su detección. La molécula o grupo **marcador** puede tener diversa naturaleza, siempre y cuando se pueda detectar su presencia de forma **directa** o **indirecta**; los marcadores más empleados son **isótopos radiactivos**, **fluorocromos** y **enzimas**. En general, dan lugar a métodos sencillos de realizar, que detectan concentraciones muy bajas de la sonda. El marcaje debe ser estable bajo diversas condiciones de temperatura, detergentes, disolventes, etc., y no afectar a la reacción de hibridación. Clásicamente se utilizaron los métodos radiactivos, pero en gran medida se han visto desplazados por otros marcadores, sobre todo para evitar los riesgos derivados de la radiación y la gestión de residuos, para aprovechar la mayor vida media de sus reactivos de marcaje y, en algunos casos, para mejorar el límite de detección.

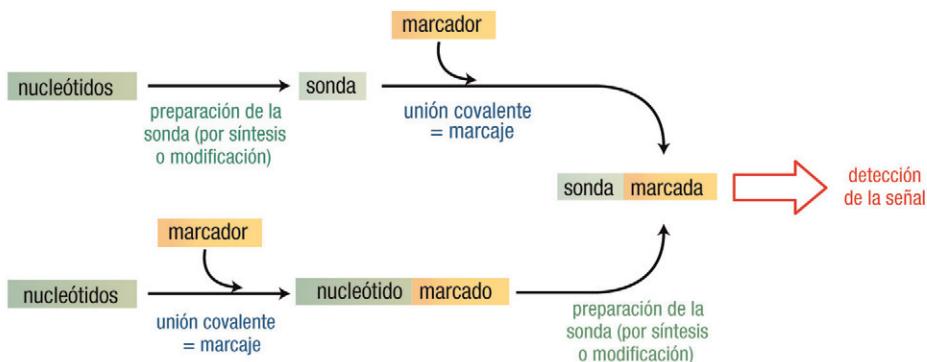
De forma genérica, pueden plantearse dos tipos de proceso de marcaje:

- *In vivo* (más correctamente, en cultivo): se hace suministrando nucleótidos radiactivos a células en cultivo, que los incorporan en sus ácidos nucleicos de nueva síntesis. Su uso es limitado; por ejemplo, para preparar DNA viral marcado en células infectadas por virus.
- *In vitro*: es más versátil. Pueden emplearse isótopos radiactivos u otros métodos alternativos. Se describen a continuación algunas de las diversas variantes metodológicas.

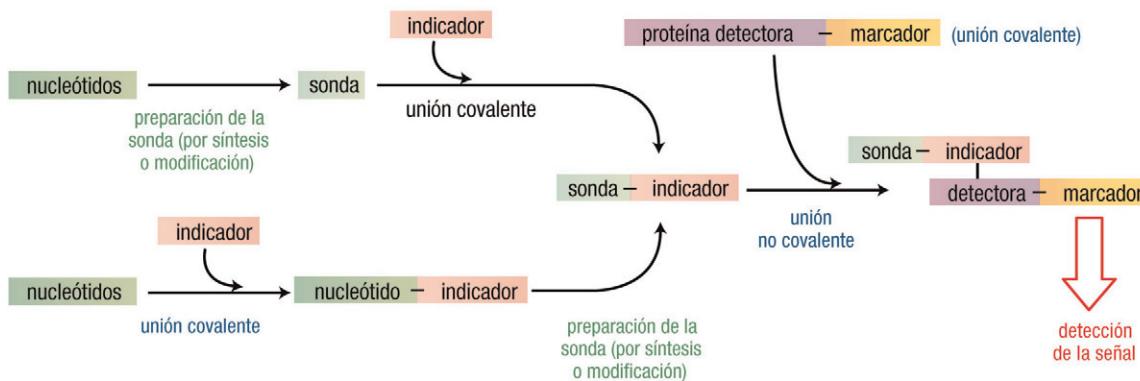
En muchas ocasiones, en especial con ribosondas y sondas sintéticas, la sonda se marca **al mismo tiempo que se prepara** incluyendo, con los nucleótidos que hay que suministrar a la reacción de síntesis, alguno de ellos marcado (con un isótopo radiactivo, un fluorocromo o cualquier tipo de grupo *indicador*), mientras que en otras el marcaje se realiza **tras haber completado la síntesis** de la sonda. Puesto que la preparación de sondas ya se ha estudiado, nos centraremos ahora en este último caso, aunque las consideraciones relativas a la detección sirven también para sondas marcadas durante su obtención.

El principio básico del marcaje es la unión química o *conjugación* del **marcador** a la sonda (en el caso radiactivo, en realidad el isótopo forma parte de la molécula de sonda, no se une a ella). Como se acaba de indicar, esto puede hacerse uniendo el marcador a la sonda ya preparada o a un nucleótido que se incorporará durante la síntesis de aquélla. Por otra parte, la sonda puede llevar unido el marcador detectable (**métodos directos** de marcaje, a continuación) o bien otra molécula no detectable, “etiqueta” o “grupo indicador” (en inglés, *reporter*), con la capacidad de ser reconocida posteriormente por una “proteína detectora”, que es la que va marcada para la detección (**métodos indirectos**, pág. 188).

Marcaje directo: el “marcador”, que se puede detectar, va unido a la sonda.
Dos alternativas, después de la síntesis de la sonda o durante ella:



Marcaje indirecto: el “marcador” unido a la sonda no es detectable directamente; se le llama **indicador**.
Para la detección, el indicador se une a una proteína detectora, que lleva el marcador.
Dos alternativas, después de la síntesis de la sonda o durante ella:

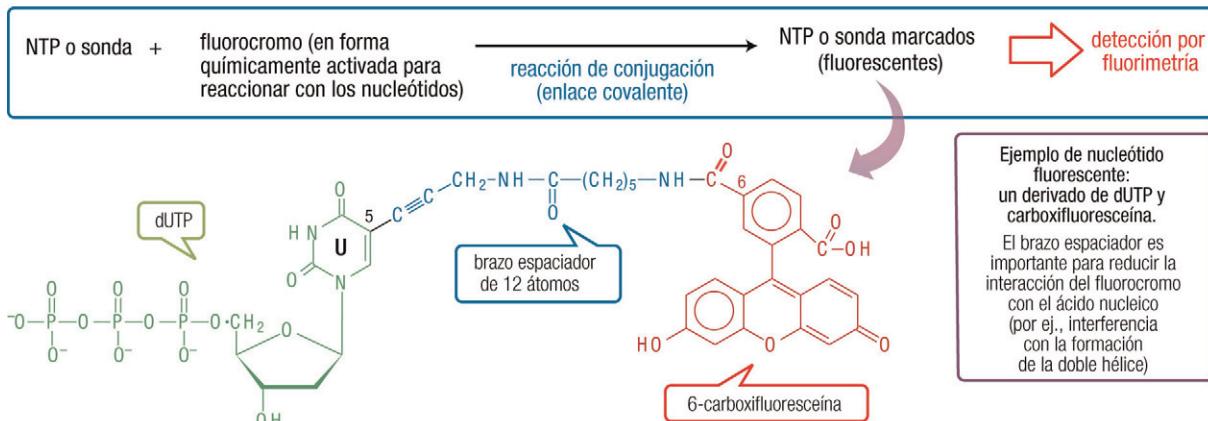


12.25

12.5. I Marcaje directo en la molécula de sonda

El grupo marcador se une de forma covalente, bien a grupos funcionales específicos o bien de forma inespecífica. Los ácidos nucleicos ofrecen menos grupos funcionales adecuados para este fin que las proteínas; básicamente, algunos átomos de las bases nitrogenadas (en especial, los no implicados en el emparejamiento con la base complementaria) y los grupos terminales 3'-OH y 5'-P.

Marcaje covalente de la sonda (por métodos químicos):



12.26

En esta categoría de métodos pueden considerarse tanto la inclusión de nucleótidos marcados durante la preparación de la sonda (ya sea por síntesis química o biológica) como el marcaje de la sonda ya terminada. Se comentan a continuación los más representativos.

12.5.1.1 Marcaje con DNA polimerasa

Se trata de métodos de **síntesis enzimática**, los más comunes. Los cuatro tipos de nucleótido (al menos uno de ellos marcado) se incorporan durante la reacción de síntesis de una cadena de DNA. Existen dos variantes principales: traslado de la mella (*nick translation*) e iniciación o cebado aleatorios (*random priming*).

a) Método de traslado de la mella

Es un método fácil y rápido, considerado el ejemplo clásico de incorporación enzimática de un marcaje. Requiere dos enzimas distintas: una desoxirribonucleasa (DNasa I pancreática) y una DNA polimerasa (DNAPol-I de *E. coli*). El resultado del marcaje es la sustitución de una gran parte (30-60%) de los nucleótidos originales de cada hebra de DNA por otros marcados con isótopo radiactivo en su posición α -5'-P.



Web 12.7. Método de traslado de la mella.

Puesto que el nucleótido marcado se incorpora a la sonda, en lugar del marcaje radiactivo pueden emplearse también nucleótidos conjugados con fluorocromos o con grupos indicadores (págs. 185 y 189).

b) Método de iniciación o cebado aleatorios

Esta alternativa, a diferencia de la anterior, requiere la desnaturalización inicial del DNA bicatenario para que hibride un oligonucleótido que actúa como cebador de la replicación por DNAPol-I de *E. coli*. Ahora bien, en este caso sólo se requiere la actividad polimerasa, por lo que se prescinde del *fragmento pequeño* y se suele utilizar el *fragmento de Klenow* (pág. 150). Esta técnica ofrece una serie de ventajas de carácter técnico respecto a la anterior, cuya consideración escapa a nuestros objetivos de carácter esencialmente conceptual.



Web 12.8. Método de cebado aleatorio.

Al igual que el método anterior, admite también nucleótidos conjugados con fluorocromos o con grupos indicadores.

12.5.1.2 Marcaje terminal

En este caso se aplica, en lugar de una síntesis, una **modificación enzimática** de las hebras de ácido nucleico existentes, mediante la incorporación de uno o unos pocos nucleótidos marcados a uno o ambos extremos de la sonda de DNA o RNA. Aunque se utilizan menos que los anteriores, son métodos útiles en ciertos casos, por ejemplo para marcar oligos monocatenarios. Se emplean dos planteamientos: el marcaje con polinucleótido quinasa y el marcaje terminal por relleno (*fill-in end-labeling*).

a) Marcaje con polinucleótido quinasa

Este método no depende de la existencia de una hebra molde y, por tanto, puede emplearse para cualquier tipo de sonda, mono- o bicatenaria, de DNA o de RNA. Consiste, en esencia, en un intercambio del 5'-P terminal de cualquier hebra por el fosfato (marcado) en posición γ de un ATP, con la intervención de polinucleótido quinasa y, opcionalmente, fosfatasa alcalina.



Web 12.9. Método de marcaje con polinucleótido quinasa.

En este caso sólo pueden usarse marcajes radiactivos, pues lo único que se incorpora a la sonda es el fosfato γ del ATP.

b) Marcaje terminal por relleno

Este método es el único que utiliza una enzima de restricción específica para formar dos extremos cohesivos (pág. 215), que luego son “rellenados” por la DNAPol-I.

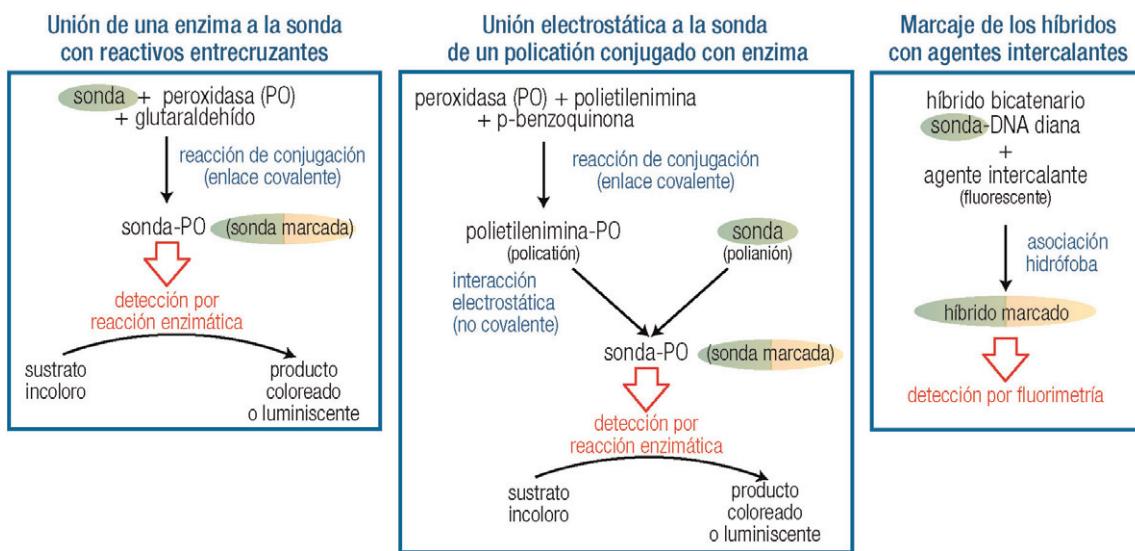


Web 12.10. Método de marcaje terminal por relleno.

Al igual que algunos de los anteriores, es posible el uso tanto de nucleótidos radiactivos como de los conjugados con fluorocromos o con grupos indicadores.

12.5.2 Marcaje externo a la sonda

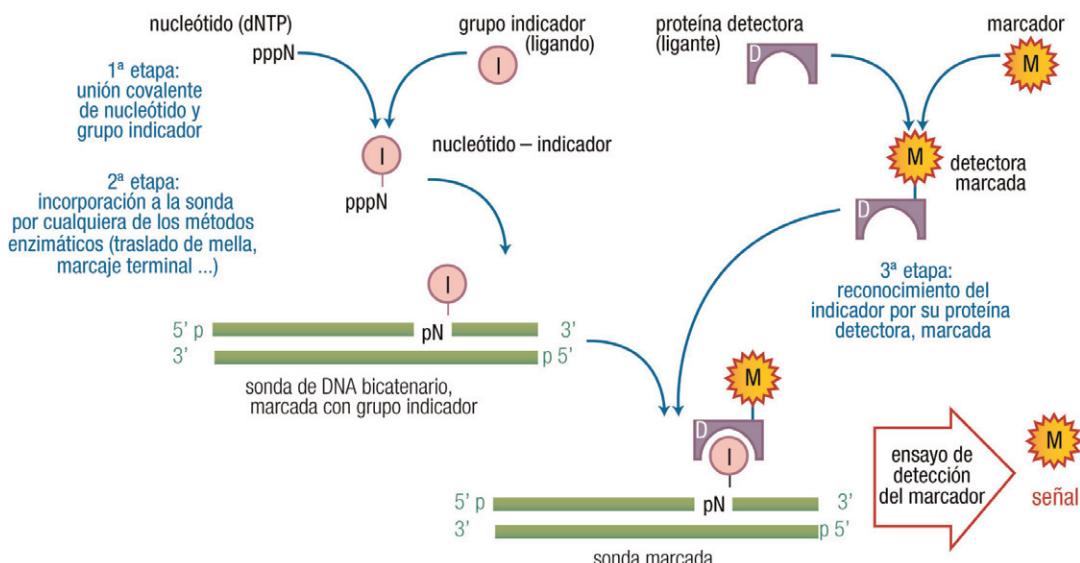
En lugar de marcar sus propios nucleótidos, es posible hacer que la sonda se una a otra molécula que a su vez lleva el marcador. Una forma es la unión de proteínas mediante reactivos entrecruzantes (formaldehído, glutaraldehído, etc.; en inglés, *crosslinkers*). También puede unirse el marcador por interacción electrostática con el DNA o emplearse agentes intercalantes (v. web 10.14), que se insertan entre los pares de bases de la doble hélice (híbrido sonda: DNA) de forma no covalente.



12.27

12.5.3 Marcaje indirecto

Como se ha mencionado (pág. 185), puede convenir que a la sonda se une un grupo no detectable, *etiqueta* o *indicador*, el cual en una segunda etapa es reconocido por otra molécula, por lo general una proteína, que incorpora el marcador detectable. Para la unión del indicador son aplicables las consideraciones ya realizadas respecto a la unión del marcador en los métodos directos. Se ilustra un ejemplo correspondiente a la modificación de los NTP:



12.28



Web 12.11. Marcaje indirecto de las sondas.

12.5.4 Tipos de marcadores, etiquetas o indicadores y su detección

12.5.4.1 Marcadores radiactivos

El marcaje radiactivo o isotópico es el que se ha empleado históricamente con más frecuencia, aunque está siendo desplazado de manera progresiva por los métodos no radiactivos. Aunque no existe consenso acerca del marcaje ideal, en la mayoría de los casos se hace con ^{32}P . Puesto que la breve vida media de este isótopo (14,2 días) crea dificultades en la preparación, comercialización y uso rutinario de los ensayos, se utilizan también otros alternativos (^{35}S , ^{3}H , ^{125}I o ^{14}C).

La detección y registro de la hibridación se hace por *autorradiografía*: el soporte con el material hibridado se pone en contacto con una película fotográfica (del tipo empleado para las radiografías), la cual queda impresionada con una imagen que muestra la presencia de la radiactividad (v. por ejemplo, el método de Southern, pág. 171). La intensidad de la señal depende de la intensidad de la radiación emitida por el isótopo y del tiempo de exposición. La detección empleando película está siendo desplazada por el uso de instrumentos que detectan la radiación y la ofrecen directamente como señal digitalizada.

12.5.4.2 Marcadores no radiactivos

En función del marcador empleado, la detección se hace mediante fluorimetría o añadiendo un sustrato específico cuya transformación por la enzima marcadora pueda detectarse empleando alguno de los siguientes principios:

- Espectrofotometría: medida de la absorción de la luz debida a los productos de la reacción, sean solubles o insolubles, en el ultravioleta o el visible (colorimetría).
- Fluorimetría: medida de la luz emitida por productos de la reacción que posean un fluoróforo (web 10.7).
- Quimioluminiscencia: medida de luz emitida como consecuencia de reacciones con esta característica (por ejemplo, la transformación de las luciferinas catalizada por luciferasa, o la oxidación del luminol).

Debe señalarse que éstos y otros procedimientos no son exclusivos del marcaje de sondas, sino que se aplican también en otros ensayos de detección de proteínas, antígenos, pequeños ligandos, etc.

Resumen de marcadores, indicadores y métodos para su detección

I Grupo <i>indicador</i> (ligando)	D Proteína <i>detectora</i> o ligante (*)	M Marcador	Detección del marcador	Sustrato de la reacción enzimática de detección
biotina	avidina, estreptavidina, o variantes (**)	isótopo radiactivo	autoradiografía	
digoxigenina	anticuerpo anti-DIG	fluorocromo	fluorimetría	
		peroxidasa (abreviada como PO, POD, HRPO, HRP)	reacción cromogénica: reacción fluorogénica: quimioluminiscencia:	ABTS TMB derivados de α -naftol derivados de dioxetano o de luminol
		fosfatasa alcalina (AP)	reacción cromogénica: reacción fluorogénica: quimioluminiscencia:	NBT+BCIP NPP AttoPhos™ HNPP+FastRedTR derivados de dioxetano
		β -galactosidasa (β -gal)	reacción cromogénica: microscopía	X-gal CPRG
		oro coloidal		

* La proteína ligante, con afinidad por el grupo indicador, va marcada para su detección, por alguno de los métodos indicados en la tabla de la derecha.

** Variantes de avidina obtenidas por ingeniería genética para modificar la afinidad, reducir la señal inespecífica, etc.

Abreviaturas:

ABTS = 2,2'-azino-di-[3-etylbenzotiazolina sulfonato (6)]
TMB = 3,3',5,5'-tetrametilbencidina
NBT = azul de nitrotetrazolilo
BCIP = 5-bromo-4-cloro-3-indolilfosfato
NPP = 4-nitrofenilfosfato
HNPP = 2-hidroxi-3-naftoico-2'-fenilanilida fosfato
Fast Red TR = 4-cloro-2-metilbenceno-diazonio
X-gal = 5-bromo-4-cloro-3-indolil- β -D-galactopiranósido
CPRG = rojo de clorofenol- β -D-galactopiranósido

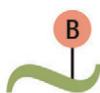
12.5.4.3 Etiquetas o indicadores

Los grupos que se usan más habitualmente son la **biotina** (una vitamina natural y coenzima) y la **digoxigenina** (un esteroide de la planta *Digitalis*). Ambos se pueden conjugar fácilmente a los nucleótidos y se unen con gran afinidad, respectivamente, a las proteínas avidina o estreptavidina y a anticuerpos anti-digoxigenina. Al igual que para los derivados fluorescentes, se dispone comercialmente de dNTP ya conjugados tanto a biotina como a digoxigenina.

La detección de las sondas que llevan biotina o digoxigenina se realiza mediante ensayos de unión, similares a los empleados para el inmunoensayo. A este fin, la proteína detectora, con afinidad por el grupo indicador, se marca previamente con un isótopo radiactivo, un grupo fluorescente o una enzima.

Algunos ejemplos de marcaje de sondas para detección indirecta

sonda biotinilada



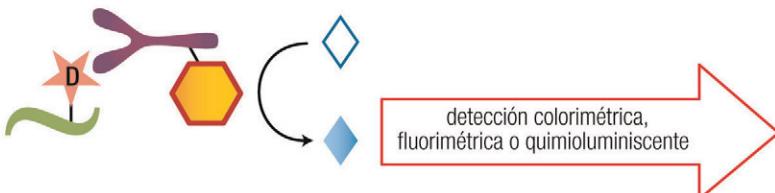
unión de avidina o estreptavidina
conjugada con fluorocromo



sonda marcada
con digoxigenina



unión de anticuerpo anti-DIG
conjugado con enzima



12.30

Clonación y células madre



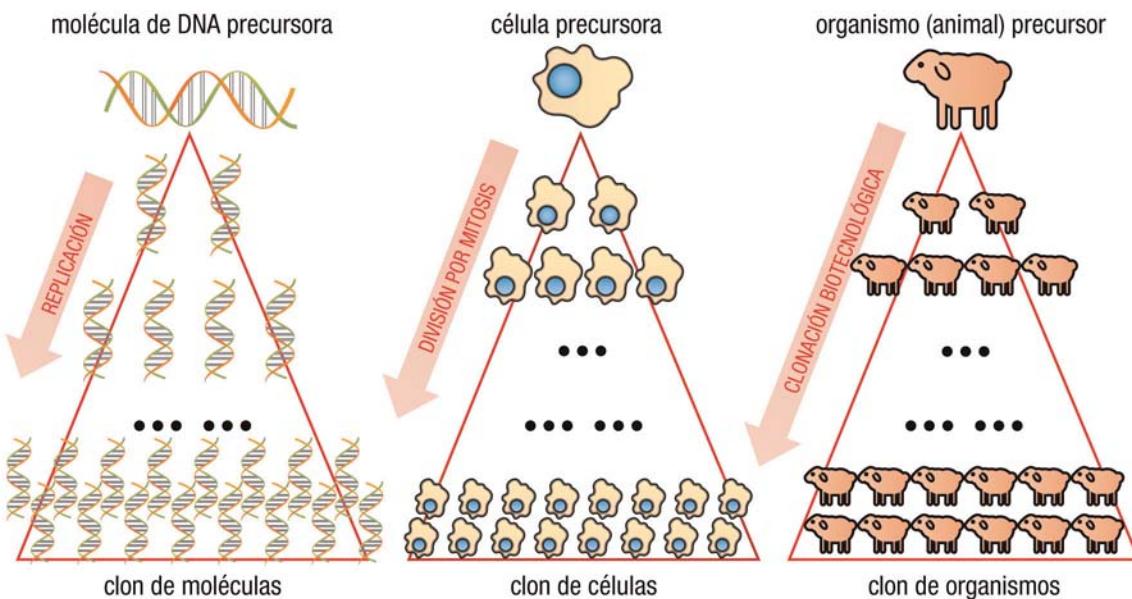
13.1 INTRODUCCIÓN GENERAL A LA CLONACIÓN	191	13.3 CÉLULAS TRONCALES O CÉLULAS MADRE	197
13.2 CLONACIÓN DE CÉLULAS	193	13.3.1 Fuentes de obtención de células troncales	198
13.2.1 Clonación por transferencia nuclear	193	13.3.2 Capacidad de diferenciación de las células	199
13.2.2 Características de la clonación de células	196	13.3.2.1 Totipotencia	199
13.2.3 Aplicaciones terapéuticas de las células obtenidas por clonación	197	13.3.2.2 Pluripotencia	200
		13.3.2.3 Multipotencia	200

13.1 INTRODUCCIÓN GENERAL A LA CLONACIÓN

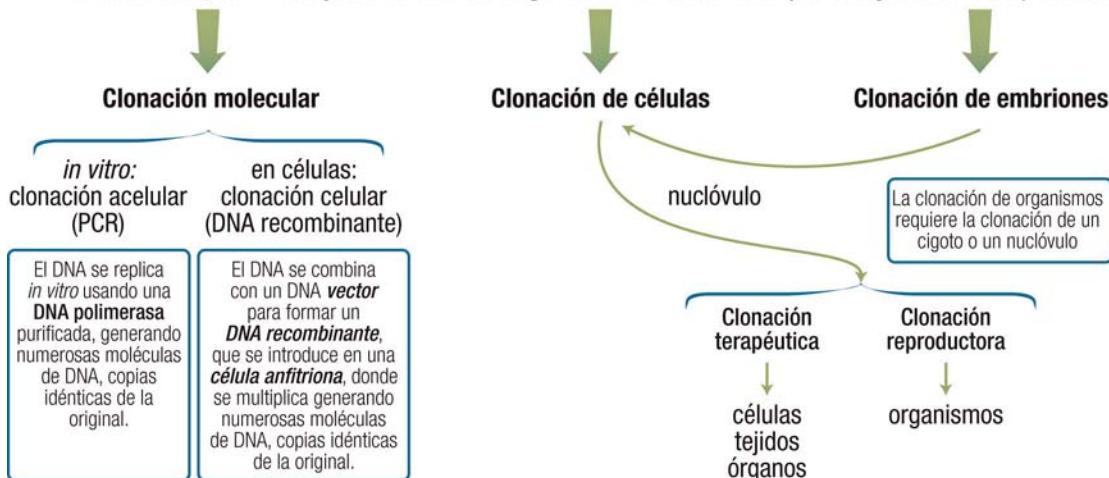
Conceptualmente, el proceso de *clonación* (también *clonado* o *clonaje*) consiste en la obtención de un *clon*, entendido como un conjunto de “elementos” genéticamente idénticos entre sí y a su precursor. Estos “elementos” pueden ser moléculas, células, tejidos, órganos u organismos pluricelulares completos. Cada componente individual de un clon contiene la misma información genética, el mismo genotipo, que el elemento de partida; por ello, se puede considerar que la clonación supone una **amplificación** genética. Ahora bien, a pesar de la definición anterior, entre los componentes de un clon también se afirma que cada uno es un *clon* de los otros (por ejemplo, cinco ovejas genéticamente idénticas forman un clon, y a la vez cada oveja es un clon de cualquiera de las demás, o todas son *clónicas* entre sí).

Debe indicarse que la clonación es un proceso fisiológico normal: son clónicas las *células embrionarias totipotentes* formadas por mitosis a partir del cigoto (v. web 10.1 y pág. 199); cada una de estas células da lugar (mediante complejos procesos de diferenciación y desarrollo) a individuos adultos formados por células somáticas clónicas, genéticamente iguales entre sí y a la célula embrionaria de partida (excepto en lo que afecta a su *diversidad* o *polimorfismo genético*, resultante de mutación, recombinación, transposición, etc.). En el nivel de organismos completos, son clónicos los gemelos monocigóticos o univitelinos, procedentes de un mismo óvulo fecundado.

El interés de la clonación radica, esencialmente, en su carácter biotecnológico como proceso de multiplicación de moléculas de DNA o RNA, células, tejidos u organismos completos, mediante mecanismos de manipulación genética. La biología en general, y la investigación biomédica en particular, están avanzando enormemente y lo harán aún más en el futuro, con la variada gama de copias de *elementos genéticos* ya empleadas para estudiar y resolver problemas tanto de carácter básico como aplicado. En particular, la clonación molecular ha sido una técnica clave para la aparición de la ingeniería genética y se emplea con fines muy diversos (aislamiento, amplificación, hibridación, secuenciación, etc.) y con aplicaciones de todo tipo a partir de DNA de cualquier procedencia, entre las que cabe citar las de interés terapéutico (producción de insulina, interferón, eritropoyetina, hormona del crecimiento, factores de la coagulación, etc.).



En cada caso, clon = conjunto de elementos genéticamente idénticos (entre sí y al elemento precursor)



Globalmente, interesa describir la clonación desde el punto de vista de su finalidad, es decir:

- **Clonación de organismos:** plantas o animales completos.
- **Clonación de células,** aisladas o formando tejidos u órganos.
- **Clonación de moléculas,** sean éstas genes o fragmentos de DNA o RNA. De acuerdo con el proceso experimental, esta clonación puede dividirse a su vez en dos tipos:
 - **Clonación acelular**, es decir, clonación de moléculas sin intervención de célula alguna. Es sinónimo de amplificación *in vitro* de moléculas de DNA o RNA mediante la **reacción en cadena de la DNA polimerasa (PCR, polymerase chain reaction)**. Se describe en el Capítulo 14.
 - **Clonación celular** (*con* células, *no de* células). Es la amplificación de moléculas de ácido nucleico mediante su introducción, asistida por *vectores*, en *células anfitrionas* en cultivo. Se emplea la **tecnología del DNA recombinante**. Se estudia en el Capítulo 15.

13.2 CLONACIÓN DE CÉLULAS

La clonación de células implica toda manipulación que conduzca a la obtención de células iguales entre sí y a la célula precursora (progenitora). No debe confundirse este concepto con el término **clonación celular**, empleado en general en el sentido de clonación *en* células, designando la obtención de un clon de moléculas de DNA (por tanto, una **clonación molecular**) mediante la tecnología del DNA recombinante, ya que requiere el uso de células “anfitrionas” para duplicar el DNA que se quiere clonar.

La clonación de células ha adquirido trascendencia en años recientes por constituir una etapa necesaria para la preparación real o potencial tanto de células troncales embrionarias, tejidos u órganos con fines aplicados en clínica (**clonación terapéutica**) como de organismos completos con diversos fines, terapéuticos o no (**clonación reproductora**). Este último objetivo posee un gran impacto en la agricultura y en la ganadería; por ejemplo, para la producción y la mejora genética.

Sin pretender llevar a cabo una descripción exhaustiva de estas metodologías, el objetivo del presente capítulo es analizar los conceptos y principios básicos de la clonación de células, en particular la concebida con fines terapéuticos. Corresponde, por tanto, analizar las peculiaridades de este planteamiento desde un punto de vista teórico, para poder después comprender su aplicación real o potencial, presente o futura, en el tratamiento de determinadas enfermedades.

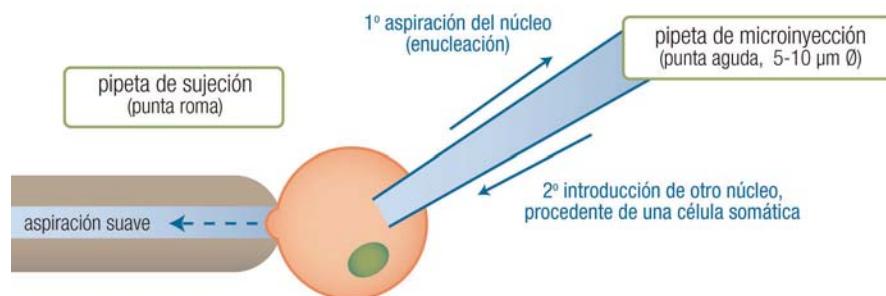
13.2.1 Clonación por transferencia nuclear

El proceso de preparación de clones celulares está íntimamente relacionado con la obtención de **células troncales**, más frecuentemente conocidas como **células madre** (pág. 197). El método más común corresponde a lo que se denomina **transferencia nuclear**, entendida como la inserción de un núcleo en una célula receptora. Éste es el método que condujo a la famosa oveja Dolly, y a muchos más animales clonados posteriormente.

El paso esencial en la clonación de células consiste en la formación, *in vitro*, de una célula precursora que dará lugar al clon celular. Debe destacarse que esta célula, a diferencia del cigoto, no comparte el genoma de dos gametos, sino que se forma mediante la transferencia del núcleo de una célula al interior de otra, previamente desprovista del suyo. La nueva célula posee, por tanto, la misma dotación genética que la célula donante del núcleo (salvo la parte de genoma mitocondrial, que corresponde a la célula receptora). El citoplasma de la célula receptora (que habitualmente es un oocito) contiene factores responsables de la reprogramación, que son proteínas citoplasmáticas capaces de interaccionar con los genes y devolverlos al patrón de expresión típico de un embrión.

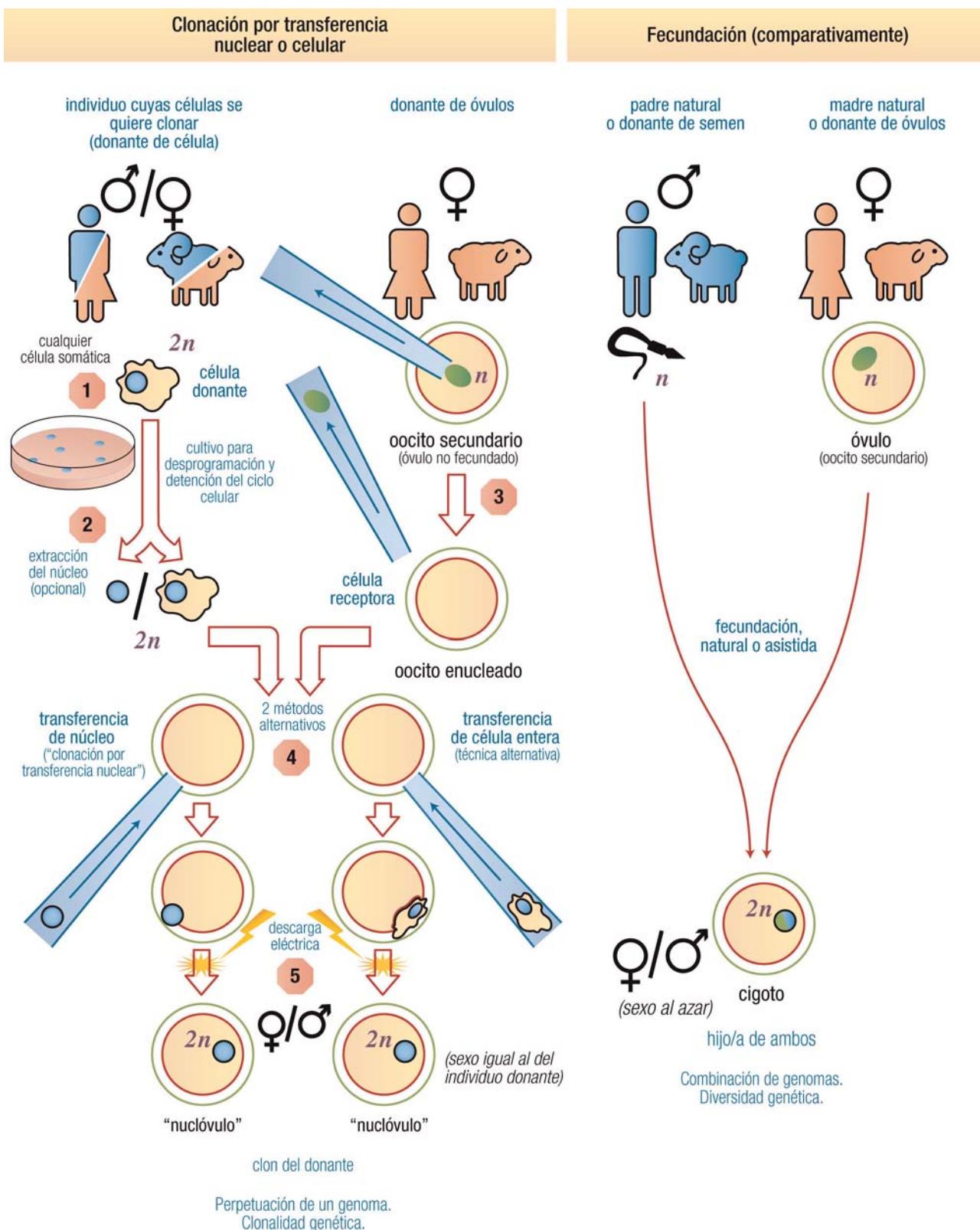
1. Como célula donante se utiliza una célula somática, tomada de cualquier tejido de un individuo normal adulto o, si se trata de un paciente, de un tejido que no se encuentre afectado por la enfermedad. Dicha célula se somete, en cultivo con nutrientes, a un proceso de desdiferenciación para convertirla en una célula similar a la embrionaria multipotencial.

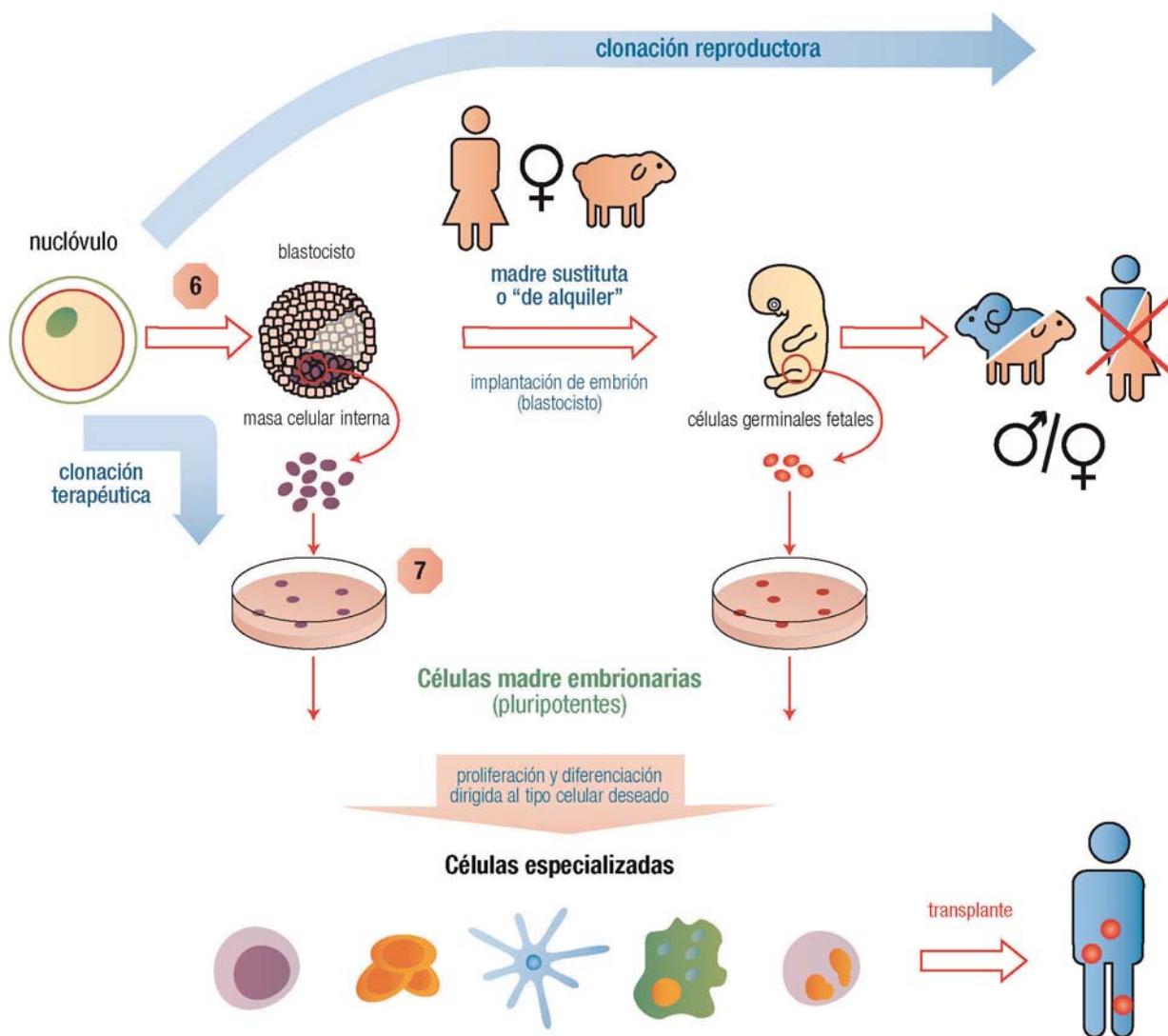
2. Se extrae entonces el núcleo de dicha célula donante mediante técnicas de micromanipulación y aspiración con una pipeta de microinyección.



13.2

3. Como célula receptora, anfitriona u hospedadora se utiliza un oocito (óvulo no fertilizado), del cual se elimina el núcleo (**enucleación**) de forma análoga al apartado anterior.
4. Mediante técnicas de microinyección se introduce el núcleo de la célula somática reprogramada en el oocito enucleado, sin alterar su citoplasma.
5. Una vez microinyectado el núcleo somático, se facilita su fusión con el oocito enucleado mediante una breve descarga eléctrica, dando lugar así a una célula híbrida, combinación de la donante y la receptora, a la que se ha denominado **nuclívulo**.
6. Se estimula el nuclívulo en cultivo para que inicie su división y para lograr que se desarrolle, primero hasta varias células y luego hasta un blastocisto (embrión muy precoz, previo a la implantación, v. web 10.1), cuya identidad genética depende, en exclusiva, del núcleo donante.
7. Se toman las células troncales pluripotentes de la masa celular interna del blastocisto (células troncales embrionarias) y se cultivan de modo adecuado para que se diferencien y generen células de uno u otro tejido: neuronal, muscular, óseo, cardíaco, etc. Una vez diferenciadas, las células pueden trasplantarse al paciente que donó el núcleo de la célula inicial (podrían también, en el caso de un individuo sano, guardarse como reserva para una eventual enfermedad).





13.2.2 Características de la clonación de células

Una vez descrito someramente el proceso de clonación, conviene detenerse a señalar algunas características esenciales, que ayuden a comprender su naturaleza, así como similitudes y diferencias con los procesos biológicos de desarrollo embrionario.

En primer lugar, se trata de un procedimiento completamente artificial para la duplicación de una célula precursora, generando una población de células genéticamente idénticas a aquélla y entre sí (un clon). Esto también ocurre, pero de forma natural, en el proceso de renovación de las células de cualquier tejido en el organismo.

Respecto al origen, puede partirse de células de cualquier tipo, que se encuentren bien libres (por ejemplo, una suspensión celular o un cultivo), bien circulantes en el organismo (como las células sanguíneas), o bien agrupadas formando tejidos sólidos u órganos. En este último caso habrá que proceder, obviamente, a su disgregación previa (pág. 124).

El proceso de clonación tiene lugar *in vitro*, tanto en lo referente a la preparación del nucléolo como a su desarrollo posterior hasta embrión temprano. El nucléolo posee cierta semejanza con un cigoto, puesto que puede dar lugar por división a otras muchas células (hipotéticamente, a un organismo completo). Sin embargo, debe destacarse que el nucléolo no comparte el genoma de dos gametos, sino que contiene los $2n$

cromosomas del núcleo somático y el citoplasma del oocito. El cigoto, por el contrario, posee n cromosomas del óvulo, n cromosomas del espermatozoide y, al menos, parte del citoplasma de ambos.

Otro aspecto que debe resaltarse es que la clonación de células se inicia con la división por mitosis del nuclóvulo, en un medio de cultivo, lo que da lugar a un cúmulo creciente de células, un clon celular –en el que todas las células son iguales genéticamente entre sí y al nuclóvulo– que sería análogo a un embrión temprano. No se conoce cómo puede producirse la diferenciación celular en este proceso de desarrollo embrionario, pero se admite que sea análoga a la embriogénesis normal en el ambiente materno. El problema que aquí se plantea deriva, precisamente, del proceso de desdiferenciación necesario para que el núcleo de una célula somática pueda dar lugar a una célula similar a las embrionarias. Se requiere una total reprogramación, muy precisa, de toda la información contenida en dicho núcleo.

13.2.3 Aplicaciones terapéuticas de las células obtenidas por clonación

El principal beneficio de la clonación, que justifica su creciente interés, técnico y práctico, es la posibilidad de aplicación terapéutica. Lo que se pretende es la obtención de células troncales embrionarias, pluripotentes, indiferenciadas, extraídas de la masa interna del blastocisto (embrión clonado) y que puedan luego generar, empleando condiciones de cultivo adecuadas, células de diferentes tejidos: corazón, hígado, piel, tejido nervioso, etc. Por ejemplo, la clonación de células sanas de un paciente permitiría obtener células madre embrionarias con capacidad para diferenciarse a células sanas del tejido afectado por la enfermedad. En realidad, la aplicación es idéntica a la descrita para células troncales naturales (no procedentes de clonación). La gran ventaja radica en que al ser las células clonadas inmunológicamente idénticas a las células donantes del paciente se evitaría el rechazo en el tratamiento celular o tisular de enfermedades del propio individuo donante.

Se han desarrollado ya aplicaciones, aunque aún no de forma generalizada, de utilización de células troncales para la regeneración de tejidos dañados. Con este fin se utilizan sobre todo las células madre de tejidos de adultos. Un ejemplo representativo es el cultivo directo de células troncales de la epidermis para producir cantidades de piel suficiente para remediar grandes quemaduras. La terapia con células madre del propio paciente podría también resolver la pérdida de funcionalidad de un órgano, hoy día dependiente de la donación de órganos o tejidos de otra persona, con la dificultad de disponibilidad y la problemática derivada de las reacciones de rechazo.

13.3 CÉLULAS TRONCALES O CÉLULAS MADRE

Como ha quedado reflejado en el apartado anterior, la clonación de células pasa por la obtención de células troncales (traducción más fidedigna de la expresión inglesa *stem cells*, aunque está más extendida la expresión alternativa, células madre). Con este nombre se hace referencia a aquellas células indiferenciadas, poco especializadas, que al dividirse son capaces de diferenciarse dando lugar a todos los tipos de células especializadas que forman parte del conjunto de tejidos y órganos de un organismo (estimado, en total, en unos 200 tipos celulares en el ser humano).

Para comprender el origen, las características y la potencialidad de las células madre es preciso conocer el proceso natural de embriogénesis (véase web 10.1). La descripción completa de éste excede los propósitos de esta obra, por lo que sólo se presentan las líneas generales imprescindibles para comprender las diferencias y el potencial de los distintos tipos de células madre. El lector interesado en profundizar deberá acudir a un texto más específico.

Aunque las células madre embrionarias se conocen desde hace más de 40 años, fue a raíz de su obtención por primera vez a partir de embriones de ratón (años 1980) y, sobre todo, desde que se consiguió aislarlas de muestras humanas y cultivarlas (año 1998), cuando se produjeron los mayores avances. La obtención de células madre vive en la última década una etapa brillante de expansión, y se ha convertido en noticia científica casi diaria, de gran impacto en la información general y en la formación profesional. De forma simultánea, este tema es objeto de un candente debate bioético. Existen opiniones de todo tipo, desde aquellas para las que es imperativo buscar opciones que no conlleven destrucción, a las que, en el otro extremo, justifican cualquier medio ante el imparable fin investigador del científico. Una posición intermedia, posiblemente mayoritaria, sería la de los que aceptan el uso de los embriones congelados, sobrantes de procesos de

fertilización en clínicas de reproducción asistida. Debe recordarse que estos embriones, en un número elevado (estimado entre 35.000 y 40.000 en España), siguen pendientes de uno de estos tres destinos finales: su destrucción transcurrido el plazo legal de utilización (5 años), su uso para investigación o su desarrollo por parejas adoptivas, una posibilidad dudosa.

13.3.1 Fuentes de obtención de células troncales

Para una comprensión correcta de las implicaciones, tanto técnicas como éticas, del trabajo con células troncales deben distinguirse, claramente, tres fuentes distintas para su obtención.

A) Células madre de embriones tempranos:

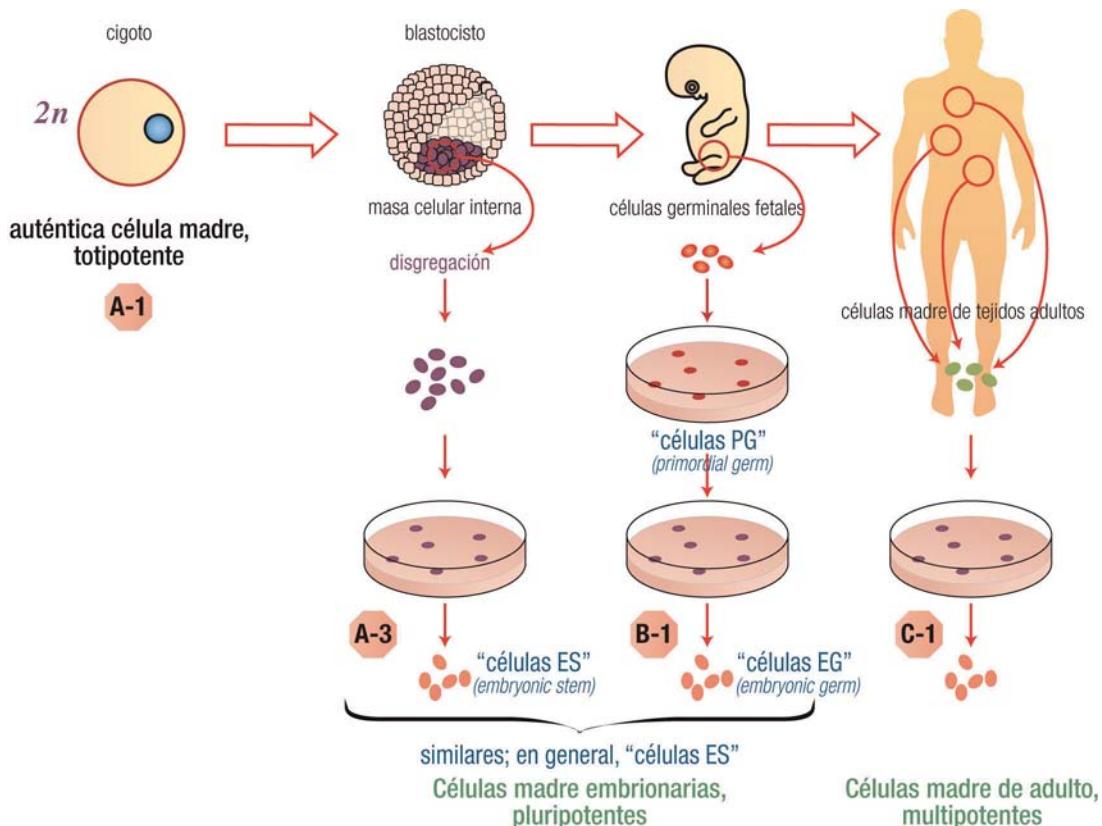
- A-1) El propio cigoto fisiológico, resultante de la fecundación (natural o asistida) del óvulo por el espermatozoide.
- A-2) El nuclóvulo, formado en laboratorio mediante transferencia del núcleo de células del propio paciente a un oocito enucleado (clonación de células).
- A-3) Células procedentes de los embriones desarrollados de forma natural (en útero) a partir de un cigoto.
- A-4) Células procedentes de embriones desarrollados *in vitro* a partir del cigoto formado por reproducción asistida.
- A-5) Células procedentes de embriones desarrollados *in vitro* a partir de un nuclóvulo.

B) Células madre germinales de fetos:

- B-1) De fetos normales animales, incluso humanos, procedentes de abortos.
- B-2) De fetos clónicos por haberse desarrollado a partir de un nuclóvulo.

C) Células madre de adulto. Obtenidas por desdiferenciación de células especializadas de tejidos de individuos adultos, jóvenes o niños:

- C-1) De tejidos normales de cualquier procedencia anatómica.
- C-2) De tejidos de animales previamente obtenidos por clonación.

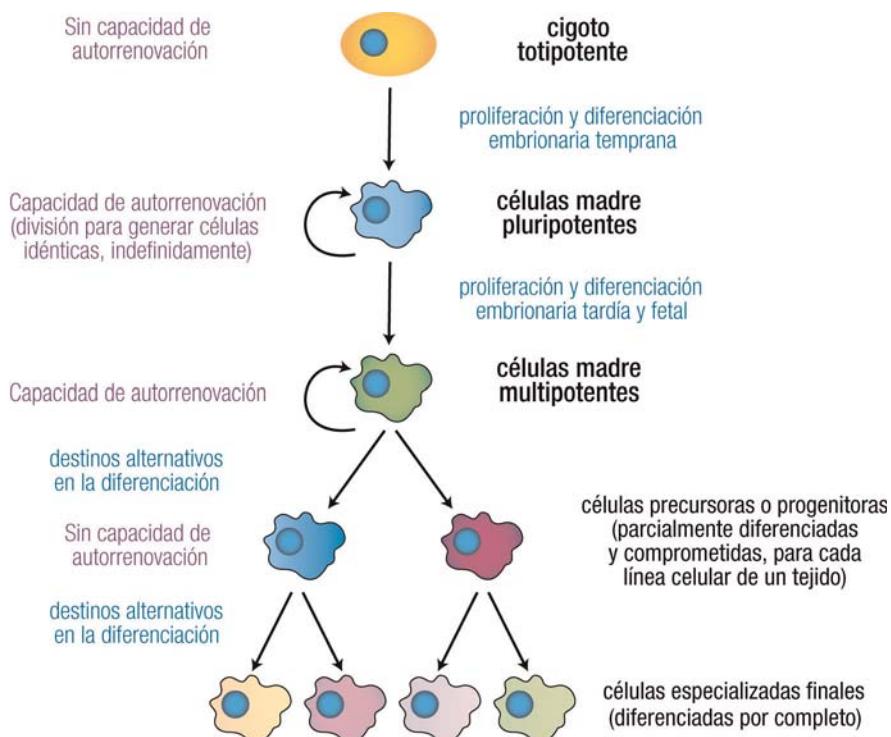


Esta variada procedencia (real o hipotética), junto con la relación de cada tipo de célula madre con el proceso en que se encuentra implicada de forma directa (fecundación, embriogénesis, formación del feto, desdiferenciación celular), constituyen parámetros, variables, que pueden determinar las innumerables posibilidades de aplicación terapéutica, descritas o aún incluso pendientes de proponerse.

13.3.2 Capacidad de diferenciación de las células

13.3.2.1. Totipotencia

El **cigoto**, formado por fecundación natural, asistida o *in vitro*, puede considerarse un embrión unicelular, cuya característica principal es la **totipotencia**. Puede definirse así como la única célula normal con capacidad de formar, mediante procesos de división, especialización y diferenciación, todos los tipos de células pluripotentes, multipotentes, precursoras y progenitoras, a partir de las cuales se originan después los distintos tipos de células especializadas del nuevo individuo. Un **nuclóvulo** obtenido por clonación también tendrá esta misma capacidad, contando con que su fusión y su estimulación posterior se consigan con éxito.



13.6

Otro aspecto molecular importante en la aparición del cigoto es el fenómeno de la **reprogramación** del genoma, que empieza a tener lugar en la fecundación. Como consecuencia, la dotación genética del cigoto o embrión unicelular es algo más que la simple suma de los materiales genéticos paterno y materno que lo originaron. El genoma del cigoto se mantiene sin variación en todas las células del organismo a las que da lugar, pero el desarrollo del organismo, el proceso de diferenciación celular individual, se debe a una selección ordenada de los genes que se expresan, mientras otros permanecen silenciados. Existe una red de señales genéticas que controla las decisiones que deben tomarse en cada momento del desarrollo. Estos mecanismos (una especie de “memoria molecular”) aparecen en el cigoto y se mantienen en todas las células derivadas; conocidos como procesos **epigenéticos**, implican modificaciones reversibles del DNA (nunca la alteración de su secuencia), fundamentalmente la metilación de los islotes CpG (citosina-guanina) (págs. 282 y 286).

13.3.2.2. Pluripotencia

Las células tempranas formadas a partir del cigoto, constituyentes de los embriones de 2, 4, 8, 16... células (v. web 10.1), se consideraban inicialmente idénticas entre sí y al cigoto inicial. Sin embargo, se han encontrado pruebas de que desde la primera división del cigoto las células comienzan a diferenciarse, diferencia que aumenta a medida que progresá la división del embrión temprano. Esto significa que las células troncales embrionarias, que se obtienen en cualquier momento de este desarrollo temprano, serán pluripotentes pero nunca totipotentes.

Las células troncales embrionarias (*células ES*, de *embryonic stem*) son, como ya se ha indicado, el conjunto de células que forman la masa celular interna del blastocisto, o embrión de cinco días. Por tanto, se obtienen a partir de los embriones tempranos formados en cultivo a partir de un cigoto o un nucléolo. Estas células troncales pueden diferenciarse para dar todo tipo de células (excluidos la placenta y otros tejidos extraembrionarios) y son, por tanto, pluripotentes. Aunque no servirán para generar un organismo completo (pues el hipotético embrión no se podría implantar), pueden utilizarse para formar cualquiera de sus tejidos y, por tanto, ofrecen potencial terapéutico. Por el contrario, las células del trofoblasto –la capa externa del blastocisto– se diferencian dando lugar a tejidos extraembrionarios, como la placenta y otros tejidos implicados en el desarrollo del feto, de modo que no pueden originar un embrión completo y no son células pluripotentes en sentido estricto.

Otra fuente de células troncales son las células germinales embrionarias (*células EG*, de *embryonic germ*), que aparecen en el embrión ya en el estado de gástrula y que son las precursoras de los futuros gametos. A diferencia de las células madre embrionarias ES, pueden obtenerse del feto o, en todo caso, de embriones ya muy desarrollados. Tanto la pluripotencia como la capacidad de dividirse de manera indefinida en cultivo son comunes a ES y EG, por lo que no se suele distinguir entre ambas, e incluso se denominan ES de forma global, a pesar de su distinto origen.

Se han establecido líneas celulares estables de células madre embrionarias, obtenidas mediante transferencia de núcleos de células somáticas, que están disponibles comercialmente. La normativa a este respecto es diversa dependiendo del país. A modo de ejemplo, en los Estados Unidos el registro oficial del NIH comprende todas las cepas de células ES aprobadas para su empleo en investigaciones con fondos públicos estadounidenses, en número variable según se consideren o no las protegidas por patente. Con esta colección se pretende evitar la destrucción de embriones y competir con las investigaciones que ya se vienen realizando con fondos privados.

13.3.2.3. Multipotencia

En principio, no se creía que las células madre con capacidades de multiplicación independiente y de diferenciación hacia un determinado tejido pudieran proceder de células no embrionarias. Más tarde se encontró que esto era posible a partir de células madre de adulto, presentes en pequeña cantidad en la mayoría de los tejidos del cuerpo. Como consecuencia, surgieron planteamientos de terapia celular, alternativos a los que implican la destrucción de embriones necesaria para obtener células ES. En todo caso, un inconveniente general es su escasa abundancia en los tejidos, que supone una dificultad para obtenerlas en cantidad suficiente.

Las células troncales de adulto se caracterizan por ser **multipotentes**, a diferencia de las células embrionarias pluripotentes. Aunque están especializadas, retienen un gran potencial de maduración y diferenciación, si bien inferior al de las células madre embrionarias. En particular, se están consiguiendo notables avances en cuanto al modo de inducir “desdiferenciación” para que estas células madre puedan dar lugar a diversos tipos celulares específicos, diferentes de aquellos propios del tejido en el que se encuentran.



Clonación acelular: reacción en cadena de la polimerasa (PCR)

14.1 PRINCIPIO DEL MÉTODO	202	14.5.3 PCR con "comienzo en caliente"	208
14.2 ETAPAS DEL PROCESO	203	14.5.4 PCR "larga"	208
14.3 CARACTERÍSTICAS	204	14.5.5 PCR "anidada"	208
14.4 AUTOMATIZACIÓN	205	14.5.6 PCR inversa	208
14.5 VARIANTES DEL MÉTODO	205	14.5.7 PCR con adaptadores	209
14.5.1 PCR en tiempo real	205	14.5.8 PCR asimétrica	209
14.5.2 RT-PCR: amplificación de RNA	206		

Se comienza en este capítulo la descripción de la **clonación molecular**, o clonación de moléculas de ácido nucleico, con la técnica que permite amplificar el DNA (o RNA) sin emplear células. Por ello, se trata de una **clonación acelular** de acuerdo con la clasificación establecida en el capítulo anterior.

La PCR (acrónimo de *polymerase chain reaction*, reacción en cadena de la polimerasa) es hoy día una herramienta imprescindible en los laboratorios de biología molecular e ingeniería genética. El objetivo de esta técnica es la amplificación directa de un gen o un fragmento de DNA, presente en muestras muy diversas, sin necesidad de una purificación previa. Cuando el objetivo es un RNA, puede aplicarse la PCR de forma indirecta, si previamente se sintetiza su DNA complementario. Se pueden emplear como muestra para hacer PCR homogenados, extractos crudos de tejido, sangre completa, mezclas de fragmentos de DNA obtenidos con enzimas de restricción, muestras resultantes de la extracción y aislamiento de DNA, etc. Sin embargo, debe resaltarse que, por razones que se comprenderán más adelante, es un requisito imprescindible que **se conozca la secuencia de una parte de la región de DNA o RNA que se quiere amplificar** (salvo en el caso de algunas variantes como la PCR inversa, pág. 208).

Por estas características, en especial la capacidad de amplificación, la PCR es un método idóneo para preparar ácidos nucleicos en una cantidad muy superior a la presente en la muestra original, tanto con fines de clonación acelular como para la detección.

Objetivo básico de la PCR:
amplificar DNA

AMPLIFICACIÓN propiamente dicha

Para disponer de cantidad suficiente como para utilizarlo, con fines diversos

DETECCIÓN

Para poderlo detectar en muestras con pequeñas cantidades del DNA diana

14.1

Las **aplicaciones** de la PCR son muy numerosas y variadas. Como ejemplos, pueden citarse: clonación acelular de fragmentos de DNA, detección de secuencias sin purificación previa, preparación de muestras para secuenciación de ácidos nucleicos, establecimiento de polimorfismos de secuencia, rastreo de mutaciones, tipificación de DNA para trasplantes, diagnóstico de enfermedades genéticas (incluido el diagnóstico prenatal), determinación de secuencias específicas de DNA relacionadas con situaciones patológicas definidas, resolución de problemas forenses o arqueológicos, estudios evolutivos, detección de microorganismos infecciosos, detección de células tumorales, amplificación de DNA para su posterior clonación celular, etc.

14.1 PRINCIPIO DEL MÉTODO

La PCR no es una técnica analítica *per se*, sino más bien una metodología, resultante de la aplicación práctica de tres conceptos, ya descritos (págs. 164, 168 y 145, respectivamente):

1. Desnaturalización del DNA para dar moléculas monocatenarias.
2. Hibridación específica de la molécula monocatenaria con un oligonucleótido. Ésta es la etapa que justifica por qué se debe conocer parte de la secuencia.
3. Replicación de la molécula monocatenaria mediante una DNA polimerasa que emplea el oligonucleótido anterior como cebador: también se puede denominar elongación o extensión del cebador, o polimerización.

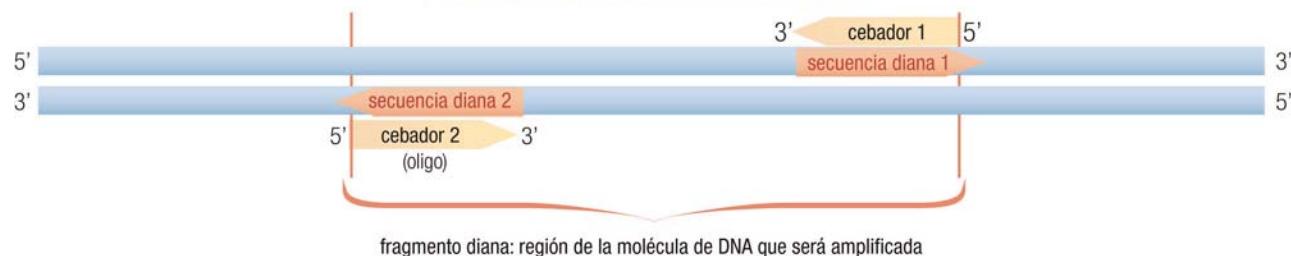
Mediante la aplicación en forma cíclica de estos tres procesos se consigue un número muy alto de copias del fragmento de ácido nucleico en un corto espacio de tiempo, como se verá a continuación.

A pesar de esta aparente complejidad, en teoría es un método sencillo, sensible y relativamente rápido para amplificar secuencias. En la práctica, se requiere un control preciso de las variables que condicionan este triple proceso (secuencia diana, cebadores, enzimas y resto de componentes), además de instrumentos adecuados (*termocicladores*) para establecer con precisión las condiciones de cada etapa y repetirlas cíclicamente (tiempo, temperatura, número de ciclos, etc.). Se consigue de esta forma amplificar secuencias de tamaños diversos, comprendidos entre 50 pb y 2,5 kb.

Para la realización práctica se precisan en la mezcla de reacción los siguientes "reactivos":

- a) Los cuatro dNTP, en exceso, como sustratos para la síntesis de las innumerables copias de DNA. Para ser reconocidos por la polimerasa, deben ir acompañados de Mg²⁺ que es, además, una coenzima requerida por la polimerasa.
- b) Dos oligonucleótidos monocatenarios (coloquialmente, "oligos"), por lo común sintéticos, de 18 a 30 nt. Sus secuencias han de ser complementarias, respectivamente, a los dos extremos 3' de la región diana, uno en cada hebra, de modo que los oligos puedan ejercer de cebadores para la replicación de las dos hebras en la región diana. Por esta razón no puede amplificarse una región de DNA si no se conoce la secuencia de sus dos extremos. De hecho, es la posición donde hibridan los cebadores la que define cuál es el fragmento que se amplifica (secuencia diana).

REGIÓN DE DNA DE LONGITUD VARIABLE



14.2

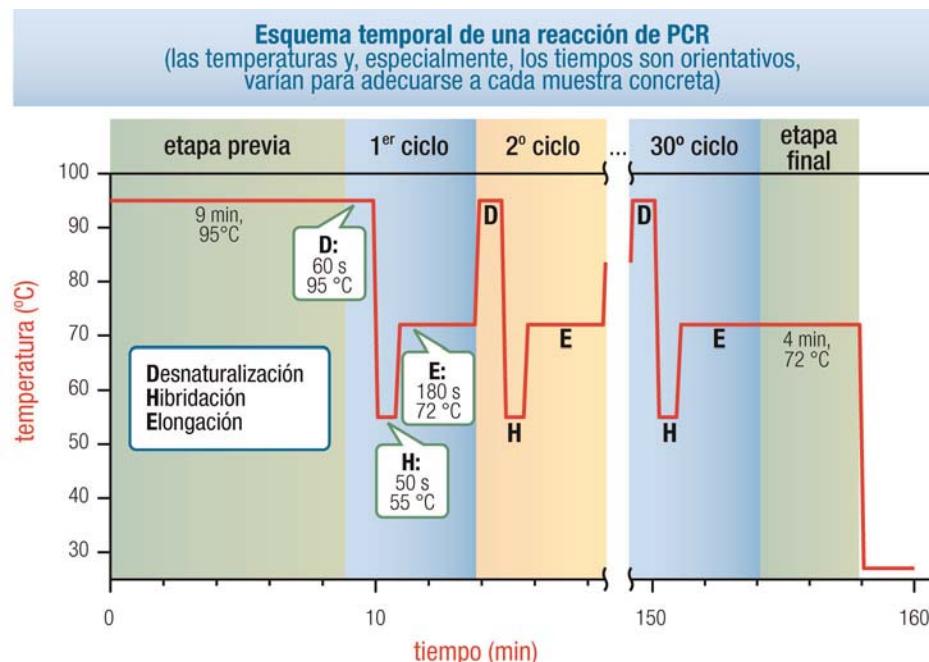
- c) Una DNA polimerasa *termoestable*, es decir, enzimáticamente activa a temperaturas relativamente altas (habitualmente en torno a 75 °C). La replicación a temperaturas elevadas impide la formación de híbridos parcialmente desemparejados y contribuye así a la especificidad y rendimiento del proceso. Por otra parte, la estabilidad de la polimerasa a temperaturas de hasta 95 °C (aunque no mantenga su actividad, no se desnaturiza) permite que recupere su actividad al enfriarse de nuevo y evita así la necesidad de reponer la enzima en sucesivos ciclos. La enzima más empleada se denomina *polimerasa Taq*, por proceder de la bacteria *Thermus aquaticus*, que vive en manantiales de agua caliente. Esta enzima tiene el inconveniente de carecer de la actividad correctora de pruebas (exonucleasa 3') (pág. 150), por lo que la frecuencia de errores es superior a la propia de la replicación (alrededor de un error por cada 5.000 nucleótidos incorporados); a pesar de ello, es suficiente para los propósitos de la técnica. Hoy día existen otras DNA polimerasas termoestables que pueden emplearse como alternativa.

14.2 ETAPAS DEL PROCESO

Como ya se ha indicado, se requiere una sucesión de ciclos (en general entre 20 y 40, de 1,5 a 5 minutos de duración cada uno) formados por desnaturización, hibridación y replicación, para conseguir una enorme amplificación del número de moléculas que contienen la secuencia de interés o diana. La duración total es cercana a 2 horas, dependiendo de las condiciones experimentales concretas.

Cada ciclo de una PCR consta, pues, de tres etapas:

- 1. Desnaturalización:** calentamiento para la separación de las dos hebras del DNA, mediante una incubación breve (30-120 s) a una temperatura comprendida entre 68 y 97 °C, que debe ser superior a la de fusión (T_m , pág. 167) de la región de DNA que quiere amplificarse.
- 2. Hibridación:** enfriamiento rápido por debajo de T_m , de forma que se permite el emparejamiento (pág. 168) del DNA monocatenario de interés con los oligos cebadores. En general se usan temperaturas comprendidas entre 37 y 65 °C que se mantienen entre 10 y 120 s.
- 3. Elongación (o replicación):** etapa de amplificación propiamente dicha (72-75 °C, 1-3 min), en la que la DNA polimerasa termoestable elonga los cebadores, empleando como molde ambas hebras originales. La replicación transcurre (pág. 147) en dirección 5' → 3' a partir del extremo 3'-OH de cada cebador, empleando como sustrato los cuatro dNTP, hasta terminar la lectura del molde o hasta que se eleve la temperatura en una nueva etapa de desnaturización (siguiente ciclo).



14.3

Además de las 3 etapas de cada ciclo, suelen añadirse una etapa previa y una final al conjunto de todos los ciclos. La previa, a elevada temperatura (incluso superior a la de las etapas de desnaturización), sirve principalmente para inactivar proteasas y nucleasas de la muestra, así como para asegurar la desnaturización completa del DNA de partida, en especial si éste tiene gran tamaño o posee regiones muy compactadas (caso de un DNA genómico completo). La etapa final, por su parte, consiste en una prolongación de la última elongación, para permitir que se completen todos los fragmentos.



Web 14.1. Mecanismo de la PCR: reacciones y resultados en los 4 primeros ciclos.

Debe destacarse que si la muestra de partida, como ocurre en el ejemplo descrito, está formada por moléculas de DNA de mayor tamaño que la región diana sólo aparecen copias exactas de la diana a partir del tercer ciclo. Si la muestra de partida hubiese sido el propio fragmento diana, todas las copias aparecidas desde el primer ciclo serían iguales a él.

Además de la amplificación global, por aumento del número de moléculas de DNA, debe notarse cómo el número de copias de la diana se incrementa mucho más rápidamente que el de otras copias más largas. Todo ello hace que en la mezcla final, al terminar la PCR, los fragmentos diana superen abrumadoramente en número a cualquier otro tipo de fragmento, incluyendo, obviamente, las moléculas de DNA originales de la muestra (que no se han amplificado). En la siguiente tabla se observa el enorme grado de amplificación conseguido, de un millón de veces en 20 ciclos y un billón en 40 ciclos. Esto es lo que permite aplicar la PCR a muestras sin necesidad de purificar su DNA, y analizar de manera directa los resultados, en los que sólo se podrá detectar el fragmento amplificado.

N.º de ciclos realizados	N.º total de copias	N.º de copias "largas"	N.º de copias diana	Dianas/total
1	2	2	0	0
2	4	4	0	0
3	8	6	2	25%
4	16	8	8	50%
5	32	10	22	69%
10	1.024	20	1.004	98%
20	$\approx 10^6$	40	$\approx 10^6$	99,996%
40	$\approx 10^{12}$	80	$\approx 10^{12}$	100%
x	2^x	$2x$	$2^x - 2x$	

14.3 CARACTERÍSTICAS

A continuación se comentan algunas características del proceso de PCR.

Rendimiento: puesto que los productos de cada ciclo sirven de molde para el siguiente, la acumulación de copias (tanto totales como dianas) es exponencial en vez de lineal. Sin embargo, en la práctica el rendimiento de cada etapa no es completo, por lo que las cifras se ven algo reducidas. A pesar de ello, siguen siendo muy elevadas. El número de copias al final de los ciclos realizados puede calcularse como:

$$N = N_0 (1 + R)^X$$

siendo:

N = número de copias obtenidas

N₀ = número de moléculas iniciales

R = rendimiento de cada ciclo (generalmente, de 70 a 85%), expresado en tanto por uno (0,70-0,85)

X = número de ciclos

Se han desarrollado modificaciones del método descrito, encaminadas a aumentar la eficiencia, sobre todo en lo relativo a la concentración de Mg²⁺ (habitualmente entre 1 y 4 mM) y al tiempo y temperatura de hibridación.

Duración: la duración de cada una de las tres etapas de cada ciclo y el número de ciclos deben optimizarse, dependiendo de la secuencia concreta que quiera amplificarse. En muchos casos la elección de estas condiciones debe hacerse de forma empírica.

Especificidad: es muy elevada. Está determinada especialmente por la secuencia de los dos cebadores utilizados y por las condiciones de hibridación. Se puede determinar la especificidad analizando en electroforesis el producto de la PCR: la aparición de una banda única indica que la técnica ha actuado con buena especificidad. Por ejemplo, si la temperatura de hibridación se reduce en exceso se producen falsos positivos debidos, fundamentalmente, a emparejamientos imperfectos del cebador con secuencias distintas de la diana, aunque similares, que resultan así también amplificadas.

Capacidad de detección (común aunque incorrectamente denominada sensibilidad): es muy alta (bajo límite de detección), tanto que la PCR puede permitir detectar una *única* molécula de DNA en casi cualquier tipo de muestra clínica, forense o arqueológica (pelo, sangre, semen, etc.). Sin embargo, esta característica supone también un elevado riesgo de contaminación por moléculas de origen ajeno a la muestra.

Fidelidad: es decir, la capacidad de copiar secuencias con precisión, sin introducir mutaciones (cambios en la secuencia de nucleótidos). Esta característica depende sobre todo de que la polimerasa empleada tenga o no

actividad correctora de pruebas (3'-exonucleasa, pág. 150). Así, la frecuencia de error oscila entre $2 \cdot 10^{-4}$ de una polimerasa *Taq* convencional y $1 \cdot 10^{-6}$ de la polimerasa *Pfu*, que sí tiene actividad 3'-exonucleasa. En todo caso, la elección de la enzima dependerá del objetivo de la PCR. No es crítica la fidelidad si sólo se pretende obtener una sonda, pero es fundamental si el objetivo es la identificación de mutaciones poco frecuentes.

14.4 AUTOMATIZACIÓN

La automatización de la PCR es sencilla y de bajo coste, dado el carácter cíclico repetitivo del proceso, el empleo de componentes estables al calor, la necesidad de añadir reactivos únicamente al comienzo y el desarrollo de sistemas electrónicos para programar los cambios de temperatura. Al mismo tiempo es imprescindible, pues sólo mediante la automatización se hace factible la realización de un número elevado de ciclos y se aumenta la precisión del proceso. Ello facilita un empleo cada vez más específico y sensible en el laboratorio clínico. Los instrumentos diseñados para el desarrollo de reacciones de PCR se denominan **termocicladores** y suelen admitir varias decenas de muestras simultáneas, con volúmenes comprendidos entre 25 y 100 µl.

14.5 VARIANTES DEL MÉTODO

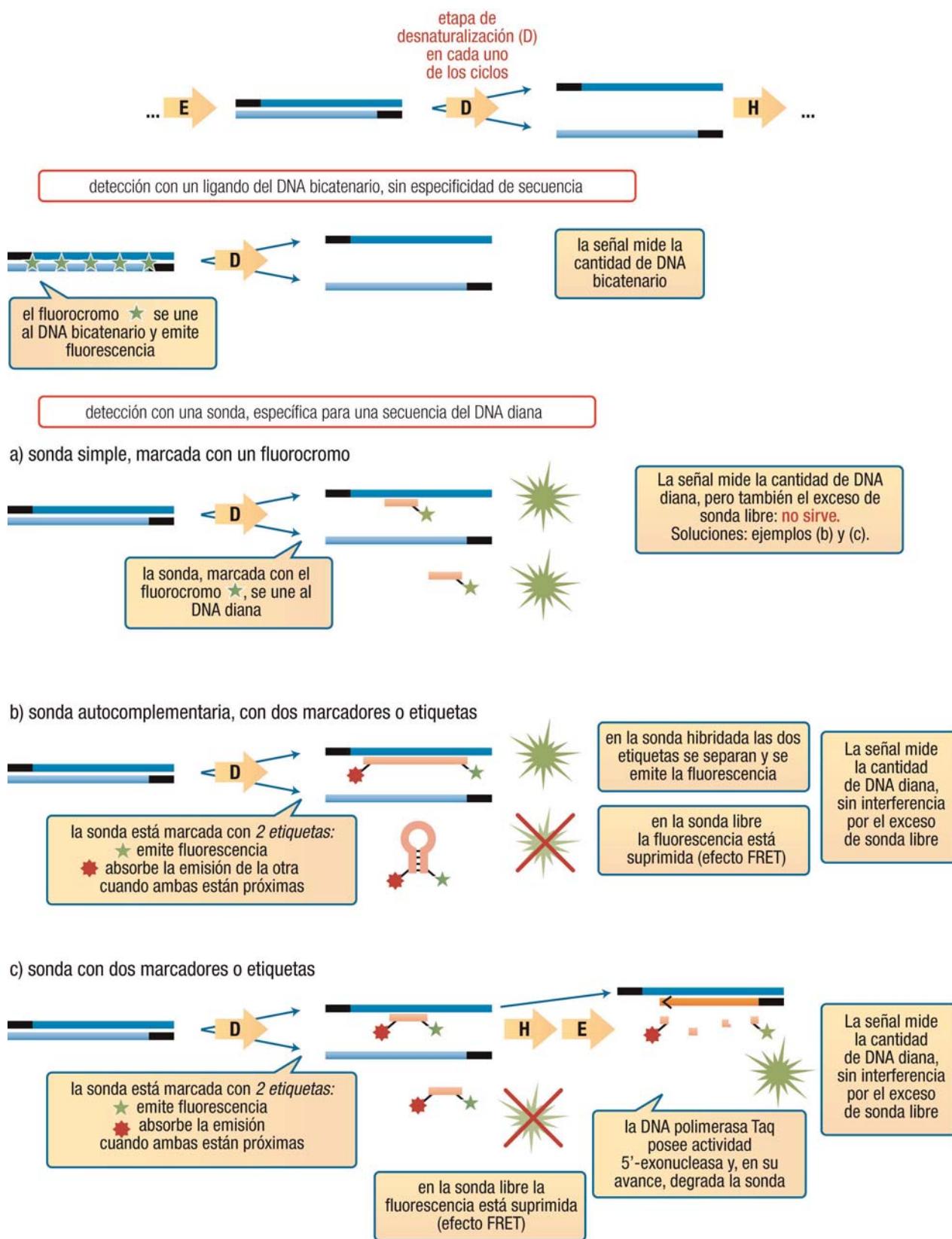
Han surgido numerosas modificaciones derivadas del método básico inicial, con el propósito de mejorar el rendimiento o la especificidad, adaptarse a muestras particulares, amplificar moléculas de RNA en lugar de DNA, producir moléculas monocatenarias, etc. A continuación, se comentan de forma breve algunas de las variantes más significativas.

14.5.1 PCR en tiempo real

La PCR, en su diseño original, es una técnica de punto final, es decir, analiza el producto obtenido tras finalizar la reacción. Este planteamiento resulta ser poco cuantitativo: muestras con diferente cantidad de DNA inicial pueden producir resultados similares tras la amplificación, y viceversa. Por supuesto, esto no es un problema para algunas aplicaciones, pero existen casos en los que es importante analizar múltiples muestras de forma comparada (notablemente, la expresión de genes bajo diferentes situaciones). Debido a esta necesidad, se ha desarrollado la técnica de **PCR cuantitativa** (qPCR) o **PCR en tiempo real**, que mide la velocidad a la que se va amplificando el DNA. (En ocasiones, se usa RT-PCR para indicar *real time PCR*, pero este acrónimo es poco adecuado, pues se confunde con la PCR con transcriptasa inversa, descrita a continuación.)

Las razones por las que la PCR tradicional no es cuantitativa comprenden, a grandes rasgos, la baja detectabilidad y la poca proporcionalidad de la señal propias de los métodos empleados para revelar el DNA (tradicionalmente, la tinción de los geles de electroforesis con etidio, un intercalante, pág. 141) y la tendencia de la reacción a saturarse en los ciclos finales (por agotamiento de los sustratos, inhibición por los productos finales y degradación de éstos, entre otros factores). El planteamiento cinético de la qPCR consiste en medir no la cantidad de producto formado al final, sino la velocidad a la que se está formando en momentos moderadamente tempranos de la reacción. En concreto, se observa la fase en la que la reacción es aún exponencial –por tanto, mantiene la proporcionalidad respecto a la cantidad inicial de DNA– y se cuantifica el número de ciclos que han sido necesarios para alcanzar una cantidad prefijada de DNA producto. Cuanto mayor sea la abundancia en la muestra, antes se alcanzará la cantidad prefijada.

La cuantificación depende de un compuesto que se une al DNA bícatenario, producto de la reacción de amplificación. Habitualmente se utilizan fluorocromos (bien intercalantes o bien ligandos en el surco menor, con mejor detectabilidad que el etidio) que se unen muy preferentemente al DNA bícatenario e incrementan de forma notable su fluorescencia. De forma automatizada, prácticamente en continuo, se mide la intensidad de fluorescencia en cada muestra al finalizar cada ciclo de amplificación, como índice del número de copias conseguidas. Obviamente, esta determinación no es demasiado selectiva, pues es independiente de la secuencia del DNA y no permite diferenciar si ese DNA bícatenario corresponde efectivamente a la amplificación selectiva del DNA diana. Por ello, se ha desarrollado una segunda generación de métodos de detección basados en el uso de sondas de hibridación (Capítulo 12), dirigidas contra una secuencia interna del DNA diana. Tras la etapa de desnaturalización de cada ciclo la sonda hibrida con el DNA amplificado. Puesto que es preciso diferenciar la fluorescencia de la sonda unida al DNA diana de la que posee la sonda libre, se han diseñado diversas estrategias para asegurar que la sonda libre no muestre señal.



14.4

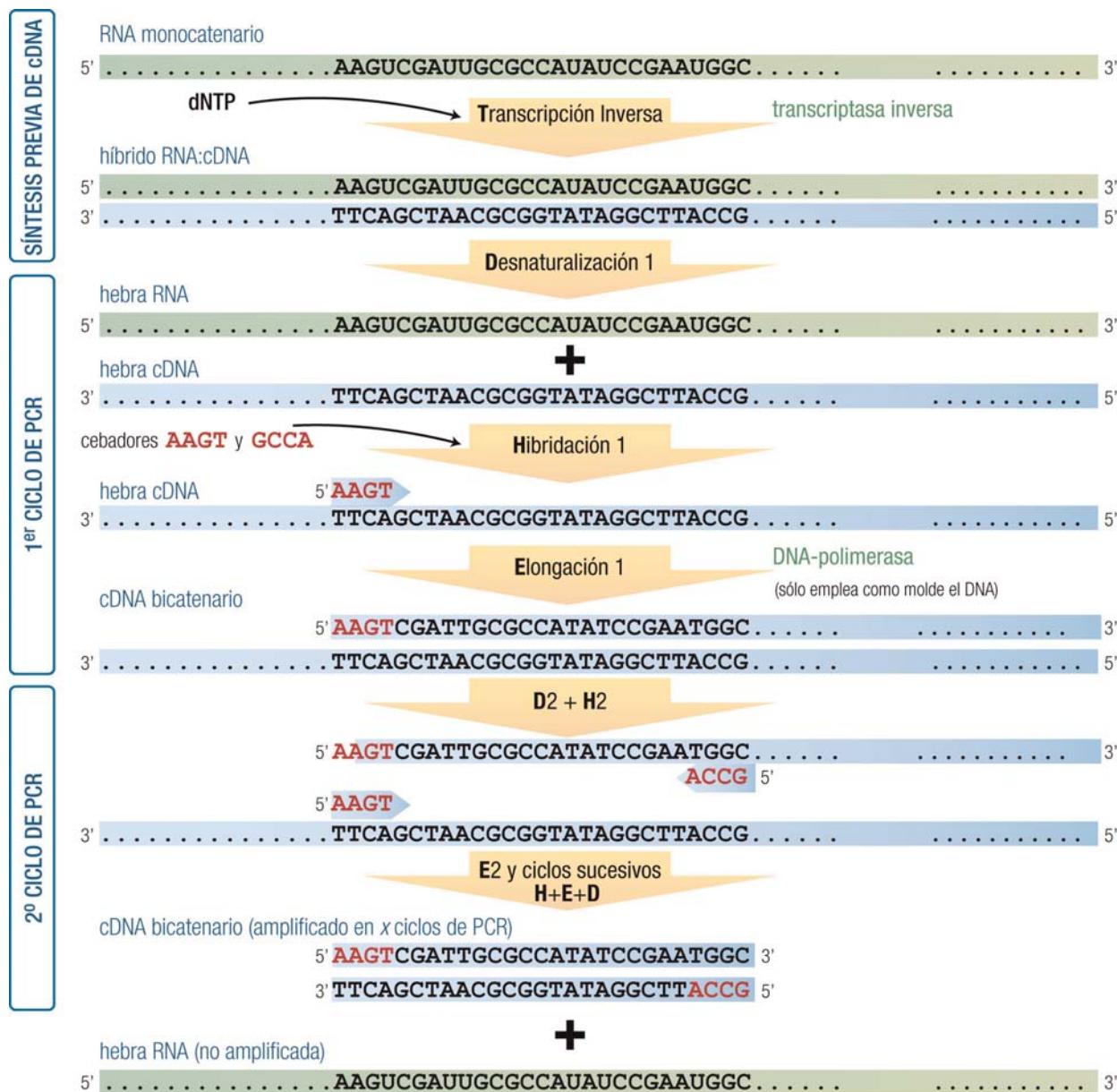
14.5.2 RT-PCR: amplificación de RNA

El nombre, "PCR con transcriptasa inversa" (iniciales de *reverse transcriptase*), indica que se trata de una amplificación de RNA (especialmente mensajero) a través de la síntesis previa de su cDNA (DNA

complementario al RNA), que después se amplifica por PCR. Es decir, no se obtienen copias del RNA de partida, sino de DNA, aunque se conserva, obviamente, la secuencia de aquél.

Es el método con mayor capacidad de detección de los disponibles para la medida de la expresión génica *in vitro*. Se utiliza preferentemente para amplificar mRNA de células asequibles con una expresión elevada de ciertos genes, aunque también se ha aplicado al análisis de transcriptos de genes expresados en grado mínimo.

La mezcla inicial contiene todos los componentes necesarios: muestra de RNA, transcriptasa inversa, DNA polimerasa, cebadores y dNTP. El proceso comienza (por ejemplo, 45 min a 48 °C) con la síntesis de una hebra de cDNA por la acción de la transcriptasa inversa, una polimerasa de DNA dirigida por RNA (pág. 149), permaneciendo el cDNA unido al molde como molécula bicatenaria híbrida RNA:cDNA. En una segunda etapa (por ejemplo, 2 min a 94 °C) se desnaturaliza la molécula bicatenaria y comienza la amplificación según el mecanismo de una PCR normal: la hebra de cDNA liberada actúa como molde para una segunda hebra de DNA y luego la molécula bicatenaria se amplifica en sucesivos ciclos.



En teoría, basta con una molécula de RNA que esté intacta entre los dos sitios de unión a los cebadores para conseguir la amplificación. Bajo el mismo principio, también puede amplificarse RNA usando la polimerasa Tth (también termoestable, aislada de la bacteria termófila *Thermus thermophilus*), que contiene al mismo tiempo actividades DNA polimerasa y transcriptasa inversa.

14.5.3 PCR con "comienzo en caliente"

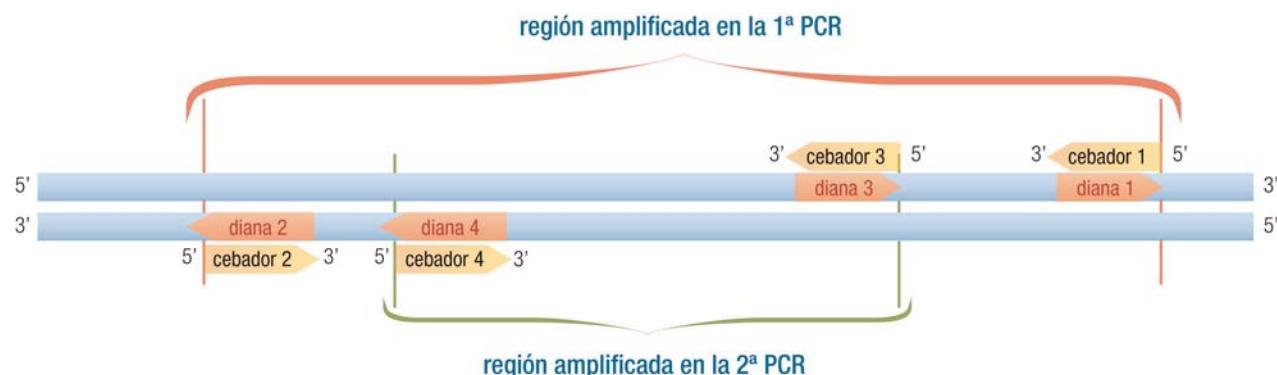
Esta modificación (*hot start PCR*) consiste en privar a la mezcla de reacción de alguno de sus componentes (frecuentemente la polimerasa) hasta que se haya alcanzado una temperatura superior a la de hibridación. De este modo, se evita la elongación de cebadores que hayan podido asociarse al comienzo con poco rigor (a secuencias parcialmente homólogas con la diana) y, como resultado, se aumenta de manera señalada la especificidad de la amplificación.

14.5.4 PCR "larga"

Denominada L-PCR (*Long PCR*) o LA-PCR (*Longer and Accurate PCR*, "PCR más larga y exacta"), su objetivo es superar los límites de la PCR convencional para amplificar con fidelidad regiones diana de gran tamaño (entre 5 y 40 kb). El principal factor limitante es la ausencia de actividad correctora de pruebas en la polimerasa *Taq* (pág. 202), por lo que se añade una cantidad menor de otra enzima con capacidad de corrección de pruebas (por ejemplo, la *Pfu*) para contrarrestar la carencia de esta actividad y, al mismo tiempo, seguir aprovechando la eficacia de elongación de la polimerasa principal (*Taq* o similar).

14.5.5 PCR "anidada"

La especificidad puede aumentarse realizando una segunda reacción de PCR, con dos cebadores nuevos que hibridan dentro del fragmento diana amplificado por el primer par, lo que da lugar a productos de PCR más cortos, pero más específicos. Las copias correctas de la primera diana contendrán ambas secuencias complementarias a los nuevos cebadores, a diferencia de los productos no específicos generados en la primera PCR (que se observan en la electroforesis como varias bandas o una banda difusa) que, por ello, no resultarán amplificados en la segunda. Este método también sirve para aumentar el factor de amplificación conseguido, al suponer dos rondas de amplificación.



14.5.6 PCR inversa

Esta variante se emplea para clonar regiones desconocidas de un DNA, situadas en posición vecinal a secuencias diana conocidas. Es decir, en lugar de amplificar la región interna, *flanqueada por* los dos cebadores (PCR convencional), se amplifica la región externa, *que flanquea a* los cebadores. Para ello es necesario cortar el DNA

a ambos lados de la región diana con una enzima de restricción (pág. 212), de tal forma que los extremos cohesivos resultantes puedan hibridar entre sí, formando una molécula circular. Ésta se cierra mediante una ligasa y se realiza la PCR con cebadores que hibridan con los **extremos 5'** de la secuencia conocida, por lo que la elongación se extenderá alrededor del círculo. Se generan copias de un DNA lineal delimitado, como en la PCR normal, por la posición de ambos cebadores.

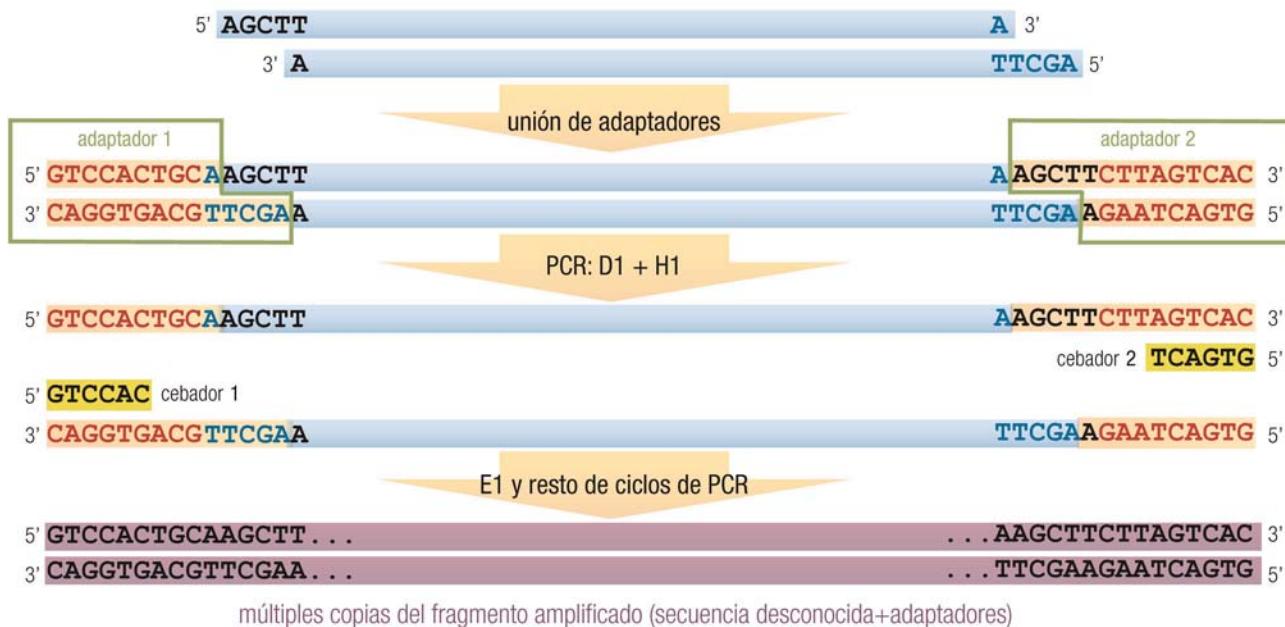


Web 14.3. Técnica de PCR inversa.

14.5.7 PCR con adaptadores

También puede amplificarse una región de DNA de secuencia desconocida ligando a sus fragmentos de restricción unos *adaptadores*, oligonucleótidos sintéticos con extremos cohesivos (pág. 215) compatibles con los generados en la muestra. A continuación, una vez desnaturizados, se añaden cebadores específicos para las secuencias 3' de los adaptadores. Se amplifica así el conjunto de adaptadores y secuencia diana, pero debe señalarse que se amplifican por igual todos los fragmentos de restricción presentes, no sólo uno.

región que se quiere amplificar, de secuencia desconocida



14.7

Web 14.4. Técnica de PCR con adaptadores.

14.5.8 PCR asimétrica

En este caso, se trata de generar **copias monocatenarias** de un DNA. En la variante más simple, se añaden cantidades muy diferentes de ambos cebadores, de modo que tras los primeros ciclos de PCR uno de ellos se agota y, disponibles ya suficientes copias del DNA diana, sólo una de sus hebras sigue amplificándose gracias al cebador más abundante.

Clonación celular: tecnología del DNA recombinante



15.1 MANIPULACIÓN EXPERIMENTAL DE MOLÉCULAS DE DNA	
15.1.1 Introducción a las nucleasas	211
15.1.2 Enzimas de restricción	212
15.1.2.1 Características generales	212
15.1.2.2 Papel biológico del sistema metilación-restricción	213
15.1.2.3 Acción de las enzimas de restricción de tipo II	215
15.1.3 Ligasas	218
15.2 CLONACIÓN CELULAR DE MOLÉCULAS DE DNA	219
15.2.1 Objetivos de la clonación de DNA	219
15.2.2 Descripción global del proceso de clonación	219
15.2.3 Preparación del DNA para su clonación	221
15.2.3.1 Insertos de DNA	221
15.2.3.2 Insertos de DNA complementario procedente de un mRNA	222
15.2.4 Preparación del vector de clonación	223
15.2.5 Formación de la molécula de DNA recombinante	225
15.2.6 Tipos de vectores	226
15.2.6.1 Plásmidos	227
15.2.6.2 Bacteriófagos	228
15.2.6.3 Cósidos	229
15.2.6.4 Cromosomas artificiales	230
15.2.7 Incorporación del DNA recombinante a la célula anfitriona	230
15.2.7.1 Tipos de célula anfitriona	230
15.2.7.2 Métodos de introducción del DNA recombinante	231
15.2.8 Propagación en cultivo	233
15.2.9 Detección y selección de los clones recombinantes	234
15.2.10 Expresión génica	236
15.3 GENOTECAS	236
15.3.1 Genotecas genómicas	237
15.3.2 Genotecas de DNA complementario	238

Continuando la descripción de la **clonación molecular**, o clonación de moléculas de ácido nucleico, se aborda en este capítulo la técnica que permite amplificar el DNA aprovechando la capacidad de replicación propia de las células. Por ello, se trata de una **clonación celular**, en el sentido ya definido previamente: clonación *con* células.

El método clásico de clonación de ácidos nucleicos es el que aprovecha la capacidad de las células para replicar el DNA. Surge gracias a la aparición de las técnicas que permiten el aislamiento del DNA y al descubrimiento en la década de 1970 de las enzimas de restricción, que hacen posible, en el laboratorio, cortar el DNA de modo controlado e incorporar los fragmentos a *vectores*, que, a su vez, se introducen en *células anfitrionas*, en las que tiene lugar su replicación. Aunque, en principio, esta tecnología sólo se planteaba para clonar el DNA, existen hoy día variantes que extienden su aplicación al RNA.

15.1 MANIPULACIÓN EXPERIMENTAL DE MOLÉCULAS DE DNA

Antes de describir el proceso de clonación del DNA, es necesario estudiar con cierto detenimiento las enzimas que permiten en el laboratorio la manipulación experimental de las moléculas de DNA. El descubrimiento, caracterización y purificación de estas enzimas –hoy en día disponibles de forma comercial como un reactivo más– fue lo que definió el concepto de **ingeniería genética** en general y, en particular, las **técnicas del DNA recombinante**. En una analogía informal, pero intuitiva, se han comparado estas enzimas con *tijeras y pegamento moleculares*. Para cortar moléculas de ácido nucleico se emplean las **nucleasas**, que son fosfodiesterasas específicas para hidrolizar el enlace entre dos nucleótidos; de entre ellas, las más útiles –con mucho– son las enzimas de restricción. Por otra parte, para unir moléculas de ácido nucleico se utilizan las **ligasas**, que restablecen el enlace fosfodiéster.

Se comenzará con el estudio de las propiedades y características de las **endonucleasas de restricción**. Aunque estas enzimas desempeñan una función propia en las células procariotas en las que se descubrieron, son

particularmente importantes por su empleo experimental, de interés tanto básico como aplicado al diagnóstico clínico. Además de su uso en la clonación molecular, se aplican en otras técnicas, como la secuenciación del genoma y el estudio de los polimorfismos (v. Capítulos 16 y 24). Antes de abordar estas enzimas se plantea un breve resumen de las nucleasas en general.

15.1.1 Introducción a las nucleasas

Se denomina de forma genérica *nucleasa* a cualquier enzima que hidroliza ácidos nucleicos; más concretamente, que tiene capacidad de escindir los enlaces fosfodiéster de la cadena polinucleotídica. Pertenecen, por tanto, a la categoría de las fosfodiesteras.

La característica más relevante es su especificidad, que puede analizarse con respecto a varios criterios:

1. Posición del corte con respecto a la cadena. Las **endonucleasas** hidrolizan enlaces entre nucleótidos internos, mientras que las **exonucleasas** separan nucleótidos terminales.
2. Posición del corte con respecto al fosfato. Actúan sólo sobre uno de los dos **enlaces éster** que forman el fosfodiéster (lado 3' o lado 5').
3. Reconocimiento de moléculas mono- o bicatenarias, o bien de cadenas formadas por ribonucleótidos o por desoxirribonucleótidos (**desoxirribonucleasas** o DNAs; **ribonucleasas** o RNAs), o incluso de cadenas híbridas DNA:RNA.
4. Reconocimiento de nucleósidos específicos a través de su **base nitrogenada**.



Web 15.1. Clasificación de las nucleasas y ejemplos.

15.1.2 Enzimas de restricción

15.1.2.1 Características generales

Las enzimas de restricción se denominan también **nucleasas de restricción**, **endonucleasas de restricción** y, en ocasiones, **enzimas restrictivas** o **restrictas**. Pertenece a este grupo una gran diversidad de endodesoxirribonucleasas, descubiertas como enzimas propias de distintas bacterias. Se dispone hoy comercialmente de un gran número de ellas con elevada especificidad, estabilidad y pureza.

Se nombran con tres letras tomadas del género y especie de la bacteria de la que se aislaron originalmente, seguidas a veces por una letra más, que identifica el serotipo (variante antigénica de la bacteria) y, finalmente, por un número romano que las identifica cuando en una misma variante se hayan encontrado varias enzimas con distinta especificidad:



15.1

La importancia de las enzimas de restricción radica en su gran **especificidad** para reconocer una secuencia corta de DNA bicatenario e hidrolizar un enlace fosfodiéster en cada hebra, siempre en la misma posición. Cada enzima se caracteriza, pues, por su **sitio** o **secuencia de reconocimiento** o de restricción, o **secuencia diana**. Los fragmentos de DNA resultantes son útiles para iniciar la clonación celular o acelular, para el análisis del DNA, para elaborar mapas físicos de restricción (pág. 248) o para la detección de polimorfismos (pág. 420).

Las diversas enzimas de restricción se suelen agrupar en tres familias, de acuerdo con sus propiedades:

Endonucleasas de restricción

	Tipo I	Tipo II	Tipo III
Actividad enzimática	Endonucleasa y metilasa en una misma proteína (en distintas subunidades)	Endonucleasa (metilasa en una proteína independiente)	Endonucleasa y metilasa en una misma proteína (en distintas subunidades)
Coenzimas o cosustratos	ATP, Mg ²⁺ , SAM (*)	Mg ²⁺	ATP, Mg ²⁺ , SAM (*)
Estructura proteica	3 subunidades: S reconocimiento, R restricción, M metilación	1 subunidad, aunque actúa como homodímero	2 subunidades: MS reconocimiento y metilación, R restricción
Sitio de reconocimiento	No palindrómico, dos partes separadas (p. ej.: TGA(N) ₈ TGCT)	Palindrómico (4-8 pb) (***)	No palindrómico
Punto de corte	Corta las 2 hebras en puntos cercanos entre sí, poco específicos, a unos 1.000 pb del sitio de reconocimiento	Corta las 2 hebras en puntos equivalentes, muy específicos, dentro de la secuencia de reconocimiento	Corta las dos hebras en puntos distantes entre sí 2 o 3 nt, situados a 24-26 pb del sitio de reconocimiento
Punto de metilación	Dentro de la secuencia de reconocimiento. En ambas hebras	Adenina o citosina muy específicas dentro de la secuencia de reconocimiento. En ambas hebras	Adenina dentro de la secuencia de reconocimiento. En una sola hebra
Interés en ingeniería genética	No	Sí	No
Ejemplos	<i>Eco K, Eco B ...</i>	Numerosos (****)	<i>Eco PI, Eco P15, Hinf III</i> ...

(*)SAM: S-adenosilmetionina, véase pág. 214

(***)Palíndromo no interrumpido, véase pág. 52

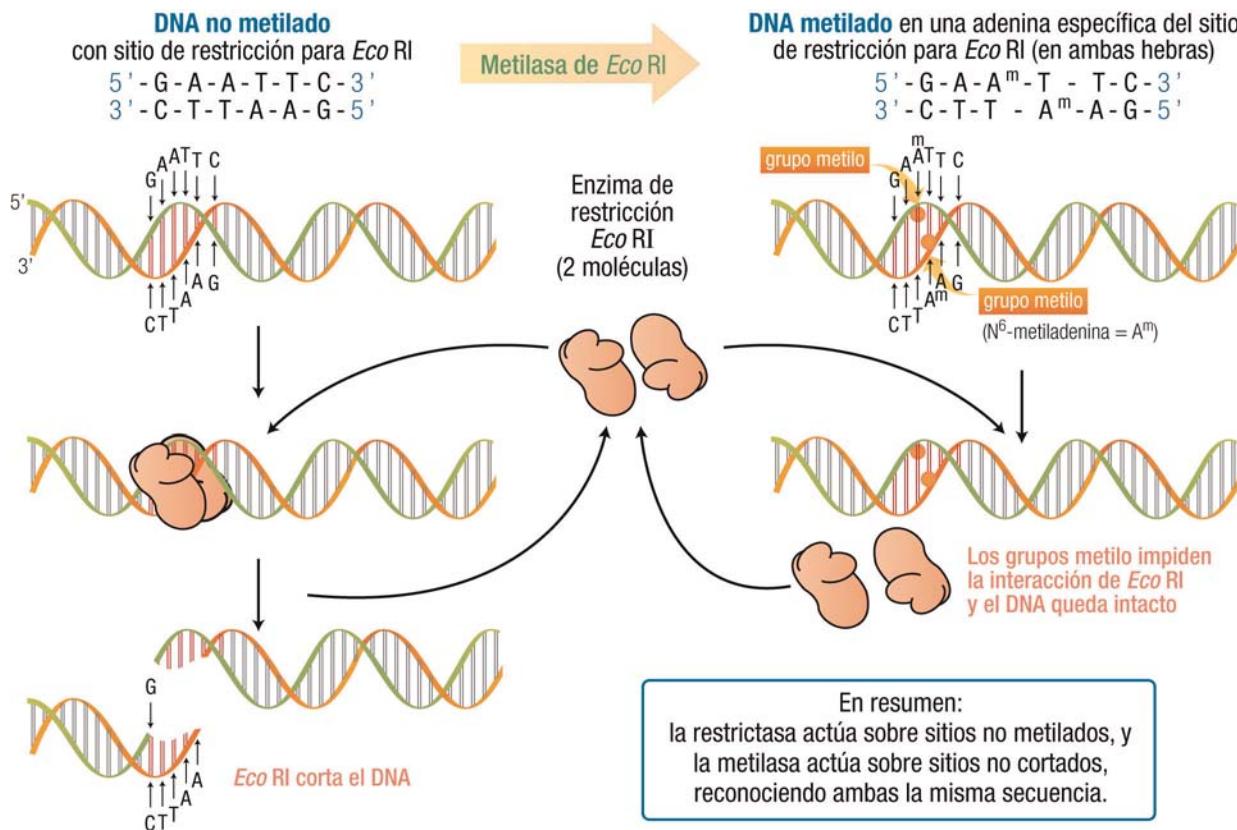
(****)Véase pág. 216

Nos concentraremos en el estudio de las enzimas de tipo II pues, al cortar en un punto muy definido dentro de su secuencia diana, son las utilizadas como herramientas en ingeniería genética.

15.1.2.2 Papel biológico del sistema metilación-restricción

La existencia natural de las nucleasas de restricción en muchas bacterias ofrece a éstas un mecanismo de defensa contra la entrada de material genético de otro organismo, concretamente de virus bacteriófagos (cuya reproducción depende de la maquinaria genética de la bacteria): se trata de los sistemas de **restricción-modificación** o **metilación-restricción**. Para cada enzima de restricción, una metilasa reconoce la misma secuencia que constituye el sitio de restricción y une covalentemente grupos metilo a determinadas bases del DNA en dicha secuencia. En el caso de enzimas de tipo II, la metilasa es una proteína independiente, mientras que las de tipo I y III poseen las actividades nucleasa y metilasa en su molécula oligomérica, como subunidades que actúan de forma coordinada (v. la tabla anterior).

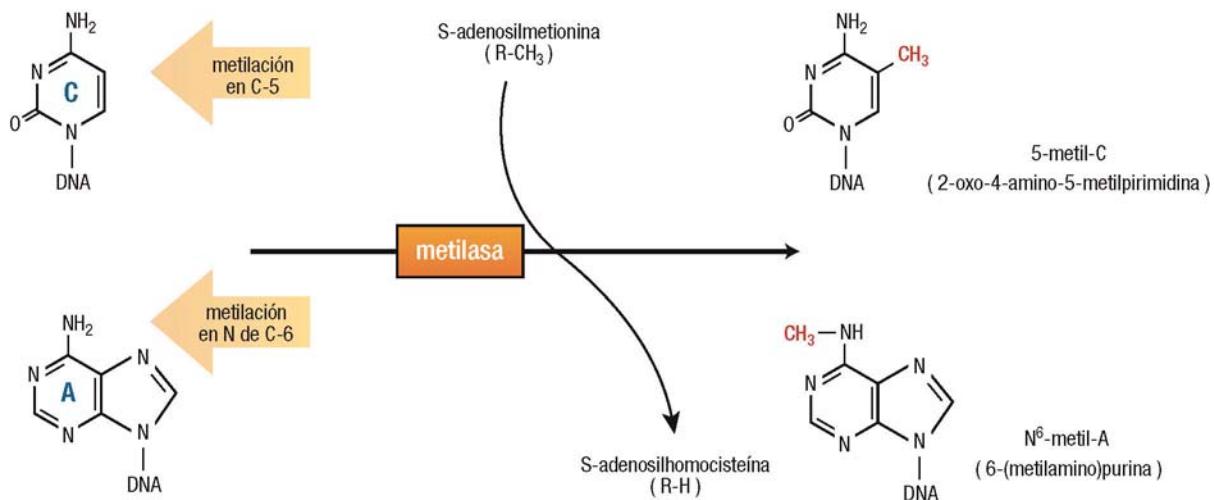
El mecanismo de defensa es el siguiente: el DNA propio de la bacteria se metila de una forma específica, característica de cada especie bacteriana, en las secuencias reconocidas tanto por la restrictasa como por la metilasa de esa especie. La metilación de las bases impide la unión de la enzima de restricción, con lo que el DNA propio no es hidrolizado. Por el contrario, al entrar en la célula DNA de otro organismo, no metilado o con un patrón de metilación diferente, este DNA puede ser degradado por la enzima de restricción, ya que carece de los grupos metilo en la secuencia diana (por razones puramente aleatorias, en cualquier genoma habrá un cierto número de secuencias de 4, 6 u 8 pb iguales a la reconocida por la enzima, luego habrá secuencias diana disponibles). A partir de este mecanismo de acción combinada surgió precisamente el nombre de enzimas de restricción: la acción de la nucleasa *restringe* la posibilidad de infección por virus.



15.2

Si está metilada una sola hebra del DNA, el sitio no es reconocido por la enzima de restricción, pero sí por la metilasa, que añade el grupo metilo a la otra hebra. Esto es lo que ocurre, por ejemplo, al final de la replicación, cuando el DNA bicatenario está formado por una hebra molde –preexistente y, en consecuencia, metilada– y otra hebra recién sintetizada, que aún carece de grupos metilo. Esta acción de las metilasas asegura que todo el DNA destinado a las dos células hijas quede metilado lo antes posible.

La metilación del DNA afecta principalmente a citosinas y adeninas de la secuencia reconocida. En la reacción, las metilasas utilizan como donador del grupo metilo el compuesto S-adenosilmetionina (pág. 19).



15.3

Debe adelantarse que la metilación del DNA interviene en eucariotas como un mecanismo de regulación de la expresión génica, aunque este proceso no tiene nada que ver con el mecanismo de protección que desempeñan los sistemas metilasa-restrictasa en procariotas. Esencialmente, en eucariotas los genes tienden a estar desmetilados en los tejidos donde se expresan y metilados donde no se expresan, en especial en las regiones de DNA denominadas “islotes CpG”, por tener esta secuencia dinucleotídica en la que C sufre la metilación. Este mecanismo es parte importante de lo que se conoce como regulación epigenética del genoma (pág. 282).

15.1.2.3 Acción de las enzimas de restricción de tipo II

Son éstas las restrictasas de estructura más simple y las mejor estudiadas por su utilidad en ingeniería genética para el corte del DNA. El sitio de restricción es una secuencia palindrómica (pág. 51) y ejerce a un tiempo de lugar de reconocimiento y de hidrólisis. Se hidrolizan de forma muy específica enlaces fosfoéster del DNA, uno en cada hebra, generándose dos fragmentos de restricción. El enlace hidrolizado es siempre el 3'-P (es decir, son endonucleasas del tipo “a”, web 15.1), por lo que el fosfato queda unido a 5', dejando en ambos fragmentos extremos 3'-OH y 5'-P. La reacción no necesita ATP, lo que la diferencia de la catalizada por las enzimas de los tipos I y III.

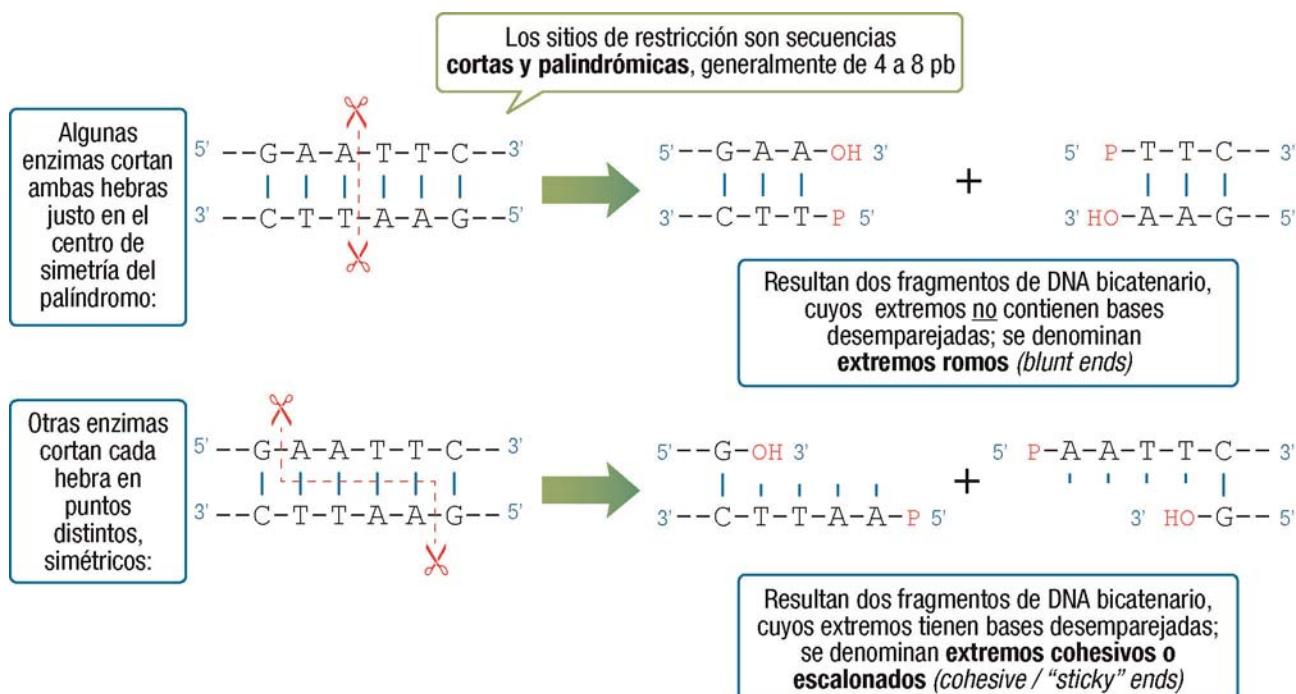
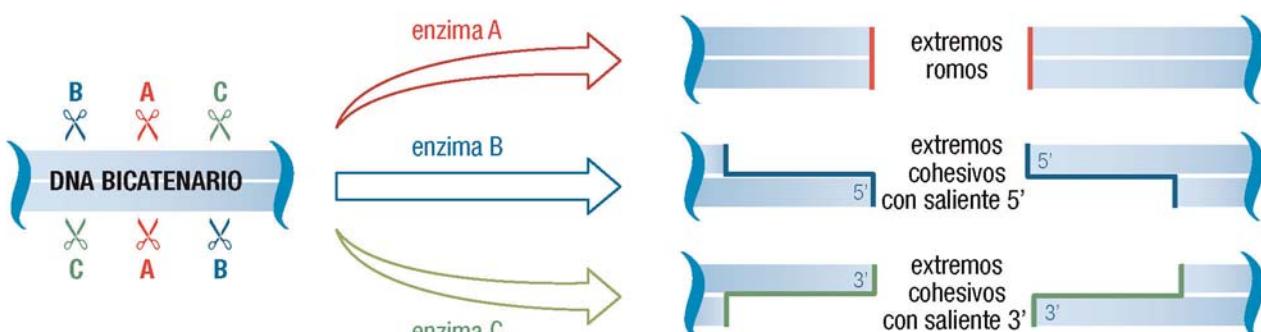


Ilustración de los distintos tipos de extremos resultantes de la acción de enzimas de restricción:



Algunas enzimas de restricción de tipo II empleadas en ingeniería genética

Nombre de la enzima	Bacteria de origen	Secuencia reconocida y puntos de corte (/ \) (N = cualquier nucleótido)	Tipo de extremos	Productos resultantes de la escisión
Acc 65I	<i>Acinetobacter calcoaceticus</i> 65	-G/G-T-A-C-C- -C-C-A-T-G\G-	Cohesivos, saliente 5'	-G -C-C-A-T-Gp ★ pG-T-A-C-C- G-
Alu I	<i>Arthrobacter luteus</i>	-A-G/C-T- -T-C\G-A-	Romos	-A-G-T-Cp pC-T- G-A-
Bam HI	<i>Bacillus amyloliquefaciens</i> H	-G/G-A-T-C-C- -C-C-T-A-G\G-	Cohesivos, saliente 5'	-G -C-C-T-A-Gp ★ pG-A-T-C-C- G-
Bgl II	<i>Bacillus globigii</i>	-A/G-A-T-C-T- -T-C-T-A-G\A-	Cohesivos, saliente 5'	-A -T-C-T-A-Gp ★ pG-A-T-C-T- A-
Eco RI	<i>Escherichia coli</i> RY13	-G/A-A-T-T-C- -C-T-T-A-A\G-	cohesivos, saliente 5'	-G -C-T-T-A-Ap pA-A-T-T-C- G-
Hae III	<i>Haemophilus aegyptius</i>	-G-G/C-C- -C-C\G-G-	Romos	-G-G -C-Cp pC-C- G-G-
Hin dIII	<i>Haemophilus influenzae</i> Rd	-A/A-G-C-T-T- -T-T-C-G-A\A-	Cohesivos, saliente 5'	-A -T-T-C-G-Ap pA-G-C-T-T- A-
Kpn I	<i>Klebsiella pneumoniae</i> OK8	-G-G-T-A-C/C- -C\C-A-T-G-G-	Cohesivos, saliente 3'	-G-G-T-A-C -Cp ★ pC- C-A-T-G-G-
Mbo I	<i>Moraxella bovis</i>	-N/G-A-T-C-N- -N-C-T-A-G\N-	Cohesivos, saliente 5'	-N -N-C-T-A-Gp ★ ★ pG-A-T-C-N- N-
Not I	<i>Nocardia otitidis-caviarum</i>	-G-C/G-G-C-C-G-C- -C-G-C-C-G-G\C-G-	Cohesivos, saliente 5'	-G-C -C-G-C-C-G-Gp pG-G-C-C-G-C- C-G-
Pst I	<i>Providencia stuartii</i>	-C-T-G-C-A/G- -G\A-C-G-T-C-	Cohesivos, saliente 3'	-C-T-G-C-A -Gp pG- A-C-G-T-C-
Pvu I	<i>Proteus vulgaris</i>	-C-G-A-T/C-G- -G-C\T-A-G-C-	Cohesivos, saliente 3'	-C-G-A-T -G-Cp pC-G- T-A-G-C-
Pvu II	<i>Proteus vulgaris</i>	-C-A-G/C-T-G- -G-T-C\G-A-C-	Romos	-C-A-G -G-T-Cp pC-T-G- G-A-C-
Sau 3AI	<i>Staphylococcus aureus</i> 3A	-N/G-A-T-C-N- -N-C-T-A-G\N-	Cohesivos, saliente 5'	-N -N-C-T-A-Gp ★ pG-A-T-C-N- N-
Tth 111I	<i>Thermus thermophilus</i> strain 111	-G-A-C-N/N-N-G-T-C- -C-T-G-N-N\N-C-A-G-	Cohesivos, saliente 5'	-G-A-C-N -C-T-G-N-Np pN-N-G-T-C- N-C-A-G-

El nombre de la enzima es una abreviatura del de la bacteria en la que se encuentra

Las secuencias diana más frecuentes tienen 6 pb, pero también son comunes las de 4 pb y 8 pb (*Not I*); se dan algunos casos con otro tamaño, por ejemplo, 9 pb (*Tth 111I*). Siempre son secuencias palíndromicas.

Tth 111I es un ejemplo de enzima con especificidad parcial de secuencia (acepta cualquier base en las tres posiciones N)

Extremos compatibles: en algunos casos, distintas enzimas de restricción generan extremos cuyo saliente tiene la misma secuencia. Ejemplo: *Bam HI*, *Bgl II* y *Mbo I* (★).

Isoesquizómeros: nombre que se aplica a dos o más enzimas que reconocen la misma secuencia; pueden cortar en el mismo punto o bien se diferencian en que el corte o la metilación asociada tienen lugar en puntos distintos dentro de esa secuencia común de reconocimiento. Ejemplos: *Mbo I* y *Sau 3 Al* (★); *Acc 65I* y *Kpn I* (★).

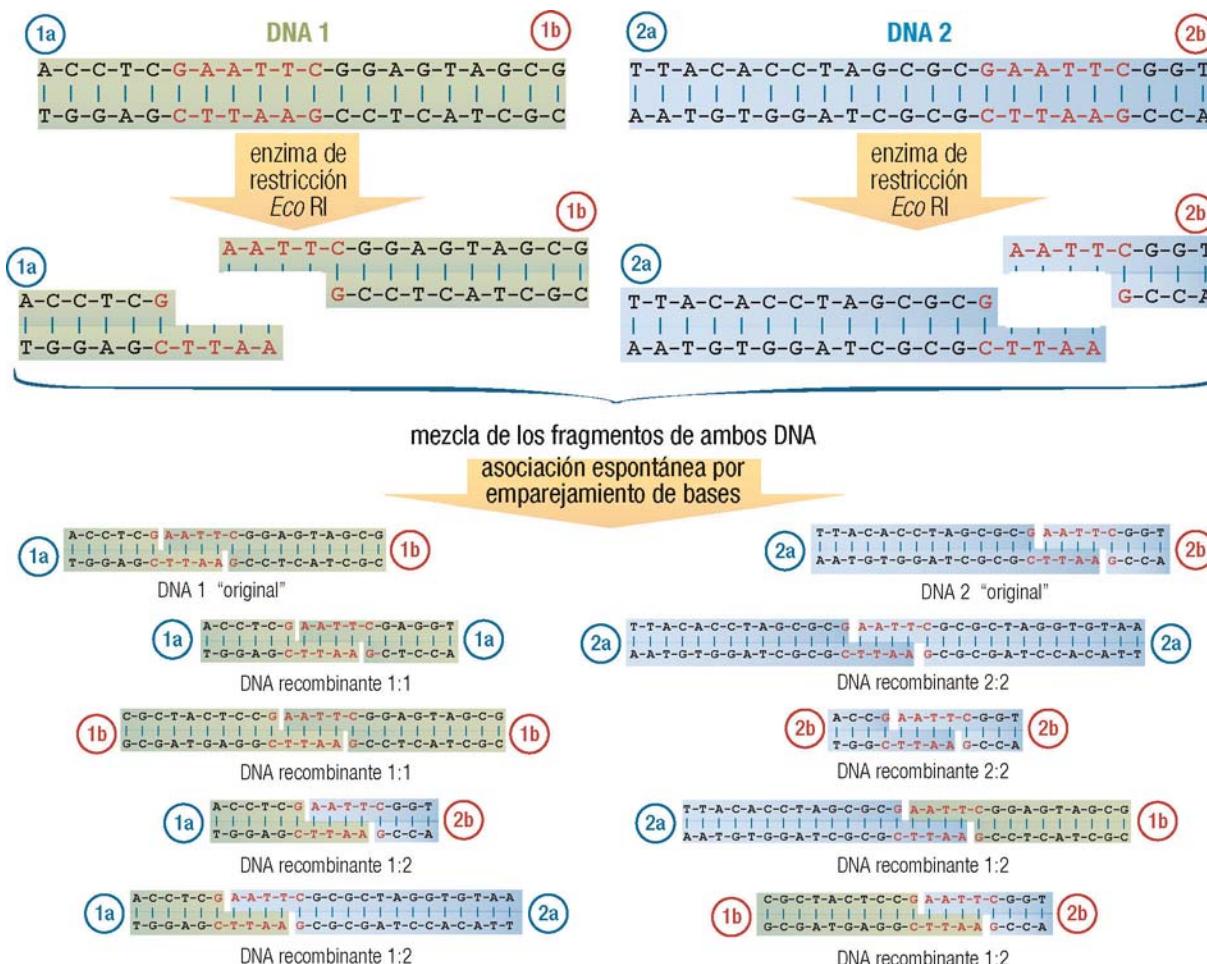
Puede observarse que, en todos los casos, los extremos de los dos fragmentos generados son idénticos, tienen la misma secuencia (puede comprobarse girando 180° uno de ellos). Esto es consecuencia de la simetría del palíndromo y de la posición también simétrica de los puntos de corte en cada hebra. Todo ello se debe a que la enzima actúa en forma de dímero, por tanto reconociendo una misma secuencia y cortando un mismo tipo de enlace en las dos hebras: por ejemplo, *Eco RI* corta el enlace (5') G—AATTG (3') en ambas hebras.



Web 15.2. Frecuencia de corte y tamaño de los fragmentos de restricción.

La formación de extremos cohesivos es de gran interés aplicado en ingeniería genética (pág. 225), pues fragmentos generados con una misma restrictasa a partir de dos DNA distintos pueden interaccionar mediante el emparejamiento de las bases de sus extremos cohesivos (dicha interacción no es posible entre extremos romos). Asimismo, pueden asociarse los fragmentos producidos por enzimas distintas que producen extremos compatibles.

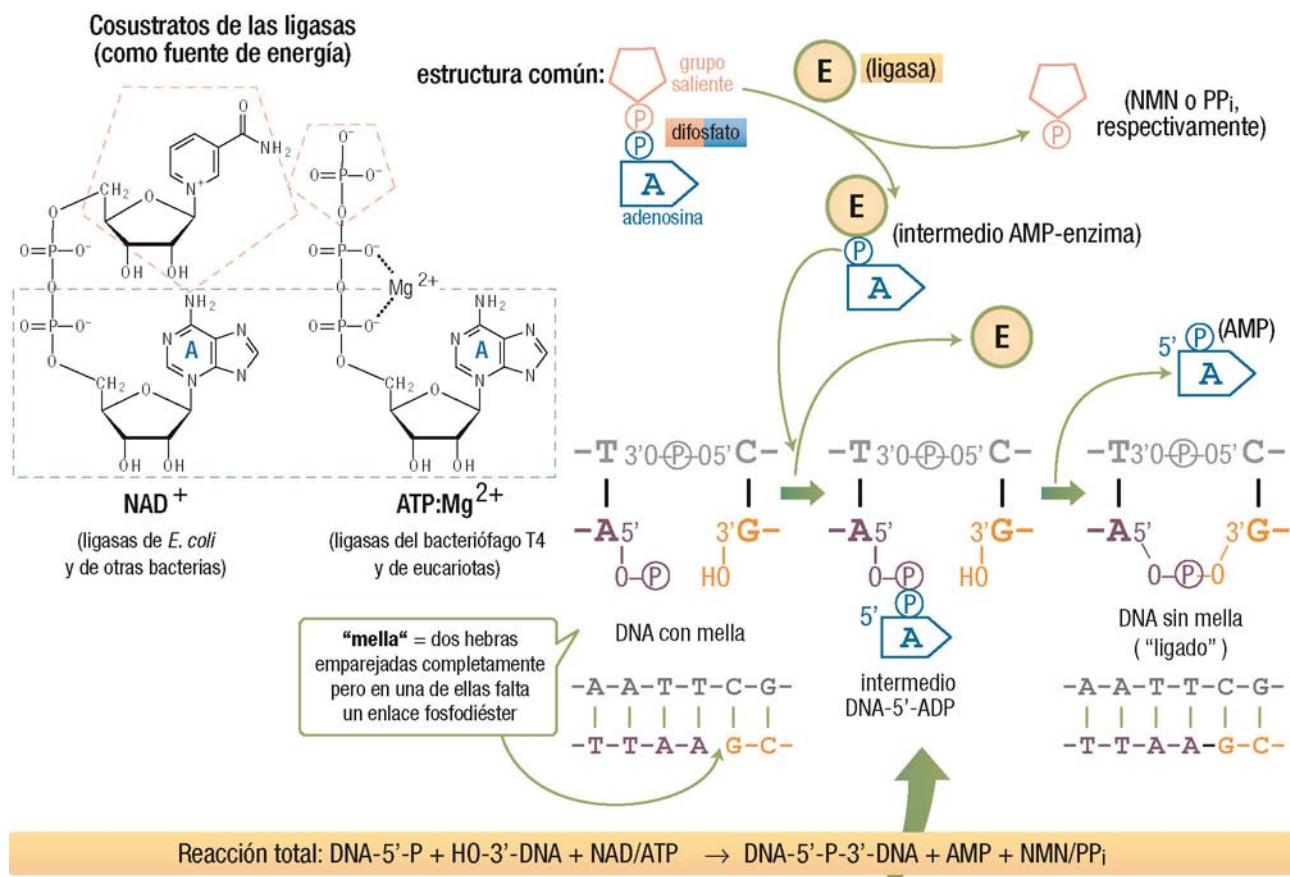
Debe resaltarse que el producto de la reasociación de los dos fragmentos en ninguno de los casos es idéntico al DNA de partida, ya que está formado por dos moléculas. La falta de enlace covalente entre los extremos de los fragmentos reasociados se indica con el nombre de “mella” o “muesca” (en inglés, *nick*) (pág. 218). Ésta se puede cerrar o “sellar” por la acción de las enzimas **ligasas** (o DNA ligasas), dando una molécula bicatenaria única. Como se verá a continuación, las ligasas catalizan la formación en cada hebra del enlace fosfoéster que faltaba, con consumo de energía. Pueden, incluso, unir o “ligar” dos fragmentos de DNA con extremos romos, aunque con mucha menor eficacia, pues no hay emparejamiento entre ellos que los asocie previamente.



15.1.3 Ligasas

Las ligasas, o DNA ligasas, desempeñan su función en las células durante el proceso de replicación, entre otros. Su capacidad para catalizar la unión de cadenas de DNA es de particular relevancia en la síntesis de la hebra retardada de DNA, cuyos fragmentos de Okazaki deben unirse entre sí (pág. 158). Por otra parte, la experimentación con ácidos nucleicos aprovecha esta actividad enzimática como herramienta de uso frecuente, para conectar moléculas de DNA, en una actuación formalmente opuesta a la de las nucleasas.

La reacción de las ligasas requiere la presencia de un grupo fosfato en el extremo 5' de una hebra de DNA y un grupo OH en el extremo 3' de la otra. Estas enzimas catalizan, pues, la formación de un solo enlace fosfoéster, parte del fosfodiéster resultante. El mecanismo químico de la reacción es el siguiente:



15.7

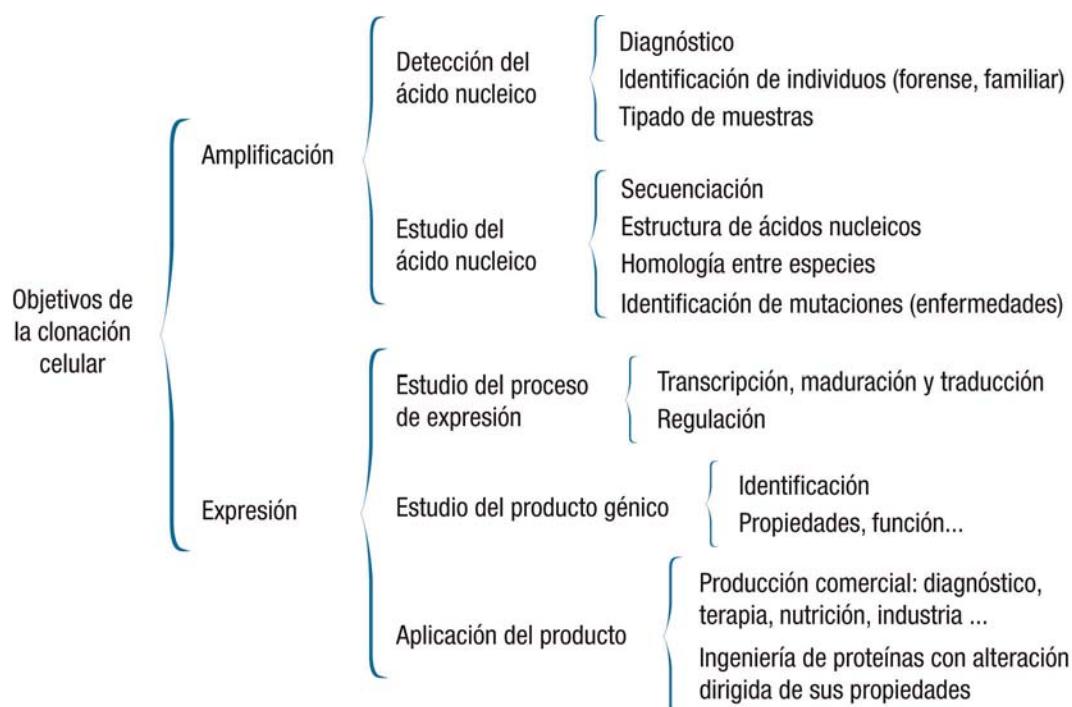
El DNA-5'-ADP es la forma activada del extremo 5'-P del DNA que permite el ataque nucleofílico del 3'-OH para regenerar el enlace fosfodiéster

La acción de las ligasas es, por tanto, contraria a la de las nucleasas (formación de enlace fosfoéster frente a su hidrólisis), pero el mecanismo de la reacción no es exactamente opuesto, debido al requerimiento energético de las primeras. Debe resaltarse que la ligasa siempre requiere un grupo 5'-P para su actuación, propiedad ésta que posee utilidad experimental (pág. 226).

15.2 CLONACIÓN CELULAR DE MOLÉCULAS DE DNA

15.2.1 Objetivos de la clonación de DNA

Como se ha presentado en los dos capítulos precedentes, la clonación cumple el propósito básico de amplificar la muestra genética de partida, generando un número elevado de copias. Para el caso de la clonación de moléculas de ácido nucleico, a la mera amplificación se añaden las aplicaciones derivadas de la expresión de su producto génico. La clonación acelular por PCR cubre básicamente el primero de dichos objetivos, mientras que la clonación celular permite, además, abordar el segundo.

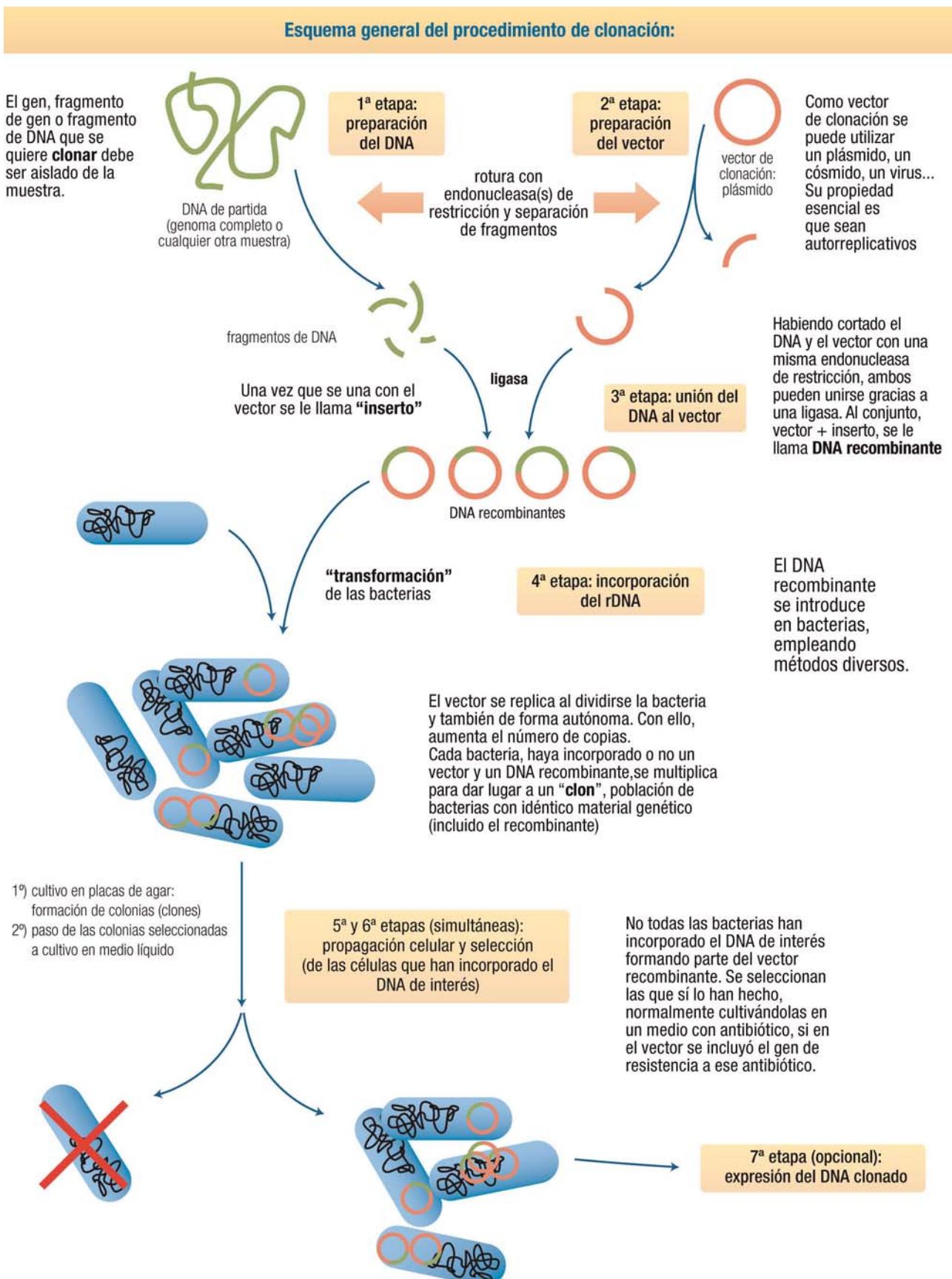


15.8

15.2.2 Descripción global del proceso de clonación

El objetivo central de esta clonación es la producción de un gran número de copias de una región de DNA (fragmentos o genes) o de cDNA (formado a partir de un mRNA). Se trata, por tanto, de un proceso de clonación de tipo **celular** (pues se consigue gracias al empleo de células, pág. 192-193) que, por contraposición a la clonación **acelular** por PCR, *in vitro*, se puede considerar que se realiza *in vivo* (estrictamente hablando, en cultivo). Se basa esta clonación, al igual que la PCR, en la realización de múltiples rondas de replicación, pero en este caso catalizadas por la DNA polimerasa propia de la célula en la que se lleva a cabo el proceso, célula “**anfitriona**”. Para ello, el fragmento de DNA o cDNA que debe clonarse (llamado **inserto**, por razones que se comprenderán después) se une a otro DNA (**vector de clonación**) y la molécula resultante (**DNA recombinante** o *recombinado*) se incorpora a la célula anfitriona donde tiene lugar la amplificación por replicación. Una vez detectada la presencia del DNA de interés y seleccionados los clones que lo han amplificado, se aísla el gen o fragmento de DNA o cDNA, que se emplea con distintos fines.

Por razones didácticas, a la vez que de enfoque experimental, puede describirse la clonación en **siete etapas**: dos de preparación de muestras, cuatro sucesivas de clonación propiamente dicha y, según los casos, una de expresión. En ocasiones, se añade la amplificación por PCR del DNA clonado (en este caso, el proceso comprendería dos clonaciones sucesivas, celular la primera y acelular la segunda). Hoy día, todas las etapas se realizan de forma sistemática, a veces incluso automatizada, al disponerse de forma comercial de las enzimas, sustratos, vectores, etc., e instrumentos adecuados para su realización.





Web 15.3. Esquema y etapas de la clonación celular de DNA.

15.2.3 Preparación del DNA para su clonación

Esta primera etapa de la clonación, la preparación inicial del ácido nucleico que se quiere clonar, es una parte esencial del proceso. Aunque por simplicidad se describe la clonación de un solo fragmento, en la realidad es difícil aislar un fragmento único de DNA y, en la práctica, se clonian de manera simultánea múltiples fragmentos que más tarde deben seleccionarse (genotecas, pág. 236). Sin embargo, los planteamientos son idénticos desde el punto de vista conceptual, por lo que se aborda en primer lugar la descripción del caso simplificado.

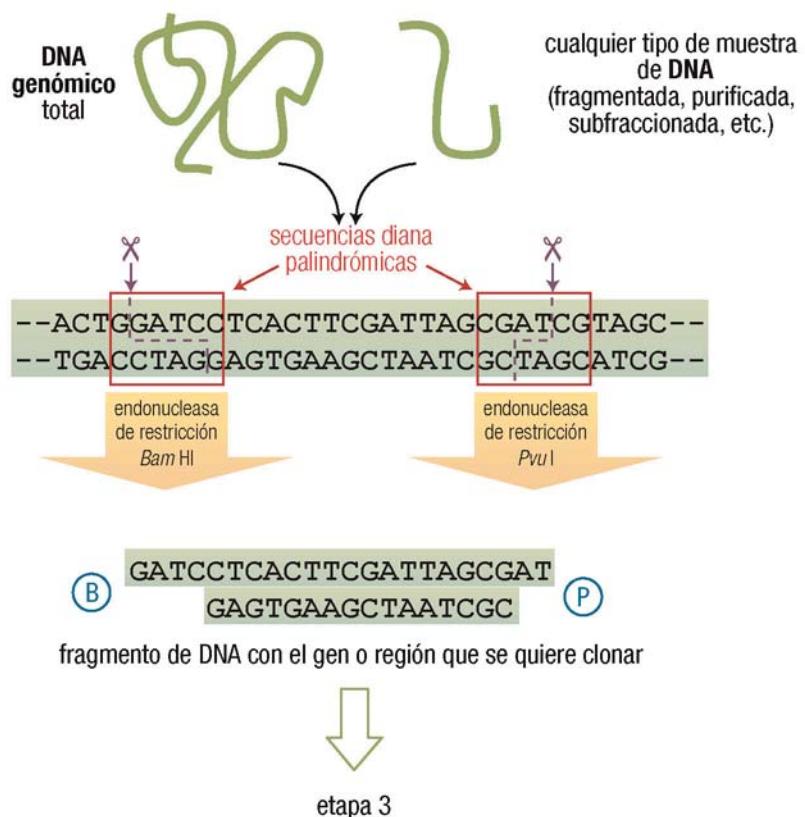
15.2.3.1 Insertos de DNA

Se parte de un lisado de células nucleadas, a partir del cual se purifica el DNA (Capítulo 10). Éste posteriormente se **fragmenta** con enzimas de restricción, para dar lugar a una mezcla cuyo contenido, en número y tamaño de fragmentos, depende del número de sitios de restricción y de su accesibilidad para las enzimas. En principio, el fragmento que se va a clonar debe aislarse, o separarse del resto, mediante alguna técnica adecuada, por ejemplo electroforesis en gel de agarosa, centrifugación en gradiente de densidad o cromatografía líquida de alta eficacia (HPLC), entre otras. En la práctica, en muchos casos es más difícil aislar un único fragmento de DNA para su clonación que clonar la mezcla de fragmentos y más tarde seleccionar el clon que contiene el fragmento de interés (págs. 234-236).

Cuando se desea expresar el fragmento clonado (etapa 7) es necesario, además, preparar un inserto auxiliar que contenga elementos reguladores de la expresión génica (salvo que dichos elementos estén ya preincorporados en el vector empleado).

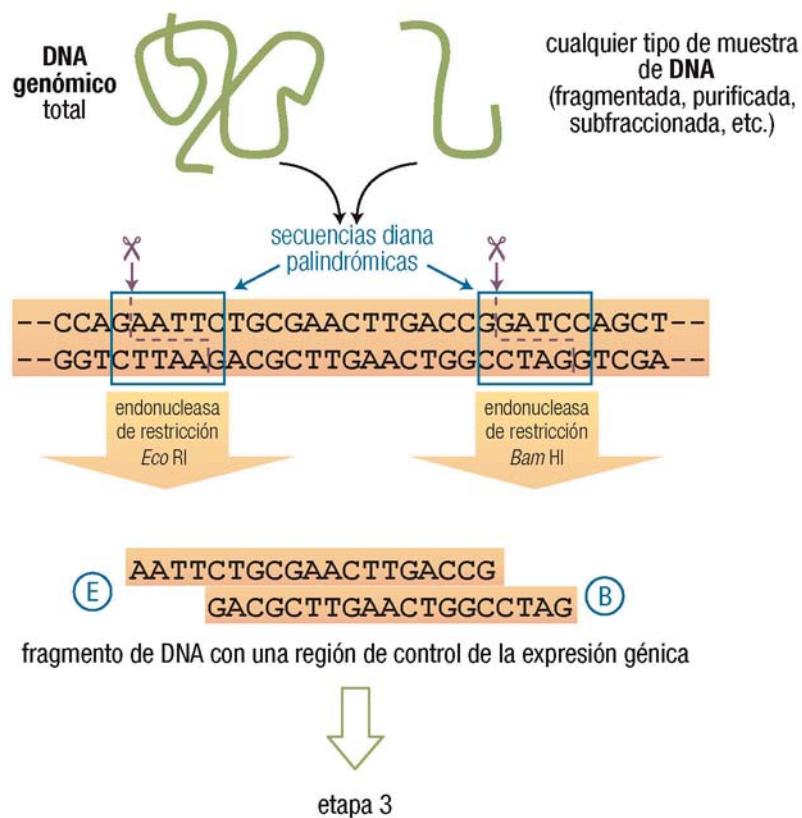
Detalles del procedimiento de clonación: preparación del DNA (etapa 1)

(A) Diseño más simple: inserto que se va a clonar (DNA de interés)



Detalles del procedimiento de clonación: preparación del DNA (etapa 1)

(B) Opcional: inserto auxiliar para el caso de clonación con expresión



15.10b

15.2.3.2 Insertos de DNA complementario procedente de un mRNA

En muchos casos interesa clonar la secuencia codificante de un gen eucariótico, para lo cual se parte de muestras de mRNA que debe convertirse en el cDNA correspondiente para poder clonarse.

En este caso no es indiferente la elección de las células de partida, pues cada tipo celular expresa un conjunto de proteínas diferente y, por consiguiente, no contiene los mismos mRNA. Una vez lisadas las células del órgano, tejido, tipo celular, estadio de desarrollo, etc., elegidos, se purifica el mRNA, generalmente aprovechando la presencia en su molécula de la *cola de poli(A)* (pág. 302). Se emplea para ello cromatografía de afinidad en columna o microesferas magnéticas de afinidad (pág. 138); en ambos casos, la matriz cromatográfica o la microesfera magnética llevan unido un oligo(dT). Cuando el mRNA que va a clonarse no es abundante, puede amplificarse mediante PCR, lo que facilita su clonado posterior en células.

Para preparar el cDNA se emplea la transcriptasa inversa, de forma análoga a lo estudiado en la RT-PCR (pág. 208):

Detalles del procedimiento de clonación: preparación del DNA (etapa 1)

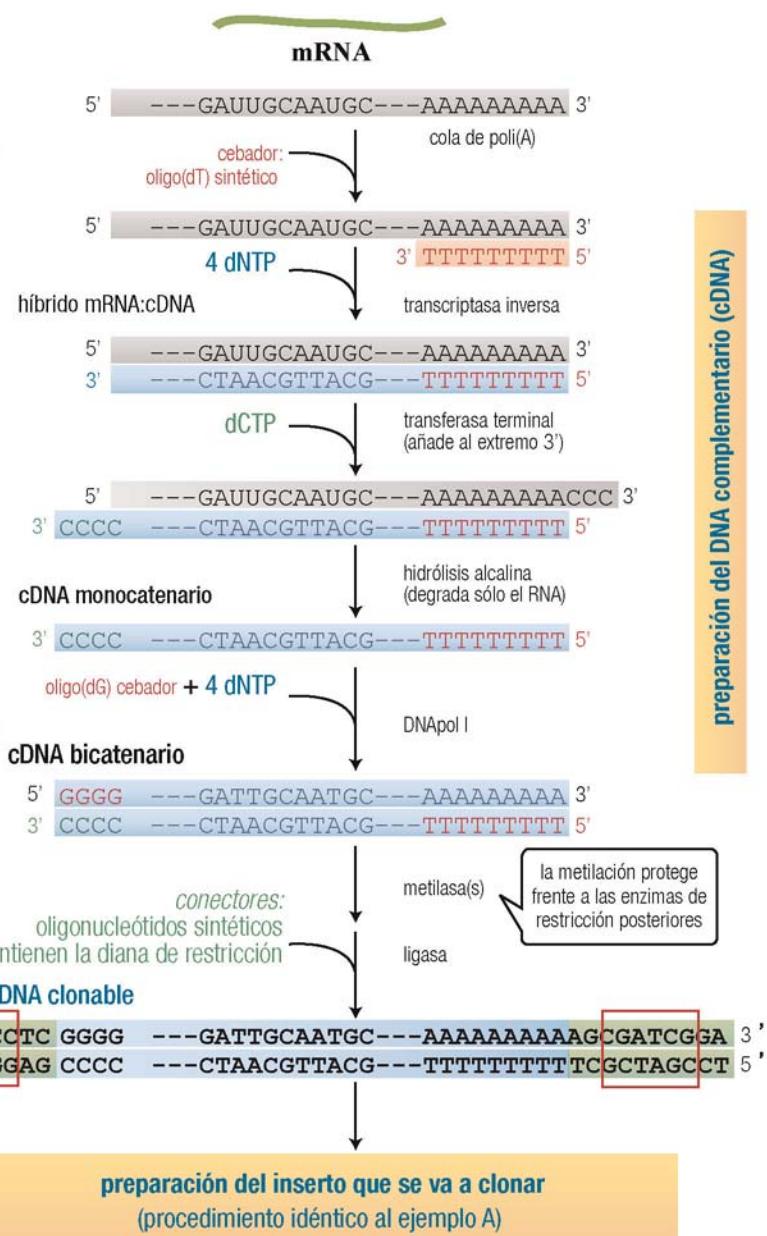
(C) Alternativa: clonación a partir de un RNA

La propia cola poli(A) característica de los mRNA eucarióticos sirve de molde para un cebador oligo(dT), de modo que sólo tendrá lugar la transcripción inversa de los mRNA

La transferasa terminal añade nucleótidos a los extremos 3', sin necesidad de molde y de forma inespecífica. Añadiendo sólo dCTP se asegura que el saliente 3' creado sea poli(dC)

En otros casos, se aprovecha una tendencia natural del extremo 3' a formar una horquilla que hace el papel de cebador sin necesidad de introducir la cola 3' poli(dC).

El cDNA obtenido sirve de molde para la síntesis de una segunda hebra, mediante la acción de la DNAPol-I de *E. coli*. El saliente poli(dC) sirve de molde a un cebador oligo(dG).



preparación del DNA complementario (cDNA)

15.11

15.2.4 Preparación del vector de clonación

El vector es un portador, una molécula de DNA cuya misión es unirse con el fragmento de DNA que se quiere clonar para facilitar su entrada en la célula anfitriona y su replicación. En general, el vector tiene un tamaño pequeño, es fácil de aislar y caracterizar, se conocen su secuencia y su mapa de restricción (pág. 248), es fácil de introducir en la célula anfitriona y una vez allí posee *capacidad de replicación autónoma*, es decir, independiente de la replicación del genoma de la célula anfitriona. Es conveniente que el vector posea la mayor variedad posible de sitios de restricción, para facilitar su unión con el fragmento de DNA que se quiere clonar, y que incluya al menos un *gen marcador* (por ejemplo, un gen de resistencia a un antibiótico) que permita

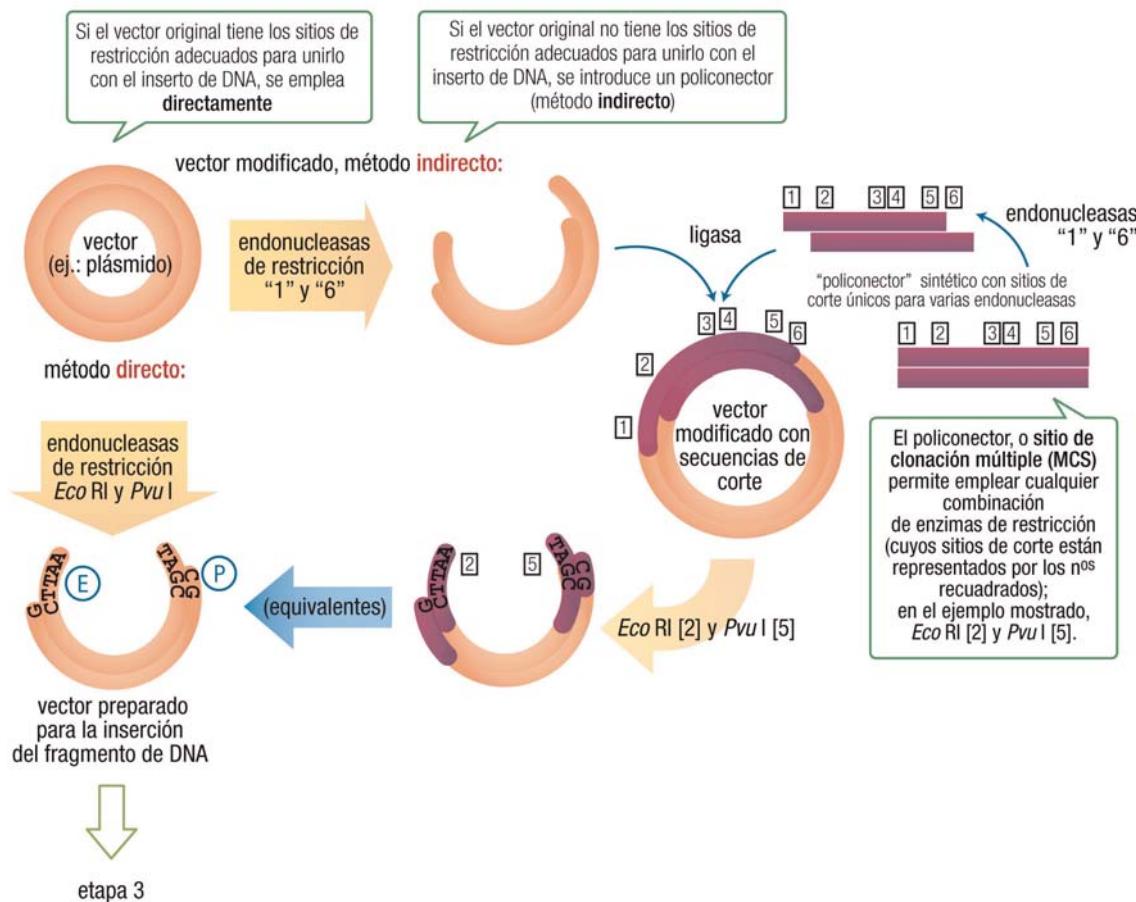
identificar o seleccionar las células que llevan el DNA recombinante (pág. 234). Lo común es que los vectores naturales no posean todas estas características, por lo que habitualmente se emplean vectores modificados (a su vez, obtenidos por técnicas de DNA recombinante, pero ya disponibles comercialmente con todas las prestaciones deseadas).

Para poder unir con facilidad el vector con el fragmento de DNA que se quiere clonar (llamado entonces *inserto*) se debe cortar el vector con las mismas enzimas de restricción que se utilizaron para escindir el DNA de la muestra, o bien con enzimas que proporcionen extremos compatibles; de ese modo, sus secuencias complementarias pueden asociarse y ser unidos por la ligasa, dando lugar a una sola molécula de DNA bicatenario, que recibe el nombre de *DNA recombinante* (rDNA).

En función de los objetivos de la clonación (amplificación y expresión), se suele hablar de dos grandes grupos de vectores. Por un lado, los **vectores de clonación**, que se emplean únicamente para amplificar el DNA de interés; también se denominan a veces **vectores de inserción** o **vectores de propagación**. Teóricamente, sólo incorporarían el inserto de interés (pág. 221). Por otro lado, se encuentran los **vectores de expresión**, que poseen, además, características que facilitan la expresión del DNA de interés por parte de la célula anfitriona. En este caso deberá incorporarse, además del DNA que se quiere clonar, el ya mencionado inserto auxiliar que proporciona una región de control de la expresión (pág. 222). En la práctica, los vectores de expresión comerciales suelen incluir ya una región de control, comúnmente un promotor potente que se induce fácilmente bajo condiciones experimentales controlables, válido para expresar cualquier inserto de interés que se sitúe tras él en el rDNA.

Aunque existen diversos tipos de vectores (pág. 226), por ahora la descripción se hará con un ejemplo sencillo, un plásmido, molécula pequeña de DNA circular.

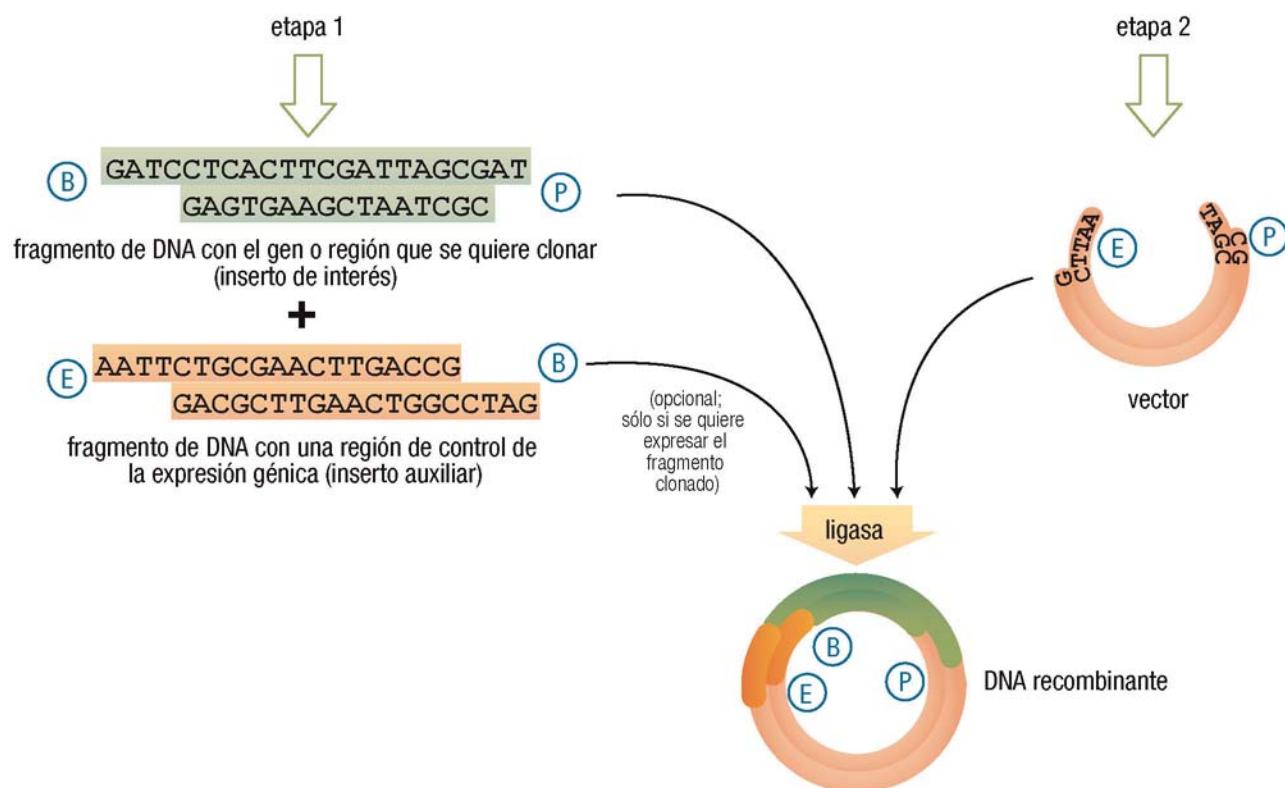
Detalles del procedimiento de clonación: preparación del vector (etapa 2)
(corte del vector para generar extremos complementarios a los del DNA)



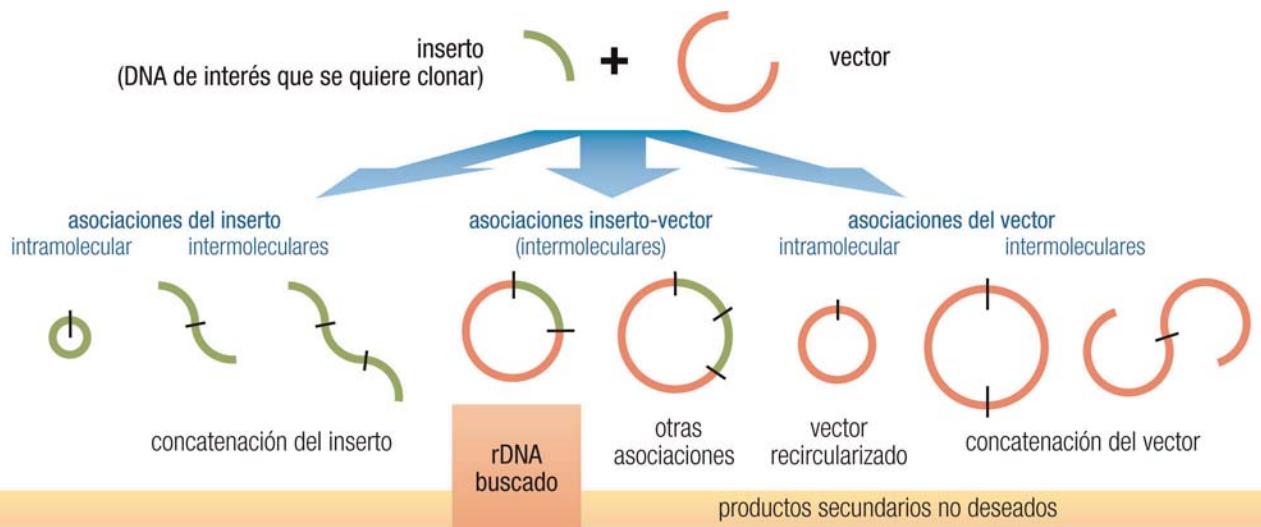
15.2.5 Formación de la molécula de DNA recombinante

El DNA que sufre la clonación se denomina recombinante porque combina dos fragmentos de DNA de orígenes diferentes. Su preparación consiste en la unión covalente, *in vitro*, del inserto de DNA al vector cortado, empleando una ligasa. La especificidad de dicha unión depende del tipo de extremos que hayan resultado al preparar el inserto y el vector: si son cohesivos, su secuencia es la específica de la enzima de restricción empleada y sólo se asocian los extremos compatibles (pág. 216); una vez asociados por emparejamiento de bases, la ligasa establece la unión covalente sellando las dos mellas. Si se trata de extremos romos, no pueden asociarse previamente, pero la ligasa los puede unir, aunque con menor eficacia, y no supone ninguna diferencia la enzima de restricción con la que se prepararon (se puede decir que todos los extremos romos son compatibles).

Detalles del procedimiento de clonación: preparación del rDNA (etapa 3) (unión del DNA que se quiere clonar al vector)

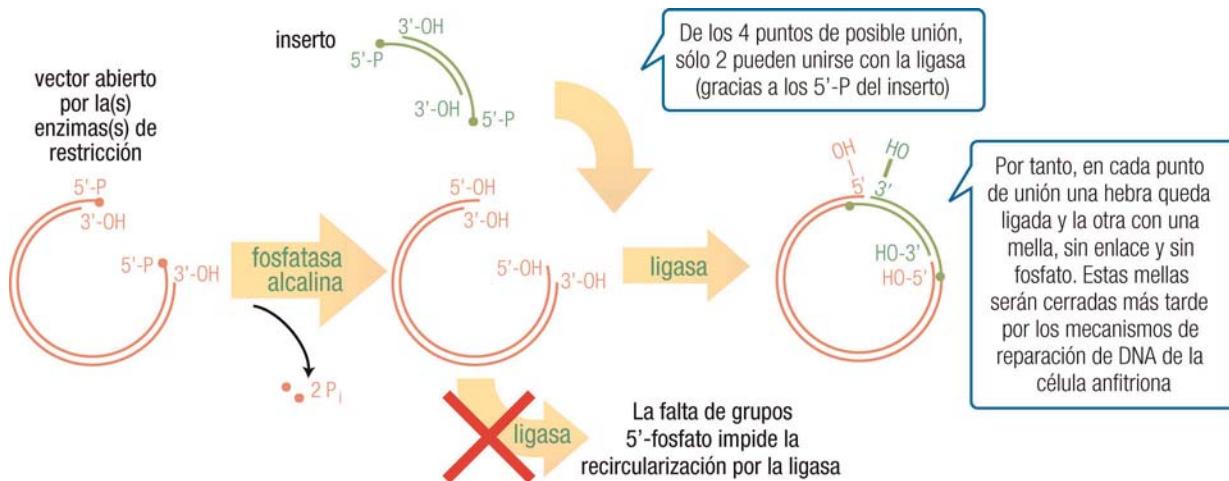


En la práctica, pueden tener lugar diversos tipos de asociación entre fragmentos de restricción con extremos compatibles, y todos ellos serán fijados por la ligasa. En primer lugar, siempre puede darse la asociación **intramolecular**, es decir, entre los dos extremos de una misma molécula (vector o inserto); si se trata de un vector circular, da lugar a su recircularización. Esta interacción está favorecida cuando la concentración de DNA es baja (pues la probabilidad de que se encuentren dos moléculas es menor). En segundo lugar, puede tener lugar una asociación **intermolecular**, tanto la que conduce al rDNA (vector + inserto) como otras que conducen a productos no deseados: vector + vector, inserto + inserto, unión de más de dos moléculas, etc.; todas ellas pueden ser circulares o lineales. La probabilidad de estas interacciones aumenta al hacerlo la concentración del DNA.



15.14

Para evitar la formación de estos productos secundarios, que disminuye el rendimiento de la clonación, se acude a estrategias experimentales, previas a la mezcla de vector e inserto, como la escisión con dos enzimas de restricción diferentes (ya expuesta, pág. 224), que evita la recircularización del vector o, como se describe a continuación, la desfosforilación del vector con fosfatasa alcalina.



15.15

15.2.6 Tipos de vectores

Los vectores empleados para la clonación de insertos de DNA son muy variados. Pueden clasificarse según varios criterios, tales como:

- Su procedencia procariótica o eucariótica.
- El tamaño de inserto que admiten.

- El tipo de molécula a partir de la que se preparan:
 - Plásmidos: bacterianos, de levadura o de plantas.
 - Virus: que infectan bacterias (bacteriófagos o simplemente *fagos*), plantas, invertebrados o vertebrados.
 - Cromosomas artificiales: derivados de elementos cromosómicos de fagos (PAC), de bacterias (BAC) o de levaduras (YAC).
 - Quimeras, es decir, moléculas formadas combinando partes de otras cuyo origen es diferente. Normalmente, quimeras de plásmido y fago: cósmidos, fagémidos, fásmidos.
- El tipo de célula anfitriona en la que se puede luego incorporar el DNA recombinante producido (etapa 4).
- El gen de resistencia incluido en el vector para su posterior detección o selección (etapa 6).

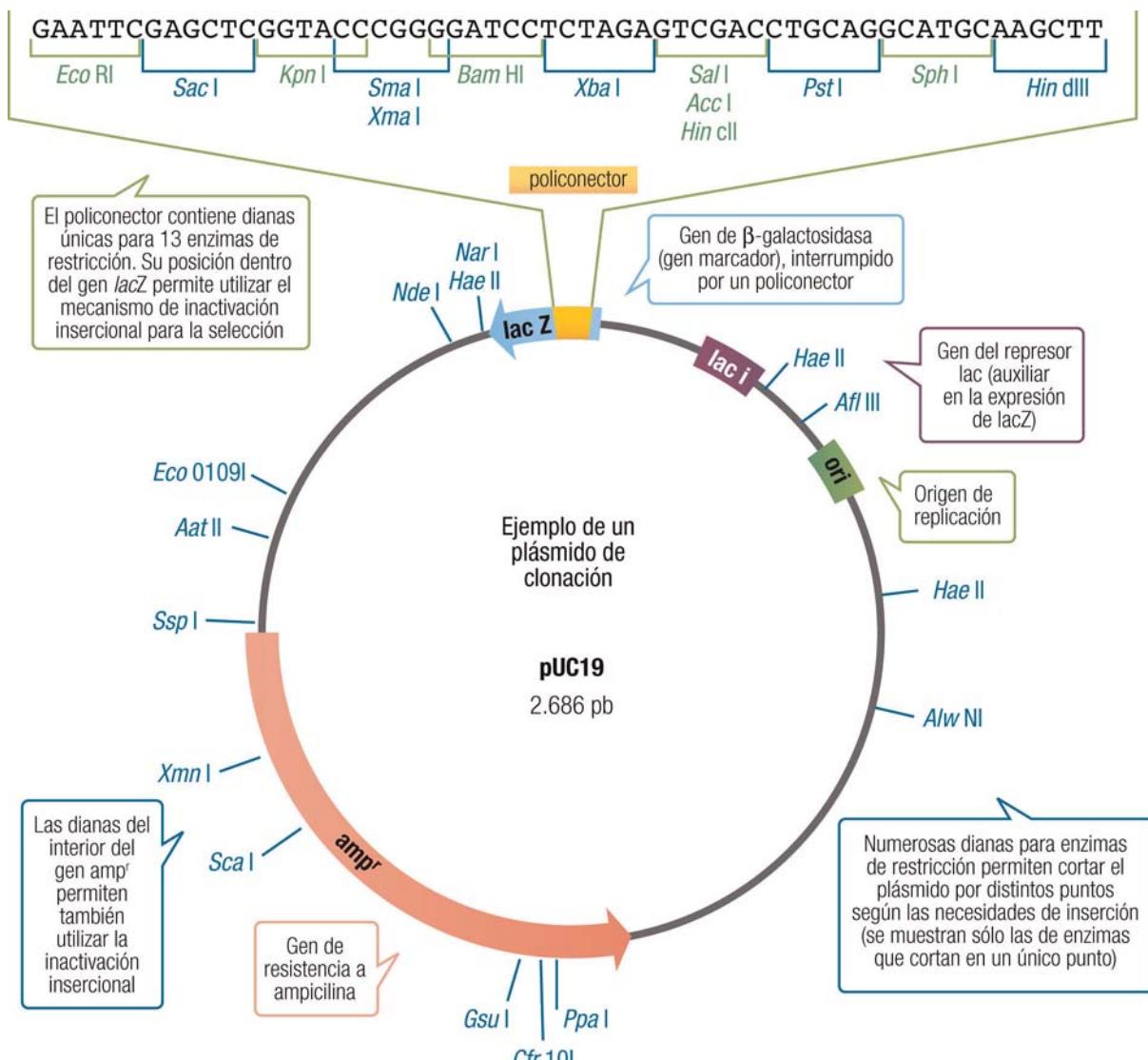
Tipo de vector	Tamaño del inserto que admite	Método de propagación
Plásmidos normales	0-10 kb	Replicación autónoma natural del plásmido
Bacteriófagos, por inserción (fago λ)	0-10 kb	Replicación propia del fago
Bacteriófagos, por sustitución (fago λ)	9-23 kb	Replicación propia del fago
Cósmidos	30-44 kb	Replicación a modo de plásmido
Bacteriófago P1	70-100 kb	
PAC (cromosomas artificiales de P1)	130-150 kb	
BAC (cromosomas artificiales de bacteria)	Hasta 300 kb	
YAC (cromosomas artificiales de levadura)	0,2-1 Mb	

Además, debe señalarse que continuamente surgen nuevos tipos de vectores como consecuencia del avance en las técnicas de ingeniería genética y las necesidades impuestas por el desarrollo acelerado de los proyectos genoma.

15.2.6.1 Plásmidos

Son moléculas de DNA bicatenario, de pequeño tamaño (de 2 a 5 kb) y estructura circular cerrada, presentes en forma libre en el citosol de numerosas bacterias (de 1 a 3.000 plásmidos/célula) y de algunos eucariotas unicelulares (por ejemplo, el plásmido μ de levaduras, que también recibe el nombre de “círculo de 2 μm ”). Se purifican fácilmente a partir del cultivo celular. Permiten la incorporación de insertos de DNA de hasta 10 kb, aproximadamente. Los plásmidos naturales contienen pocos genes propios, ninguno esencial para la viabilidad de la célula, por lo que pueden ser modificados sin afectarla. Para su replicación autónoma (independiente de la del cromosoma bacteriano, aunque utilizando las mismas enzimas), se requiere una secuencia origen de replicación. Mientras que en cada división celular se produce una sola copia del cromosoma por cada célula hija, el plásmido (en su caso, el rDNA formado a partir de él, etapa 3) sufre replicaciones múltiples y origina varias copias por célula (entre 20 y 50). Los plásmidos son así replicones (unidades de replicación, pág. 153) que se transmiten y mantienen de manera estable como círculos extracromosómicos.

El empleo de un plásmido como vector de clonación se favorece cuando incluye genes marcadores, cuya presencia pueda detectarse o cuya expresión confiera a la célula anfitriona portadora del rDNA alguna propiedad especial (pág. 234). Por ello, se suelen utilizar plásmidos artificiales o sintéticos, obtenidos a su vez por técnicas de DNA recombinante y disponibles comercialmente. Baste como ejemplo el plásmido pUC19.

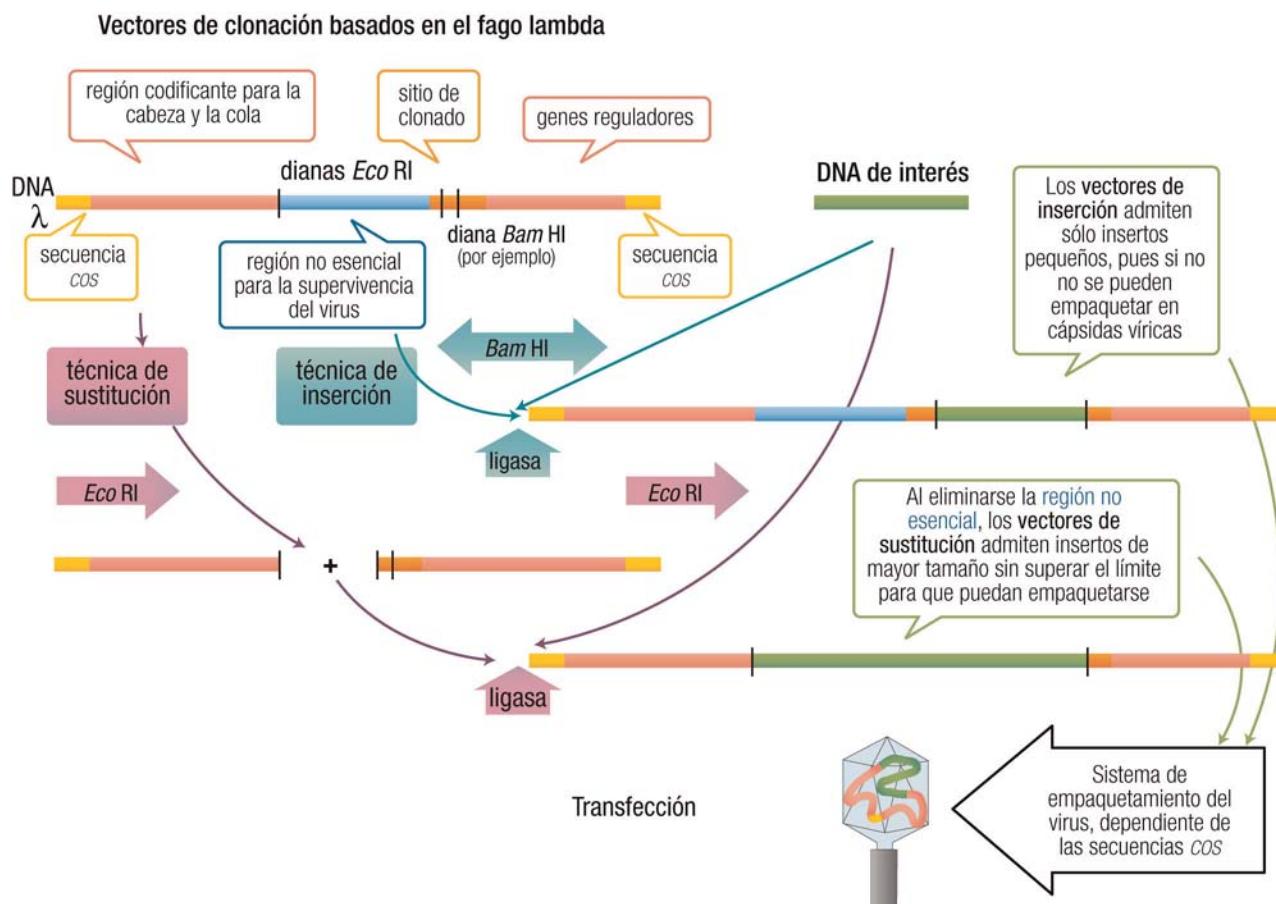


15.2.6.2 Bacteriófagos

Los virus en general se replican mediante la introducción de su genoma en células procarióticas o eucarióticas (infección). Esta propiedad biológica resulta útil para conseguir una alta eficiencia de clonación, empleando como vectores virus modificados. Cada especie de virus infecta a células específicas y, dependiendo del tamaño del genoma vírico, admite insertos más o menos grandes, pero en general mayores que los que se pueden insertar en plásmidos. Así, se usan **bacteriófagos** para clonar en bacterias, **baculovirus** para células de insecto, y virus SV40 o retrovirus para células de mamífero. La proliferación de los virus recombinantes se consigue en un cultivo de las células anfitrionas respectivas, en las que el virus se reproduce. El uso de virus es especialmente adecuado para la clonación en células de mamífero, por su elevada eficacia de introducción del rDNA, pero dada su complejidad y por razones de extensión no se tratará en este capítulo.

Los virus empleados más comúnmente como vectores son los parásitos bacterianos **fago M13** y **fago λ (lambda)**. El tamaño del vector recombinante se ve limitado por la necesidad de que pueda empaquetarse dentro de la cápsida. El inserto se puede introducir de la misma forma que para los vectores plasmídicos, es

decir, cortando inserto y DNA del fago con las mismas enzimas de restricción y ligando para obtener un rDNA; esto se denomina *técnica de inserción* y existen vectores específicos para ella. Cuando el tamaño del DNA de interés que se quiere clonar es mayor, se puede incorporar eliminando previamente la parte del genoma vírico que no se requiere para la infección y replicación; se habla entonces de *técnica de sustitución* o reemplazamiento.



15.17

15.2.6.3 Cósvidos

Son vectores sintéticos, quimeras que combinan características de plásmidos y fagos para aunar ventajas de ambos tipos de vectores. El fragmento plasmídico proporciona las características de tamaño pequeño (5-7 kb), circularidad, un origen de replicación, algunos sitios de restricción donde se puede insertar el DNA, uno o más marcadores seleccionables y la propagación en bacterias como si fuesen plásmidos. Del fago λ se toman las secuencias "cos" (del inglés *cohesive sites*), que proporcionan al cósmido la capacidad, propia del DNA del fago, de empaquetarse en cápsidas víricas *in vitro*.

Los cósmidos son útiles para clonar insertos de gran tamaño (hasta 45 kb), en especial en la preparación de genotecas genómicas de organismos superiores, ya que disminuyen en gran medida el número de clones necesarios para que todo el genoma esté representado en la genoteca (pág. 238). Los métodos de introducción del cósmido recombinante (etapa 4) dependen del tamaño del inserto. Si éste es pequeño, se puede introducir por transformación (las células se hacen *competentes* para incorporarlo en forma de plásmido recombinante, pág. 231). Si el tamaño es relativamente grande, la transformación se complica y debe acudirse al empaquetamiento previo en forma de fagos, aprovechando las secuencias *cos*.

15.2.6.4 Cromosomas artificiales

Este tipo de vectores se ha diseñado en especial para adaptarse al gran tamaño de insertos requeridos en el caso del genoma humano y de otros mamíferos, y para la clonación en células eucarióticas.

a) Cromosomas artificiales bacterianos

Suelen denominarse habitualmente BAC (*bacterial artificial chromosomes*). Se trata de vectores sintéticos derivados del plásmido F o del fago P1. Se caracterizan, en primer lugar, por aceptar grandes insertos de DNA, entre 100 y 300 kb. En comparación con otros tipos de vectores, los BAC contienen genes que reducen su capacidad de copia, es decir, de replicación en la célula anfitriona del rDNA formado (etapa 4). Como resultado, se consigue mayor estabilidad de los insertos de DNA de origen eucariótico, que en cósmidos sufren con frecuencia reordenamientos internos. Gracias a estas dos características, los BAC son vectores adecuados para la elaboración de genotecas eucarióticas (pág. 236).

b) Cromosomas artificiales de levadura

Los YAC (*yeast artificial chromosomes*) son vectores sintéticos que contienen las tres regiones relacionadas con la funcionalidad de un cromosoma (pág. 93): secuencias constituyentes de un centrómero, de dos telómeros y de un origen de replicación. Además, incorporan en general genes marcadores seleccionables (etapa 6) y un sitio de clonado múltiple (etapa 2) en el que admiten un inserto de gran tamaño, que puede alcanzar 1 Mb, lo que constituye una característica fundamental. En ocasiones, este tipo de vectores se denominan ARS/CEN/TEL, por los 3 tipos de secuencias que contienen. La secuencia centromérica permite la segregación correcta de las copias hijas del cromosoma durante la división de la célula (pág. 101), mientras que los telómeros estabilizan, como ya se ha estudiado, los cromosomas a lo largo de sucesivas divisiones (pág. 158). El origen de replicación o secuencia de replicación autónoma (ARS, *autonomously replicating sequence*) permite la formación de copias del cromosoma artificial de forma independiente a la del genoma propio de la célula anfitriona. Los YAC se utilizan en particular para preparar genotecas de genomas muy grandes, o para mantener un gen completo en un único clon.

15.2.7 Incorporación del DNA recombinante a la célula anfitriona

La replicación de las moléculas de rDNA (clonación) requiere que éstas se introduzcan en una célula anfitriona (receptora, hospedadora u hospedera; el término *célula huésped*, usado con frecuencia, no transmite inequívocamente el significado). Para que esta incorporación, que en general tiene lugar por mecanismos poco conocidos, sea eficaz es preciso emplear células especializadas (con un genotipo seleccionado), adecuadas para el vector de clonación utilizado y con capacidad para dividirse rápidamente.

15.2.7.1 Tipos de célula anfitriona

Con frecuencia, las células anfitrionas son procarióticas, principalmente cepas de *E. coli*, porque son fáciles de obtener, manipular y multiplicar, y porque se conoce bien su genética. El proceso de introducción del rDNA plasmídico se denomina *transformación* (pág. 231). La o las moléculas de rDNA plasmídico incorporadas se replican luego varias veces dentro de la célula, de acuerdo con las propiedades naturales de los plásmidos. El inserto de DNA puede permanecer unido al plásmido (lo más frecuente) o recombinarse con el DNA del cromosoma. Si se integra en el vector más de un tipo de DNA –por ejemplo, dos genes o un gen y un regulador de la transcripción– el proceso de incorporación del rDNA a la célula anfitriona recibe el nombre de *cotransformación*.

Las células anfitrionas eucarióticas son más difíciles de preparar y mantener, por lo que sólo se utilizan con fines específicos, fundamentalmente para el análisis de la regulación génica, la expresión de una proteína eucariótica, etc. De entre ellas, las más manejables en cultivo son las **levaduras** que, aunque mantienen algunas características similares a procariotas (células individuales, crecimiento rápido y económico, manipulables,

etc.), tienen una organización celular propia de eucariotas y son de interés biotecnológico por su capacidad de expresión y maduración de proteínas.

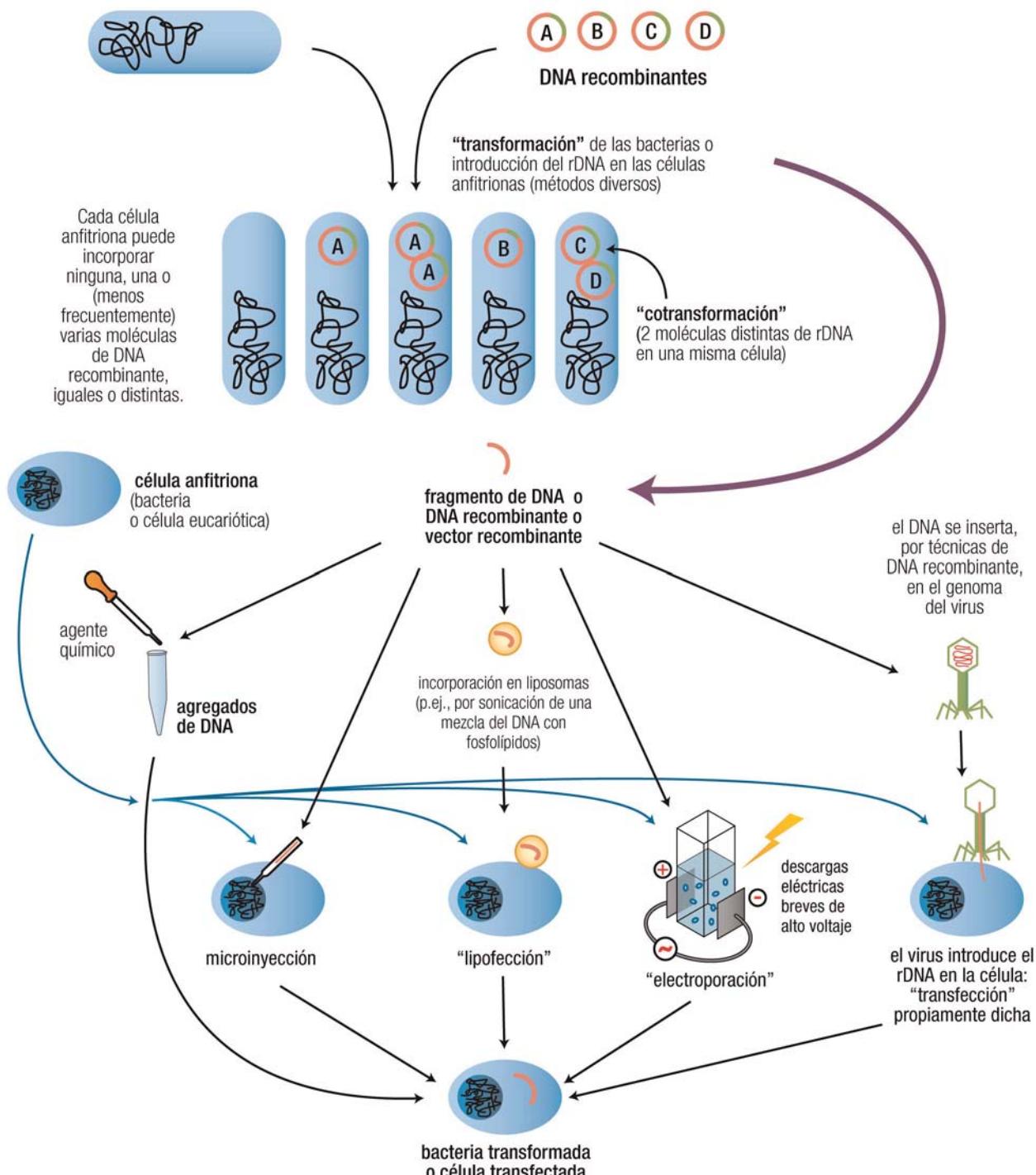
Entre las **células animales** que pueden establecerse en cultivo, se pueden citar las células HeLa de carcinoma epidérmico humano, los fibroblastos NIH de ratón y BHK de riñón de hámster, los oocitos de ratón, etc. En general, pueden dividirse multitud de veces conservando los mismos caracteres, por lo que su rendimiento de clonación es elevado.

15.2.7.2 Métodos de introducción del DNA recombinante

Puesto que habitualmente la membrana plasmática muestra una permeabilidad selectiva y no permite el paso de las moléculas de rDNA, se acude a diversos métodos para facilitar su captación por alteración transitoria de la permeabilidad celular (métodos pasivos) o para introducir el rDNA directamente (métodos activos). En muchos casos, los mecanismos son poco conocidos y los métodos resultan más o menos eficaces dependiendo del tipo de vector y de célula. El proceso suele denominarse *transformación* o *transfección*, aunque estrictamente el primer término corresponde a la introducción en bacterias anfitrionas y el segundo al empleo de virus como vectores. Los métodos utilizados pueden dividirse también en físicos y químicos.

- El tratamiento más frecuente es de tipo físico-químico y se dice que la célula se hace **competente**. Se comienza con una incubación en presencia de CaCl_2 a 0 °C, seguida de un calentamiento brusco hasta 37-43 °C, lo que induce a las bacterias a aceptar moléculas de DNA. Aunque el mecanismo de este proceso no se ha podido confirmar, posiblemente la neutralización de la carga del DNA por parte de los iones calcio posibilite el paso del DNA a través de la membrana celular. Se dispone de cepas establecidas que son especialmente susceptibles a este proceso.
- La incorporación pasiva de DNA se favorece cuando sus moléculas se han agregado o coprecipitado por interacción con **cationes**, habitualmente fosfato cálcico o DEAE-dextrano (dieltilaminoetil-dextrano, un polímero poliacidólico).
- Otros agentes químicos, en especial el polietilenglicol (PEG), aumentan la permeabilidad de la membrana celular, con lo que se incorpora más fácilmente el rDNA, también de forma pasiva.
- El DNA se puede incorporar en liposomas, que luego con facilidad se fusionan con la membrana celular y liberan su contenido al interior (incluyendo el DNA aunque, en realidad, parece más bien asociarse a la superficie exterior hidrófila del liposoma).
- Un método con elevada eficacia es el empleo de vectores de tipo **vírico**, que por su vía natural de infección introducen en la célula anfitriona su DNA, en este caso recombinante. Para este tipo de incorporación se ha acuñado el término *transfección*.
- Aplicando **descargas eléctricas** en forma de pulsos breves de alto voltaje se consigue la apertura de poros transitorios en la membrana, por los que puede pasar el rDNA. Este método de *electroporación* es bastante habitual para células eucarióticas, aunque de resultados muy variables.
- Un método curioso y de aplicación creciente, la “**biolística**”, consiste en el **bombardeo** de las células anfitrionas con microesferas (de tungsteno u oro) recubiertas con el rDNA. Se emplean dispositivos adecuados (“pistolas de genes”) que impulsan esos microproyectiles mediante un gas a presión, y pueden aplicarse sobre suspensiones celulares, células en cultivo o incluso *in vivo* (sobre la piel, hojas, etc.). Algunos de los proyectiles impactan y penetran en las células, cuya membrana se resella rápidamente, quedando el DNA incorporado en su interior.
- En casos en los que la eficiencia de suministro del rDNA debe ser máxima se acude a la **microinyección** directa del DNA en la célula o en su núcleo, empleando micromanipuladores y agujas de vidrio (micrjeringas o micropipetas, de punta extremadamente fina) (pág. 194), bajo observación al microscopio. Obviamente, el inconveniente es el escaso número de células que se pueden tratar por este método manual, por lo que en general se utiliza sólo sobre oocitos y cigotos. Dado su carácter directo, se puede incluso prescindir del vector y suministrar el fragmento de DNA deseado sin modificar.

Detalles del procedimiento de clonación: incorporación del rDNA (etapa 4)



Agente promotor de la incorporación	Nombre específico del método	Carácter	
Ca ²⁺ (generalmente como fosfato)	Transformación, células competentes	Químico	Pasivo
DEAE-dextrano		Químico	Pasivo
PEG (polietilenglicol)		Químico	Pasivo
Liposomas	Lipofeción	Físico	Pasivo
Vectores virales	Transfección	Físico	Activo
Descargas eléctricas	Electroporación	Físico	Activo
Microproyectiles	Biolística	Físico	Activo
Inyección directa con micropipetas	Microinyección	Físico	Activo

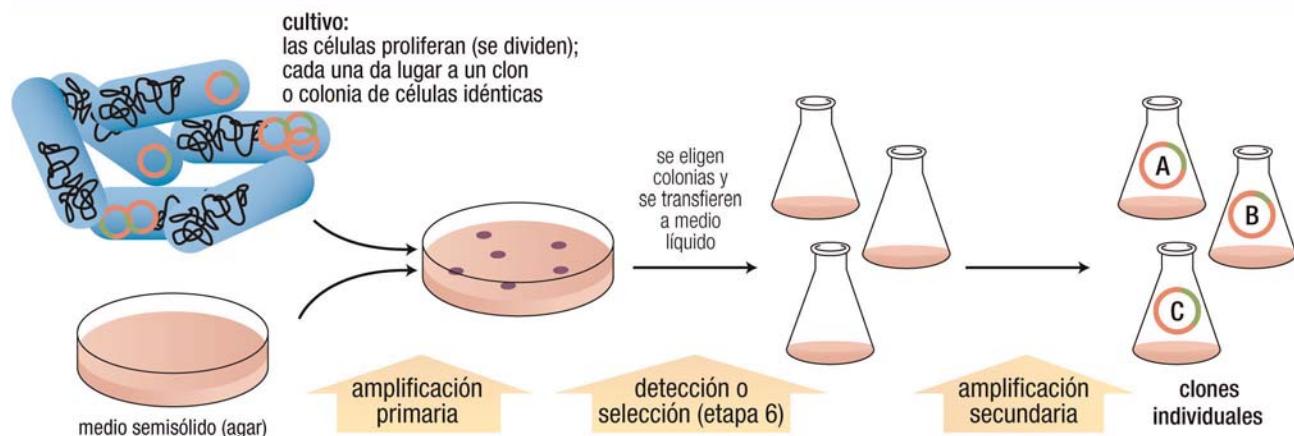
15.2.8 Propagación en cultivo

La **división** de las células anfitrionas en cultivo producirá, junto con la replicación de su propio genoma, la formación de copias del rDNA que haya podido incorporar y, por tanto, la **clonación** propiamente dicha.

El proceso de incorporación del DNA a la célula anfitriona posee una tasa de eficacia muy baja. Esto significa que sólo una pequeña fracción de las células presentes incorporan rDNA y, en consecuencia, por razones estadísticas las que lo hacen captan una única molécula. Esto tiene dos consecuencias importantes. En primer lugar, es esencial en la práctica aplicar métodos de detección y selección de las células que han incorporado rDNA (etapa 6). En segundo, la progenie de cada célula –una *colonia* en el cultivo celular– contiene copias de una única molécula de DNA: el clon celular es también un clon de DNA, no hay mezcla de moléculas de DNA de secuencia diferente. Es importante, pues, propagar las células en tal dilución que asegure que cada colonia esté físicamente separada del resto.

En la práctica, suele hacerse una primera siembra celular sobre una placa de agar para que crezcan las colonias y una resiembra posterior de éstas en medio líquido, en la que se forman poblaciones con múltiples copias de cada célula original (clones celulares). Además, dependiendo de las propiedades del vector, el rDNA puede replicarse de forma rápida e independiente del genoma de la célula anfitriona, con lo cual el número de copias del DNA recombinante es aún mayor.

Detalles del procedimiento de clonado: propagación celular (etapa 5)



Simultáneamente con esta etapa de propagación suele tener lugar la selección de clones recombinantes y la expresión opcional del gen clonado.

15.2.9 Detección y selección de los clones recombinantes

Dado el gran número de células empleadas en la clonación, y la diversidad de productos que pueden originarse en el proceso de formación del DNA recombinante, es esencial poder diferenciar qué células (o colonias a las que han dado lugar en cultivo) han incorporado realmente el DNA recombinante deseado. Esta *detección* se puede combinar con una *selección*, es decir, conseguir que sólo las células adecuadas sobrevivan en el cultivo. Esto facilita mucho la tarea de propagación celular, al reducir el número de clones que se deben seguir cultivando. En función del método utilizado, es posible distinguir células que han incorporado el vector (recombinante o no), que han incorporado rDNA o incluso que han incorporado el rDNA con el inserto dispuesto correctamente.

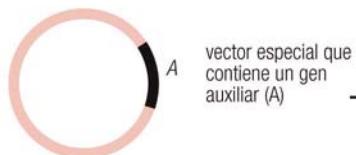
Se pueden considerar tres categorías de métodos de detección:

- **Métodos de hibridación:** detección de una secuencia del DNA clonado, o del mRNA transcrita a partir de él, por hibridación con una sonda marcada (pág. 170).
- **Métodos inmunoquímicos:** detección de la proteína codificada por el gen clonado, mediante un anticuerpo específico.
- **Métodos genéticos o de selección fenotípica:** en ciertos casos, la expresión del propio inserto produce cambios apreciables en la célula anfitriona, que sirven para reflejar su clonación con éxito. Sin embargo, la estrategia más común es la detección y selección basadas en la presencia de **genes marcadores** en el vector:
 - **Genes que codifican una proteína detectable**, habitualmente una enzima para la que se dispone de un ensayo. Por ejemplo, la β -galactosidasa (gen *lacZ* de *E. coli*) transforma el sustrato sintético X-gal (5-bromo-4-cloro-3-indolil- β -D-galactopiranósido) en un producto de color azul.
 - **Genes de resistencia a un antibiótico**, que codifican una proteína que permite el crecimiento de la célula anfitriona en presencia de un determinado antibiótico. Por ejemplo, el gen *amp^r* codifica una β -lactamasa capaz de degradar la ampicilina, una penicilina semisintética. El gen de resistencia a tetraciclina (*tetA*) codifica una proteína de membrana que actúa como una bomba, extrayendo este antibiótico de la célula.
 - **Genes que complementan defectos nutricionales**. En este caso, se emplean cepas mutantes de la célula anfitriona, que no pueden crecer en ausencia de un determinado nutriente. Por ejemplo, clonando en levaduras con un defecto metabólico en la ruta de biosíntesis de triptófano, se emplea un vector que contiene el gen *TRP1* correcto, con lo que sólo las células transformadas crecen en un medio carente de Trp.

Los planteamientos experimentales son muy diversos y en general combinan varios de estos marcadores. Por razones didácticas, se recogen unos ejemplos que ilustran el uso de varios genes marcadores, aunque no se ajusten a la práctica real del laboratorio. Un concepto importante a este respecto es el de *inactivación insercional*, consistente en la inclusión de un policonector o sitio de clonado múltiple (págs. 224 y 228) en el interior de un gen marcador. Como consecuencia, la acción de éste se ejerce sólo en las células que han incorporado vector sin inserto.

Detalles del procedimiento de clonación: detección y selección de clones recombinantes (etapa 6)

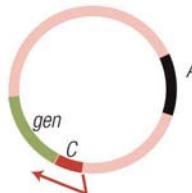
Ejemplo de un vector para clonación:
contiene el inserto de DNA (gen de interés), una región de control y un gen auxiliar (gen marcador o gen de resistencia a antibióticos)



fragmento de DNA con el gen que se quiere clonar

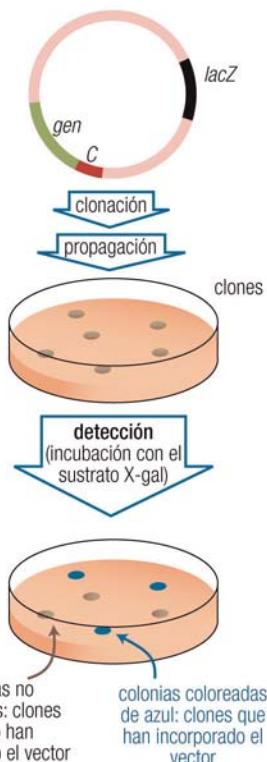
fragmento de DNA con una región de control (promotores de transcripción)

restricción y ligamiento



la región de control favorece y permite regular la expresión del gen de interés

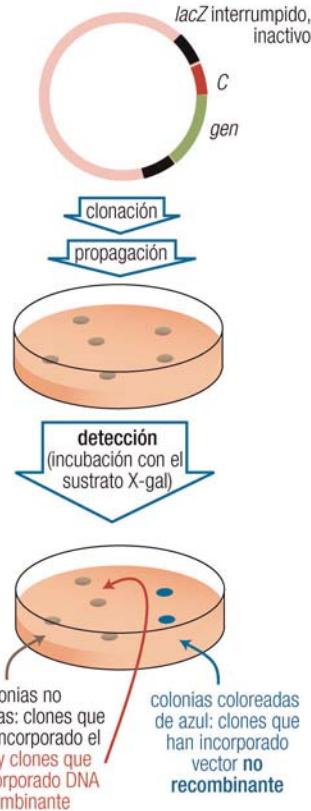
Detección de clones:
el gen A sería un gen marcador, codifica una proteína que se puede detectar



Selección de clones:
(también sirve como detección)
el gen A sería un gen de resistencia a antibiótico



Detección de clones por inactivación insercional: el gen A sería un gen marcador, codifica una proteína que se puede detectar



15.20

www

Web 15.4. Fundamento y ejemplos del uso de genes marcadores para la detección y selección de clones.

15.2.10 Expresión génica

En algunos casos se pretende no sólo la amplificación del DNA clonado, sino también su expresión, bien para estudiar este proceso o bien para investigar o aplicar el producto (RNA o proteína). Ésta es la gran diferencia entre los objetivos o aplicaciones de la clonación acelular *in vitro* (PCR) y la clonación celular, que abre la puerta a una gran área aplicada: el uso de las *proteínas recombinantes*. Hablamos de la posibilidad de *fabricar* proteínas, en cierta medida “a la carta”, con aplicaciones en terapia, diagnóstico, sensores, detergentes, producción de alimentos y otros muchos ámbitos.

Lógicamente, el DNA clonado será en estos casos un gen y, por ello, con frecuencia se tratará de un cDNA. Para conseguir esta expresión se emplean *vectores de expresión*, distintos de los vectores de clonación porque presentan características que favorecen la transcripción del gen una vez incorporado a la célula anfitriona. Habitualmente, un vector de expresión incluye un **promotor**, que es una región reguladora de la transcripción (pág. 288). (Si el vector no lo incluye, debe insertarse a la vez que el inserto del gen, pág. 222.) Es frecuente emplear promotores inducibles, ajenos al organismo del que procede el gen clonado pero que se activan bajo condiciones experimentales controlables, como al añadir un nutriente u otro tipo de compuesto, disparando entonces la transcripción del gen clonado.



Web 15.5. Ejemplo de clonación con vectores de expresión.

Una de las aplicaciones particularmente útiles, y de uso cada vez más frecuente con fines de investigación, es la producción de las llamadas *proteínas de fusión*, proteínas recombinantes que combinan en un mismo polipéptido la proteína de interés y una proteína marcadora o etiqueta que, por sus propiedades, permitirá el seguimiento y la detección de la primera.

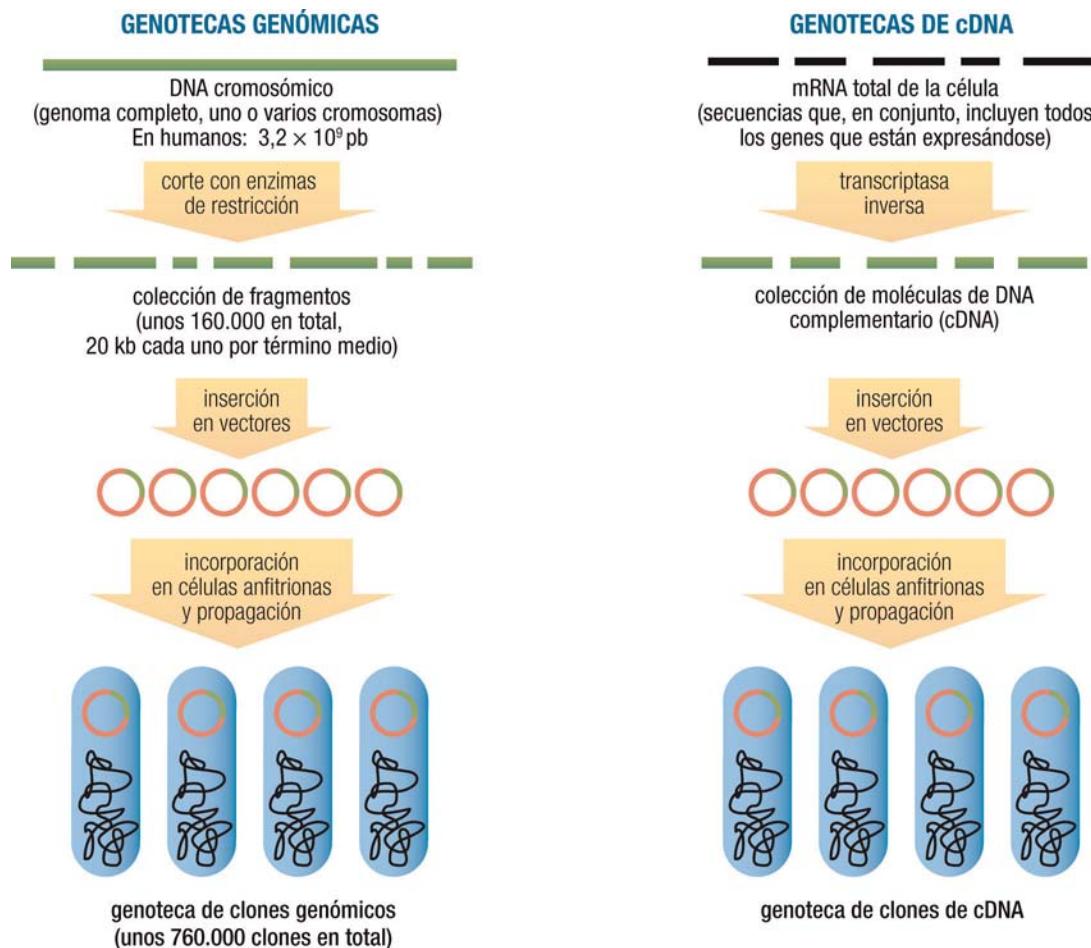


Web 15.6. Preparación de proteínas de fusión.

15.3 GENOTECAS

Se llama genéricamente **genoteca** o **biblioteca de DNA** a una colección de clones preparada con todos los fragmentos de DNA obtenidos previamente por escisión del genoma de una especie o de un tejido. Es decir, se hace la clonación en paralelo de todos los fragmentos de DNA que componen un genoma. En el caso de la clonación celular, cada uno de estos fragmentos se presenta integrado, bajo la forma de rDNA, en un vector incorporado a uno de los clones celulares que componen la genoteca. En la clonación acelular se presentan como fragmentos individuales independientes (al no existir vector ni célula anfitriona).

Dependiendo de si las muestras de partida para la clonación son fragmentos del DNA genómico o moléculas de mRNA (págs. 221-223), se distingue entre **genotecas genómicas** y **genotecas de cDNA**. Es importante indicar que la información proporcionada por ambas genotecas no es idéntica, pues sólo las primeras incluyen tanto las regiones codificantes como las no codificantes; además, las genotecas de cDNA son diferentes según el tejido particular con el que se han preparado, pues sólo contienen las regiones codificantes que se expresan en ese tejido, tipo celular, etapa del desarrollo o diferenciación, entorno bioquímico, etc.

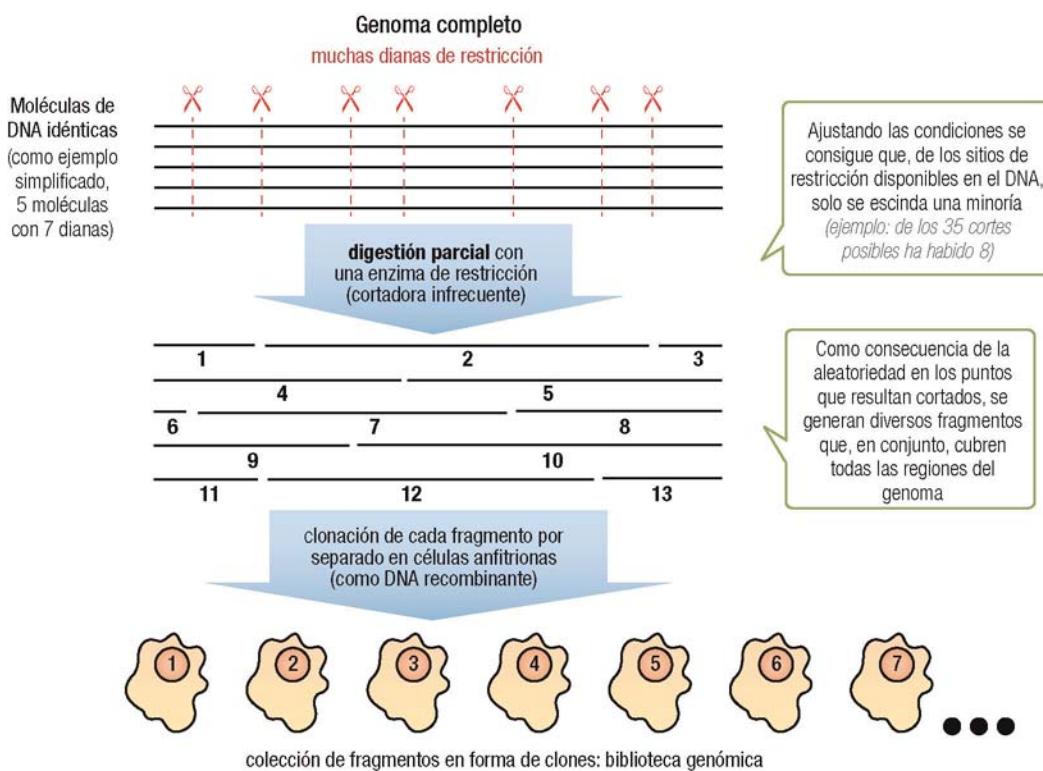


15.21

Una vez preparada una genoteca, ésta se convierte en el punto de partida para la búsqueda de cualquier región del DNA o de genes concretos, y para su estudio mediante clonación adicional (que recibe el nombre de *subclonación*), secuenciación, expresión, etc. Con respecto al uso de las genotecas como fuente para otros estudios, el problema genérico que se plantea es la localización en la genoteca del clon que contiene el DNA buscado. Para ello se acude a los distintos métodos de detección y selección, previamente comentados al describir la clonación de un solo fragmento (pág. 234).

15.3.1 Genotecas genómicas

El concepto original de genoteca correspondía estrictamente al de genoteca genómica, representativa del genoma completo. El número de clones componentes debe ser, en la práctica, suficiente para abarcar todo el genoma, incluyendo tanto el DNA codificante como el no codificante (Capítulo 9). La preparación del DNA de partida se realiza por **escisión** del genoma con enzimas de restricción bajo condiciones de rotura parcial, es decir, que no permitan la actuación exhaustiva de la enzima sobre todos los sitios de restricción presentes en el DNA. La rotura se produce en cada molécula aleatoriamente sobre el conjunto de sitios de restricción, de forma que aparecen un número menor de fragmentos y éstos son de mayor tamaño, respecto a lo teóricamente posible, y diferentes de una a otra molécula. También se favorece este resultado usando enzimas “cortadoras infrecuentes”, aquellas cuya secuencia diana aparece pocas veces en el genoma.



15.22

A partir del tamaño del genoma completo y el tamaño medio de los insertos se puede calcular el mínimo número de clones necesario para que la genoteca represente al genoma completo. Sin embargo, el carácter aleatorio de la preparación de fragmentos y del proceso de clonación hace que, para tener certeza de que una genoteca contiene cualquier región del genoma, se requiera en la práctica un número de clones muy superior al teórico. Este número de clones dividido por el número de clones teóricos define el parámetro conocido como **equivalente genómico**, que debe alcanzar valores de 10 para asegurar un 99% de probabilidad de cobertura completa.

Relación entre tamaños de genoma e inserto y número de clones teóricos y prácticos

Organismo	Tamaño del genoma (Mb)	Tamaño de inserto (kb)	N.º de clones independientes (teóricos)	N.º de clones en la práctica	Equivalente genómico
Bacterias	~ 4	20	$4.000 / 20 = 210$	~900	$900 / 210 \cong 4$
Levaduras	~14	20	$14.000 / 20 = 700$	~3.500	$3.500 / 700 \cong 5$
Humanos	~3.200	20	$3.200.000 / 20 = 160.000$	~760.000	$760.000 / 160.000 \cong 4,7$

15.3.2 Genotecas de DNA complementario

A diferencia del caso anterior, las genotecas de cDNA representan de manera exclusiva a la parte codificante del genoma. Son especialmente útiles en eucariotas, donde una célula expresa una fracción muy pequeña de su genoma (menos del 3% en humanos), lo que reduce el tamaño del material genético necesario para construir la genoteca y, por tanto, el número de clones. El material de partida, el mRNA de la célula, comprende menos de 20.000 moléculas diferentes, con un tamaño medio de 2 a 3 kb. Por otro lado, dado que el número de copias de los distintos tipos de mRNA es muy variable en eucariotas, su representación en la genoteca no será uniforme. Puesto que los cDNA carecen de señales de expresión, cuando se pretenda la expresión del gen clonado es imprescindible incluir en el vector de expresión un potente promotor adecuado para la célula anfitriona, de modo que ésta exprese la proteína codificada por el cDNA respectivo.

Genómica, cartografía del genoma y secuenciación



16.1 INTRODUCCIÓN	239		
16.2 MAPAS GENÉTICOS	240		
16.2.1 Fracción de recombinación como medida de distancia en un mapa genético	241	16.3.2.2 Mapa físico con enzimas de restricción	248
16.2.2 Obtención del mapa genético	243	16.3.2.3 Mapas físicos de STS y EST	248
16.3 MAPAS FÍSICOS	244	16.3.2.4 Mapas de cíntigos	250
16.3.1 Obtención de mapas físicos a baja resolución	244	16.4 SECUENCIACIÓN	252
16.3.1.1 Métodos basados en empleo de líneas celulares somáticas híbridas	244	16.4.1 Método químico de secuenciación del DNA	252
16.3.1.2 Obtención de mapas mediante ensayos de hibridación con sondas fluorescentes	246	16.4.2 Método enzimático de secuenciación del DNA	252
a) Mapas por hibridación <i>in situ</i> (FISH) cromosómica	246	16.4.2.1 Síntesis de un DNA complementario de longitud variable	253
b) Mapas por citometría de flujo	247	16.4.2.2 Separación de los fragmentos y análisis	254
16.3.2 Mapas físicos de alta resolución	247	16.4.2.3 Automatización	255
16.3.2.1 Mapa físico por hibridación <i>in situ</i> (FISH) de alta resolución	247	16.4.2.4 Variantes del método	256
		16.4.3 Métodos de secuenciación masiva	257
		16.4.4 Secuenciación del RNA	257
		16.5 EL PROYECTO GENOMA	257

16.1 INTRODUCCIÓN

Se conoce como *genómica* el conjunto de estrategias y tecnologías empleadas para la caracterización molecular del genoma y, en general, el estudio del genoma de un organismo en su conjunto (por contraposición a cuestiones que afecten a genes individuales). Una buena parte de la genómica corresponde a la cartografía, u obtención del **mapa de un genoma**, que a su vez puede considerarse como la *estructura fina* del material genético de cualquier especie. Por tanto, constituye también la base de una biología molecular e ingeniería genética modernas aplicadas a la sanidad.

De acuerdo con estos objetivos, se pueden distinguir dos **áreas** dentro de la genómica, consecutivas, en apariencia independientes pero totalmente relacionadas entre sí:

En primer lugar, la **genómica estructural**, que corresponde al estudio de las técnicas y estrategias necesarias para obtener tanto los mapas físicos del genoma, posibles a distintos niveles de resolución, como la secuencia completa del DNA genómico, o mapa físico a la resolución máxima, el nucleótido individual.

En segundo lugar, la **genómica funcional**, correspondiente a la caracterización del **proteoma**, es decir, el conjunto de todas las proteínas originadas por el genoma, bajo los diversos patrones posibles de expresión génica a proteínas, modificación y degradación de éstas y desarrollo de su función, todo ello dentro de cada tipo de célula, momento celular, condiciones particulares *in vivo*, etc. Corresponde, en definitiva, al estudio funcional del DNA codificante. Por analogía, se denomina **proteómica** al conjunto de técnicas y estrategias utilizadas para el estudio del proteoma. En contraste con la genómica estructural, la proteómica progresó con un desfase, y el número de organismos modelo para los que se dispone de datos es más limitado. Aunque los avances conseguidos en los últimos 10 años son extraordinarios, aún se necesita tiempo para conocer al completo la función de todos los genes humanos y sobre todo para dar utilidad práctica, en su caso, a las proteínas, tanto con fines básicos como aplicados en favor de la humanidad (por ejemplo, a la biomedicina).

Los Proyectos Genoma poseen tal envergadura que sus objetivos no se pueden alcanzar con el simple planteamiento de buscar métodos para la etapa de secuenciación, a pesar de su creciente eficiencia. De hecho, el progreso conseguido ha supuesto la implicación de numerosas iniciativas y metodologías que superan con

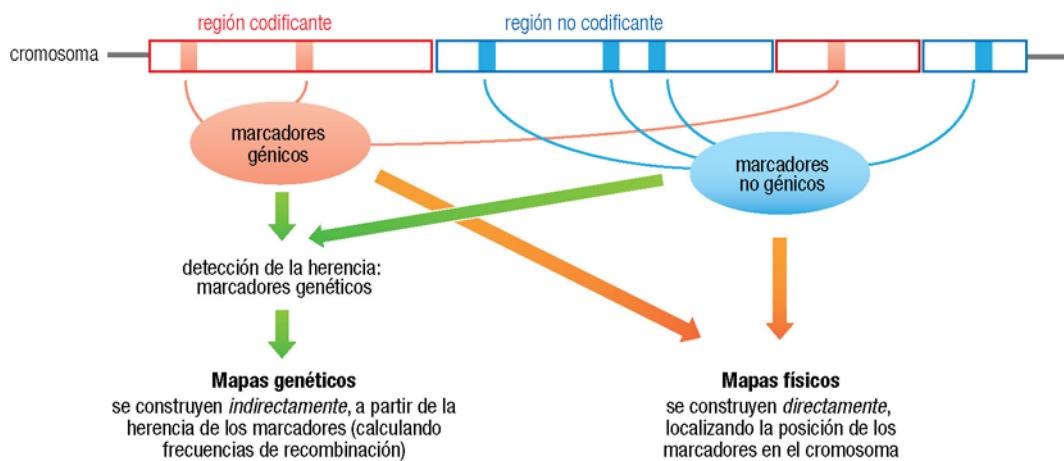
creces dicha etapa. Ante la imposibilidad material de recoger toda esta información (por otro lado, ya disponible de forma especializada día a día en internet), se presentan en este capítulo las líneas básicas para un enfoque “lógico” del conocimiento del genoma.

La elaboración del mapa, o cartografía, de un genoma, tanto del ser humano como de otros organismos, se ha abordado a lo largo del tiempo en varias etapas consecutivas, en niveles crecientes de resolución. Éstos se corresponden, a grandes rasgos, con tres grandes estrategias metodológicas, complementarias entre sí: **cartografía genética, cartografía física y secuenciación**. Además, ha sido necesario el desarrollo de técnicas informáticas para recoger, almacenar, distribuir y analizar el creciente volumen de datos obtenidos.

Fases del Proyecto Genoma		Características
Cartografía	Mapa genético o de ligamiento	Proporciona la menor resolución. Sitúa de forma aproximada en los cromosomas diversos loci (de genes y de otros marcadores o secuencias identificables). El cálculo de la distancia y, por tanto, la ubicación relativa de genes o marcadores, se basa en la frecuencia con la que los pares de loci se transmiten conjuntamente a la descendencia
	Mapa físico	Asigna localizaciones a los marcadores, definidas por distancias en pares de bases. Destacan entre sus variantes los mapas de restricción (ordenación en el cromosoma de las dianas para enzimas de restricción)
Secuenciación		Equivale a la obtención del mapa genómico en su nivel más detallado. Su objetivo es establecer la secuencia nucleotídica completa del genoma. Se lleva a cabo en paralelo y con el apoyo de la información proporcionada por los mapas genético y físico. Hace uso de planteamientos y métodos sumamente específicos, completamente automatizados e informatizados
Trabajo posterior	Refinamiento	Se resecuencian algunas regiones para corregir errores o completar los tramos difíciles que no se habían podido analizar. Se añade la “anotación”: sobre la secuencia se marcan las posiciones para las que se dispone de información sobre función biológica: identificación de genes, regiones de control, exones, mutaciones relacionadas con enfermedades, etc.
	Variabilidad individual	Ánalisis del genoma de diferentes individuos para ubicar las posiciones de variabilidad, normal o patológica, en la secuencia. Polimorfismos genéticos y mutaciones
	Genómica comparada	Secuenciando otras especies y comparando resultados se observan los tramos fuertemente conservados, indicativos de secuencias génicas o, al menos, funcionales. La comparación ofrece además información útil sobre el desarrollo evolutivo del genoma

16.2 MAPAS GENÉTICOS

En un principio, los mapas genéticos consistían en una representación gráfica de la ordenación y ubicación de los genes en cada cromosoma. Sin embargo, hoy día se incluyen en estos mapas no sólo genes, sino también otras regiones del DNA, de modo que para englobar a ambos se emplea la expresión **marcador genético**; este papel puede desempeñarlo cualquier característica física o molecular del DNA que difiera entre individuos y que sea detectable en el laboratorio, para poder seguir su patrón de herencia. Como marcadores genéticos, por tanto, se incluyen tanto regiones que forman parte de un gen (**marcadores génicos**) como segmentos no codificantes (**marcadores no génicos**). La gran mayoría de marcadores genéticos en los mapas actuales son marcadores no génicos.



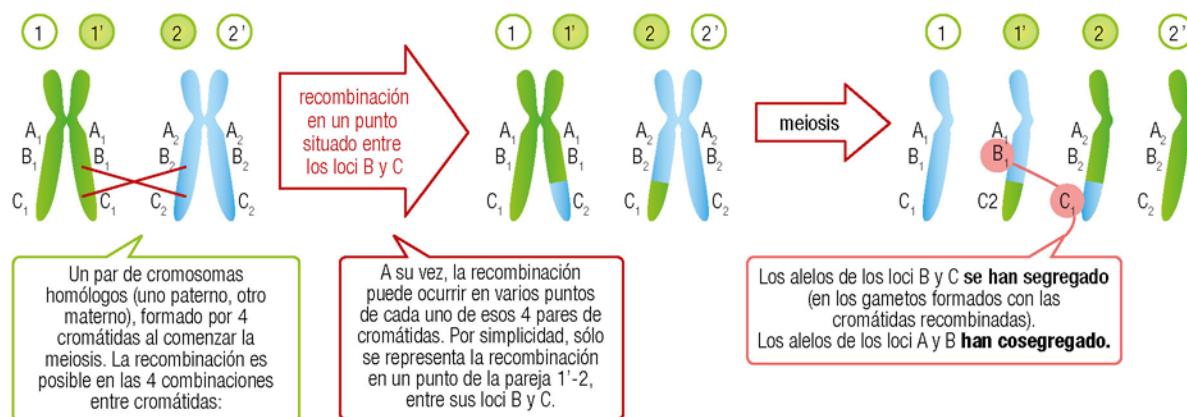
16.2.1 Fracción de recombinación como medida de distancia en un mapa genético

Durante la meiosis tiene lugar el *sobrecruzamiento*, o intercambio recíproco de fragmentos de DNA entre cromosomas homólogos (*recombinación homóloga o meiótica*, pág. 108). Debe destacarse que la recombinación nunca tiene lugar entre cromosomas diferentes, sino únicamente entre los dos miembros de un par de cromosomas homólogos. Los mecanismos moleculares de la recombinación no se describen en esta obra; puede consultarse para ello un texto de genética. Aquí interesan, esencialmente, las consecuencias del proceso. Por un lado, que los cromosomas *recombinantes* que aparecen en cada gameto ya no son idénticos a los originales (paterno y materno) y, por otro, que la variación genética resultante depende, obviamente, de la posición del sobrecruzamiento en cada par de cromosomas.

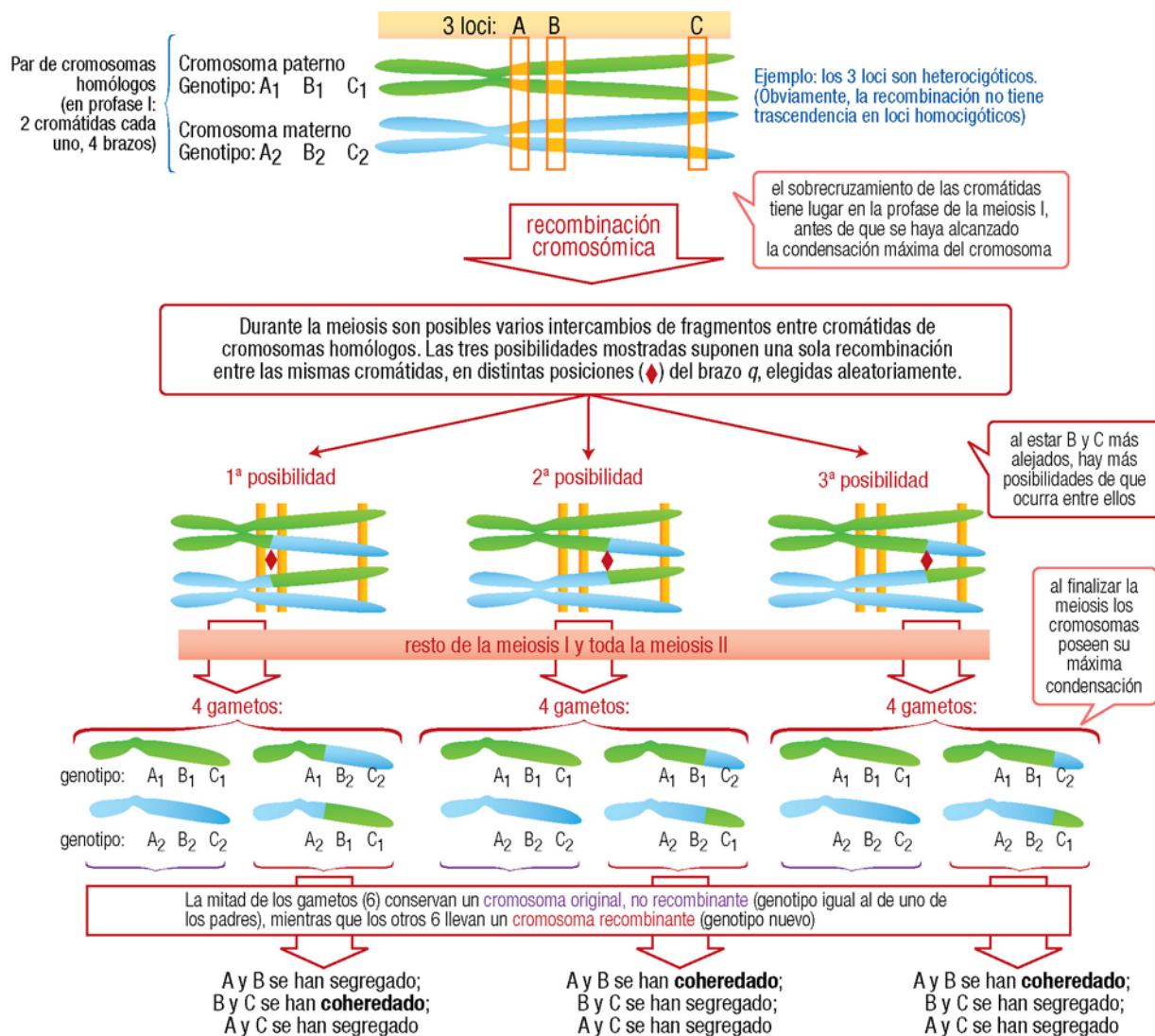
Como fase previa a la obtención de mapas genéticos, deben recordarse algunos términos básicos (v. también pág. 97 y web 8.2):

"Segregación" de cromosomas homólogos (y cromátidas) = separación de los cromosomas homólogos en la meiosis I (para formar parte de los gametocitos secundarios) y de las cromátidas en la meiosis II (para ir a parar a gametos diferentes). Cada par se segregá de forma **independiente** de los demás. Este proceso es independiente de que exista o no recombinación.

"Segregación" de los alelos = separación de los alelos de loci diferentes de un mismo cromosoma durante la meiosis (B y C). Sólo puede darse si ha **habido recombinación** entre esos loci, y se manifiesta en el momento de la segregación de los cromosomas y las cromátidas. El término "**cosegregación**" de los alelos (A y B) puede parecer矛盾, pues indica que éstos permanecen juntos, es decir, no sufren segregación entre sí.



La característica que define la cartografía genética es el análisis de la fracción o frecuencia de **recombinación**, definida como la probabilidad de que los alelos presentes en dos loci diferentes (correspondientes a la posición de los marcadores considerados) se separen o *segreguen* como consecuencia de la recombinación meiótica. Dicho de otro modo, es la frecuencia con la que dos alelos determinados aparecen recombinados en la descendencia. Como se discute a continuación, esta fracción depende de la **distancia entre los dos loci** en el cromosoma, lo que permite emplearla para establecer el mapa.



Conclusiones de interés, relacionadas con la construcción de mapas genéticos:

- a) Si dos loci están muy próximos en el cromosoma (por ej., A y B):

Es difícil que se produzca recombinación en una posición intermedia entre ambos; por tanto, hay **baja frecuencia de recombinación** (2 de un total de 12 gametos = 16,7%).

Con alta probabilidad, los dos alelos quedan en un mismo cromosoma y cromátida: *se coheredan*.

Los alelos que se coheredan con frecuencia máxima (es decir, que casi nunca se segregan) se dice que están **ligados**. Lógicamente, los loci de cromosomas diferentes nunca están ligados, o tienen un **ligamiento genético nulo**.

A un conjunto de alelos que siempre se transmiten conjuntamente (son alelos ligados) se lo denomina **haplotipo**. Habitualmente, ello se debe a que sus loci están todos localizados en una región pequeña del cromosoma.

- b) A **mayor separación entre los loci** en el cromosoma (por ej., B y C o, más aún, A y C):

Existe mayor probabilidad de recombinación en posiciones intermedias; por tanto, hay **mayor frecuencia de recombinación** (para B y C: 4 de 12 = 33,3%; para A y C: 6 de 12 = 50%).

Al sufrir la recombinación, los dos alelos quedan en distinto cromosoma, por lo que hay **menor probabilidad de que se hereden juntos**.

Nota: la terminología es aún más variada de la que se ha empleado aquí; por ejemplo, una alta frecuencia de recombinación es sinónimo de herencia conjunta infrecuente, coherencia infrecuente, segregación de alelos frecuente y cosegregación infrecuente. Esto puede dificultar, a veces, la comprensión del significado de “frecuencia de recombinación”.

16.4

16.2.2 Obtención del mapa genético

Una vez conocidas las posiciones relativas de un número elevado de marcadores genéticos, se puede estudiar la frecuencia de recombinación con respecto a ellos de cualquier nuevo gen o secuencia y así situar su locus en el “mapa” definido por aquéllos. De esa forma, se dice que se ha “cartografiado” su posición en el cromosoma. Una vez localizado, el nuevo gen o secuencia pasa a engrosar el conjunto de marcadores que forman el mapa.

La escala en los mapas genéticos se mide en unidades **centimorgan** (cM); 1 cM corresponde, por definición, a la distancia que separa dos loci que experimentan recombinación en el 1% de las meiosis. Empíricamente, en el genoma humano esta distancia corresponde aproximadamente a 10^6 pb (1 Mb). A partir de la frecuencia de recombinación se deducen las distancias relativas entre varios loci y se puede establecer su ordenación en el cromosoma, comenzando así a precisar el mapa genético.



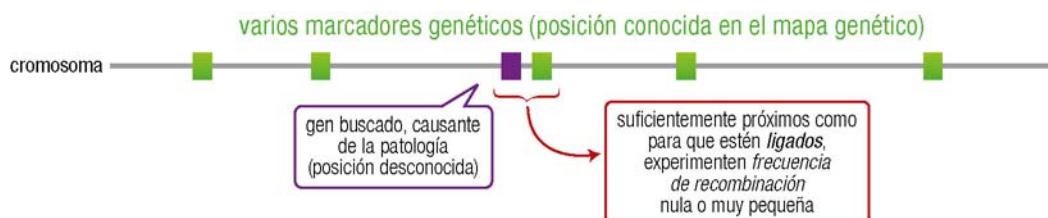
Web 16.1. Construcción del mapa genético.

Experimentalmente, los mapas genéticos requieren estudios familiares con árboles genealógicos, en los que se examina, para alguno de los marcadores, la combinación de alelos que se va heredando. Éstos se analizan en los distintos individuos a lo largo de varias generaciones. La presencia de una u otra forma alélica se detecta por los procedimientos, variados, que permitan distinguirlas: bien a través de su influencia sobre el fenotipo (caracteres observables, aparición de enfermedad, etc.) o bien mediante técnicas moleculares de detección de secuencias (hibridación *in situ* por fluorescencia, hibridación de Southern, amplificación por PCR, incluso secuenciación actualmente). A partir de la información de la herencia de alelos se calculan las frecuencias de recombinación entre éstos.

En contraste con organismos modelo experimentales –como plantas, levadura, *Drosophila* o ratón–, en los seres humanos es difícil determinar mediante estudios familiares las frecuencias de recombinación de loci génicos. Esto se debe, por un lado, a la escasez de casos en los que se transmiten de forma simultánea dos alelos representativos, por ejemplo los que causan dos enfermedades. Por otra parte, la baja tasa de reproducción del ser humano dificulta el cálculo fiable de dicha frecuencia. Para superar esta limitación, en la práctica se acude a **marcadores polimórficos**, es decir, los que presentan en la población un elevado número de alelos (formas alternativas), con lo cual son muy frecuentes los individuos heterocigóticos y la recombinación podrá detectarse en un alto porcentaje de la descendencia (debe recordarse que la recombinación de loci homocigóticos no

puede apreciarse). Los marcadores polimórficos son en general no génicos, no relacionados con enfermedades o con genes, sino simplemente secuencias de DNA detectables en una posición fija del cromosoma. Entre ellos destacan los basados en el DNA mini- y microsatélite que presentan polimorfismo en el número de repeticiones (VNTR, pág. 421). El número de marcadores polimórficos caracterizados, especialmente los microsatélites, ha crecido de forma exponencial a lo largo del tiempo. En años más recientes se han empleado cada vez más como marcadores los polimorfismos de un solo nucleótido (SNP, págs. 418 y 424). Hoy día se dispone de un mapa para cada cromosoma con la ubicación de un elevado número de marcadores.

La utilidad del mapa genético se ilustra, por ejemplo, con la posibilidad de localizar el gen responsable de una enfermedad hereditaria, aun sin conocer la identidad de dicho gen ni la base molecular de la enfermedad. Para ello, se busca, entre todos los que componen el mapa, un alelo de un marcador que esté presente en los individuos afectados, pero no en los sanos, lo que indica que el locus de dicho marcador está ligado al del gen buscado. Por consiguiente, el gen debe estar muy próximo a la posición conocida del marcador en el mapa. Esta posibilidad de ubicación de genes se potencia si se elabora el mapa con marcadores dispuestos a lo largo de todo el genoma; se asegura así, por puro azar, que siempre existe algún marcador suficientemente próximo al gen como para que ambos estén ligados. De esta forma se ha podido encontrar, por ejemplo, la localización cromosómica de los genes responsables de la fibrosis quística, la anemia drepanocítica, la enfermedad de Tay-Sachs, el síndrome del cromosoma X frágil y la distrofia miotónica.



16.5

16.3 MAPAS FÍSICOS

Como se ha indicado, estos mapas miden de manera directa distancias físicas (es decir, en pares de bases) entre genes u otros marcadores de DNA en el genoma. Las estrategias metodológicas son muy diversas y conducen a distintas variantes de mapas y niveles de resolución. Dependiendo de la escala o nivel de resolución, conocida como “unidad cartográfica o de mapa”, se distinguen tres estrategias: baja resolución (con una precisión de varias Mb), alta resolución (por debajo de 10^5 bases) y máxima resolución (secuencia nucleotídica, 1 pb).

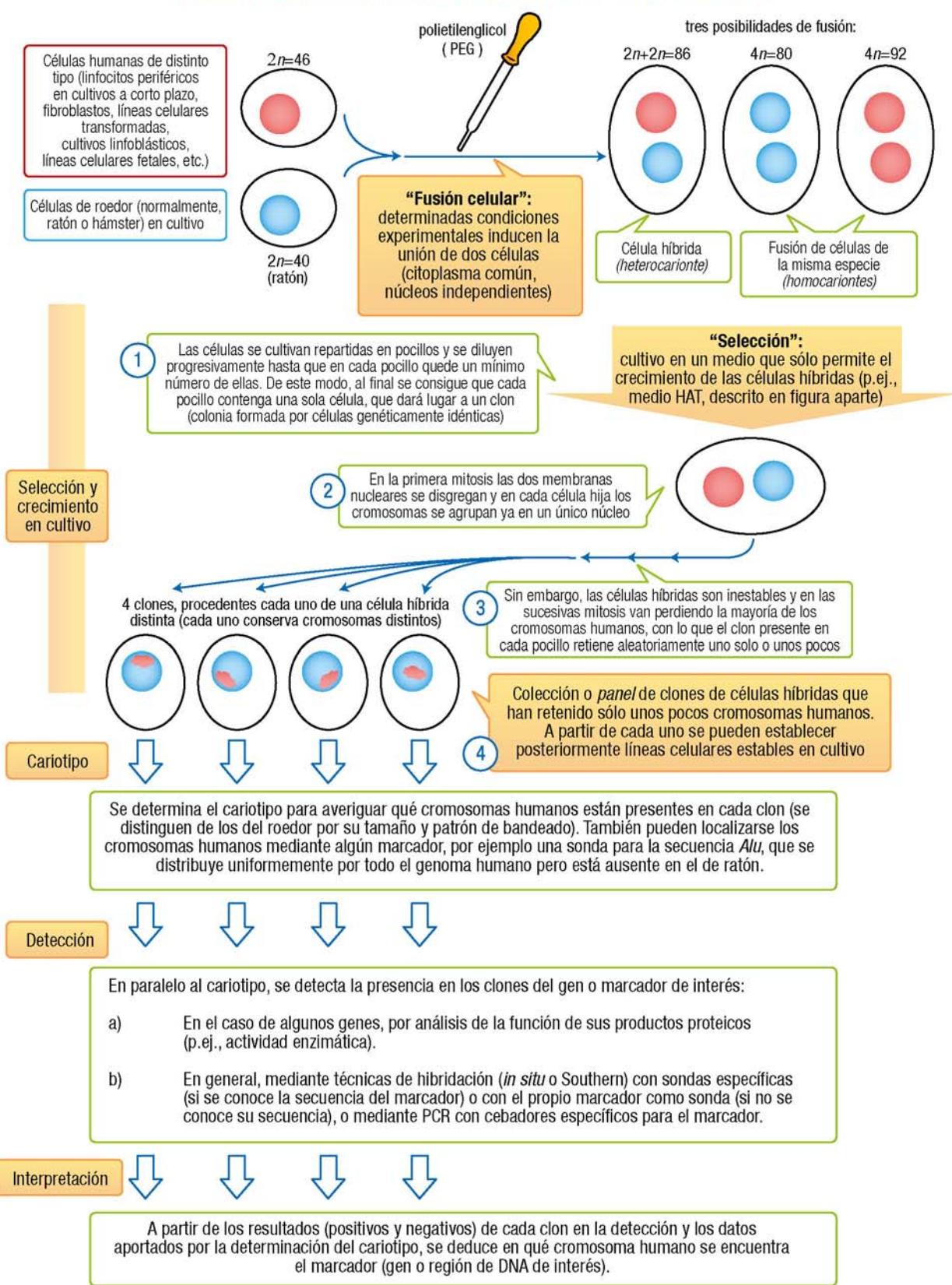
16.3.1 Obtención de mapas físicos a baja resolución

Este nivel cartográfico corresponde a la asignación de cada marcador a un cromosoma o región cromosómica. Se utilizan distintos planteamientos que, además, implican muestras diferentes, según se describe a continuación. En general, el marcador se detecta mediante el empleo de sondas o de cebadores de PCR, ambos específicos para la secuencia de aquél; en el caso de un marcador génico, puede acudirse también a la detección de su producto funcional (proteína o RNA).

16.3.1.1 Métodos basados en líneas celulares somáticas híbridas

Consisten en la búsqueda del marcador (génico o no) en líneas celulares híbridas, preparadas de tal forma que contienen material genético de dos orígenes diferentes: el genoma completo de un roedor (aunque con una mutación útil, explicada más adelante) y una parte del genoma humano que se quiere cartografiar, que puede consistir en varios cromosomas, uno solo o un fragmento de cromosoma. Disponiendo de una colección (o *panel*) de tales clones celulares híbridos, cada uno con partes distintas del genoma humano, se puede ensayar la presencia del marcador en cada uno de ellos y así ubicar en qué cromosoma o parte de éste se encuentra. Repitiendo el proceso para distintos marcadores se llega a construir un mapa físico.

TÉCNICA DE CARTOGRAFÍA MEDIANTE HÍBRIDOS DE CÉLULAS SOMÁTICAS



Ejemplo de cartografía física de un DNA de interés en cromosomas humanos. Para facilitar el análisis, se simplifica el ejemplo a 4 clones que sólo han retenido los cromosomas 1-5

		clon 1	clon 2	clon 3	clon 4
determinación del cariotipo humano presente en cada clon	cromosoma 1 cromosoma 2 cromosoma 3 cromosoma 4 cromosoma 5 resto de cromosomas	no sí no no no no	sí no no sí no no	no no sí sí no no	no no sí no sí no
resultado de la detección sobre cada clon (actividad enzimática, hibridación, etc.)		-	+	+	-

Conclusión:
la región de DNA que se está detectando (gen o DNA de interés) debe estar en el cromosoma 4

16.7

Un aspecto sumamente importante para la cartografía mediante células somáticas híbridas es la capacidad de seleccionar los híbridos, es decir, que en cultivo sólo sobrevivan estas células. (Esta misma estrategia se emplea en la preparación de anticuerpos monoclonales.) El sistema de selección más utilizado hace uso del **medio HAT**, así llamado por contener hipoxantina, aminopterina y timidina.



Web 16.2. Selección de las células híbridas empleando medio HAT.

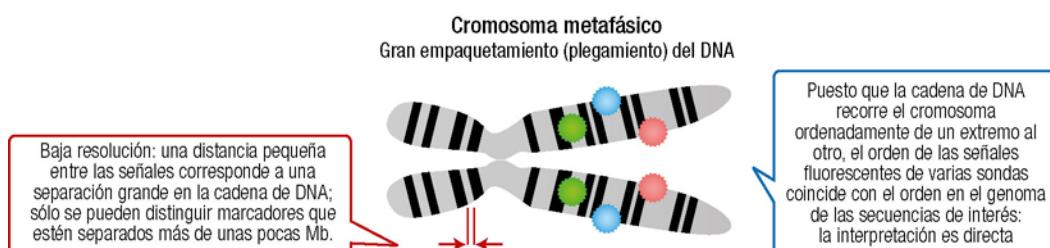
La preparación de híbridos a partir de células somáticas humanas sólo permite identificar el cromosoma en el que está situado el DNA de interés. Se puede alcanzar mayor resolución cartográfica empleando híbridos similares preparados con una porción menor del genoma humano: *híbridos monocromosómicos*, conteniendo un solo cromosoma, e *híbridos subcromosómicos*, con un fragmento de cromosoma. Con estos últimos se construye lo que se denomina “mapa de asignación regional”, es decir, la localización de cada marcador en una región determinada de un cromosoma.

16.3.1.2 Obtención de mapas mediante ensayos de hibridación con sondas fluorescentes

La ubicación de marcadores para formar un mapa físico de baja resolución puede también conseguirse realizando ensayos de hibridación con sondas que contengan la propia secuencia del marcador. Esta técnica adquirió gran aplicabilidad en la cartografía con el desarrollo de las sondas fluorescentes (**hibridación *in situ* por fluorescencia** o **FISH**, pág. 177). La sonda se marca con un fluorocromo, bien de forma directa o a través de un grupo “indicador” (por ejemplo, biotina o digoxigenina, págs. 188-189). Para identificar el cromosoma o la región del mismo donde hibrida la sonda se puede acudir a la observación bajo el microscopio de fluorescencia o a la citometría de flujo (pág. 127).

a) Mapas por hibridación *in situ* (FISH) cromosómica

La hibridación se realiza sobre preparaciones para el microscopio de cromosomas metafásicos, previamente desnaturalizados. La identidad de los cromosomas se revela por las técnicas habituales de establecimiento del cariotipo (forma, tamaño y bandas bajo tinción, pág. 90). El DNA diana (marcador) se manifiesta bajo el microscopio de fluorescencia como puntos luminosos dobles, debido a la hibridación de la sonda con las dos cromátidas (pág. 168). Se alcanza con esta técnica un nivel de resolución cromosómico o subcromosómico, y es posible localizar o cartografiar de forma simultánea varios marcadores empleando varias sondas marcadas con fluorocromos distintos.



16.8

b) Mapas por citometría de flujo

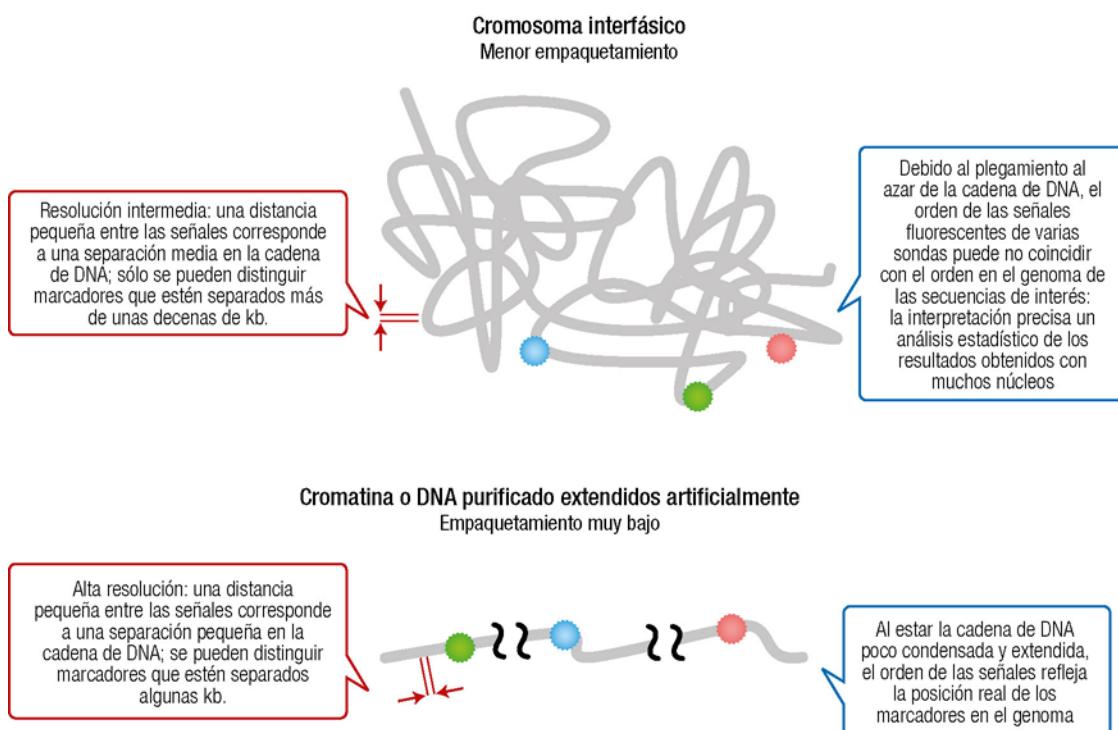
La identificación del cromosoma en el que se encuentra cada marcador puede también realizarse si la sonda se hibrida con muestras que contienen cada una un solo tipo de cromosoma. Éstas pueden conseguirse aplicando la clasificación de células activada por fluorescencia (FACS, pág. 127), adaptada a la cartografía del modo siguiente: los cromosomas se tratan con fluorocromos que se intercalan entre las bases nitrogenadas apiladas del DNA; los de mayor longitud pueden incorporar mayor cantidad de ese agente intercalante, por lo que adquieren más fluorescencia. Esto permite diferenciar los cromosomas en la etapa de análisis o caracterización, proporcionando lo que se ha llamado un *cariograma de flujo*, así como separarlos en la etapa de clasificación y obtener muestras, cada una con un solo tipo de cromosoma. Finalmente, estos cromosomas separados pueden emplearse, por ejemplo, en hibridación en formato *dot-blot* (pág. 171) con la sonda específica para el marcador buscado.

16.3.2 Mapas físicos de alta resolución

Estos métodos, complementarios a los anteriores, permiten alcanzar una resolución inferior a 1 Mb, incluso hasta llegar a unos pocos nucleótidos, lo que ya supone el **nivel molecular** de resolución; de ahí que sean los métodos de cartografía física más importantes. Sirven de enlace entre los mapas físicos anteriores y los de secuenciación o mapas de resolución total a escala de un nucleótido (pág. 252). Se plantean principalmente dos tipos de abordaje: hibridación y enzimas de restricción.

16.3.2.1 Mapa físico por hibridación in situ (FISH) de alta resolución

La baja resolución asequible sobre cromosomas metafásicos puede aumentarse mediante la hibridación sobre cromosomas menos compactados.



16.3.2.2 Mapa físico con enzimas de restricción

Antes de que se descubrieran las endonucleasas de restricción y se extendiera su uso en el laboratorio de investigación, se emplearon otras nucleasas para secuenciar RNA (pág. 257), siguiendo el mismo enfoque utilizado previamente para secuenciar polipéptidos, es decir, la fragmentación de la molécula por puntos de características conocidas y la reconstrucción del conjunto mediante solapamiento de los fragmentos. Este método no dio los resultados esperados con DNA, debido a su mayor longitud y, sobre todo, a la menor especificidad de las DNAsas, por lo que sólo se consiguió secuenciar algunos oligodesoxirribonucleótidos.



Web 16.3. Reconstrucción de la secuencia mediante alineación de fragmentos.

La ventaja de estos planteamientos iniciales es que sirvieron de pauta para la confección de los actuales mapas físicos del genoma a partir de fragmentos de DNA originados por **enzimas de restricción**. Un mapa de restricción para una región determinada del genoma está constituido por una relación ordenada de sitios que pueden cortarse con enzimas de restricción concretas, con indicación de las distancias entre ellos (en pb). A diferencia de los mapas genéticos, para su elaboración no afecta la funcionalidad de la región de DNA (su carácter génico) ni la presencia en ella de mutaciones (salvo que éstas alteren precisamente el sitio de restricción). Por el contrario, sí importan algunas alteraciones que puede sufrir el DNA, como translocaciones, delecciones o inserciones.

La resolución o escala de los mapas de restricción depende de la separación entre los sitios de corte o, dicho de otro modo, de la frecuencia con que la secuencia diana reconocida por la enzima aparece en el genoma (*frecuencia de reconocimiento* del genoma por la enzima). Por ejemplo, asumiendo que en un DNA de gran longitud las cuatro bases queden distribuidas aleatoriamente, cuanto más corta sea la secuencia del sitio de restricción, más probabilidad habrá de encontrarla, por puro azar, es decir, mayor será la frecuencia de reconocimiento (v. web 15.2).

La construcción de un mapa de restricción se simplifica si el número de sitios de restricción es reducido. Esto es especialmente necesario para cartografiar moléculas grandes de DNA, en las que las secuencias específicas para las enzimas de restricción habituales se hacen muy numerosas. Para estos fragmentos grandes la cartografía sólo se puede abordar empleando enzimas “cortadoras infrecuentes” y electroforesis en campo pulsante (pág. 142). En cualquiera de los casos, el revelado de los geles y la hibridación Southern con diversas sondas permiten deducir la distribución de los sitios de restricción, o “cartografía” del mapa.

Para facilitar la comprensión del método de obtención de mapas físicos de restricción, se proponen a continuación unos ejemplos prácticos. En el primer caso, se emplean dos enzimas (A y B), que actúan de forma aislada o conjunta (experimento de digestión doble). En el segundo ejemplo, las dos enzimas actúan de forma sucesiva. En el tercero, la digestión enzimática es idéntica al primero, pero se emplea la hibridación para enriquecer la información obtenida.

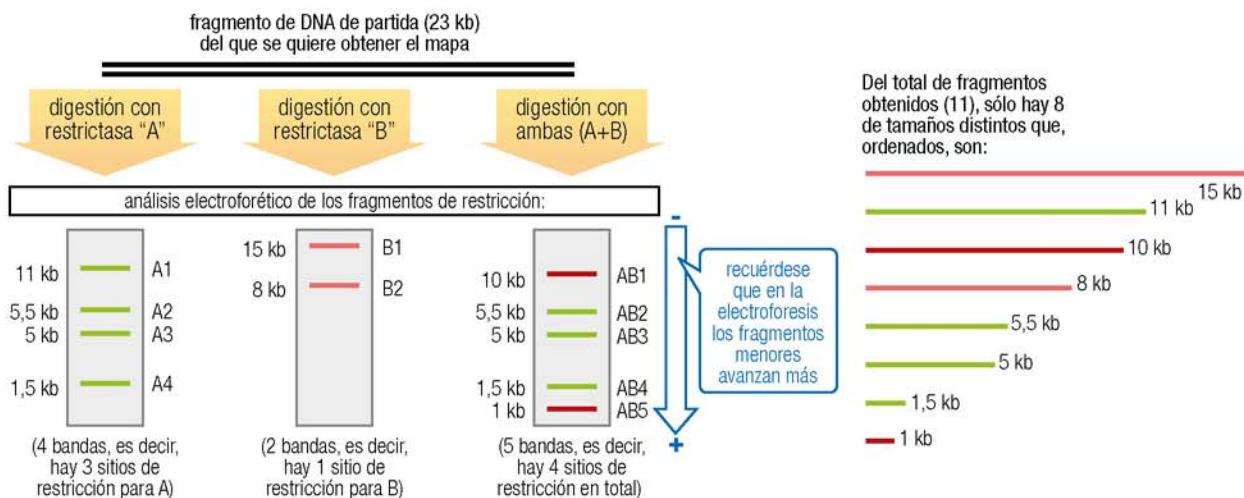
Finalmente, debe resaltarse que, puesto que un mapa de restricción no identifica la posición de genes, su utilidad es tanto mayor cuando más pueda relacionarse con un mapa genético. Para establecer esta relación es necesario encontrar mutaciones que afecten a los sitios de restricción y determinar su ligamiento con alteraciones que afectan a genes, como las que producen enfermedades u otros cambios fenotípicos. Ello permite situar ese gen dentro del mapa de restricción (y otros que puedan haberse localizado respecto a aquél en un mapa genético); éste es un ejemplo de cómo se complementan los diversos abordajes de cartografía para ir construyendo el mapa completo del genoma.

16.3.2.3 Mapas físicos de STS y EST

Cada vez con mayor frecuencia, para la construcción de mapas físicos se utilizan marcadores no génicos (pág. 240). Dos de los tipos de marcador más extendidos son los **STS** (*sequence-tagged sites*), también llamados “secuencias de referencia”, “sitios marcados únicos” o “sitios de secuencia marcada”, y los **EST** (*expressed sequence tags*), también denominados “etiquetas o marcas de secuencias expresadas”. En ambos casos, se trata de regiones cortas del genoma (entre 50 y 500 pb) que se han podido secuenciar y para las que se han desarrollado cebadores de PCR (Capítulo 14), lo que permite su amplificación y consiguiente detección. Reciben el nombre de STS los marcadores encontrados en el DNA genómico, mientras que los EST se identifican en moléculas de cDNA, resultantes de la transcripción inversa de mRNA. En consecuencia, los STS son marcadores que sirven para elaborar mapas físicos de todo el genoma, mientras que los EST sólo sirven para mapas de la parte del genoma de

①

Ejemplo de aplicación de las enzimas de restricción para investigar la posición de determinadas secuencias de DNA (cartografía física), dentro de un fragmento de genoma de interés (por ejemplo, por la presencia de un gen)



Resultados:

- En las 3 digestiones, la suma de los tamaños de las bandas es de 23 kb (igual al fragmento inicial)
- Sólo 3 bandas de la digestión con A coinciden en la digestión doble
- Las 2 bandas de la digestión con B no aparecen en la digestión doble

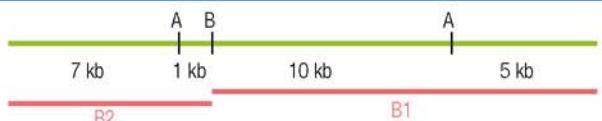
Interpretación:

- En la digestión doble aparecen 2 bandas nuevas, de 10 y 1 kb, cuyos tamaños suman el de la banda A1 desaparecida.

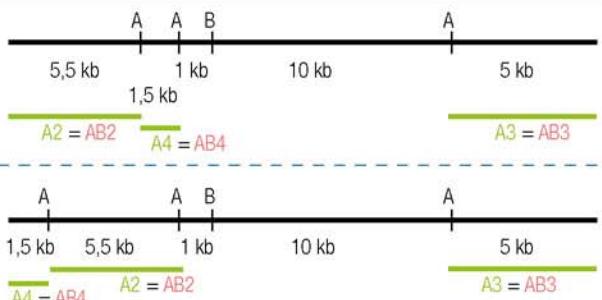
Conclusión: el fragmento A1 es escindido por B en AB1 y AB5; es decir, hay un sitio de restricción B intermedio, situado a una distancia de 1 kb de un sitio A y 10 kb de otro sitio A:



- Los 2 fragmentos de B (8 y 15 kb) pueden estar situados sólo de una forma, a ambos lados del punto de corte B anterior:



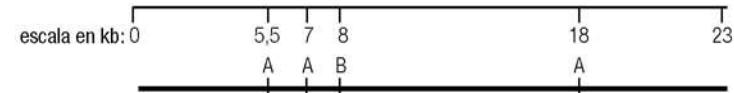
- La asignación de los fragmentos A2, A3 y A4 (iguales a AB2, AB3 y AB4) es la siguiente: La región de 7 kb en el mapa anterior sólo puede corresponder a la suma de A2 y A4. La región de 5 kb debe corresponder a A3. Dependiendo del orden de colocación de estos 2 fragmentos, surgen **dos posibles soluciones**, que no pueden discernirse con los resultados obtenidos:



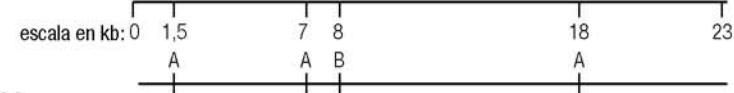
Conclusión final:

Mapa de restricción

solución 1:



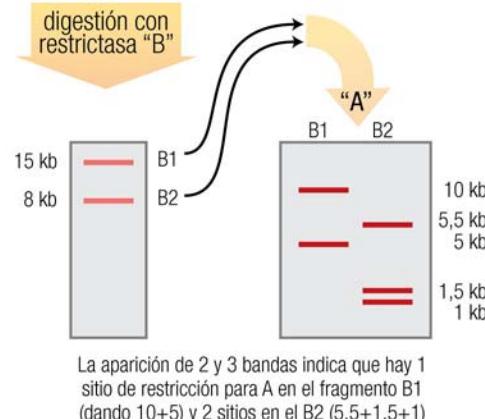
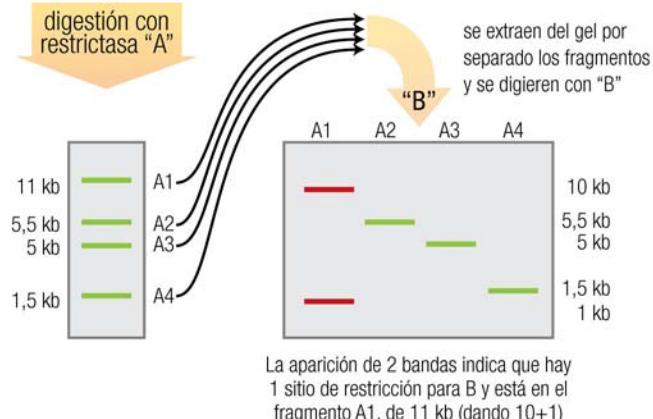
solución 2:



(obviamente, la misma molécula puede representarse de derecha a izquierda, con lo que los valores que figuran en la escala serían otros pero las distancias son las mismas: 0 5 15 16 17,5 23 y 0 5 15 16 21,5 23 respectivamente)

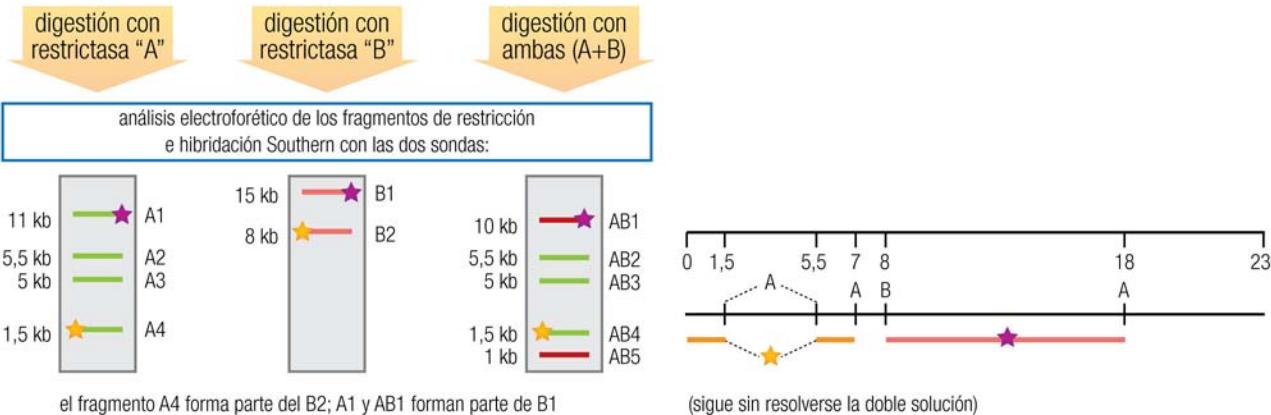
(2)

La deducción del mapa se puede facilitar si, en lugar de la doble digestión simultánea, se hace la digestión con ambas enzimas de forma sucesiva, extrayendo del gel cada uno de los fragmentos de una digestión simple y tratándolo después con la otra enzima



(3)

La deducción del mapa (del caso 1) se puede facilitar también si se dispone de sondas que hibriden con el DNA en estudio



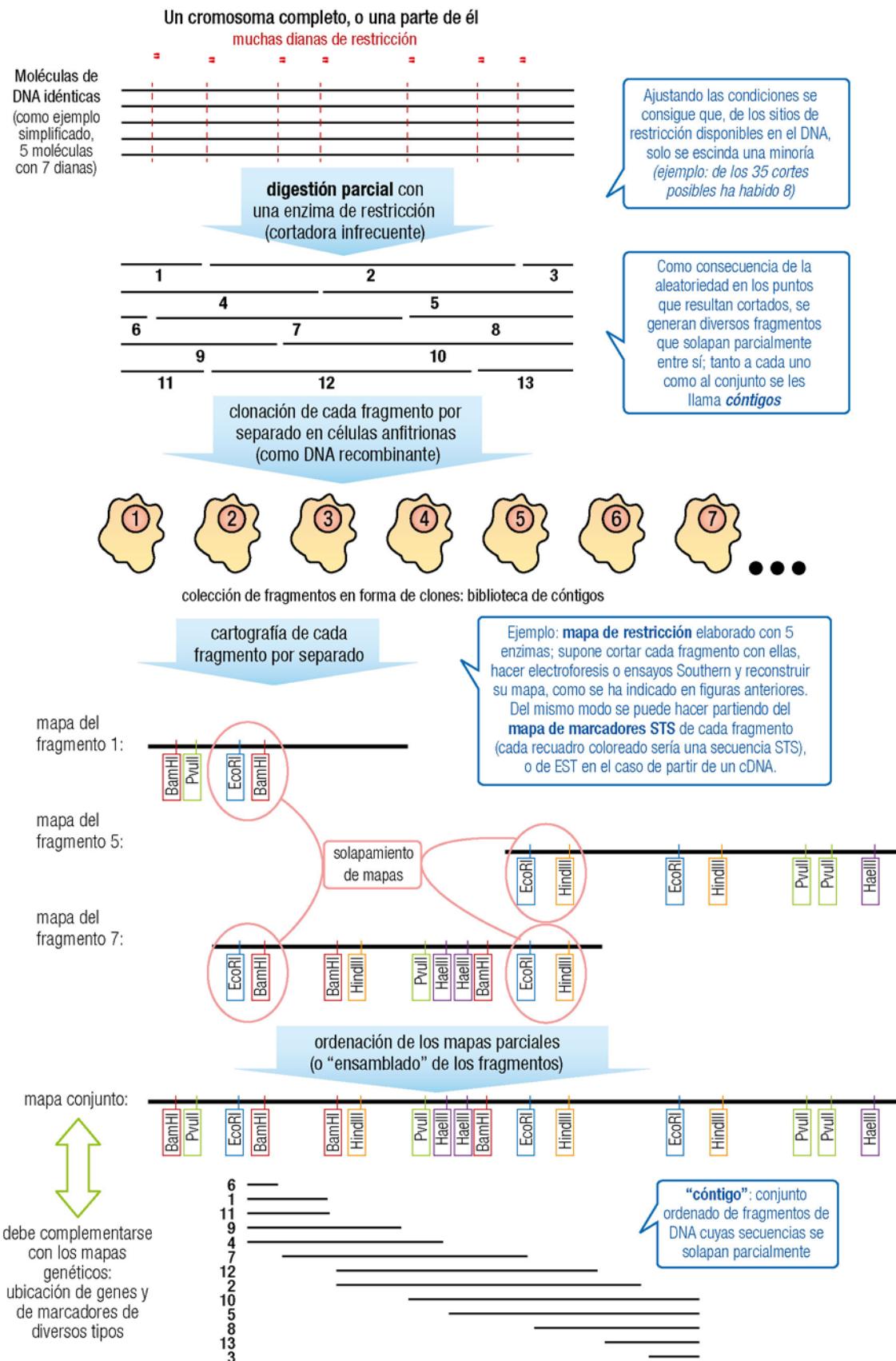
16.11

mayor interés aplicado, los genes responsables de la síntesis de proteínas. Entre otras cosas, estos últimos se han empleado para estimar el número de genes del genoma humano suprimir cita.

A pesar de que la posición de estos marcadores en el genoma se desconoce inicialmente, se puede usar una colección de ellos para caracterizar los clones componentes de una genoteca (pág. 238) –genómica en el caso de los STS, genoteca de cDNA con los EST–, mediante la deducción del solapamiento de clones, o construcción de *mapas de contigos*.

16.3.2.4 Mapas de contigos

Se denomina así a los mapas que se obtienen mediante el solapamiento de fragmentos de DNA. Estos fragmentos proceden en general de la digestión parcial con enzimas de restricción de otros fragmentos mayores, cada uno de los cuales se ha obtenido a su vez por fragmentación de los 24 tipos diferentes de cromosomas (humanos) de la muestra. El planteamiento de la cartografía es idéntico en ambos niveles, y similar al descrito para las nucleasas. Para la deducción del orden o solapamiento de los diferentes fragmentos se acude a la comparación de los marcadores presentes en cada uno, que constituyen generalmente mapas de restricción o mapas de STS o EST.



16.4 SECUENCIACIÓN

Los esfuerzos iniciales para secuenciar ácidos nucleicos dieron pocos resultados. Hasta la década de 1970 era difícil y costoso secuenciar incluso fragmentos de sólo 5 a 10 nucleótidos. Fueron los métodos químico de Maxam y Gilbert (1977) y enzimático de Sanger (1980) los que abrieron el camino a dicho gran objetivo. Se consiguió secuenciar moléculas mayores de DNA y se definieron algunas secuencias concretas como las del fago φX174 (5,4 kb), mitocondria humana (16,5 kb), fago lambda (48,5 kb) y virus animal de Epstein-Barr (170 kb).

Estos dos métodos pioneros de la secuenciación tienen en común algunos aspectos, como la utilización de fragmentos del DNA original escindidos con enzimas de restricción, el marcaje del DNA de forma radiactiva o química (con fluoróforos) y el empleo de métodos electroforéticos para separar los fragmentos generados, entre otros. El auge de los Proyectos Genoma llegó cuando se consiguieron los avances tecnológicos que permitieron acelerar la tasa de secuenciación, aún aplicando de manera básica el método de Sanger. A continuación se han desarrollado métodos de “nueva generación” y de “secuenciación masiva”, algunos basados en principios radicalmente diferentes a los dos anteriores, que suponen un nuevo salto en el progreso de los proyectos de secuenciación de genomas de forma rápida y económica.

La muestra de DNA de partida para la secuenciación procede en general de clonación celular, de amplificación por PCR o de la digestión controlada con una enzima de restricción de un genoma, un cromosoma u otro fragmento de DNA. Con frecuencia, las muestras se amplifican mediante PCR para disponer de suficiente cantidad. Asimismo, es conveniente la purificación para eliminar DNA contaminante en la muestra (por ejemplo, DNA bacteriano) y reactivos procedentes de la purificación o de otros procesos previos (como cebadores y nucleótidos remanentes de la PCR). Se puede acudir también a recuperar el DNA contenido en una banda de electroforesis, extrayéndolo del gel.

16.4.1 Método químico de secuenciación del DNA

Este procedimiento, que se utilizó mucho pero ha quedado desplazado por el método enzimático, se basa en la hidrólisis química y en un diseño original aplicable a secuencias de DNA de menos de 250 nucleótidos. Se suele conocer como **método de Maxam y Gilbert**. Puede describirse en tres etapas: marcaje del DNA, hidrólisis química selectiva del DNA y análisis de los productos.



Web 16.4. Método de Maxam y Gilbert.

16.4.2 Método enzimático de secuenciación del DNA

El método enzimático o **de Sanger** es la base de las variantes que se han utilizado exhaustivamente durante años para la secuenciación a gran escala del DNA. Sólo en la década posterior a la finalización del Proyecto Genoma Humano ha comenzado a ser desplazado por métodos más modernos. En él no se degrada el DNA como en el método químico, sino que se acude a la interrupción controlada de la síntesis de una hebra complementaria durante una replicación *in vitro*. Esta síntesis, catalizada por una DNA polimerasa, define el carácter enzimático del método. La comprensión de los principios del método y el gran cambio introducido con la utilización de fluorocromos son de interés tanto por sí mismos como para la interpretación de métodos posteriores.

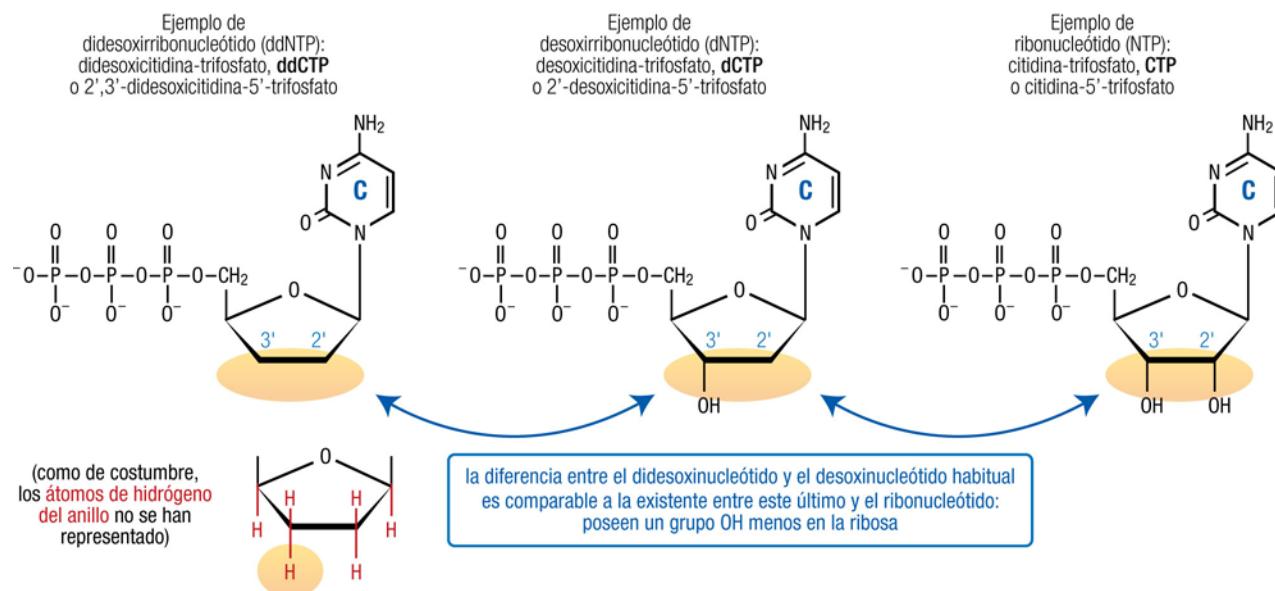
El método se apoya en dos conceptos fundamentales: la síntesis a cargo de una DNA polimerasa y el uso de nucleótidos terminadores. Dado que la actividad 5'-exonucleasa de las DNA polimerasas naturales no es conveniente aquí, se suelen usar variantes, como el “fragmento de Klenow” o fragmento grande de la DNA-polimerasa I de *E. coli* (pág. 150), o formas alteradas artificialmente de la DNA polimerasa del fago T7 (algunas comercializadas con el nombre de “secuenasa”, Sequenase®). También se emplean en ocasiones transcriptas inversas o la DNA polimerasa *Taq* (polimerasa termoestable de la arquea *Thermus aquaticus*, conocida por su aplicación en PCR, pág. 202). Estas alternativas intentan, además de evitar la actividad 5'-exonucleasa, conseguir mayores procesividad, actividad enzimática y eficacia sobre regiones homopoliméricas y regiones con tendencia a adoptar estructura secundaria. Todo ello ha conducido a la posibilidad de secuenciar cada vez fragmentos más largos y con menor margen de error.

Asimismo, han ido surgiendo distintas variantes del método inicial en lo relativo al tipo y posición de marcaje de los nucleótidos, en especial al aparecer los fluorocromos como alternativa a los isótopos radiactivos. Para simplificar, y puesto que el fundamento no varía, se expone a continuación una de ellas, que emplea cuatro fluorocromos distintos y ofrece la posibilidad de automatización. El método puede describirse en tres secciones: síntesis enzimática, análisis de los fragmentos y automatización. Al final se indicarán los matices que supone el empleo de variantes alternativas del método.

16.4.2.1 Síntesis de un DNA complementario de longitud variable

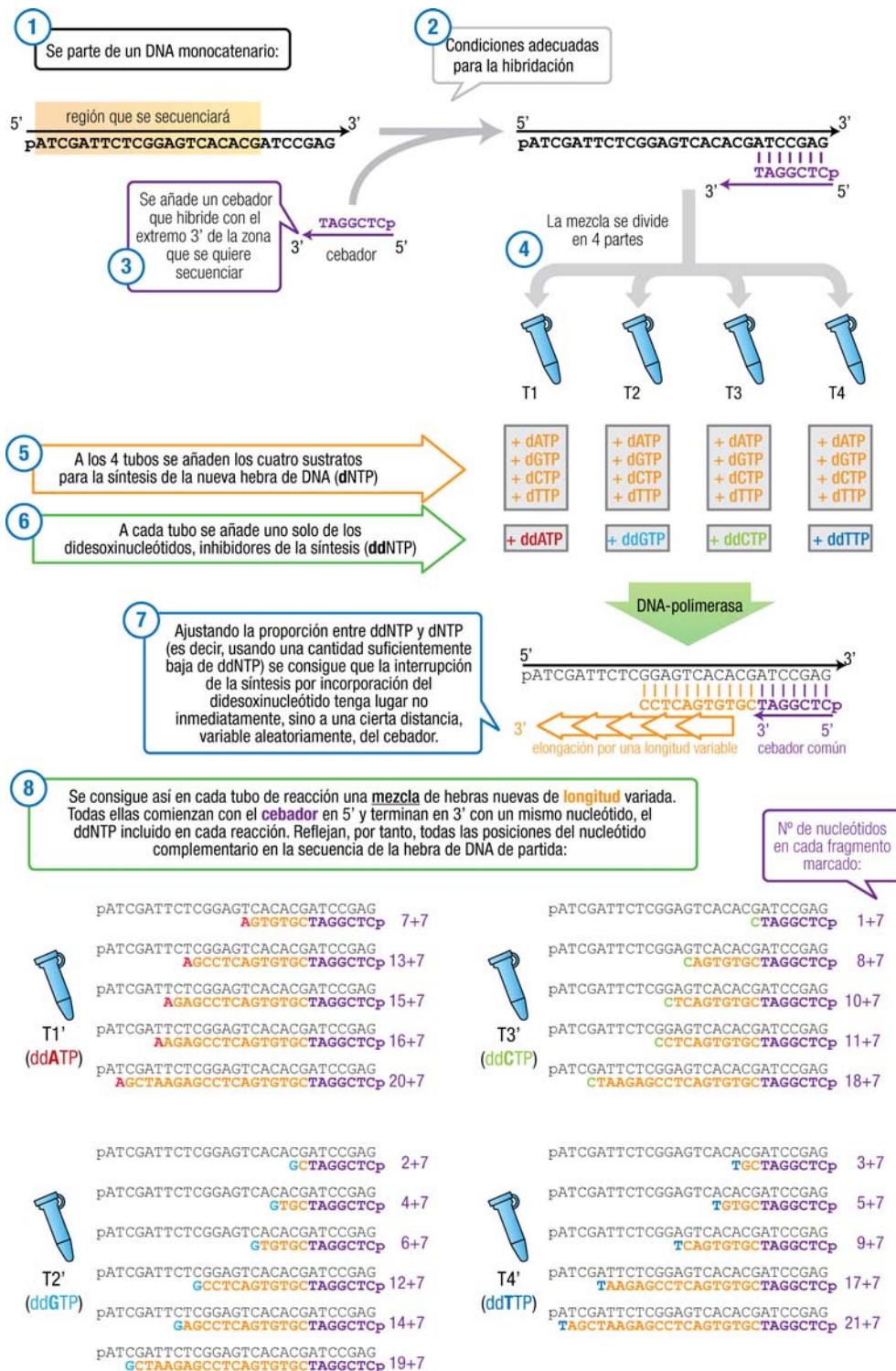
Dado que se emplea una DNA polimerasa para sintetizar hebras complementarias a la que se quiere secuenciar, una primera característica del método es la necesidad de un cebador, un oligonucleótido diseñado para que se hibride con el extremo 3' de la región que se quiere secuenciar.

La segunda, y principal, característica de este método de secuenciación es el uso de **didesoxinucleótidos**, análogos estructurales de los desoxinucleótidos pero que provocan la detención de la reacción de síntesis de DNA. Por ello, el método se conoce también como **secuenciación didesoxi o de terminación de cadena**.



Web 16.5. Estructura de didesoxinucleótidos.

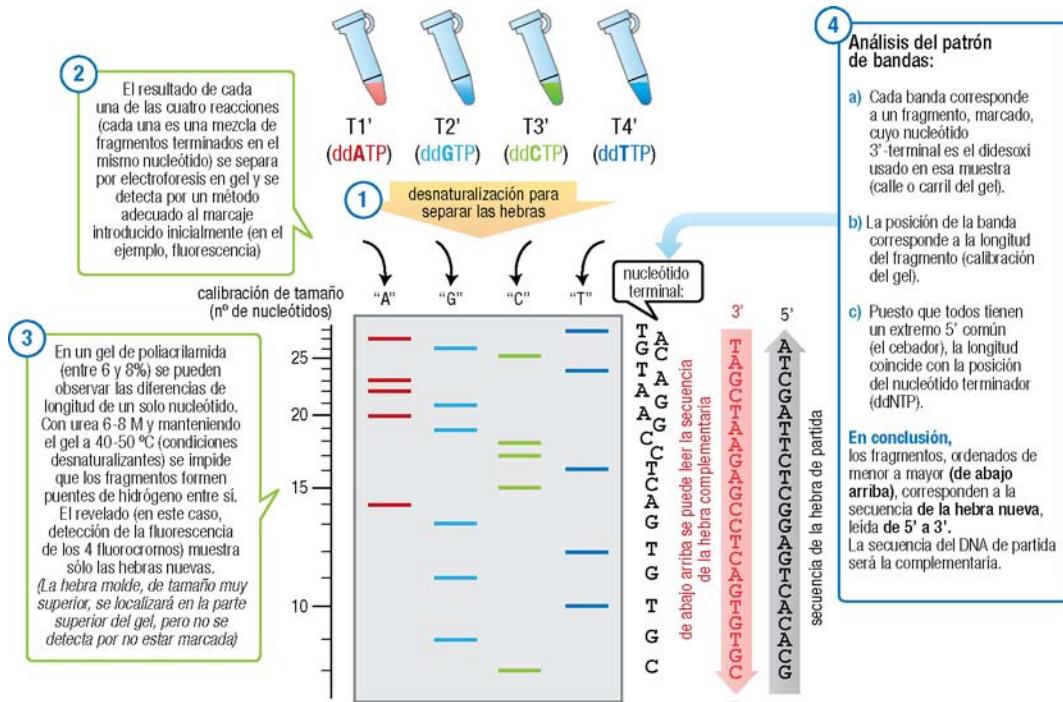
La razón por la que se emplean didesoxinucleótidos es que **compiten con los desoxinucleótidos** en la reacción de síntesis. La DNA polimerasa utiliza el ddNTP como sustrato, por su analogía estructural con el dNTP, y queda incorporado a la hebra en crecimiento como ddNMP, pero a partir de él la cadena no se puede elongar debido a la falta de grupo -OH en 3', que impide que forme enlace fosfodiéster con el siguiente nucleótido. Los ddNTP, por tanto, **detienen la síntesis** de la molécula a la que se incorporan.



16.14

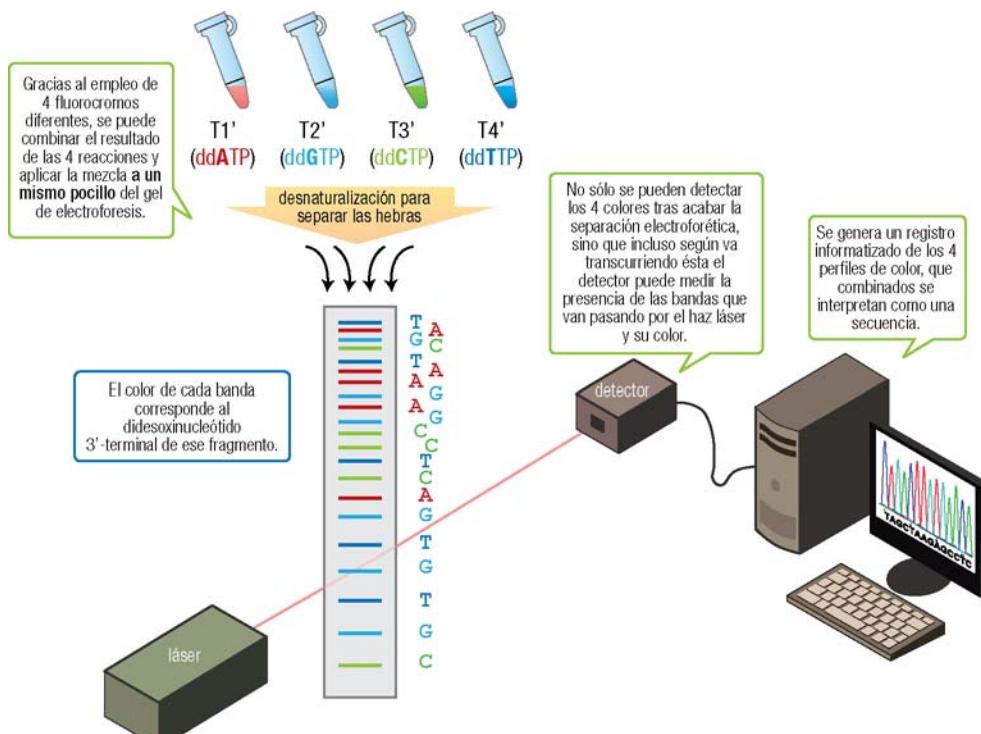
16.4.2.2 Separación de los fragmentos y análisis

Para analizar las 4 mezclas de moléculas obtenidas, se separan todos ellas de acuerdo con su tamaño, empleando electroforesis. El tamaño de cada molécula marcada indica directamente la posición en la secuencia del didesoxinucleótido responsable de la interrupción de su síntesis.



16.4.2.3 Automatización

La posibilidad de usar un fluorocromo distinto en cada una de las 4 reacciones de síntesis permite automatizar el método, de forma que se “lean” simultáneamente las hebras marcadas componentes de las 4 mezclas. Además, esta lectura de la fluorescencia se puede hacer en continuo, con lo que la electroforesis sigue transcurriendo, las moléculas de menor tamaño, ya detectadas, salen por el extremo inferior del gel, y se sigue consiguiendo la separación de moléculas mayores en el gel, de modo que se amplía el número de nucleótidos que es posible secuenciar en un solo experimento.



A título ilustrativo, de los numerosos fluorocromos disponibles se indica un grupo de cuatro que se ha usado con frecuencia para la secuenciación (todos como N-hidroxisuccinimidilésteres, forma activa que permite su unión a un grupo amino del nucleótido).

Nombre común	Nombre completo	Propiedades de fluorescencia		
		λ_{ex} (nm)	λ_{em} (nm)	Color emitido
FAM	(5 o 6)-carboxifluoresceína	494	518	Rojo anaranjado
JOE	(5 o 6)-carboxi-4',5'-dicloro-2',7'-dimetoxifluoresceína	520	548	Rojo-púrpura
TAMRA	(5 o 6)-carboxitetrametilrodamina	553	576	Púrpura
ROX	(5 o 6)-carboxi-X-rodamina	578	604	Azul verdoso



Web 16.6. Estructura y propiedades espectrales de fluorocromos para la secuenciación.

16.4.2.4 Variantes del método

El método que se ha ilustrado es representativo, una de las variantes posibles, pero sirve para comprender el diseño y los fundamentos, válidos para todos. El marcaje de la hebra recién sintetizada puede conseguirse de otras maneras, aplicándolo a otros componentes de la reacción en la que se sintetizan dichas hebras:

(los nros corresponden al esquema de la síntesis)	Método de marcaje de la hebra nueva					
	En el didesoxinucleótido (extremo 3')		En el cebador (extremo 5')		En los desoxinucleótidos (a lo largo de la hebra)	
	fluorescencia	radiactividad	fluorescencia	radiactividad	fluorescencia	radiatividad
② cebador, con la misma secuencia en los 4 tubos (fracciones en que se divide la muestra)	sin marcar		marcado con fluorocromos distintos (nota 1)	marcado con el mismo fluorocromo (nota 1)	marcado con un isótopo (nota 1)	sin marcar
⑤ 4 dNTP, en cada tubo los 4	sin marcar		sin marcar			3 sin marcar, 1 marcado (con 1 fluorocromo o 1 isótopo) (nota 2)
⑥ 4 ddNTP, en cada tubo uno solo	marcados con fluorocromos distintos	marcados con un mismo fluorocromo	marcados con un mismo isótopo	sin marcar		sin marcar
detección de fluorescencia	✓	✓		✓	✓	✓
autoradiografía			✓		✓	✓
electroforesis en 4 calles independientes y lectura tras terminar		✓	✓		✓	✓
electroforesis en 1 calle y lectura simultánea (posibilidad de automatización)	✓			✓		

corresponde al ejemplo ilustrado en las figuras

totalmente comparable al ejemplo ilustrado, salvo que la división de la muestra en 4 debe hacerse antes de añadir el cebador

método original de secuenciación

Nota 1: Empleando nucleótidos marcados durante su síntesis química o bien por modificación enzimática del cebador completo, mediante los métodos de marcaje terminal (véase el marcaje de sondas).

Nota 2: Basta con uno de los dNTP marcado con isótopo radiactivo ($\alpha^{32}P$, $\alpha^{33}P$, $\alpha^{35}S$) o con fluorocromo, que se incorporará en varios puntos a la hebra en crecimiento. Nótese que, a diferencia del método químico de secuenciación, en este caso el isótopo debe estar en el fosfato α (no en el γ), que es el único que se conserva incorporado en el DNA.

16.4.3 Métodos de secuenciación masiva

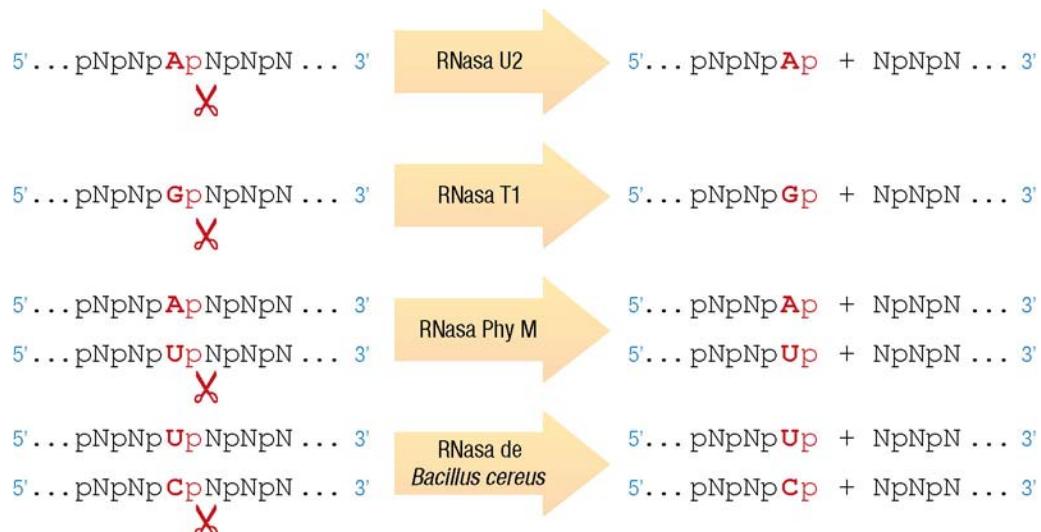
Bajo este epígrafe se pretende recoger las técnicas de secuenciación diseñadas después de la consecución del Proyecto Genoma Humano, que abandonan la estrategia del método enzimático de Sanger. En muchos casos algunas de estas técnicas se han denominado “de última generación” o “de próxima generación”, términos muy atrayentes pero que, obviamente, pierden su sentido de forma rápida. La indicación “masiva”, si bien no aporta información precisa, transmite la idea de que una de las principales características es la capacidad de procesar cada vez mayor número de muestras en paralelo, en un tiempo cada vez menor. Los detalles de estas nuevas aproximaciones están en ocasiones oscurecidos por los intereses comerciales y de patentes, así como sometidos a continua evolución, por lo que se intentará proporcionar una descripción global de sus principios.



Web 16.7. Evolución de los métodos de secuenciación masiva.

16.4.4 Secuenciación del RNA

La secuenciación del RNA es, en general, más difícil que la del DNA, entre otras causas por su menor estabilidad. Es posible deducirla a partir de la secuencia del DNA del que se transcribe, o también sintetizar un cDNA a partir del RNA de interés y secuenciar aquél. Sin embargo, a veces es necesario secuenciar directamente el RNA, sobre todo cuando se desea determinar las posiciones de nucleósidos modificados, característicos de los tRNA y rRNA. Para ello, se ha desarrollado un procedimiento que emplea distintas RNAsas que cortan en el lado 3' de nucleótidos específicos, con lo que se puede obtener una serie de fragmentos que reflejan las posiciones de cada base. Siguiendo un planteamiento análogo al de la secuenciación química del DNA se puede entonces deducir la secuencia del RNA.



Asimismo, al menos en teoría, es posible aplicar el método de secuenciación didesoxi a RNA. Lógicamente, en lugar de la DNA polimerasa usada para secuenciar el DNA, se precisa una transcriptasa inversa, capaz de emplear el RNA de interés como molde para sintetizar una hebra de DNA complementario (pág. 207), partiendo de la misma mezcla de cuatro dNTP y de un ddNTP distinto en cada tubo.

16.5 EL PROYECTO GENOMA

La consecución del mapa completo del genoma humano comenzó en la segunda mitad del siglo XX con la introducción de los mapas genéticos y físicos. A mediados de los años 1980 se concibió el *Proyecto Genoma Humano (PGH)*, un enorme esfuerzo internacional de investigación dirigido a conseguir la secuencia de todo el DNA del genoma humano (los 3.200 millones de pares de bases del genoma haploide). El proyecto comenzó en 1990 y se planteó con una duración de 15 años, pero los avances tecnológicos permitieron culminarlo en poco más de 10.

El desarrollo experimental del PGH consta, a grandes rasgos, de las siguientes etapas:

1. Obtención de las muestras de DNA (comúnmente a partir de los leucocitos de muestras de sangre).

2. a) En la aproximación del PGH de fondos públicos, escisión del DNA de forma controlada en fragmentos, empleando enzimas de restricción. Al principio se originan fragmentos relativamente grandes, de unos 150 kb, y luego, a partir de ellos, una segunda escisión produce fragmentos más pequeños. (Esta doble etapa facilita la ordenación posterior de los datos y la reconstrucción de la secuencia completa.)
- b) En la aproximación del proyecto privado (Celera Genomics), la metodología bautizada como *shotgun sequencing* (que algunos traducen como “en tiro de escopeta”) consiste en la fragmentación del DNA al azar, con una repetición mucho mayor de los fragmentos. El orden de éstos, desconocido por completo, se deduce mediante algoritmos de cálculo basados en el múltiple solapamiento de sus secuencias.
3. Clonación de los fragmentos, mediante la inserción de cada uno en un vector (plásmido, BAC, YAC, etc.) e incorporación a una célula anfitriona.
4. Aislamiento del DNA a partir de cada clón y, en su caso, amplificación por PCR para obtener suficiente cantidad de cada fragmento.
5. Secuenciación completa de cada fragmento de DNA (repetida en promedio unas 7 veces), mediante los métodos automatizados derivados del de Sanger.
6. Ordenación de los fragmentos secuenciados para establecer su posición original en el genoma.

A finales del año 2000, justo antes de publicarse los resultados del PGH, se disponía de la secuencia genómica de 42 especies, todas con genomas relativamente pequeños y simples. Diez años después, el número de genomas secuenciados supera los 4.000, con un total de $125 \cdot 10^9$ pares de bases.

Genomas con secuencia completa

Finales de 2000	Febrero de 2001	Finales de 2010
38 bacterias		4.000 bacterias y virus (5 Gb en total)
1 hongo (<i>Saccharomyces cerevisiae</i>)		
2 invertebrados (<i>Caenorhabditis elegans</i> y <i>Drosophila melanogaster</i>)	Proyecto Genoma Humano	250 eucariotas (120 Gb en total)
1 planta (<i>Arabidopsis thaliana</i>)		

Gb = 10^9 pares de bases

Saccharomyces cerevisiae es la levadura de la cerveza y del pan

Caenorhabditis elegans es un gusano nematodo de 1 mm de longitud

Drosophila melanogaster es la mosca de la fruta o del vinagre

Arabidopsis thaliana es una planta herbácea con flor, con una altura de 10 a 30 cm

En junio de 2000 la Casa Blanca anunció la consecución (anticipada casi 5 años sobre lo previsto inicialmente) de la secuencia completa del genoma humano. A principios de 2001 se publicaron los resultados de lo que se llamó el “borrador” de la secuencia –“secuenciación y análisis iniciales” del genoma–, que cubría alrededor del 90% de las regiones de eucromatina (pág. 88), aún con 250.000 brechas y numerosos errores. Poco después se abordó la secuenciación de otros vertebrados, como ratón, perro, rata, chimpancé y vaca, así como algunos marsupiales, monotremas y aves. Hoy día se dispone de muchos más, e incluso de secuencias parciales de especies extintas como el mamut lanudo y el Neandertal.

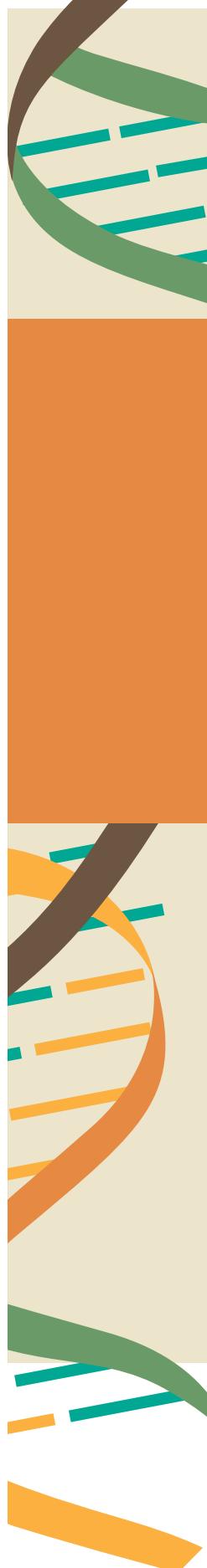
En 2004 se publicó la secuencia “final” o “casi completa” del genoma humano, de alta calidad: cubre aproximadamente el 99,7% de las regiones de eucromatina, con sólo 300 brechas y un error en cada 10^5 nucleótidos. Además de una mejor definición de la secuencia, en este período se produjo la anotación, principalmente gracias a la genómica comparada. Las publicaciones oficiales en revistas científicas se extendieron hasta el año 2006. Las brechas, que no se han conseguido secuenciar, comprenden 28 Mb de eucromatina con secuencias repetitivas y 200 Mb de heterocromatina (centrómeros grandes y brazos cortos de los cromosomas acrocéntricos, pág. 89).

Toda la secuencia obtenida y la información relacionada (conocida como *anotaciones*, que incluyen identificación de genes, secuencias expresadas, polimorfismos y mutaciones conocidos, etc.) está disponible pública y gratuitamente a través de las bases de datos del proyecto en internet.

Actualmente el mayor énfasis se está dedicando a completar la identificación de genes y al estudio de las pequeñas diferencias del genoma en cada persona (polimorfismo genético, Capítulo 24). Por ejemplo, en enero de 2008 se puso en marcha el “Proyecto de los 1.000 Genomas”, que pretende establecer un catálogo de variación genética de los seres humanos secuenciando el genoma de, al menos, mil individuos de distintos grupos étnicos.

Sección III

Expresión génica





Transcripción

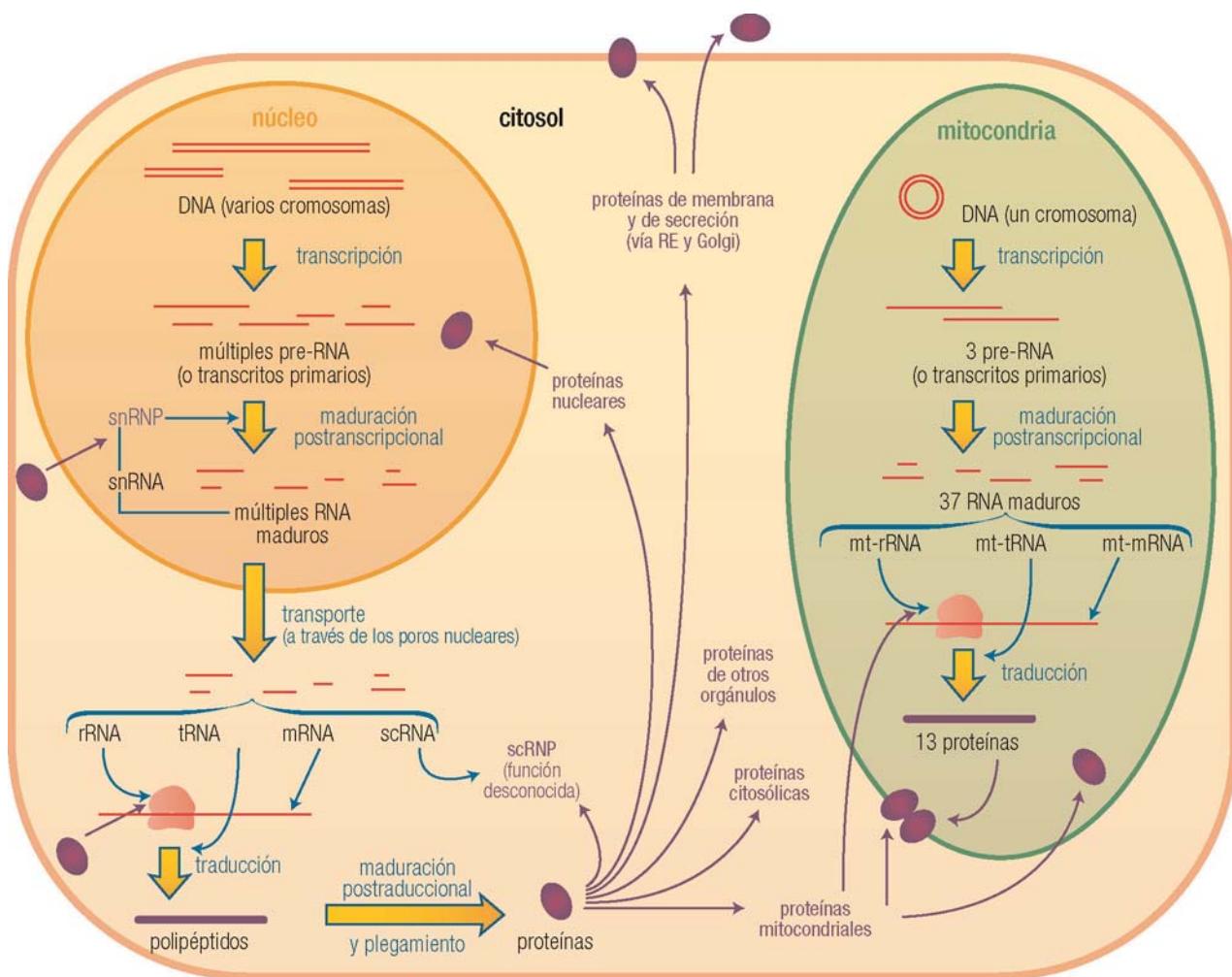
17.1 INTRODUCCIÓN Y CONCEPTOS GENERALES	261
17.1.1. La transcripción en el flujo de información genética	261
17.1.2. Transcripción inversa	263
17.1.3. Concepto de gen	263
17.1.3.1 Definición clásica, no molecular	263
17.1.3.2 Definición molecular y funcional	263
17.1.3.3 Situaciones particulares	265
17.1.3.4 Nomenclatura de los genes	265
17.2 CARACTERÍSTICAS GENERALES DE LA TRANSCRIPCIÓN	266
17.2.1. Carácter secuencial y multifocal	266
17.2.2. Carácter no simultáneo y monodireccional	266
17.2.3. Sólo se transcribe una pequeña parte del genoma	267
17.2.4. Terminología de las cadenas	267
17.2.5. Orientación y numeración	267
17.2.6. Resumen comparado con la replicación	268
17.2.7. Relación con la condensación del DNA	268
17.3 ENZIMOLOGÍA DE LA TRANSCRIPCIÓN	269
17.3.1. Requerimientos	269
17.3.2. Reacción de síntesis de RNA	269
17.3.3. RNA polimerasas	269
17.3.3.1. Especificidad respecto al tipo de gen transcripto y RNA formado	270
17.3.3.2. Localización y cantidad	270
17.3.3.3. Susceptibilidad a inhibidores	270
17.3.3.4. Estructura de las RNA polimerasas	271
17.4 ETAPAS EN EL PROCESO DE TRANSCRIPCIÓN	272
17.4.1. Iniciación	272
17.4.1.1 Formación del complejo de iniciación	272
17.4.1.2 Desnaturalización del promotor	273
17.4.1.3 Formación de los primeros enlaces fosfodiéster	273
17.4.1.4 Liberación del promotor: transición de iniciación a elongación	273
17.4.2. Elongación o alargamiento del RNA	274
17.4.3. Terminación	275
17.5 TRANSCRIPCIÓN DEL GENOMA MITOCONDRIAL	276
17.6 INHIBIDORES DE LA TRANSCRIPCIÓN	276
17.6.1. Antibióticos de tipo I	276
17.6.2. Antibióticos de tipo II	277
17.6.3. Antibióticos de tipo III	278

17.1 INTRODUCCIÓN Y CONCEPTOS GENERALES

17.1.1 La transcripción en el flujo de información genética

La transcripción es el proceso encargado de la síntesis de una molécula de RNA a partir de la información genética contenida en la región codificante de un DNA. Es decir, de dar lugar a una “copia” de RNA (con secuencia no idéntica, sino complementaria y antiparalela) a partir de una de las hebras del DNA empleada como “molde”. Este proceso constituye el segundo paso del esquema clásico de transmisión de la información genética (pág. 3).

Se trata de un proceso enzimático catalizado en todos los organismos por una enzima RNA polimerasa (polimerasa de RNA dependiente de DNA, o transcriptasa, pág. 147). La presencia del OH 2' en la ribosa, que hace al RNA más reactivo (pág. 32), puede explicar que el papel del RNA como material genético (probable en todos los organismos al inicio de la evolución) se haya desplazado al DNA, una molécula más estable.



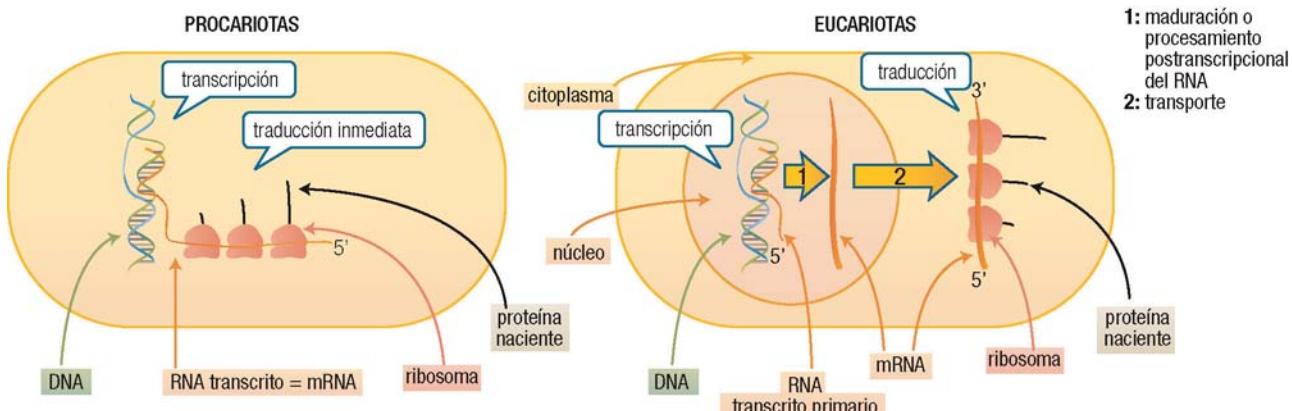
17.1

El proceso de transcripción es notablemente similar en los diversos organismos. Existen, sin embargo, diferencias relevantes en cuanto a lo que ocurre después con los RNA recién sintetizados:

En **procariotas**, la molécula de mRNA resultante de la transcripción es ya funcional, no experimenta transformaciones adicionales y se utiliza de inmediato como molde para la traducción (Capítulo 21), en el mismo compartimento subcelular (puesto que no hay membrana nuclear). Existe por ello una *asociación temporal y espacial* íntima entre transcripción y traducción; de hecho, un mRNA puede empezar a traducirse antes de haberse completado su síntesis por transcripción. Por el contrario, la mayoría de los tRNA y rRNA deben sufrir maduración, análoga a la de los RNA eucarióticos.

En **eucariotas**, el RNA resultante de la transcripción se denomina **transcrito primario**, y experimenta en el núcleo la *maduración o procesamiento postranscripcional* (Capítulo 19). Los RNA maduros se transportan después al citosol para participar en la traducción (Capítulo 21). Existe, por tanto, una *separación espacial y temporal* entre transcripción y traducción, que supone una regulación más compleja de la expresión génica, contribuyendo a la riqueza en variedad y función propia de las células eucarióticas. Este control se lleva a cabo en gran medida por regulación de la actividad de la RNAPol-II (Capítulo 18). El hecho de que esta enzima no se une de forma específica por sí misma a las secuencias promotoras en el DNA es otro signo de la complejidad respecto a procariotas.

17.2



17.1.2 Transcripción inversa

Como excepción, los retrovirus, que poseen un genoma formado por RNA, utilizan éste como molde para la síntesis de un DNA complementario (cDNA), gracias a una polimerasa de DNA dependiente de RNA, una transcriptasa inversa o retrotranscriptasa (págs. 149 y 207). Por esta razón, el esquema general del flujo de información genética se considera actualmente que es bidireccional en su primera parte (pág. 4).

17.1.3 Concepto de gen

17.1.3.1 Definición clásica, no molecular

En términos genéticos clásicos, el gen se define como la unidad elemental de la herencia, la región física y funcional que controla una característica hereditaria concreta, la portadora de la información genética de una generación a la siguiente, la que gobierna, en definitiva, las características de un rasgo particular. Es un concepto, pues, estrechamente asociado al del fenotipo. Por entonces (1910-1930) se consideraba, asimismo, al gen como la mínima unidad de recombinación, mutación y función.

17.1.3.2 Definición molecular y funcional

En términos moleculares, la definición de gen ha ido evolucionando de manera sucesiva desde la visión original, para adaptarse a los nuevos descubrimientos (v. también pág. 110).

En una fase previa, la definición de gen se basó en planteamientos muy simples, relacionados con los avances producidos con el tiempo:

Hipótesis **un gen-una enzima**. Basada en el papel del gen como responsable de la síntesis de enzimas, primeras moléculas con funciones catalíticas en las células.



Hipótesis **un gen-una proteína**. Basada en la noción original de que todas las enzimas son proteínas (hoy se sabe que algunas moléculas no proteicas, las ribozimas, también poseen funciones catalíticas) pero hay numerosas proteínas no enzimáticas.

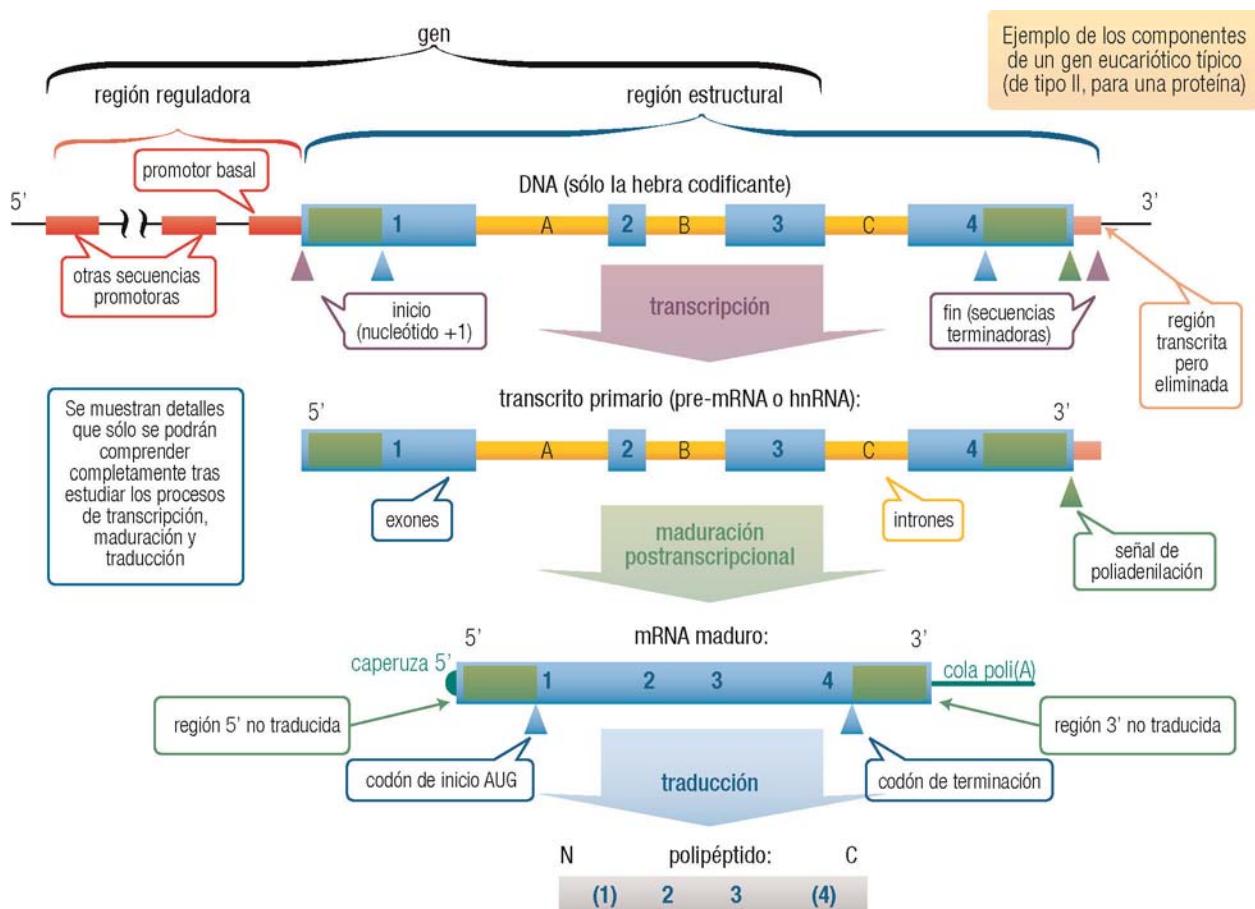


Hipótesis **un gen-un polipéptido**. Basada en que muchas proteínas están formadas por varios polipéptidos, cada uno codificado por un gen diferente.

Como consecuencia, la definición básica, la considerada normalmente como válida y aceptada, es:

Un gen es aquella región del genoma que contiene la información necesaria para sintetizar una molécula de polipéptido.

Este concepto debe ser, sin embargo, matizado y ampliado para dar entrada a dos nuevas características de un gen cualquiera: existencia de dos tipos de secuencias con funciones diferentes, una estructural o realmente codificadora y otra reguladora de la expresión, y presencia en la región estructural de regiones codificantes y no codificantes.



17.3

La **región estructural** define la secuencia del producto génico. Comprende a su vez dos tipos de regiones, en función de su capacidad de expresión: **intrones**, o regiones no codificantes presentes en el interior del gen (pág. 298), y **exones**, que incluyen todas las secuencias codificantes, así como las no codificantes de ambos extremos del gen. El conjunto de exones e intrones de la región estructural se transcribe para dar lugar a un RNA denominado **precursor** o **transcrito primario**. Éste requiere un proceso adicional, posterior a la transcripción, para dar las moléculas funcionales de RNA (Capítulo 19). La mayor parte de los transcriptos formados en procariotas y eucariotas sufre dicho proceso, que recibe el nombre de **maduración postranscripcional**; la excepción son los mRNA procarióticos, que se sintetizan directamente en su forma funcional. Como parte de ese proceso se pierden los intrones, y los exones se unen linealmente, hasta constituir el RNA maduro o funcional (pág. 303).

La **región reguladora**, sin función codificante, suele estar situada cadena arriba (es decir, hacia 5', págs. 267-268) de la región estructural. Contiene distintas **regiones promotoras**, encargadas de interaccionar con los factores de transcripción proteicos para regular positiva o negativamente el inicio de la transcripción (pág. 288).

Otra consideración importante para una definición molecular actual del gen es la formación a partir del DNA de dos tipos de productos génicos: RNA y proteínas funcionales. Para la formación de los primeros se requiere la transcripción y, en su caso, la maduración postranscripcional. Sólo uno de los tipos de RNA, el mensajero, sirve como punto de partida para sintetizar el segundo producto genético, la proteína funcional. Ello requiere, además de la transcripción y maduración postranscripcional, la traducción o síntesis del polipéptido

y, en la mayoría de los casos, el plegamiento proteico, necesario para su distribución hacia el lugar de acción y para el ejercicio de su función.

Como consecuencia, la definición molecular se ha ampliado: **un gen es el conjunto de regiones del DNA de cualquier tipo, estructurales (intrones y exones) y reguladoras, necesarias para codificar y expresar un producto génico, sea éste un RNA maduro de cualquier tipo o una proteína funcional.**

Se observa que ambas definiciones moleculares están basadas en la hipótesis *un gen-un polipéptido*; pueden extenderse, aunque no sea frecuente hacerlo, a:

Hipótesis **un gen-un RNA**, como primer producto génico.



Hipótesis **un gen-un producto génico funcional**.

Por último, debe señalarse que para el caso de los ribovirus y retrovirus, cuyo genoma está formado por RNA, la definición del gen debe hacerse de forma equivalente como región de RNA, y no de DNA.

17.1.3.3 Situaciones particulares

Algunos genes poseen características que requieren forzar las definiciones establecidas; en esos casos, se debe acudir a la interpretación más funcional del concepto moderno de gen, más que a los detalles de la definición. Se indican algunos ejemplos.

- Dos genes pueden ser *solapantes*, es decir, compartir una misma región de DNA (pág. 266), bien porque cada uno está codificado en una hebra distinta (y, por tanto, sus secuencias son completamente diferentes, pág. 42) o bien incluso en la misma hebra (lo cual no implica que las proteínas codificadas comparten secuencia de aminoácidos, debido a la existencia de tres *marcos de lectura*, pág. 314).
- Algunos genes dan lugar a varios productos (pág. 298). Un ejemplo es el gen de los rRNA, que se transcribe en un RNA precursor que por maduración se fragmenta dando varios rRNA funcionales (págs. 307-308). Otro caso está en el DNA mitocondrial, cuyos 3 transcriptos primarios maduran para dar cada uno varios RNA (pág. 309). Finalmente, muchos genes procarióticos se transcriben a un mRNA único que, sin embargo, da lugar por traducción a varias proteínas (mRNA policistrónicos, págs. 324-325)
- En otros casos, una misma secuencia de DNA puede dar lugar de manera alternativa a dos productos génicos, a través de variantes de su expresión génica; por ejemplo, por un diferente procesamiento posttranscripcional o postraduccional (Capítulos 19 y 22). Éste es un fenómeno cuya relevancia se percibe hoy como extraordinaria: se ha estimado que el 90% de los genes de proteínas humanas experimentan ayuste alternativo (pág. 311)
- De igual manera, en numerosos casos la proteína está formada por varias subunidades; en este caso, no existe un gen para la proteína, sino para cada uno de esos polipéptidos (por ejemplo, no se habla del gen de hemoglobina, sino de los genes de la globina α y la globina β).
- En casos excepcionales, como las inmunoglobulinas, un gen único sufre reorganizaciones internas en su secuencia antes de transcribirse, de forma que da lugar directamente a una variedad de polipéptidos (los distintos anticuerpos con especificidad para antígenos diversos).

17.1.3.4 Nomenclatura de los genes

Los genes se nombran en general con tres letras y en cursiva (por ejemplo, gen *ras*). Suele escribirse todo en mayúsculas en el caso de genes silvestres eucarióticos, sólo la primera mayúscula si es un gen silvestre procariótico, y todo en minúsculas si es un gen mutado. Cuando hay varios genes con el mismo nombre, es frecuente completar con una letra mayúscula en procariotas o con números en eucariotas.

La proteína codificada por el gen recibe a veces el mismo nombre, pero sin cursiva y en mayúsculas bien todo o la inicial (proteína **Ras**). En ocasiones, para evitar la confusión, se añade una “p” al final del nombre.

17.2 CARACTERÍSTICAS GENERALES DE LA TRANSCRIPCIÓN

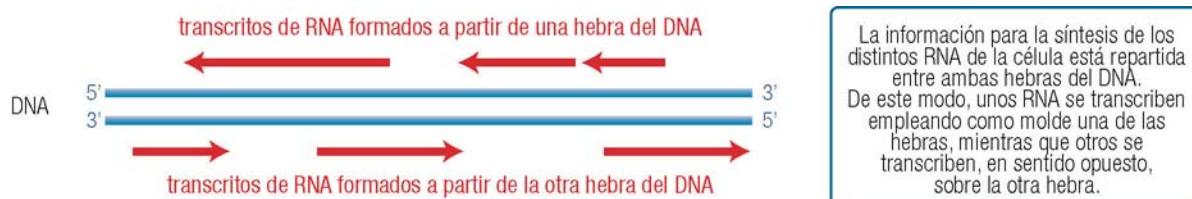
Al comenzar el estudio de la replicación se expusieron una serie de descripciones relativas a su carácter simultáneo, secuencial, bidireccional, mono- o multifocal y semiconservador. Con fines comparados, se puede describir el proceso de transcripción como no simultáneo, secuencial, monodireccional y multifocal. (El carácter conservador de la replicación no tiene aquí equivalente conceptual, pues el producto RNA no comparte nada con el molde DNA.) A continuación se desarrollan en detalle algunas de estas características; el resto son implícitas en los apartados siguientes.

17.2.1 Carácter secuencial y multifocal

La síntesis de RNA comienza en paralelo, pero no coordinadamente, en numerosos puntos de la molécula de DNA que se emplea como molde (un origen para cada gen: multifocal). A partir de ellos, al igual que ocurre en la síntesis de DNA, progresiva mediante la adición sucesiva de nucleótidos individuales a la cadena en crecimiento (secuencial).

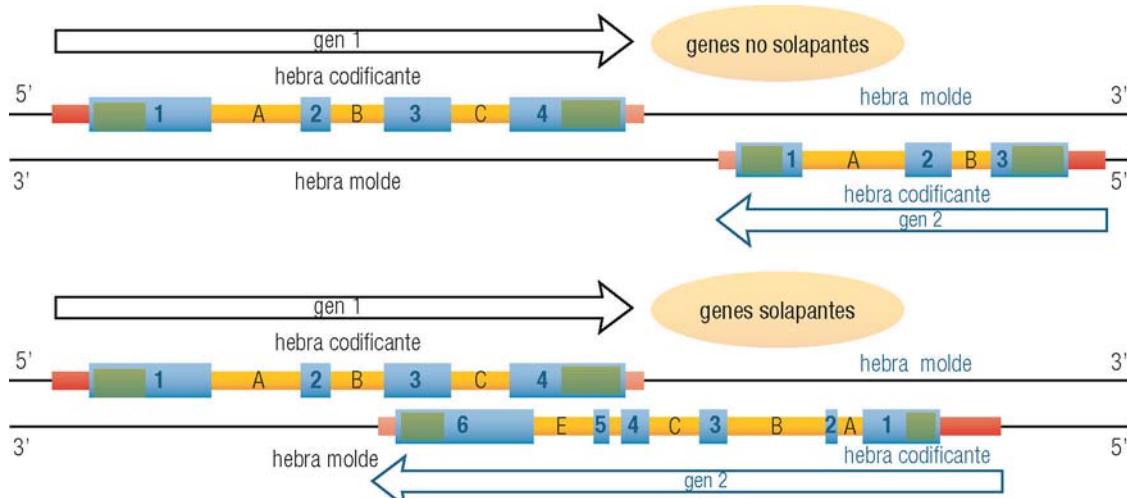
17.2.2 Carácter no simultáneo y monodireccional

A diferencia de la replicación, en la que las dos hebras se copian simultáneamente (pág. 148), la transcripción nunca se produce a un tiempo en las dos hebras. De hecho, no es sólo una cuestión de momento, sino que lo común es que sólo se transcriba una hebra; se afirma también que el proceso es **asimétrico**. Además, también al contrario que la replicación, la transcripción de distintas regiones del DNA no se realiza de forma coordinada en todo el genoma, sino independiente para cada gen. Dado el sentido único $5' \rightarrow 3'$ de la síntesis, la transcripción tiene lugar en sentido contrario en cada hebra.



17.4

A pesar de lo mencionado, en algunos casos sí se transcriben las dos hebras de una región de DNA, pero de forma independiente y no simultánea, portando mensajes genéticos diferentes; se dice entonces que contienen dos **genes solapantes** (pág. 265).



17.5

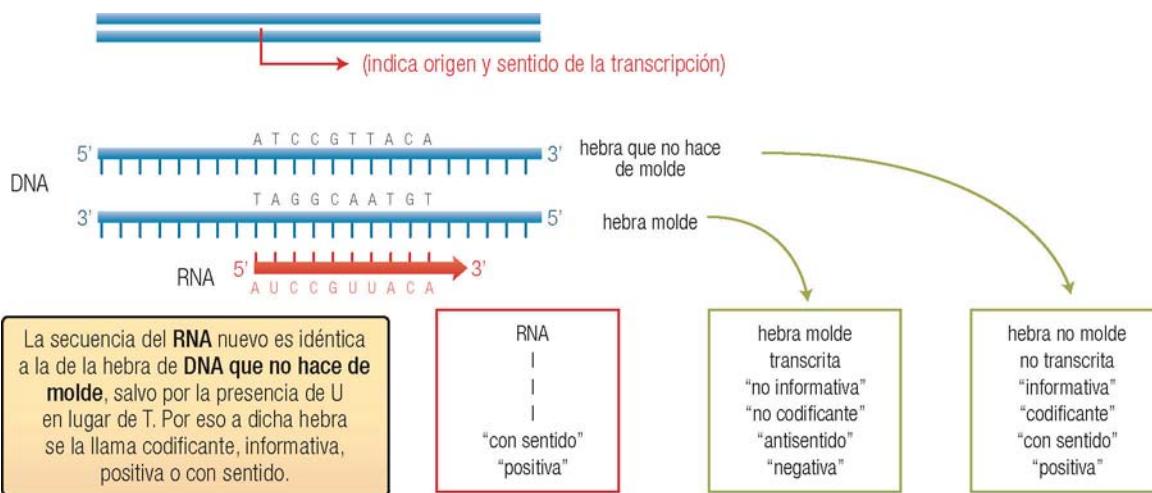
17.2.3 Sólo se transcribe una pequeña parte del genoma

Como ya se indicó, una gran mayoría del DNA genómico de eucariotas nunca se transcribe (DNA no codificante, pág. 109). Asociadas a este carácter no funcional o simplemente no transcribible aparecen copias defectuosas de genes (pseudogenes y fragmentos de genes, pág. 115), estructuras cromosómicas que no se transcriben, así como gran cantidad de *DNA intergénico*. La parte que sí se transcribe corresponde a los exones de las regiones estructurales de los genes, sean de secuencia única o repetida, y a sus intrones (aunque éstos, estrictamente, son DNA no codificante). Además, dentro del DNA que se puede transcribir, células diferentes transcriben distintas regiones (distintos genes), atendiendo a las señales de diferenciación celular y tisular, o de adaptación al medio, a situaciones metabólicas diversas, etc. Todo ello forma parte del complejo proceso de control de la expresión génica (Capítulo 18).

En este contexto parece adecuado insistir en que no es válida actualmente la definición de gen como “secuencia de DNA que se transcribe”, porque excluye la región del genoma que controla la transcripción (pág. 263).

17.2.4 Terminología de las cadenas

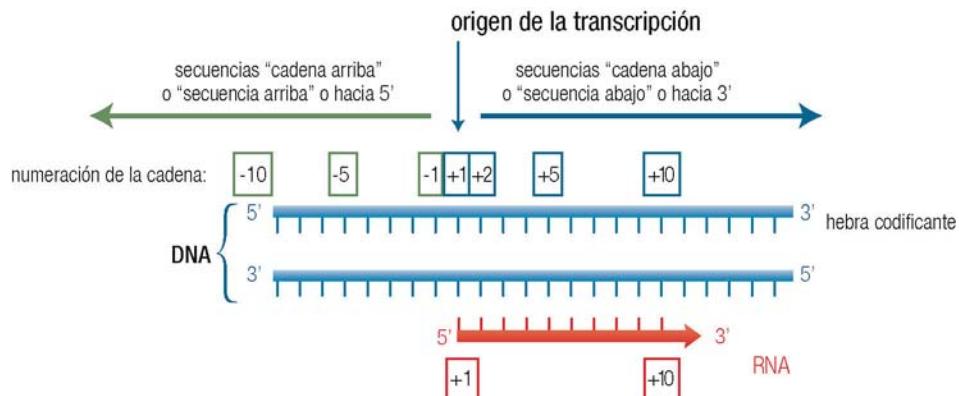
Para una descripción didáctica de la transcripción es aconsejable trazar en horizontal el DNA bicatenario que se va a transcribir; la hebra superior (extremo 5'-P a la izquierda) es la que no sirve de molde, mientras la inferior, antiparalela (extremo 5'-P a la derecha), se convierte una vez girada en el molde sobre el que se empareja el RNA sintetizado. Para la nomenclatura de ambas hebras del DNA, la referencia es la de RNA, que o bien es la molécula funcional o –en el caso del mRNA– contiene el mensaje que originará, por traducción, la secuencia de aminoácidos de la proteína. La síntesis del RNA, que se marca a menudo con línea y flecha, tiene lugar en sentido 5' → 3' del RNA naciente. Esto equivale a decir que la hebra molde de DNA se transcribe (se “lee”) en dirección 3' → 5'.



Finalmente, puede observarse que al ser la nueva cadena de RNA complementaria a la hebra de DNA molde, ambas aparecen formando un híbrido RNA:DNA. Como se verá, éste desaparece a medida que progresá la síntesis, de modo que al final de ésta el RNA queda libre.

17.2.5 Orientación y numeración

La posición de un nucleótido en cualquier gen o región de DNA relacionada con el proceso de la transcripción siempre se refiere a la hebra codificante, la que no hace de molde. Por convenio, su orientación es la del sentido de transcripción. Los nucleótidos de esta hebra se numeran a partir del punto de inicio de la transcripción, al que se asigna el valor +1 (no existe el nucleótido número cero).



17.7

17.2.6 Resumen comparado con la replicación

Es conveniente sintetizar las características recién expuestas, así como otras que se estudiarán a continuación, comparativamente con el proceso de replicación.

Reacción de la DNA polimerasa		Reacción de la RNA polimerasa
1 Requiere una estructura desenrollada.	=	Requiere también una estructura desenrollada.
2 La enzima toma instrucciones de una hebra de DNA (molde).	=	También necesita como molde una hebra de DNA.
3 Las dos hebras del DNA actúan como molde para la síntesis del nuevo DNA, simultáneamente. Cada DNA bicitenario formado conserva una de las hebras del DNA original.	≠	Las dos hebras del DNA pueden actuar como molde del RNA, pero no simultáneamente. Tras realizarse la copia de una sola hebra (la molde), se conserva íntegro el DNA bicitenario.
4 Utiliza cuatro dNTP como sustrato: dATP, dGTP, dTTP y dCTP.	≠	Utiliza cuatro NTP como sustrato: ATP, GTP, UTP y CTP.
5 No es autoiniciadora. Para iniciar la síntesis requiere RNA como cebador, que se empareja por complementariedad de bases y antiparalelamente sobre el DNA molde.	≠	Sí es autoiniciadora, no requiere cebador. Se inicia en o junto a la región promotora, con un NTP (generalmente ATP) emparejado al molde. El NTP mantiene su 5'- trifosfato y su 3'-OH libre.
6 La síntesis comienza por el extremo 3'-OH del cebador. La enzima asociada al DNA molde hace que este 3'-OH libre ataque nucleofílicamente al α -P del primer dNTP, ya emparejado. Se forma el 1º enlace fosfodiéster, se rompe el fosfoanhídrido α - β y se libera PP _i . Su hidrólisis favorece la reacción de síntesis.	=	La síntesis comienza por el 3'-OH del 1º NTP emparejado. La enzima asociada al DNA molde hace que este 3'-OH ataque nucleofílicamente al α -P del 2º NTP, ya emparejado. Se forma el 1º enlace fosfodiéster, se rompe el fosfoanhídrido α - β y se libera PP _i , cuya hidrólisis favorece la reacción de síntesis.
7 Requiere Mg ²⁺ como cofactor enzimático para que la actividad sea óptima.	=	Igualmente requiere Mg ²⁺ como cofactor enzimático.
8 Se producen emparejamientos definitivos A=T y C=G entre la hebra de DNA nuevo y la de DNA molde.	≈	Se producen emparejamientos temporales A=U y C=G entre RNA y DNA molde.
9 Es iterativa (elongación). Se repite la adición de los dNMP por ataques nucleofílicos del 3'-OH del DNA nuevo sobre el α -P-5' de cada dNTP emparejado. La síntesis es en dirección 5'→3' de la nueva cadena de DNA, es decir, en dirección 3'→5' del DNA molde.	=	El proceso es iterativo. La adición de los NMP es por ataques nucleofílicos del 3'-OH del RNA nuevo sobre el α -P en 5' de cada NTP emparejado. Igualmente, la síntesis es en dirección 5'→3' del RNA naciente, es decir, en dirección 3'→5' del DNA molde.

17.8

17.2.7 Relación con la condensación del DNA

Al igual que en la replicación eucariótica (pág. 155), un problema especialmente importante es el grado de condensación del DNA al inicio del proceso de transcripción. En los cromosomas metafásicos eucarióticos, en los que la condensación es máxima, está totalmente impedida la transcripción. Ésta tiene lugar durante la interfase, pues sólo ocurre sobre la conformación de fibras de 10 nm, de mínima condensación (pág. 85) y, por tanto, mayor accesibilidad a las RNA polimerasas y a otras proteínas necesarias para el

proceso. La descondensación máxima nunca se da simultáneamente a todo lo largo del cromosoma, lo que explica la correlación observada de forma experimental entre actividad transcripcional y posición de las regiones descondensadas (eucromatina, pág. 88).

17.3 ENZIMOLOGÍA DE LA TRANSCRIPCIÓN

17.3.1 Requerimientos

- Sustratos:** se utilizan como sustratos el conjunto de los cuatro ribonucleósidos-trifosfato: ATP, GTP, CTP y UTP. De cada uno de ellos queda incorporado en el nuevo RNA la parte NMP de la molécula.
- Cofactores:** para una actividad óptima se requiere un ion metálico divalente como cofactor, asociado a los NTP y a la polimerasa. Aunque *in vitro* este papel pueden desempeñarlo tanto Mn^{2+} como Mg^{2+} , es este último el que actúa *in vivo*.
- Molde o plantilla:** al igual que en la replicación, el orden correcto de incorporación de los nucleótidos viene determinado por su complementariedad de bases con la secuencia de una de las hebras de DNA, que actúa como *molde* o *plantilla*.
- Cebador:** a diferencia de la replicación, la síntesis de RNA comienza simplemente a partir de dos nucleótidos. No es necesaria la presencia de un *cebador*. Puede expresarse esta situación diciendo que la transcripción sí es autoiniciadora.

17.3.2 Reacción de síntesis de RNA



Web 17.1. Antecedentes históricos: síntesis de RNA por la polinucleótido fosforilasa.

Como ya se ha indicado, todos los organismos, procariotas y eucariotas, sintetizan el RNA de acuerdo con una reacción catalizada por la RNA polimerasa (que es una polimerasa de RNA dirigida por DNA, pág. 149). El mecanismo de la reacción es básicamente idéntico al de la DNA polimerasa (pág. 149). La reacción global es la siguiente:



El descenso de energía libre de la reacción, para cada nucleótido añadido, es $\Delta G^\circ = -2,6 \text{ kcal/mol} = -9 \text{ kJ/mol}$. Es decir, es ligeramente exergónica y espontánea. Al igual que en la DNA polimerasa, la reacción está asistida termodinámicamente por la ruptura del enlace fosfoanhídrido del NTP y por la hidrólisis del PP_i catalizada por la pirofosfatasa. Se asegura así que la reacción transcurra de forma irreversible en la dirección de síntesis.

La adición de nucleótidos a la hebra de RNA nueva y las características de la reacción se exponen comparativamente con las de la replicación:



Web 17.2. Mecanismo de la reacción de síntesis de RNA.

17.3.3 RNA polimerasas



Web 17.3. RNA polimerasas en procariotas.

En el núcleo de las células eucarióticas existen tres RNA polimerasas diferentes (I, II y III), que sintetizan cada una distintos tipos de RNA. Sus características diferenciales les confieren una especialización, ausente en procariotas, en cuanto a su labor de síntesis de RNA. Además, cada una emplea distintos factores de transcripción que se unen a promotores situados en distinta posición (Capítulo 18). Por otro lado, la transcripción en los orgánulos (mitocondrias y cloroplastos) corre a cargo de una única RNA polimerasa para los tres tipos de RNA (mRNA, tRNA y rRNA).

RNA polimerasas eucarióticas

	Transcrito primario que sintetiza	Producto génico final	Localización	Actividad enzimática	Inhibidor
RNApol-I	pre-rRNA grande	rRNA de 28S, 18S y 5,8S	Nucléolo	50-70%	Ninguno
RNApol-II	hnRNA (pre-mRNA) pre-snRNA pri-miRNA	mRNA snRNA miRNA	Nucleoplasma	20-40%	α -amanitina (fuerte)
RNApol-III	pre-tRNA pre-rRNA pequeño pre-snRNA U6 pri-miRNA	tRNA rRNA de 5S snRNA U6 miRNA	Nucleoplasma	3-10%	α -amanitina (débil)
RNApol Org	rRNA, mRNA y tRNA de la mitocondria o el cloroplasto		Mitocondrias y cloroplastos		Rifampicina

17.3.3.1 Especificidad respecto al tipo de gen transcripto y RNA formado

Cada RNApol transcribe un grupo de genes y da lugar a transcritos primarios que originarán por maduración postranscripcional distintos tipos de RNA maduros. Por ahora se comenta sólo el primero de estos aspectos.

- La RNApol-I transcribe el gen del precursor grande de RNA ribosómico (45S), precursor único que da lugar a 3 de los 4 tipos de rRNA componentes del ribosoma citoplasmático (págs. 63 y 76).
- La RNApol-II transcribe todos los genes de proteínas, sintetizando los RNA nucleares heterogéneos (hnRNA), que son los precursores de los mRNA (por ello, el hnRNA está constituido por moléculas de tamaños muy variados, de lo que procede su nombre). La misma RNApol-II transcribe también la mayoría de genes de los RNA nucleares de tamaño pequeño, en concreto, todos los snRNA de tipo U excepto el U6 (pág. 304), y algunos microRNA. Puesto que RNApol-II transcribe la gran mayoría de genes, los que codifican proteínas, esta enzima se considera representativa de las polimerasas eucarióticas.
- La RNApol-III, por su parte, transcribe los genes de todos los tRNA, sintetizando sus pre-tRNA respectivos, así como el gen del precursor pequeño del RNA ribosómico –que madurará para dar lugar al cuarto tipo de rRNA (5S)–, el gen del snRNA U6 y los de otros RNA pequeños, incluidos algunos microRNA (págs. 69 y 309).

17.3.3.2 Localización y cantidad

De acuerdo con su especificidad, las RNApol se localizan en el nucleoplasma (RNApol-II y III) o en el núcleo (RNApol-I). Éste es una región densa del núcleo, de tinción más intensa, rica en RNA y proteínas, no rodeada por membrana alguna ni observable durante la mitosis. Su existencia y sus características son el resultado de la propia transcripción de los genes de rRNA, proceso muy activo para proporcionar la elevada cantidad de ribosomas que necesita la célula.

La cantidad de cada tipo de RNApol (medida como fracción de la actividad enzimática total) refleja la demanda de sus productos génicos respectivos: es máxima para RNApol-II, intermedia la de RNApol-II y casi residual en RNApol-III.

17.3.3.3 Susceptibilidad a inhibidores

El estudio y empleo de inhibidores de las RNA polimerasas ofrece interés experimental, para poder analizar sus actividades por separado, así como aplicado, por ser una de las estrategias comunes para combatir las infecciones bacterianas (antibióticos que inhiben la transcripción, pág. 276).

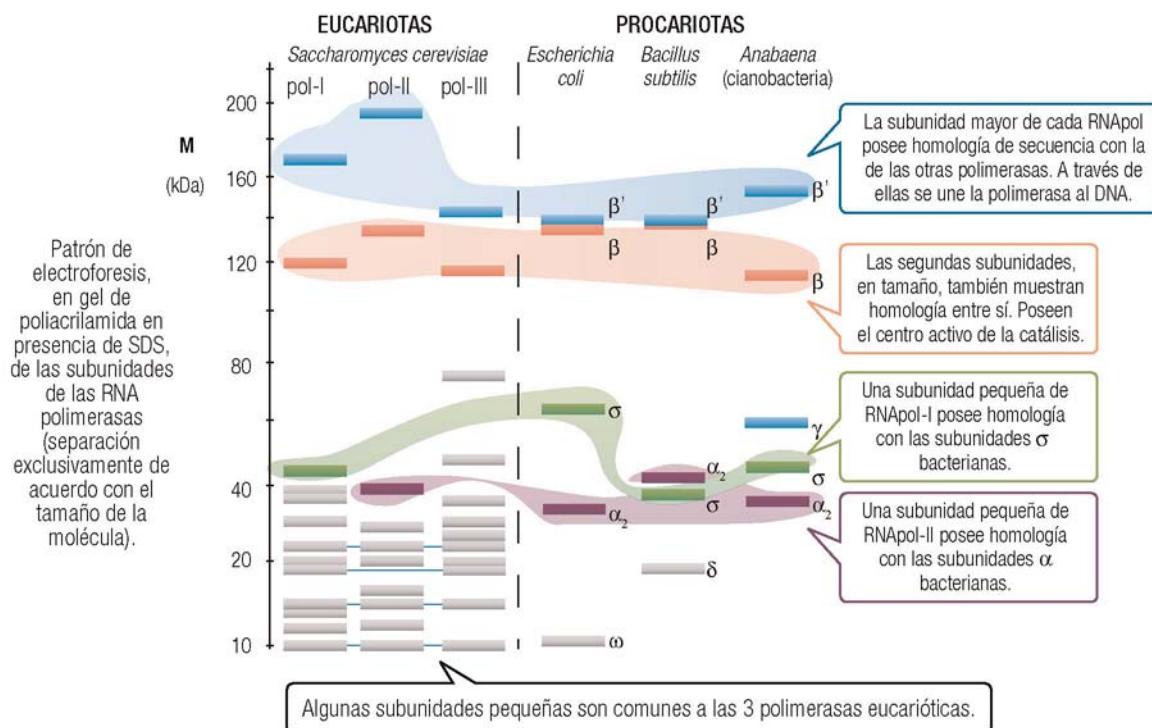
La diferencia más significativa que permite diferenciar de manera experimental las 3 polimerasas nucleares es su sensibilidad a la α -amanitina, una toxina de *Amanita phalloides* (pág. 277). En células animales este compuesto no afecta a la RNAPol-I, inhibe fuertemente (a baja concentración, 10^{-8} M) a la RNAPol-II y débilmente a la RNAPol-III (a mayor concentración, entre 10^{-5} y 10^{-6} M, aunque de forma variable según la especie).

En cuanto a la RNAPol mitocondrial, no se ve afectada por la α -amanitina, pero sí por la **rifampicina** (págs. 276-277), que no afecta a las polimerasas nucleares. En la respuesta a ambos inhibidores la polimerasa mitocondrial se comporta igual que la RNAPol de células procarióticas (un dato más que apoya el probable origen endosimbiótico de las mitocondrias, pág. 10).

17.3.3.4 Estructura de las RNA polimerasas

Las RNA polimerasas eucarióticas están formadas por entre 10 y 15 subunidades o cadenas polipeptídicas, y alcanzan una masa cercana a 600 kDa. (El número de subunidades es impreciso debido a que algunos factores de transcripción (Capítulo 18) permanecen asociados de forma más o menos estable con algunas polimerasas, lo que hace difícil precisar si son subunidades de la RNAPol o proteínas independientes.)

Las 3 subunidades de mayor tamaño son homólogas en las 3 polimerasas, así como con las bacterianas. Constituyen el centro de la enzima (enzima núcleo, enzima mínima, polimerasa central) y contienen tanto el sitio de unión al DNA como el centro catalítico. Algunas de las subunidades menores son también comunes a las 3 polimerasas y resultan esenciales para su acción; el resto probablemente modula dicha acción.



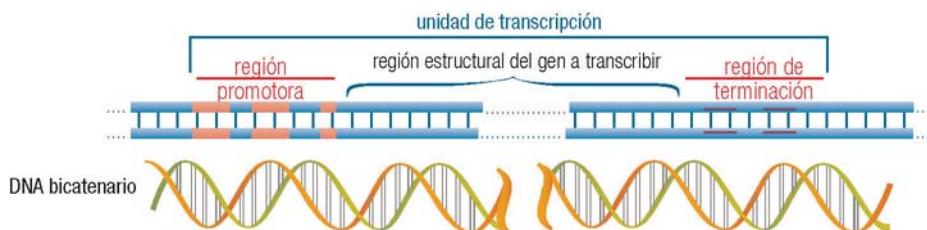
Web 17.4. Estructura de algunas RNA polimerasas. (v. también web 17.3)

La subunidad mayor de la RNAPol-II contiene un dominio C-terminal (denominado CTD, de *C-terminal domain*) que sobresale de la proteína (se le llama la “cola CTD” de la polimerasa) y se encuentra próximo al punto de salida del RNA naciente. El CTD posee múltiples repeticiones de la secuencia heptapeptídica Tyr-Ser-Pro-Thr-Ser-Pro-Ser (repetida 52 veces en la RNAPol humana). Este dominio experimenta fosforilación reversible en los residuos 2 y 5 de serina de cada repetición, adquiriendo así una elevada carga negativa, lo

que se relaciona con su implicación en la transición de la fase de iniciación a la de elongación (pág. 273); su proximidad al canal de salida del RNA puede explicar esta actividad. Asimismo, ejerce de punto de unión para *reclutar* varios de los factores de iniciación y de elongación.

17.4 ETAPAS EN EL PROCESO DE TRANSCRIPCIÓN

Para que tenga lugar la transcripción es necesaria la presencia en la hebra molde de la llamada **unidad de transcripción**, que incluye la secuencia del DNA que se ha de transcribir más las dos secuencias consenso que la flanquean (denominadas de forma genérica **promotor** y **terminador**). Se reserva el término **operón** para la unidad de transcripción en procariotas.



17.11

Al igual que la replicación, la transcripción transcurre en las tres fases de iniciación, elongación y terminación, con notables diferencias entre eucariotas y procariotas en cuanto al reconocimiento de las secuencias promotora y terminadora.

17.4.1 Iniciación

Es, con mucho, la etapa más compleja de la transcripción (y, por ello, en la que tiene lugar la mayor parte de la regulación, Capítulo 18). Requiere la separación de las hebras del DNA en un pequeño tramo cercano a la secuencia del **promotor basal**, para permitir que la RNA polimerasa sintetice un fragmento corto de RNA (de unos 10 nucleótidos) por emparejamiento con la hebra molde. La región de DNA que sufre este desenrollamiento parcial recibe el nombre de **burbuja de transcripción**; ésta se forma en el inicio y posteriormente, durante la elongación, se desplaza a lo largo del DNA junto con la RNAPol. Para la iniciación también se requiere la unión al DNA de proteínas llamadas factores de inicio de la transcripción o, más comúnmente, **factores de transcripción** (TF). Tanto éstos como los promotores se estudiarán en detalle más adelante (Capítulo 18).

En procariotas y en las **mitocondrias** es frecuente que un mismo promotor inicie la **cotranscripción** de varios genes contiguos. Por el contrario, en el genoma **nuclear de eucariotas**, a partir de cada promotor sólo se transcribe un gen. Distintos genes, en especial dependiendo de cuál sea la RNAPol, utilizan distintos tipos de secuencias promotoras, a las que se unen distintos factores de transcripción. El proceso de iniciación que se describe como típico de eucariotas utiliza promotores que tienen la secuencia TATA, pero no la Inr (llamados por ello de tipo TATA+ Inr-), cuyas características se describen posteriormente (pág. 289). La secuencia TATA no la reconoce la RNAPol directamente, sino a través de la mediación de los TF. Es importante resaltar que, tanto en procariotas como en eucariotas, la iniciación requiere la presencia de la RNAPol completa u **holoenzima** (es decir, la RNAPol bacteriana con la subunidad σ y las RNAPol eucarióticas con todas sus subunidades y con los factores de transcripción).

La iniciación puede subdividirse en cuatro subetapas, que sólo se describen aquí para los *genes de clase II*, los transcritos por la RNAPol-II. Los genes de clases I y III se transcriben por mecanismos similares, salvo por la identidad y peculiaridades de sus polimerasas, promotores y factores de transcripción (págs. 294-296).

17.4.1.1 Formación del complejo de iniciación

Este complejo plurimolecular es **imprescindible para que la RNAPol pueda actuar**. Supone la unión de varios TF a las regiones promotoras del DNA y la incorporación de la enzima, que participa en su forma con dominio CTD no fosforilado (pág. 271). Se han propuesto dos modelos para el mecanismo de formación de este

complejo, pero ambos conducen al mismo resultado; sus detalles se describen posteriormente, cuando se estudien los promotores y los TF (pág. 291).

- El *modelo del holoenzima*, en el que la RNAPol-II se une previamente a los TF y luego el conjunto se une al promotor.
- El *modelo escalonado*, en el que los TF se van asociando a la región promotora uno a continuación del otro y en un cierto orden, y la polimerasa se une al final.

17.4.1.2 Desnaturalización del promotor

El complejo de iniciación produce en la región de inicio de la transcripción el desenrollamiento y separación de las dos hebras del DNA (una desnaturación local), esencial para la exposición de las bases como molde a la RNAPol. Esta zona de hebras separadas, **burbuja de transcripción o complejo abierto**, es el lugar en el que va a tener lugar la síntesis de la cadena de RNA. Esta etapa requiere, al igual que el inicio de la replicación (pág. 154), una actividad **helicasa** de DNA, presente ahora en dos de los factores de transcripción, TFIIH y TFIIH, con un gasto energético (hidrólisis del ATP). Otro factor, el TFIIIE, estabiliza la región desnaturizada del promotor (papel en cierto modo similar al ejercido por la RPA en la replicación, pág. 155).

WWW

Web 17.5. Formación de la burbuja de transcripción en torno a la RNA polimerasa.

Web 17.6. Iniciación, elongación y terminación de la transcripción.

17.4.1.3 Formación de los primeros enlaces fosfodiéster

El enlace fosfodiéster inicial se forma entre dos ribonucleótidos, a partir de sus NTP en forma libre. Uno de ellos, que con frecuencia es ATP, se empareja con el nucleótido +1 en la hebra molde de DNA. El otro se empareja con el nucleótido +2 de la misma hebra.



Este primer enlace corresponde, por tanto, al extremo 5' del RNA naciente. La reacción se repite sobre el extremo 3' del dinucleótido resultante, y así sucesivamente hasta formar un oligorribonucleótido de unas 10 bases, que permanece unido temporalmente a la hebra molde de DNA como híbrido RNA:DNA. Una vez que se produzca la transición a la etapa de elongación, la polimerasa continuará añadiendo nucleótidos a la cadena de RNA al tiempo que se va desplazando sobre la hebra molde (recorriéndola en sentido 3' → 5').

17.4.1.4 Liberación del promotor: transición de iniciación a elongación

WWW

Web 17.7. Transición de iniciación a elongación en procariotas.

En eucariotas la transición viene marcada por la fosforilación del dominio CTD de la RNAPol-II (pág. 271); en la fase de iniciación participa la forma no fosforilada, mientras en la de elongación lo hace la fosforilada. Esta fosforilación la llevan a cabo el factor de transcripción TFIIH (pág. 291) y el factor de elongación P-TEFb. Como consecuencia, la RNAPol-II se desprende de la mayor parte de los TF (lo que se denomina *liberación del promotor*) y comienza a desplazarse a lo largo del DNA.

17.4.2 Elongación o alargamiento del RNA

La RNA polimerasa continúa alargando la cadena de RNA mientras avanza por el DNA, desplazando junto con ella la burbuja de transcripción. El complejo de elongación ya está sólo formado por el DNA abierto, la RNAPol y el RNA naciente. Durante este avance continuo ocurre lo siguiente (v. web 17.6):

- Por delante de la polimerasa se separan las dos hebras del DNA por ruptura de los puentes de hidrógeno (avance de la burbuja). La tensión de superenrollamiento resultante debe ser compensada por topoisomerasas (págs. 80 y 155).
- La polimerasa alarga el RNA añadiendo nucleótidos al extremo 3' del RNA naciente. El último tramo permanece dentro de la burbuja, emparejado de manera transitoria como híbrido RNA:DNA.
- Por detrás de la polimerasa, la cadena de RNA nuevo se va separando de la hebra molde de DNA, saliendo de la burbuja y quedando como RNA monocatenario. En el DNA se vuelven a emparejar las dos hebras y se produce el re-enrollamiento de la doble hélice, de nuevo asistido por topoisomerasas.

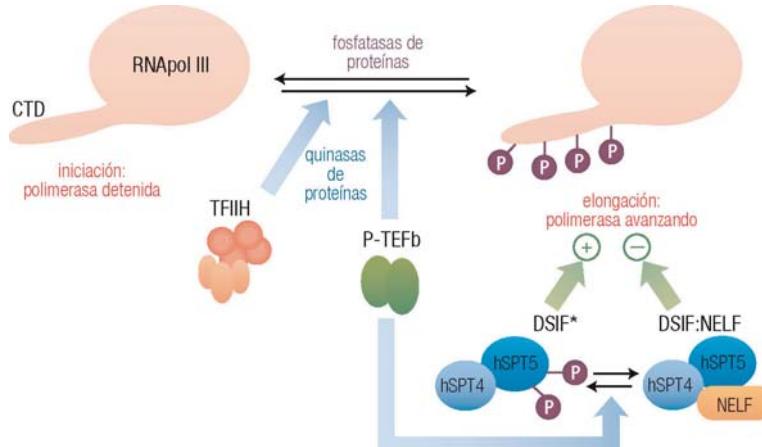
Las RNA polimerasas poseen una procesividad muy elevada (pág. 152), que les permite sintetizar la cadena completa de RNA de una vez, sin separarse del molde. En contraste con la replicación, ello se consigue sin proteínas accesorias, pues la propia RNAPol abraza la cadena de DNA molde.

En cuanto a la tasa de error en la incorporación de nucleótidos correctamente emparejados al molde, es próxima a uno por cada 10.000, lo cual es unas mil veces superior a la tasa de error de la replicación. Los errores en la transcripción son más tolerables porque el producto RNA tiene una vida corta y su información genética no se perpetúa en el organismo. La RNAPol posee cierta capacidad de corrección de pruebas (pág. 150), bien mediante la reversión de la reacción de polimerización (actividad pirofosforolítica) o bien mediante un centro activo separado con actividad hidrolítica, ribonucleasa. Esta última se ve estimulada por el factor de elongación TFIIS.

Para que la elongación sea eficiente se requiere, además, la participación de hasta 16 proteínas llamadas factores de elongación. Varios de ellos se unen a la cola CTD de la RNAPol cuando está fosforilada. Su actuación está mediada por varios mecanismos:

- (a) Estimulación de la actividad enzimática de la polimerasa: factores ELL, CSB y SII (o elonguina).
- (b) Mantenimiento del estado fosforilado del CTD: además del factor de iniciación TFIIH, también lo fosforila el factor de elongación P-TEFb. Cuando el CTD está desfosforilado, se inhibe la elongación.
- (c) Regulación de la procesividad de la polimerasa: TFIIS (antes llamado SII) reduce las pausas transitorias de la RNAPol, que pueden ser causa de una terminación prematura. También interviene DSIF.

La eficacia de la elongación depende de un equilibrio entre acciones de estimulación y freno. El principal elemento que regula la procesividad es la proteína DSIF, un heterodímero formado por hSPT4 y hSPT5. Cuando DSIF se une a la proteína NELF (*negative elongation factor*) frena la elongación. Pero cuando DSIF es fosforilado en su subunidad hSPT5, se separa de NELF y activa la elongación. La principal quinasa responsable de fosforilar hSPT5 es precisamente P-TEFb (*positive transcription elongation factor b*), la misma que también fosforila el CTD de la RNAPol.



Para completar tan complejo panorama, la presencia junto a la RNAPol de algunos de estos factores de elongación facilita el reclutamiento de otras proteínas necesarias para la maduración postranscripcional del mRNA y para la remodelación de la cromatina (retirada temporal de los nucleosomas, imprescindible para el acceso de la RNAPol a las hebras desemparejadas). La elongación y la terminación de la transcripción y la maduración del mRNA están íntimamente interconectadas y necesitan estar coordinadas con precisión.

17.4.3 Terminación

El complejo de transcripción responde a señales específicas para finalizar el proceso, separándose el RNA formado (v. web 17.6). Los detalles de esta etapa están aún poco caracterizados. En todo caso, la etapa de terminación sí está relacionada con las señales que marcan la poliadenilación del extremo 3' de los mRNA. El resultado final es la separación de los componentes del complejo de transcripción y, en el caso de la RNAPol-II, su desfosforilación para que pueda ensamblarse a un nuevo complejo de iniciación y comenzar otro ciclo de transcripción.

Finalmente, e independientemente de los mecanismos precisos, es importante adelantar que el extremo 3' del RNA funcional no corresponde al punto donde la polimerasa deja de elongar, debido a que durante la maduración postranscripcional se corta la molécula de RNA recién transcrita (pág. 302). Por esta razón es menos importante precisar la terminación en eucariotas que en procariotas, y también se explica la dificultad del estudio de las secuencias y mecanismos de terminación eucarióticos (pues es difícil aislar los transcriptos primarios, de corta vida).



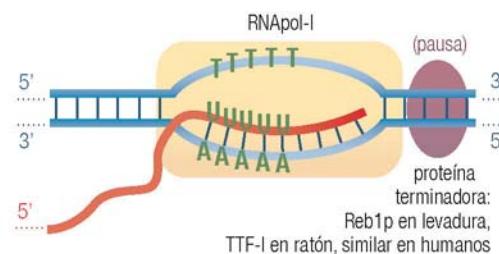
Web 17.8. Mecanismos de terminación en procariotas.

La terminación se produce como consecuencia de la combinación de dos factores o mecanismos: la **detención o pausa** de la RNA polimerasa y la **desestabilización del híbrido RNA:DNA**. La influencia de estos dos mecanismos es distinta para cada una de las tres polimerasas eucarióticas. A diferencia de procariotas, parece ser un hecho común la ausencia de regiones palindrómicas del RNA que participen en la terminación de la transcripción.

Terminación de RNAPol-I en eucariotas:

mecanismo similar a la terminación intrínseca de procariotas.

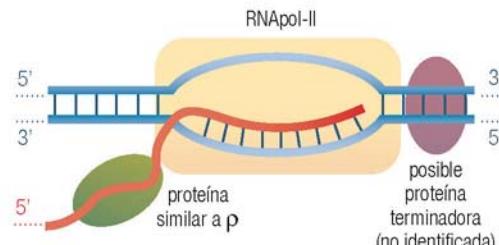
- Una "proteína de pausa" o "proteína terminadora" se une a secuencias específicas del DNA y detiene el avance de la RNAPol.
- Una vez detenida la polimerasa, la presencia de una región rica en T en la hebra codificante facilita la separación del RNA.
- No hay región palindrómica que forme horquillas.



Terminación de RNAPol-II en eucariotas:

mecanismo similar al dependiente de ρ en procariotas.

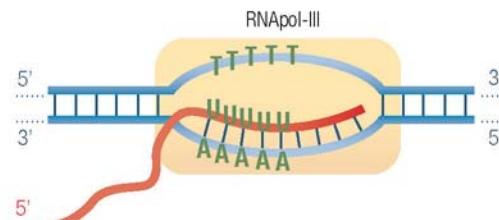
- En múltiples sitios dentro de una amplia "región de terminación" la RNAPol se detiene, lo que puede deberse a secuencias específicas en el DNA o a "proteínas de pausa" que se unen al DNA.
- Una proteína, similar a ρ (de hecho, la propia ρ bacteriana produce el mismo efecto *in vitro*), se une al RNA induciendo su separación.
- No hay región palindrómica que forme horquillas.



Terminación de RNAPol-III en eucariotas:

mecanismo similar a la terminación intrínseca de procariotas.

- La RNAPol se detiene en secuencias específicas del DNA (no se han encontrado proteínas terminadoras).
- Una vez detenida la polimerasa, la presencia de una región rica en T en la hebra codificante facilita la separación del RNA.
- No hay región palindrómica que forme horquillas.



17.5 TRANSCRIPCIÓN DEL GENOMA MITOCONDRIAL

Como ya se ha indicado, el cromosoma circular mitocondrial humano, y de mamíferos en general, lo transcribe una única RNAPol que está formada por un solo polipéptido y que, para la iniciación, requiere la presencia de un factor de transcripción mtTFA. El mtDNA se transcribe a partir de 3 lugares de iniciación, dos para la hebra pesada (H1 y H2) y uno para la ligera (L), produciéndose así tres transcritos primarios diferentes (dos de ellos complementarios y correspondientes a toda la longitud del cromosoma, 16,5 kb). Los 3 promotores se encuentran situados en la región no codificante de control o bucle D (págs. 161 y 416). Aún se desconoce si el inicio en H1 y H2 depende de un mismo promotor o de dos independientes (“principal” y “secundario”), uno de ellos todavía desconocido.



Web 17.9. Transcripción del genoma mitocondrial.

A partir del primer lugar de iniciación de la hebra pesada (H1) se transcribe el DNA que codifica los tRNA^{Phe}, rRNA 12S, tRNA^{Val} y rRNA 16S; la transcripción se detiene tras este último, mediante la participación de un factor de terminación (mTERF). Cuando la transcripción comienza en el segundo punto de inicio H2, localizado inmediatamente tras el tRNA^{Phe}, el transcripto primario (L2) continúa más allá del punto de terminación anterior, completando el resto del cromosoma. Por tanto, el lugar de iniciación define qué tipo de RNA, ribosómicos o mensajeros, deben sintetizarse; la selección del punto de inicio está, a su vez, regulada por las hormonas tiroideas. Gracias a esta doble unidad de transcripción la mitocondria sintetiza mucha mayor cantidad de RNA ribosómicos que del resto de productos génicos (unas 60 veces más). (Compárese con la situación análoga en el genoma nuclear, conseguida gracias a la abundancia de la RNAPol-I, especializada en transcribir rRNA, pág. 270.)

A partir del promotor de la hebra ligera se sintetiza una sola molécula (que es una hebra pesada, H) que sólo codifica 8 tRNA y un mRNA.

Los 3 transcritos primarios sufren luego maduración para dar lugar cada uno a varios rRNA, tRNA y mRNA maduros (pág. 309); del transcripto H se origina también el RNA cebador para la replicación de la hebra H (pág. 161).

17.6 INHIBIDORES DE LA TRANSCRIPCIÓN

Existe una serie de compuestos que inhiben la transcripción al actuar sobre la RNA polimerasa a través de distintos mecanismos. Poseen interés desde el punto de vista experimental y por su aplicación como antibióticos, por mostrar algunos distinta actividad sobre las células procarióticas y eucarióticas.

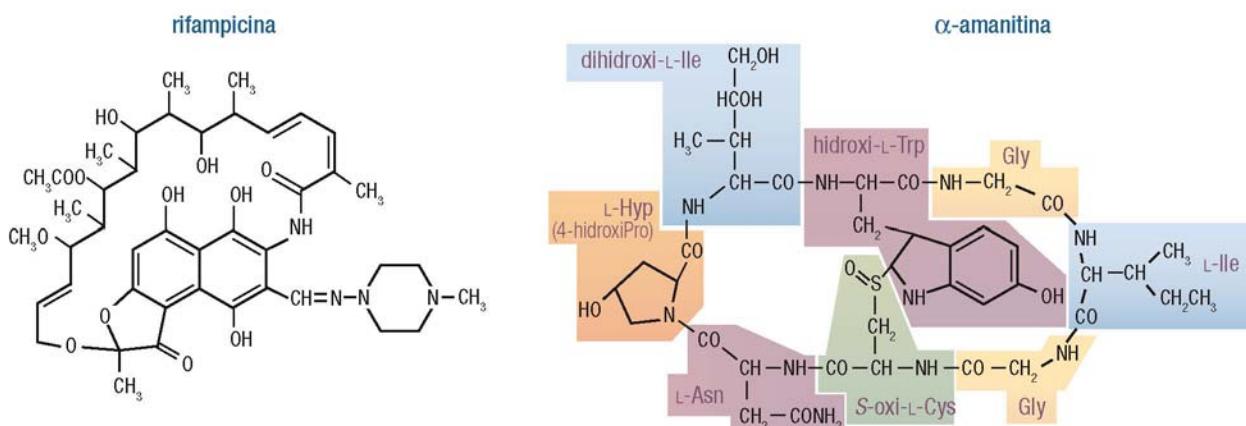


Web 17.10. Estructura tridimensional de antibióticos inhibidores de la transcripción.

17.6.1 Antibióticos de tipo I

Este grupo está formado por los compuestos que se unen a la RNA polimerasa y la inactivan de modo directo. Pertenecen a este grupo:

- **Rifamicinas** (producidas por *Streptomyces mediterranei*) y sus derivados semisintéticos como la **rifampicina**, que se unen específicamente a la subunidad β de la RNAPol bacteriana, impidiendo el inicio de la transcripción. La permanencia sobre el promotor de la enzima inactiva impide la unión de otras moléculas de RNAPol no modificadas. Sin embargo, las rifamicinas no tienen efecto sobre las polimerasas nucleares eucarióticas, por lo que se utilizan como agentes bactericidas frente a diversos tipos de infecciones. La RNAPol mitocondrial también se ve inhibida, aunque con menor potencia que la procariótica.
- **Estreptolidigina** (producida por *Streptomyces lydicus*), que también se fija a la subunidad β , pero inhibe la elongación.



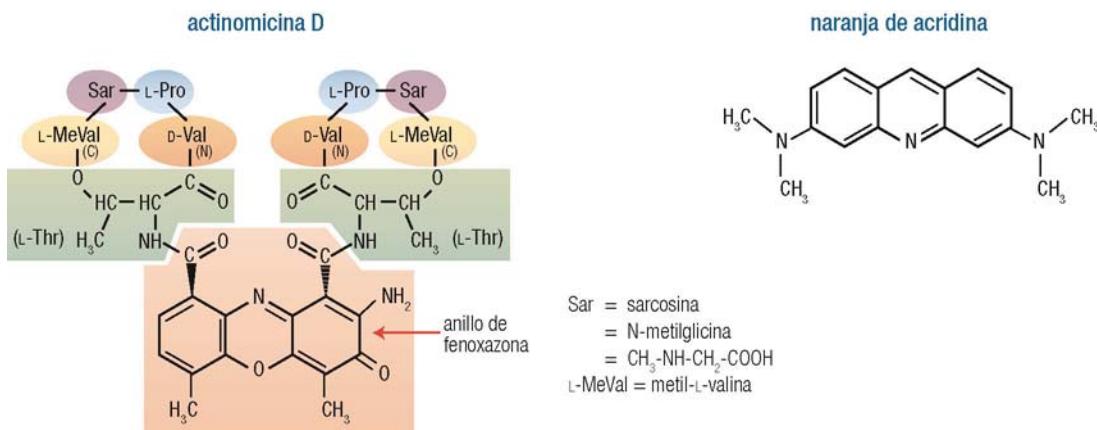
17.15

- La α -amanitina es un inhibidor específico de la etapa de elongación de la síntesis de RNA en eucariotas y no en procariotas. Su efecto inhibidor se ha empleado como un aspecto diferenciador entre las tres RNAPol eucarióticas (pág. 270). Este agente tóxico es un mecanismo de defensa de la seta venenosa *Amanita phalloides* para evitar su ingestión. La intoxicación no provoca síntomas inmediatos, sino el deterioro de la función hepática en los días siguientes, debido a la falta de producción de nuevos mRNA para la síntesis continua de proteínas. Su estructura es de un octapéptido bicíclico con varios aminoácidos infrecuentes.

17.6.2 Antibióticos de tipo II

En este caso, la transcripción se inactiva de manera indirecta, pues el antibiótico se une al DNA impidiendo bien la unión de la polimerasa o bien su avance. Por ello, se inhiben por igual todas las RNAPol, así como las DNAPol (es decir, la replicación). Éste es el caso de la **actinomicina D**, sintetizada por varias especies de *Streptomyces*, formada por dos pentapeptidos cílicos unidos a un sistema anular plano de fenoxazona. Éste se intercala en la doble hélice del DNA, entre dos pares de bases G≡C sucesivos, deformando el DNA. Al mismo tiempo, las dos estructuras pentapeptídicas cílicas abrazan a la doble hélice, situándose en el surco menor. Esta alteración impide el movimiento de la DNAPol a lo largo del molde. La actinomicina D es muy tóxica, pues impide la formación de los complejos abiertos, bloqueando el paso de cualquier tipo de polimerasa. Al inhibir la elongación en células intactas y en extractos celulares, es muy útil para identificar procesos celulares que dependen de la síntesis de RNA.

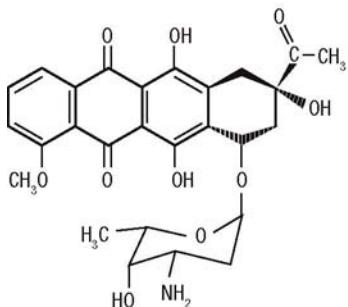
Los intercalantes como la actinomicina deforman la geometría de la doble hélice. En concreto, su inserción entre los pares de bases apilados los separa alrededor de 0,7 nm y desenrolla parcialmente la doble hélice (unos 23°). Dicha distorsión coloca al DNA en una conformación más parecida a la hélice A y cambia la forma de los surcos, por lo que las polimerasas no reconocen correctamente el molde.



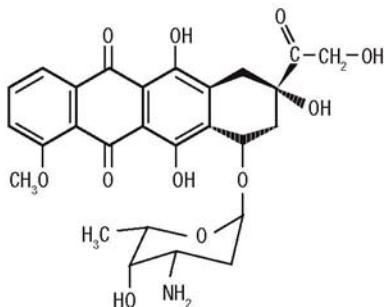
17.16

Otros compuestos intercalantes son la **acridina** y sus derivados, que inhiben la síntesis de RNA de una forma similar, y la **daunomicina** y sus derivados, que se emplean como quimioterapéuticos en diversos tipos de cáncer, ya que actúan preferentemente sobre células con crecimiento rápido. Por último, el **bromuro de etidio**, de uso común en el laboratorio para la tinción de DNA en electroforesis (pág. 141), es también un agente intercalante (aunque no empleado como antibiótico), lo que justifica su toxicidad, que debe tenerse en cuenta durante su manipulación.

daunomicina o daunorrubicina



doxorrubicina o adriamicina



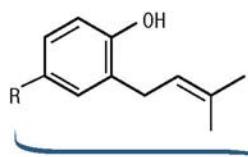
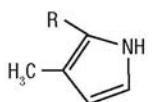
17.17

17.6.3 Antibióticos de tipo III

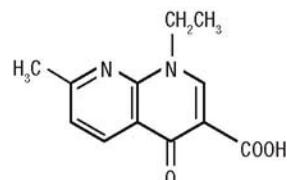
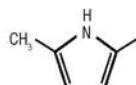
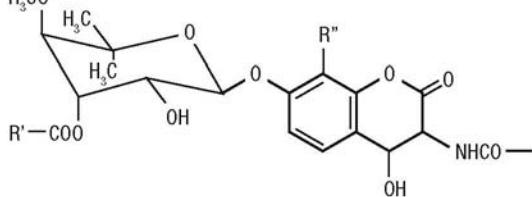
Los antibióticos de tipo III son también inhibidores indirectos, que actúan en este caso impidiendo la acción de la **DNA girasa**, una de las topoisomerasas procarióticas, que introduce superenrollamientos negativos en el DNA (pág. 156). Se comprueba así que éstos son necesarios para la separación de las hebras de DNA, lo que permite la replicación y la transcripción. Los principales antibióticos en este grupo son la **cumermicina**, la **novobiocina** y los del grupo de las quinolonas, como el **ácido nalidíxico** y sus derivados fluorados **norfloxacina** y **ciprofloxacina**.

La girasa es un tetrámero A₂B₂; la novobiocina se une a la subunidad B, bloqueando la unión e hidrólisis del ATP, mientras que las quinolonas se unen a la subunidad catalítica A, inhibiéndola.

novobiocina y clorobiocina

cumermicina A₁

ácido nalidíxico

 $R =$ novobiocina: $R' = \text{NH}_2 -$
clorobiocina y cumermicina A₁: $R' = \text{Cl} -$ novobiocina y cumermicina A₁: $R'' = \text{CH}_3 -$
clorobiocina: $R'' = \text{Cl} -$

17.18



Control de la expresión génica: pretranscripcional y transcripcional

18.1 INTRODUCCIÓN GENERAL A LA REGULACIÓN DE LA EXPRESIÓN GÉNICA	279	
18.2 CONTROL PRETRANSCRIPCIONAL	281	
18.3 REGULACIÓN EPIGENÉTICA	282	
18.3.1 Metilación del DNA	282	
18.3.2 Modificación de las histonas	284	
18.3.2.1 Acetilación de las histonas	284	
18.3.2.2 Metilación de las histonas	285	
18.3.2.3 Otras modificaciones de las histonas	286	
18.3.3 Integración de las señales epigenéticas	286	
18.3.4 Cambios epigenéticos en el desarrollo	286	
18.3.5 Cambios epigenéticos en el cáncer y otras enfermedades	287	
18.4 REGULACIÓN DE LA TRANSCRIPCIÓN	288	
18.4.1 Elementos que actúan en <i>cis</i> y <i>trans</i> como reguladores de la transcripción	288	
		18.4.2 Promotores de los genes de clase II 288
		18.4.2.1 Secuencias o elementos basales del promotor 289
		18.4.2.2 Secuencias o elementos proximales 289
		18.4.2.3 Secuencias o elementos distales 290
		18.4.3 Factores de transcripción de clase II 290
		18.4.3.1 Factores de transcripción generales 291
		a) Ensamblaje consecutivo de los factores generales (modelo "escalonado") 291
		b) Modelo de la holoenzima 292
		18.4.3.2 Factores de transcripción proximales 292
		18.4.3.3 Factores de transcripción inducibles 293
		18.4.3.4 Interacción de los factores de transcripción con la polimerasa 293
		18.4.4 Regulación de la transcripción de genes de las clases I y III 295

18.1 INTRODUCCIÓN GENERAL A LA REGULACIÓN DE LA EXPRESIÓN GÉNICA

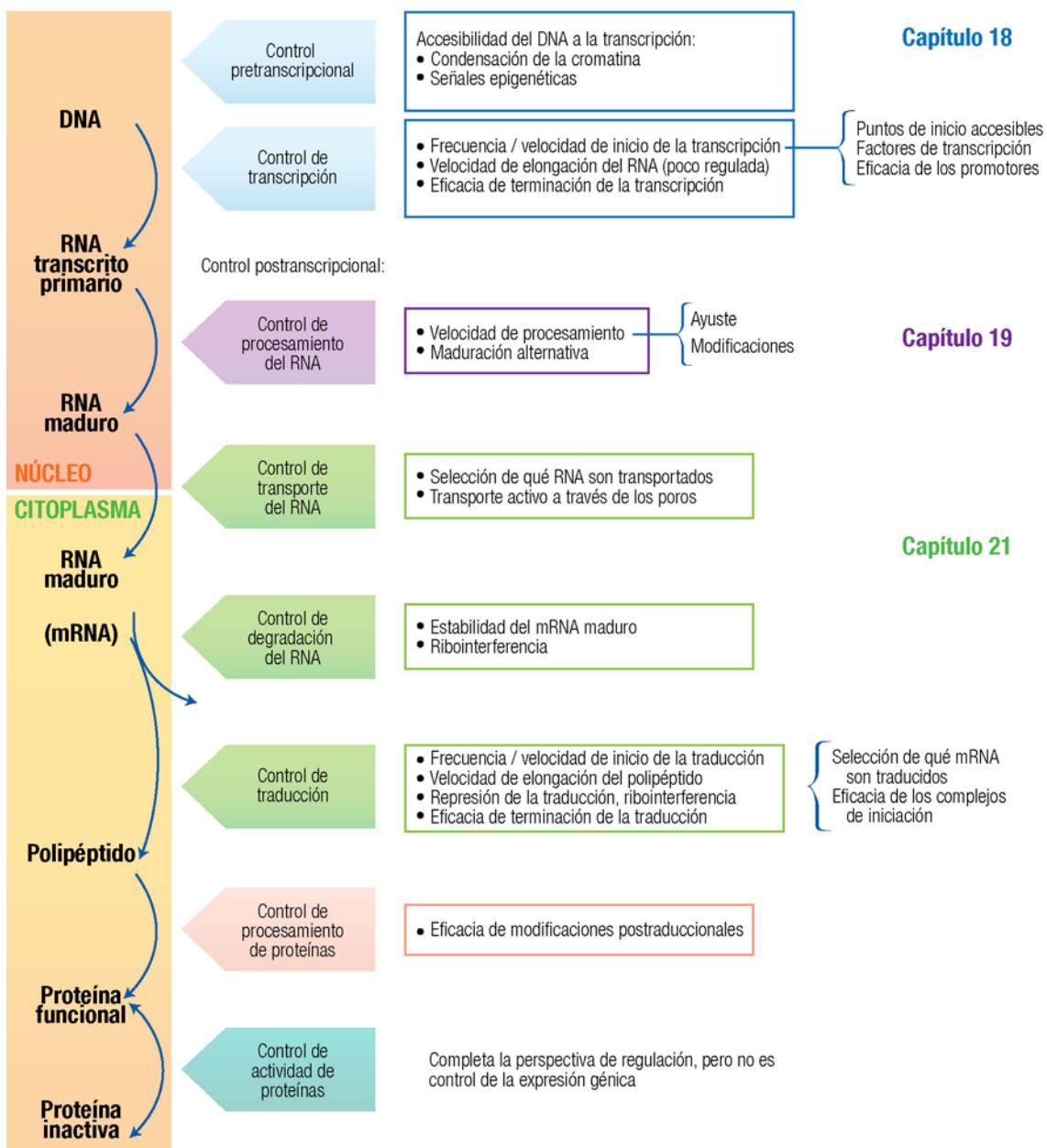
La visión global del control de la expresión génica es especialmente difícil en los eucariotas, particularmente en organismos pluricelulares. Por ejemplo, aunque todas las células humanas contienen el mismo genoma (salvo mutaciones somáticas), existen claras diferencias frente a las células procarióticas. Es todo un reto al conocimiento comprender cómo las células de organismos complejos, procedentes del desarrollo de una sola (el cigoto) y conteniendo un DNA idéntico, son capaces de expresar funciones tan diferentes y únicas en cada órgano o tejido particular. Esta enorme flexibilidad en el control de sus genes viene determinada por los siguientes tipos de *expresión génica diferencial*:

- Distintos tipos celulares en cada especie y, dentro de éstas, en cada órgano o tejido. En el ser humano se calculan unos 200 tipos celulares, agrupados en “familias” según su función específica: células epiteliales, contráctiles, sanguíneas, del sistema inmunitario, sensoriales, neuronales, etc.
- Variedad en el número de genes en cada especie. En el ser humano existen entre 20.000 y 25.000 genes.
- Proporción de genes expresados en cada tipo celular. Una célula humana sólo expresa en cada momento entre 10 y 20% del total de genes (es decir, entre 2.000 y 5.000).
- Influencia sobre la expresión de cada gen de los estadios del ciclo de división celular y del desarrollo del individuo, necesidades metabólicas o de otro tipo, respuesta a agentes externos, etc.

Dentro de estas posibilidades, es de destacar la variación de la expresión génica durante el proceso de diferenciación celular, es decir, la adquisición de propiedades específicas por parte de las diversas células de un organismo, que tiene lugar especialmente durante la embriogénesis. Ello hace que existan genes que no responden a cambio fisiológico alguno o que, por el contrario, están sujetos a fuerte control como parte del desarrollo, de la organización de células en tejidos y de los tejidos en organismos completos.

La expresión de los genes de un organismo superior puede regularse en distintos puntos del proceso de expresión génica. Desde el punto de vista patológico es también importante resaltar que la alteración de cualquiera de los pasos implicados en la expresión génica puede causar enfermedades.

NIVELES DE REGULACIÓN DE LA EXPRESIÓN GÉNICA EN EUCAΡΙOTAS



El estudio en profundidad de todos estos puntos de control debe hacerse en tratados especializados. En este texto sólo trataremos de desglosar de forma paulatina lo esencial, cada uno de ellos dentro de su correspondiente proceso de expresión génica (Capítulos 18, 19 y 21).

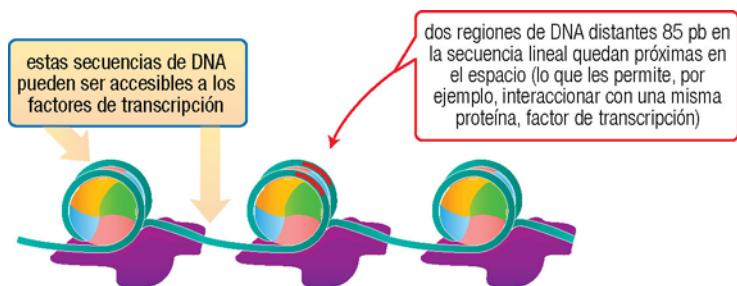
18.2 CONTROL PRETRANSCRIPCIONAL

El inicio de la transcripción requiere que la maquinaria molecular responsable (RNA polimerasa y factores de transcripción, pág. 288) pueda acceder físicamente a las bases nitrogenadas de la hebra de DNA que se ha de transcribir. Ello supone la descondensación previa del cromosoma durante la fase G₁ del ciclo de división celular (pág. 100) y, particularmente, la descondensación más completa de las regiones correspondientes a cada gen que deba transcribirse en un momento y circunstancias determinados.

La transcripción no puede tener lugar sobre DNA en estados de condensación superiores a la fibra de 10 nm (cadena de nucleosomas, pág. 85). En ocasiones, el DNA se puede transcribir cuando se encuentra formando complejos con las histonas (nucleosomas), pero también puede ser necesario que éstos se desensamblen para permitir el acceso al DNA. En general, la accesibilidad del DNA a la transcripción, o **actividad transcripcional de la cromatina**, está relacionada estrechamente con el grado de condensación (pág. 88) y puede valorarse experimentalmente, por ejemplo, midiendo la sensibilidad a la acción de DNAsas.

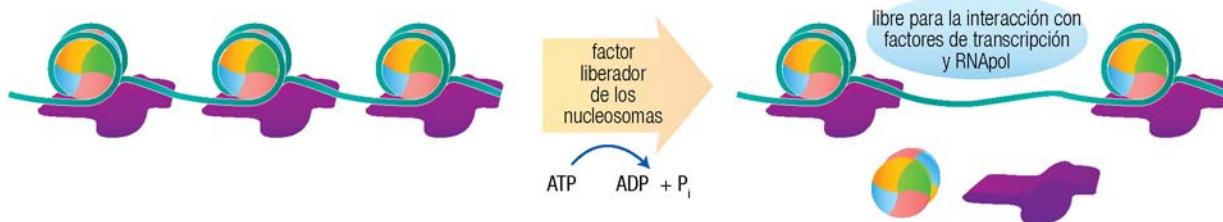
Se puede considerar que el acceso al DNA necesario para la transcripción es una combinación de tres mecanismos:

- a) **Posición precisa de los nucleosomas.** El reconocimiento de las secuencias promotoras del DNA por los factores de transcripción (pág. 288) puede tener lugar, al menos en algún caso, gracias a una posición muy específica de los nucleosomas, de forma que tales secuencias queden en la región internucleosómica (DNA espaciador, pág. 86) o, si forman parte del nucleosoma, queden orientadas hacia su superficie externa.



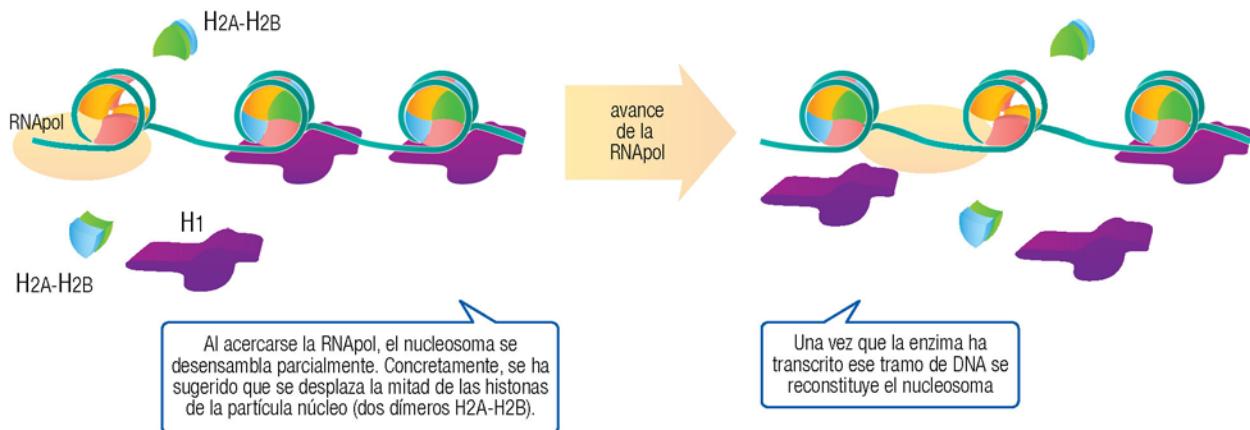
18.2

- b) **Retirada de los nucleosomas.** Tanto el reconocimiento de las secuencias promotoras por los factores de transcripción como la unión y avance de la RNAPol pueden requerir la liberación de las histonas, o al menos el desplazamiento de posición de los nucleosomas, para hacer el DNA accesible. Se han encontrado proteínas capaces de efectuar este “remodelado de la cromatina” gracias a la hidrólisis de ATP.



18.3

- c) **Avance compatible con la presencia de nucleosomas.** La transcripción puede también tener lugar en regiones de DNA con nucleosomas íntegros; posiblemente, la RNA polimerasa se desplace a lo largo de la fibra de 10 nm, transcribiendo una hebra gracias a un desensamblado parcial y reversible de los nucleosomas a su paso.



Un factor importante en los mecanismos b y c anteriores es la modificación covalente experimentada por las histonas. Tal modificación se considera parte de los mecanismos epigenéticos de regulación, que se describen a continuación.

18.3 REGULACIÓN EPIGENÉTICA

El control de la transcripción debe plantearse de acuerdo con dos componentes, ambos relacionados –aunque de modo distinto– con la estructura del genoma: la secuencia de bases (regulación genética) y las modificaciones experimentadas por el DNA y las histonas sin afectar a la secuencia (regulación epigenética).

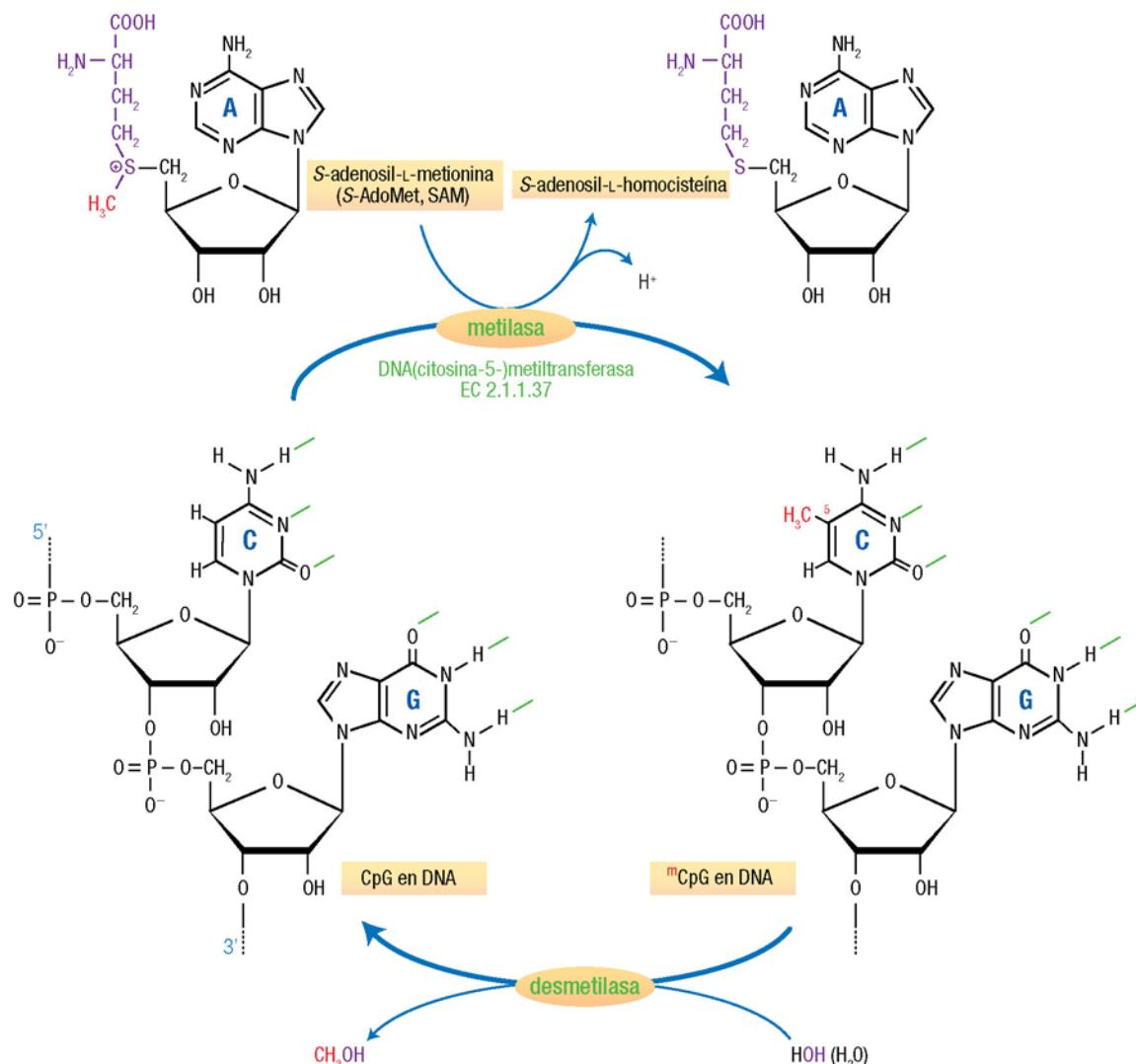
Se puede describir la epigenética como el estudio del conjunto de cambios en el patrón de expresión génica que no implican alteración de la secuencia del DNA y son heredables. De forma más concreta, corresponde al conjunto de marcas moleculares y mecanismos que señalan y perpetúan la actividad funcional de las diferentes regiones de la cromatina –a través de su estado de condensación (pág. 88).

Se trata, pues, de una regulación ejercida a nivel **pretranscripcional**, que tiene consecuencias sobre el fenotipo y es importante para explicar la diversidad morfológica y funcional de células que poseen un mismo genoma. Bajo el concepto de epigenética se incluyen la metilación del DNA en residuos de citidina, diversas modificaciones covalentes reversibles de las histonas (en particular en las colas amino-terminales de H3 y H4, pág. 87), la remodelación de los nucleosomas y la reorganización de la cromatina a mayor escala.

18.3.1 Metilación del DNA

En el DNA de vertebrados una parte de los residuos de citidina (en torno al 4% en los seres humanos) se encuentran metilados. Esta metilación aparece de manera casi exclusiva sobre la secuencia dinucleotídica CG, que además es especialmente abundante en las regiones promotoras de los genes, constituyendo los denominados **islotes CpG**. Gran parte de los genes humanos poseen islotes CpG en su promotor y su primer exón.

Los residuos de citidina presentes en secuencias CG pueden experimentar la metilación en el carbono 5 de la citosina, de forma específica catalizada por **DNA metiltransferasas (DNMT)** (o simplemente **metilasas**). La enzima actúa de manera similar a las metilasas que forman parte de los sistemas de metilación-restricción en bacterias (pág. 213) y de la maduración del mRNA en eucariotas (pág. 301), empleando *S*-adenosilmetionina como coenzima donadora de los grupos metilo (pág. 19) y, en este caso, con especificidad por citidina adyacente a guanosina.

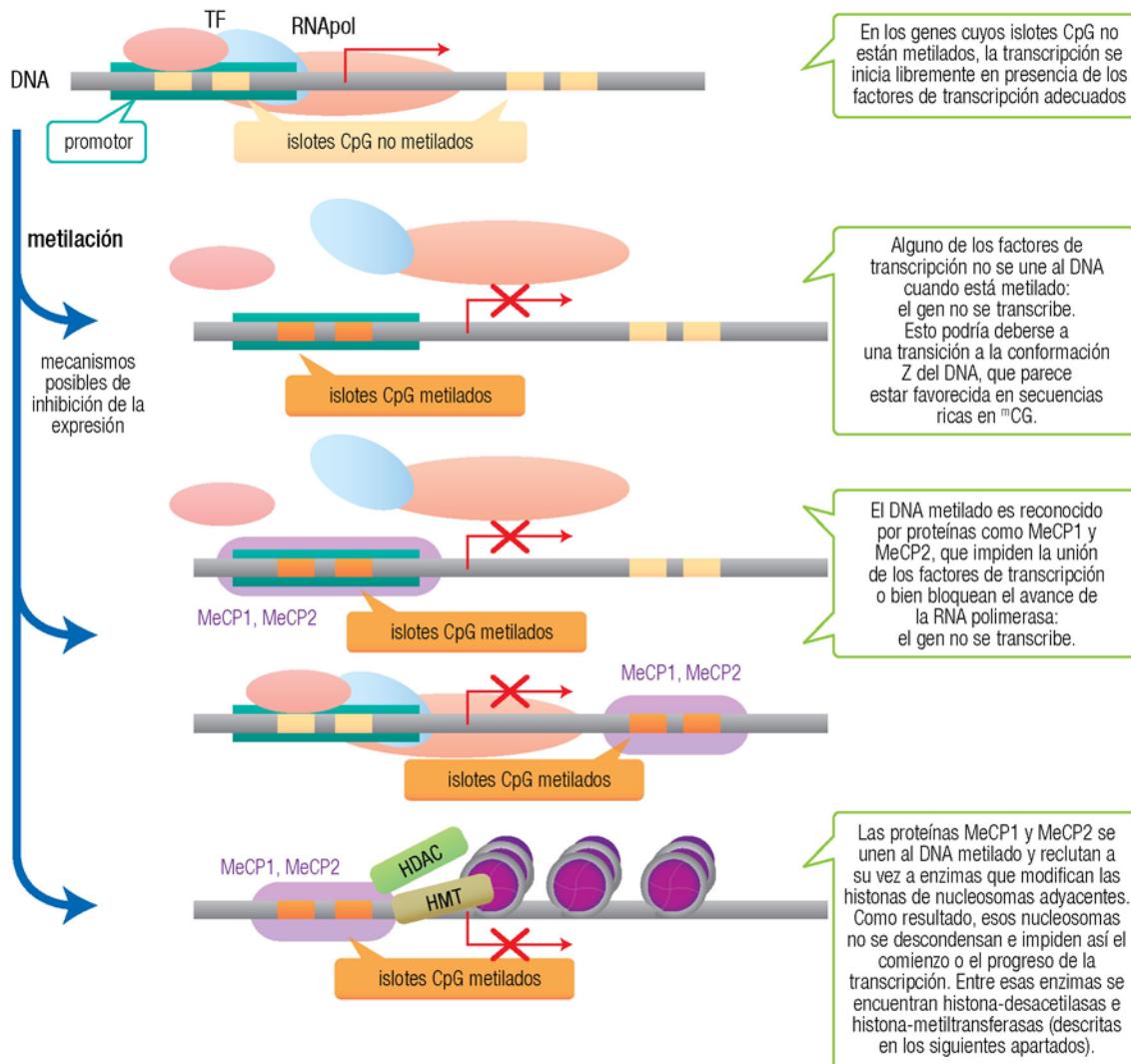


18.5

La reacción opuesta, responsable del proceso de desmetilación, la cataliza una **desmetilasa**, que reconoce las secuencias metiladas ${}^m\text{CpG}$ y elimina el grupo metilo de la citosina. También es posible que se elimine la base completa (mediante una **glicosilasa**) y luego se reponga la citosina no metilada gracias a los mecanismos de reparación del DNA (pág. 402).

Se ha observado que los genes que se transcriben de forma activa en un tejido poseen isletos CpG sin metilar (o isletos *hipometilados*), mientras que en los tejidos donde el gen no se expresa sí están metilados (o *hipermetilados*). Por tanto, la metilación de la citosina desempeña un papel en la regulación de la expresión génica.

El bloqueo de la transcripción de un gen como consecuencia de la presencia de grupos metilo en la molécula de DNA se puede explicar mediante varios mecanismos, no excluyentes:



18.6

18.3.2 Modificación de las histonas

Las histonas, componentes esenciales de la cromatina (pág. 84), experimentan varios tipos de modificación postraduccional una vez que ya están formando parte de los nucleosomas. Dichas modificaciones covalentes son reversibles y están reguladas con precisión mediante parejas de enzimas con acción opuesta. La principal diana de modificación son los residuos aminoácidos de las colas amino-terminales de las histonas H3 y H4, que sobresalen de la estructura central del nucleosoma y son por ello fácilmente accesibles (v. web 7.4).

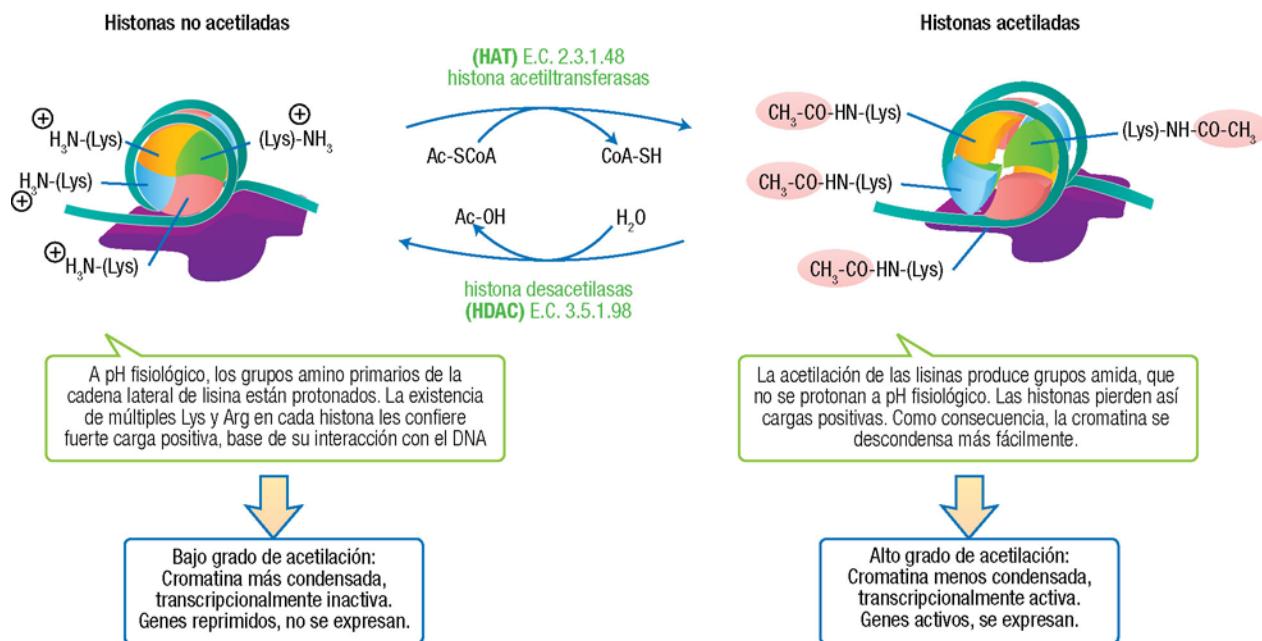


Web 18.1. Reacciones, enzimas y coenzimas implicadas en la modificación de las histonas.

18.3.2.1 Acetilación de las histonas

Existen enzimas en la célula que unen grupos acetilo a ciertos residuos de lisina de las histonas. Varias de estas histona acetiltransferasas (HAT) son factores de transcripción (en especial, coactivadores, pág. 294). Asimismo, existen histona desacetilasas (HDAC) que catalizan la reacción opuesta de eliminación del grupo acetilo.

18.7



Los genes activos, los que se están expresando, corresponden a regiones con histonas acetiladas, mientras que en las zonas de cromatina transcripcionalmente inactiva las histonas no están acetiladas. Como ejemplo, el cromosoma X inactivo en la mujer, en forma de heterocromatina (pág. 88), está hipoacetilado, lo que indica que esta forma de las histonas es un requisito para la máxima condensación de la cromatina, mientras que el otro cromosoma X, transcripcionalmente activo, tiene un grado de acetilación normal.

Además del efecto directo del cambio de carga sobre la interacción entre histonas y DNA nucleosómico, los residuos de lisina acetilada los reconocen ciertas proteínas, actuando así como señal para la asociación de más factores relacionados con la condensación de la cromatina. Por ejemplo, uno de los TAF que forman el factor de transcripción TFIID y alguna histona acetiltransferasa poseen una región denominada **bromodominio**, que se une a la lisina acetilada.

18.3.2.2 Metilación de las histonas

Otra modificación común es la adición de grupos metilo a los residuos de lisina y arginina. La reacción, con S-adenosilmetionina como donador del grupo metilo, la catalizan **histona metiltransferasas (HMT)**; se han identificado múltiples enzimas diferentes, que son específicas para modificar residuos concretos de una histona en particular y asimismo en cuanto al número de metilos que pueden añadir (hasta 3 metilaciones sucesivas en la lisina, dos en la arginina). La modificación se revierte por la acción de **histona desmetilasas**.

En este caso, no existe una clara alteración de la carga eléctrica, sino que el residuo mono-, di- o trimetilado supone una marca o señal que reconocen algunas proteínas. La consecuencia es diferente dependiendo de cuál sea el residuo metilado (y del número de metilos). Por ejemplo, las metilaciones de Lys4, Arg17 y Lys36 de la histona H3 activan la transcripción, mientras que las metilaciones de Lys9 y Lys27 la reprimen. Los residuos de lisina metilada son también reconocidos por proteínas, en este caso con una estructura llamada **cromodomino** (el nombre deriva de *chromatin organization modifier*, modificador de la organización de la cromatina).

18.3.2.3 Otras modificaciones de las histonas

Las histonas pueden sufrir otros tipos de modificación en sus residuos aminoácidos, menos frecuentes o con un papel no tan claramente definido, pero que también contribuyen a controlar el estado de condensación del material genético y, en consecuencia, su accesibilidad para la transcripción.

- Fosforilación en residuos de serina; de forma similar a la metilación, introduce marcas que reconocen otras proteínas.
- Ubiquitinación en lisinas (es decir, unión de la proteína ubiquitina, pág. 377). Por ejemplo, se ha observado que la ubiquitinación de H2A y H2B descondensa el nucleosoma.
- Sumoilación en arginina y lisina (adición de radicales SUMO, *small ubiquitin-related modifier*). Las proteínas SUMO son similares a la ubiquitina, pero no marcan para la degradación.
- ADP-ribosilación (unión de ADP-ribosa).

18.3.3 Integración de las señales epigenéticas

Las modificaciones epigenéticas no tienen lugar de forma aislada, sino que actúan de forma coordinada. La metilación del DNA puede ser la señal que provoca la modificación de las histonas, mientras que en otras ocasiones sucede al contrario; además, muchas reacciones sobre las histonas tienen como resultado la captura o reclutamiento de las enzimas responsables de otras modificaciones adicionales. Se puede resumir esto como una acción sinérgica de las diferentes señales epigenéticas. De acuerdo con ello, las marcas epigenéticas definen compartimentos o dominios en el genoma: activos (eucromatina), potencialmente activos (heterocromatina facultativa) y silenciados (heterocromatina constitutiva) (pág. 88). Como ejemplo:

- En la eucromatina predominan la hipometilación del DNA, la fosforilación de la histona H3, la acetilación de H3 y H4, y la metilación de H3 y H4.
- La heterocromatina constitutiva se caracteriza por la metilación del DNA, la hipoacetilación de H3 y H4, y la trimetilación de Lys9 de H3 y Lys20 de H4.
- En la heterocromatina facultativa abundan los islotes CpG metilados, la hipoacetilación de H3 y H4, la trimetilación de Lys27, la dimetilación de Lys9, la metilación de Lys36 (todas en H3) y la mono- o dimetilación de Lys20 en H4.

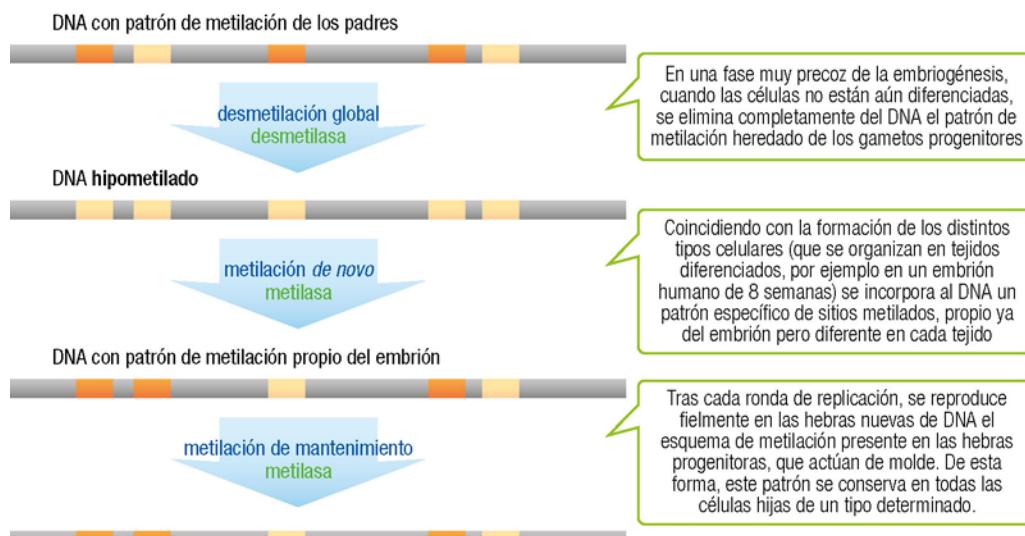
Como ejemplos de situaciones reguladas por mecanismos epigenéticos pueden mencionarse:

- La inactivación de uno de los cromosomas X en las mujeres (pág. 88), asociada con hipermetilación de todo su DNA y modificación de sus histonas.
- La supresión en ciertos casos de la expresión de uno de los dos alelos de un gen en cromosomas homólogos, conocida como sellado génico o impronta genómica (en inglés, *genomic imprinting*).
- La expresión de genes específicos de tejido durante el proceso de diferenciación.
- En general, el control de la organización de la cromatina y de la estabilidad genómica. Por ejemplo, la completa condensación de las regiones próximas al centrómero es clave para el reparto correcto de los cromosomas durante la mitosis.

18.3.4 Cambios epigenéticos en el desarrollo

En el cigoto –incluso antes, en las células germinales primordiales que darán lugar a los gametos– se pierden todas las marcas epigenéticas del genoma que luego, al progresar la diferenciación, deben recrearse de forma selectiva en cada célula del embrión temprano; esto se denomina **reprogramación**, y está estrechamente relacionado con el carácter totipotente, pluripotente o multipotente de cada estadio celular (pág. 199).

Como manifestación de la regulación mediante metilación del DNA, cada tejido presenta un patrón característico de metilación en sus células. Éste se consigue y conserva durante el desarrollo embrionario, mediante tres procesos sucesivos de distribución de los grupos metilo.



18.8

De forma similar, se reproduce el patrón específico de modificación de las histonas (lo que se ha llamado el *código histónico*). Para que el DNA se replique en cada división celular las histonas deben separarse; una vez que ha pasado la polimerasa, las moléculas originales de histonas (incluyendo sus marcas epigenéticas) se reasocian con el DNA repartiéndose entre las dos cadenas hijas y se complementan con histonas de nueva síntesis. La interconexión mencionada entre las diversas marcas epigenéticas permite que las nuevas histonas reciban rápidamente las modificaciones equivalentes a las antiguas, en cada región del genoma.

18.3.5 Cambios epigenéticos en el cáncer y otras enfermedades

Las alteraciones en la regulación de las señales epigenéticas desorganizan la regulación de la expresión de los genes, por lo que son causa de numerosas enfermedades. Se ha identificado un componente epigenético, por ejemplo, en el lupus, el síndrome del cromosoma X frágil, la aterosclerosis, la inmunodeficiencia y el síndrome de Rett. Destaca, obviamente, la implicación en procesos cancerosos, dado que la pérdida del control transcripcional conduce fácilmente a la proliferación celular desordenada y a la aparición de fenotipos alterados, característicos de este grupo de enfermedades (Capítulo 26).

Los procesos de metilación y desmetilación del DNA se han podido investigar gracias a la observación en células cancerosas de alteraciones en el patrón de metilación del DNA, posiblemente responsables de la modificación en el programa de expresión génica característica del proceso tumoral, aunque no su causa única. En esas células el DNA está globalmente **hipometilado**, pero al mismo tiempo algunos genes están **hipermetilados** (en concreto, los *genes oncosupresores*, pág. 456). Esta aparente contradicción es el resultado, por mecanismos complejos, de una elevada expresión de ambas enzimas, la metilasa y la desmetilasa. La hipometilación provocaría la expresión de genes que suelen estar silenciados (incluido el de la desmetilasa), lo que conduce posiblemente a algunos de los desajustes causantes del cáncer, mientras que la hipermetilación de los genes oncosupresores bloquea su expresión, eliminándose su función natural, el freno de la proliferación celular, y conduciendo a la multiplicación anárquica de las células que genera el tumor (Capítulo 26).

El conocimiento creciente de las señales epigenéticas ofrece un potencial prometedor en la detección precoz del cáncer y en la definición de marcadores con valor pronóstico para diferenciar tipos de cáncer, así como en la terapia, pues ya se dispone de varios inhibidores específicos para algunas de las enzimas que modifican el DNA y las histonas.

18.4 REGULACIÓN DE LA TRANSCRIPCIÓN

El mecanismo de control que se ha descrito tradicionalmente como regulador de la transcripción se basa en el **componente genético**, la secuencia de bases del DNA (por contraposición al componente epigenético). En efecto, la expresión de la mayoría de los genes eucarióticos se controla en el inicio de la transcripción, por mecanismos conceptualmente muy similares a los encontrados en bacterias. Esta regulación se centra en la interacción de los promotores con los factores de transcripción (TF). A fin de simplificar nuestro estudio, se prestará atención preferente al control por promotores y TF de la síntesis del mRNA por la RNApol-II, es decir, la transcripción de genes de clase II. Más adelante se describirá de forma breve el control de la transcripción de los genes de clases I y III.

18.4.1 Elementos que actúan en *cis* y en *trans* como reguladores de la transcripción

Debido a la magnitud del genoma eucariótico, el número de unidades de transcripción es mucho mayor que en procariotas. Además, cada una de ellas está controlada por varios promotores, con lo que en total el número de éstos es aún superior (decenas o centenares de miles de promotores en un genoma). Es decir, las regiones promotoras están más extendidas y tienen más elementos de control en eucariotas que en procariotas.

Un **promotor** o **secuencia promotora** es una región de DNA que regula el inicio de la transcripción de un gen. Corresponde, por tanto, a lo que hemos denominado región reguladora del gen (pág. 264). Actualmente suelen recibir el nombre también de “**factores que actúan en *cis***”, o “**factores *cis***”, simplemente por encontrarse en la misma molécula de DNA cuya expresión regulan. Existen numerosos promotores, diferentes para los genes de las clases I, II y III, así como algunos de ellos distintos para cada gen.

Una característica de la transcripción eucariótica es que su regulación se debe a la afinidad específica por los promotores de ciertas proteínas denominadas factores de inicio de la transcripción, o simplemente **factores de transcripción** (TF). A cada promotor se une un número variable de factores de transcripción, que actúan favoreciendo o dificultando la unión y la actividad de la RNApol o de otros factores de transcripción y, en consecuencia, determinando la posición y la eficacia del inicio de la transcripción. Por tanto, los TF tienen más responsabilidad que la propia RNApol en el reconocimiento de la secuencia del promotor y se encargan de la regulación de la transcripción. Su diversidad es aún mayor que la de promotores, lo que permite la gran flexibilidad existente en el control de la expresión de los genes. Los TF se denominan también “**factores que actúan en *trans***” o “**factores *trans***”, pues sus genes están en posición alejada y no relacionada con aquella región génica cuya expresión regulan. Su capacidad de unión a la molécula de DNA, casi siempre reconociendo ciertas secuencias de bases, depende habitualmente de que las moléculas de TF contienen motivos estructurales ligantes del DNA, como dedos de zinc, homeodomelios, etc. (pág. 58).

La especialización de las RNApol eucarióticas para transcribir sólo determinados genes (pág. 270) viene determinada por la regulación de estos genes mediante secuencias promotoras y factores de transcripción diferentes. Antes de describir los pormenores del control de la transcripción de los genes de clase II conviene recopilar la terminología empleada para las moléculas que intervienen. Se distingue así entre promotores para genes de la clase I, de la clase II y de la clase III. Los factores de transcripción respectivos se nombran con las iniciales TF seguidas del tipo de RNApol con la que pueden actuar (I, II o III) y una letra que los identifica (por ejemplo, TF_{II}A, TF_{II}B, TF_{II}D son 3 factores de transcripción para la RNApol-II).

Polimerasa (enzima)	Genes que transcribe (DNA)	Producto génico (RNA o proteína)	Promotores (DNA)	Factores de transcripción (proteínas)
RNApol-I	“de clase I”	rRNA de 28S, 18S y 5,8S	“de clase I”	TF _I
RNApol-II	“de clase II”	proteínas y snRNA	“de clase II”	TF _{II}
RNApol-III	“de clase III”	tRNA, rRNA de 5S y RNA pequeños	“de clase III”	TF _{III}

18.4.2 Promotores de los genes de clase II

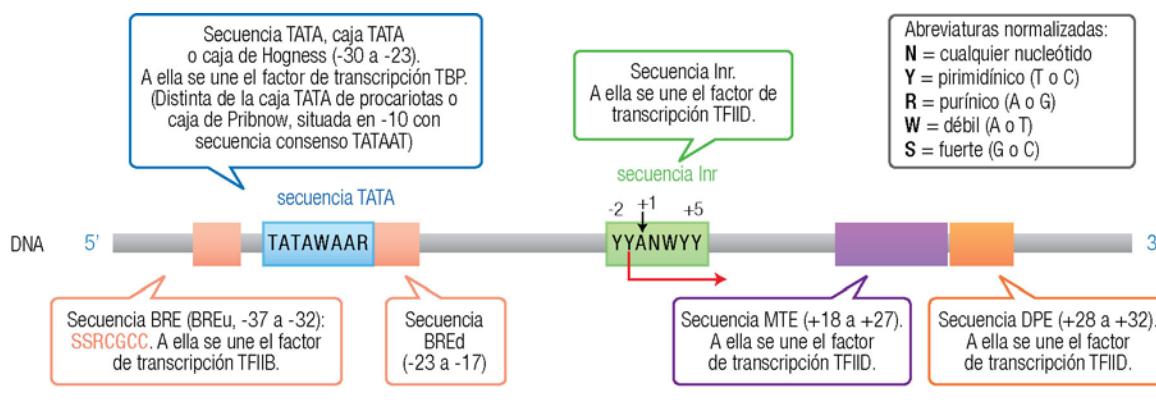
La regulación de los genes de clase II (transcritos por la RNApol-II) posee máximo interés por ser responsable de la síntesis de los mRNA y, por tanto, de la síntesis de todas las proteínas. Sus promotores se localizan

habitualmente en dirección 5' del origen de transcripción (cadena arriba) y muestran una mayor variación de secuencia que los promotores de clases I y III.

Los promotores en general, pero de forma característica los de clase II, se pueden dividir en 3 categorías, de acuerdo con su acción y con su ubicación. Su función se comprenderá mejor al estudiar los factores de transcripción que se unen a ellos.

18.4.2.1 Secuencias o elementos basales del promotor

El promotor basal comprende las secuencias que *definen el punto de inicio de la transcripción y son imprescindibles para que ésta comience*. Es una región situada en posición adyacente al origen de transcripción y se extiende, según los casos, entre la posición -37 y la +32. Dentro de esta región se ubican –dependiendo del gen– distintas combinaciones de unas pocas secuencias características, presentándose habitualmente entre 2 y 4 de ellas.



Inr = *initiator*; elemento iniciador

BRE = *TFIIB recognition element*; elemento de reconocimiento para el TFIIB. (BREu = *upstream*; BREd = *downstream*)

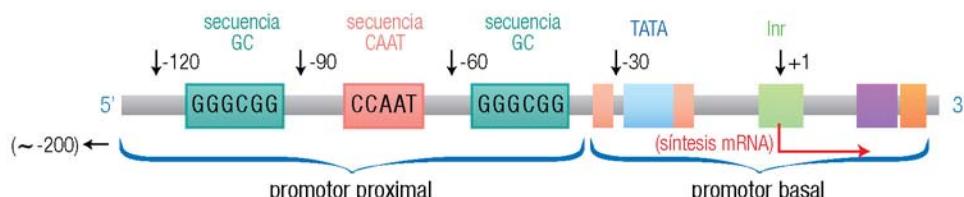
MTE = *motif 10 element*; elemento del motivo 10

DPE = *downstream promoter element*; elemento promotor de cadena abajo

18.9

18.4.2.2 Secuencias o elementos proximales

En general, el promotor basal no es suficiente por sí mismo para provocar el inicio de la transcripción, sino que se requiere la intervención adicional de otra región conocida como promotor proximal. Éste está cercano al promotor basal, pero más alejado cadena arriba del origen, comúnmente entre las posiciones -30 y -200. Los elementos proximales no especifican la posición de inicio, sino que *determinan la frecuencia con la que se produce el inicio de la transcripción*. Ello es posible gracias a que sobre ellos se unen diversos factores de transcripción que favorecen la interacción de la RNAPol-II con el DNA en el punto de inicio y su actividad enzimática. Aunque en general las secuencias son más variadas que las basales, las dos más características son la caja o secuencia CAAT, entre -60 y -80, y la caja GC, que aparece en copias múltiples y con cualquier orientación a ambos lados de la caja CAAT.

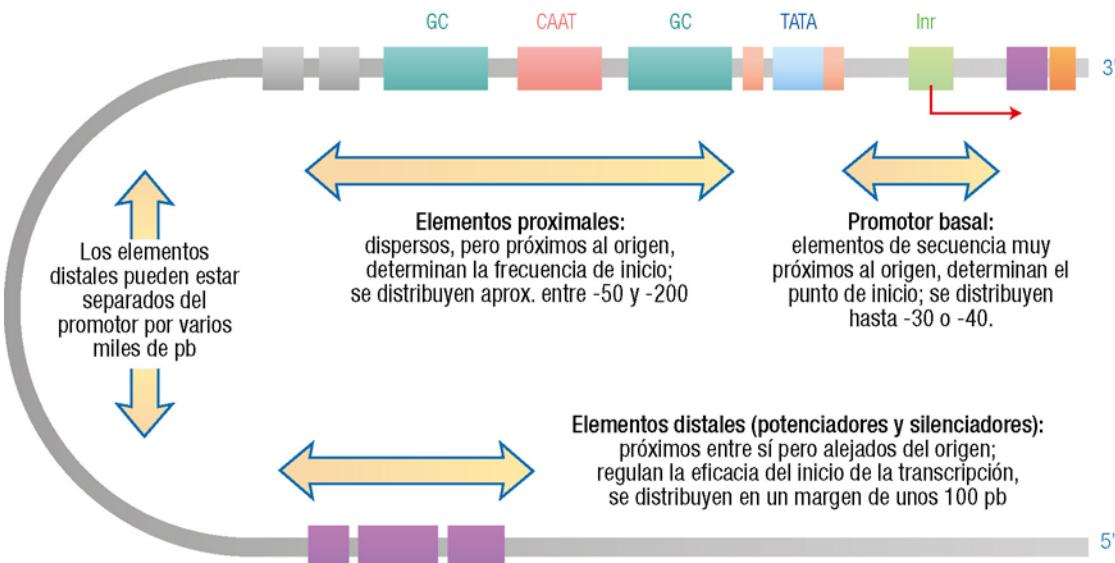


18.10

18.4.2.3 Secuencias o elementos distales

La expresión de algunos genes sufre una regulación aún más compleja, que depende de secuencias situadas a gran distancia del punto de inicio, incluso a varios miles de pb, cadena arriba o cadena abajo. Esto ocurre en especial para los **genes inducibles**, los que no se expresan continuamente en la célula, sino *cuya expresión sufre una regulación amplia y precisa como respuesta a diversas señales* (por ejemplo, las hormonas esteroideas y tiroídeas). En general se denominan **secuencias promotoras distales**, por encontrarse alejadas del origen, o también, de acuerdo con su función, **elementos específicos de regulación, módulos de control o elementos de respuesta**.

Algunos autores prefieren incluir como promotores únicamente los elementos basales y proximales, dejando a los distales como secuencias reguladoras. No obstante, atendiendo al concepto funcional de gen (pág. 263) parece apropiado contemplar estas secuencias promotoras también como parte del gen.



Estas secuencias promotoras son muy variadas y específicas para cada gen. Otra particularidad es que actúan tanto activando la transcripción como frenándola. De acuerdo con ello, se clasifican en tres tipos:

- **Potenciadores**, activadores, intensificadores, amplificadores o estimuladores (*enhancers*). Pueden aumentar mucho la velocidad de inicio de la transcripción (y, por consiguiente, la del proceso completo) conseguida a partir de promotores basales y proximales situados en la misma molécula de DNA.
- **Silenciadores o inhibidores** (*silencers*). Tienen el efecto opuesto, en general debido a que los factores de transcripción que se unen a ellos compiten con la acción de aquellos específicos para los promotores potenciadores y proximales.
- **Aisladores** (*insulators*). En esta categoría se incluyen ciertas secuencias que impiden que la acción de potenciadores y silenciadores se extienda más allá de ellas. Se ha propuesto que los factores que se unen a ellos forman lazos en la cromatina, que quedan así aislados del resto evitando que los factores de transcripción que se unen a los promotores distales se desplacen más allá. Más que meras barreras físicas deben considerarse como reguladores, pues interaccionan de manera específica con diversas proteínas.

Quizá parezca extraño que secuencias tan alejadas puedan actuar sobre el inicio de la transcripción. En realidad, esta lejanía es relativa, considerando el giro de la cadena en la doble hélice y el enrollamiento de ésta en torno a los nucleosomas. Además, la molécula de DNA (con o sin nucleosomas) es flexible y regiones distantes pueden acercarse sin problemas.

18.4.3 Factores de transcripción de clase II

La RNApol-II interacciona con gran variedad de factores de transcripción (TFII) dependiendo del gen, habitualmente de forma simultánea con un número considerable de ellos. Estos factores se pueden clasificar en

generales, proximales e inducibles, de acuerdo con su unión a uno de los 3 tipos de secuencias o elementos promotores recién estudiados. Se estudiarán a continuación con cierto detalle los que actúan sobre el promotor TATA, mejor conocidos.

18.4.3.1 Factores de transcripción generales

Reciben este nombre los TF que reconocen los elementos basales de un promotor. Son proteínas muy conservadas evolutivamente, que promueven la formación alrededor del punto de inicio de un complejo plurimolecular que incluye el promotor basal, los propios TF generales y la RNAPol-II. Se puede afirmar que al reconocer el promotor basal actúan de mediadores para que se fije la polimerasa, definiendo así el punto de inicio de la transcripción y activando a la enzima para que comience a sintetizar el RNA.

Se conocen las funciones de algunos de los factores de transcripción generales asociados a la RNAPol-II:

- **TFIID** es el más significativo de todos ellos. Este factor determina que la RNAPol-II actúe a una cierta distancia del promotor, definiendo así el punto de inicio de la transcripción (nucleótido +1). Además, la interacción asimétrica con **TFIIB** (descrita más adelante) define el sentido de la transcripción. **TFIID** tiene mayor tamaño incluso que la polimerasa (cerca de 800 kDa), pues está formado por dos tipos de componentes:
 - Una molécula de **TBP**, o “proteína ligante de TATA” (*TATA binding protein*), que confiere a **TFIID** la capacidad para reconocer la secuencia promotora. Su unión a ésta es el primer paso en la formación del complejo de inicio, con la particularidad de unirse en el surco menor del DNA (a diferencia de la mayoría de proteínas, que se unen al DNA en el surco mayor) y así provocar la flexión de la molécula de DNA.
 - Varias moléculas de **TAF_{II}** o “factores asociados a TBP” (*TBP-associated factors*), diferentes entre sí y que pueden variar para la unión a promotores diferentes. Son subunidades de 30 a 250 kDa que forman parte del **TFIID** en número de 9 a 12, habitualmente. El papel de los TAF va más allá de una activación genérica: muchos son proteínas reguladoras y específicas de un tejido, otros son mediadores en la interacción con otras proteínas reguladoras. Algunos TAF reconocen de manera directa las secuencias **Inr**, **DPE** y **MTE**.

A través de TBP, el **TFIID** posee especificidad para unirse a los promotores con caja TATA. A través de los diferentes TAF, otras variantes de **TFIID** pueden unirse a promotores carentes de caja TATA.

- **TFIIC** es particularmente importante por su doble actividad enzimática, como helicasa y como proteína quinasa, residentes en distintas subunidades (por ello, en ocasiones se denomina **TFIIC/TFIIC**). La helicasa (**H**) está implicada en la separación de las hebras del DNA en el sitio de inicio, necesaria para el acceso de la polimerasa a las bases, mientras que como quinasa (**K**) fosforila el dominio C-terminal de la subunidad mayor de la RNAPol-II (CTD, pág. 271), provocando en ésta un cambio conformacional que induce el comienzo de su desplazamiento a lo largo del DNA, la separación de algunos TF y, en definitiva, la transición entre el inicio y la elongación de la transcripción. Este factor de transcripción es también singular porque interviene en un proceso independiente, la reparación del DNA por escisión de nucleótidos (pág. 401).



Web 18.2. Detalles de la interacción molecular de TBP con la caja TATA en el DNA.

Para la formación del complejo de inicio se han propuesto dos mecanismos, cuyo resultado final es el mismo:

a) Ensamblado consecutivo de los factores generales (modelo "escalonado")

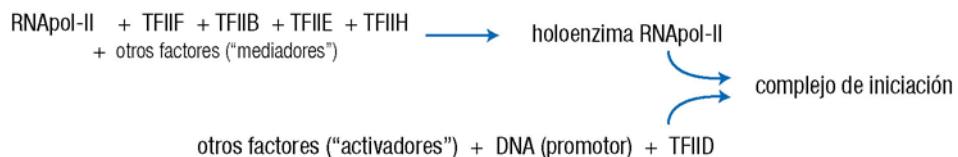
Según esta hipótesis, los TF y la RNAPol-II se van asociando a la región promotora en un cierto orden uno tras otro; se considera que cada factor va estableciendo un complejo que posee la estructura adecuada para que se una el siguiente. La formación del complejo de iniciación comienza con la unión de **TFIID** al DNA en la secuencia TATA (u otra promotora) y, a continuación, se van asociando los factores **TFIIC** y **TFIIB**, que permiten la unión de la polimerasa junto con el factor **TFIIF**. Finalmente se incorporan **TFIE** y **TFIH**.



Web 18.3. Asociación de los TFII y la RNAPol-II sobre el DNA para formar el complejo de inicio.

b) Modelo de la holoenzima

En esta hipótesis alternativa, también apoyada por estudios *in vitro*, la RNAPol-II se asocia previamente a varios factores de transcripción, formando una enzima activa u **holoenzima**, que luego se une al DNA en la secuencia promotora definida por la unión del TFIIID.

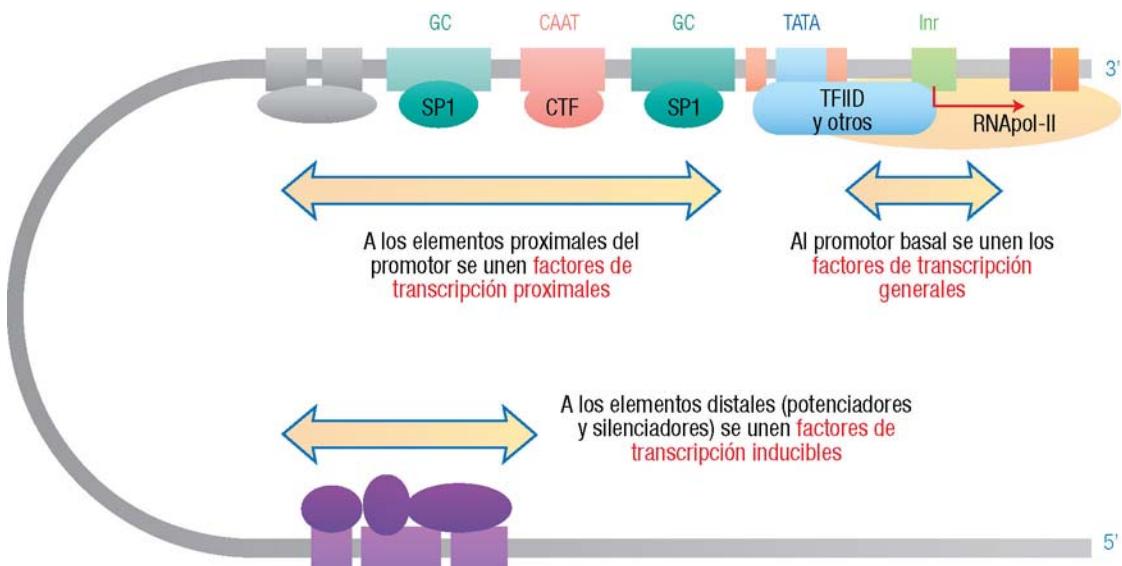


El resultado es el mismo complejo de iniciación que en el modelo escalonado, que permite el comienzo de la transcripción, y la polimerasa luego lo abandonaría y avanzaría igual que se ha indicado en aquel modelo.

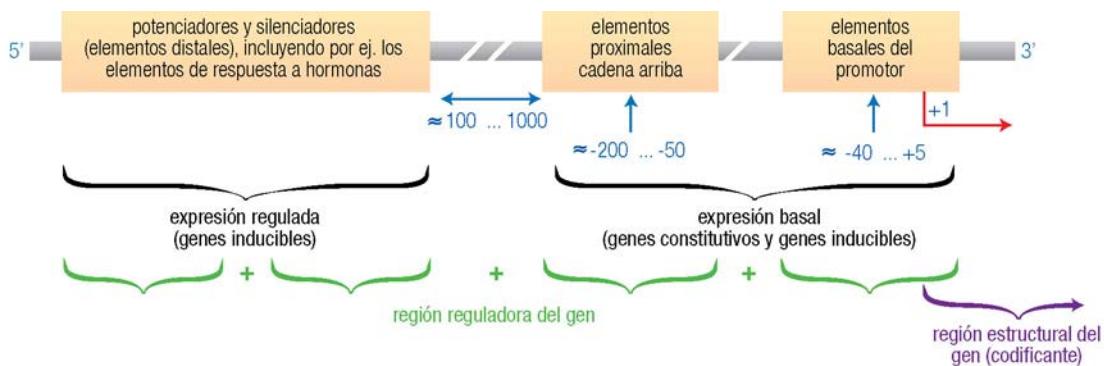
18.4.3.2 Factores de transcripción proximales

También llamados factores ligantes cadena arriba, son proteínas que reconocen específicamente los elementos proximales del promotor, contribuyendo a aumentar la eficiencia de formación del complejo de inicio o favoreciendo su actividad sobre la polimerasa. Cada promotor en particular requiere un conjunto preciso de estos factores para su expresión plena.

Como ejemplos, el factor de transcripción CTF (o NF1) interacciona con la caja CAAT, siendo aparentemente el máximo responsable de la eficiencia del promotor, mientras que el factor de transcripción SP1 se une a las cajas GC (gracias a 3 dedos de zinc) y contacta con TFIIID. En tipos celulares concretos existen otros factores que se unen a las variantes respectivas de caja CAAT.



Los promotores que sólo contienen elementos reconocidos por TF generales y proximales controlan los denominados **genes constitutivos**, los que se expresan de una forma constante, a la misma velocidad en todo momento de la vida celular. En general, se asume que este concepto es sinónimo de **genes de mantenimiento (housekeeping genes)**, los que se expresan en todas las células del organismo; el nombre se debe a que codifican los productos que aseguran las funciones indispensables para la vida de la célula. Ambos conceptos incluyen tanto genes que se transcriben con eficiencia, porque sus productos génicos se precisan en forma abundante, como genes que se transcriben a baja velocidad porque sus productos se necesitan en pequeña cantidad, pero siempre de forma continua en ambos casos.



18.14

18.4.3.3 Factores de transcripción inducibles

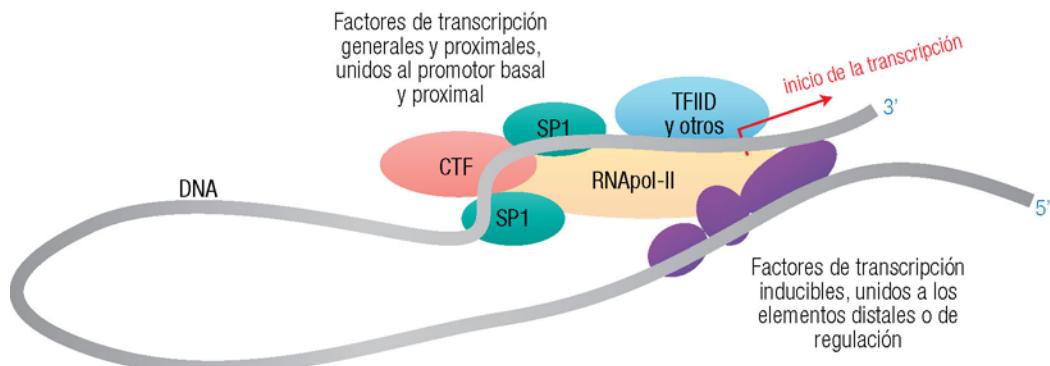
Este tercer grupo de factores de transcripción lo constituyen aquellos que reconocen los elementos distales del promotor. Allí actúan de forma similar a los factores proximales, es decir, facilitando la formación y la actividad del complejo de inicio de la transcripción o, por el contrario, dificultándolas. Regulan, por tanto, la transcripción tanto de forma positiva como negativa (de acuerdo con ello, sus promotores serán potenciadores o silenciadores, pág. 290).

A diferencia de los generales y proximales, estos factores de transcripción no están presentes en la célula de forma continua, sino que se sintetizan o se activan en momentos específicos o en tejidos particulares. Como consecuencia, la expresión de los genes cuya transcripción controlan está igualmente regulada en tiempo y lugar, es decir, se expresan de forma activa en ocasiones y no lo hacen en absoluto en otras; con frecuencia, en muchos tipos celulares un gen concreto no se expresa nunca. Por ello reciben el nombre de *genes inducibles, regulables o de expresión diferencial* (por contraposición a los genes constitutivos).

Como ejemplo característico se pueden citar los receptores nucleares para hormonas esteroideas o tiroideas. La unión de la hormona cambia la conformación del receptor, que actúa como factor de transcripción aumentando su afinidad por ciertos promotores distales, denominados **elementos de respuesta a hormonas (HRE)**. Como consecuencia, se activa o se bloquea, según los casos, la transcripción de genes específicos; los productos proteicos de estos genes representan para la célula el mecanismo de respuesta a la hormona.

18.4.3.4 Interacción de los factores de transcripción con la polimerasa

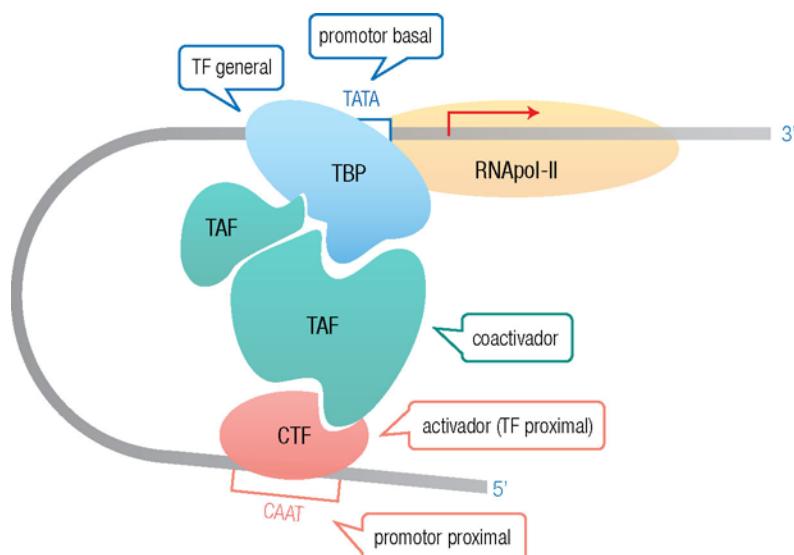
Cabe preguntarse cómo los factores que se unen a promotores alejados del origen de transcripción pueden participar en el complejo de inicio (mecanismos de regulación a distancia). Para explicarlo se debe considerar la flexibilidad de la cadena de DNA, que puede permitir la aproximación física de los tres tipos de secuencias promotoras a pesar de su lejanía en la secuencia lineal de DNA.



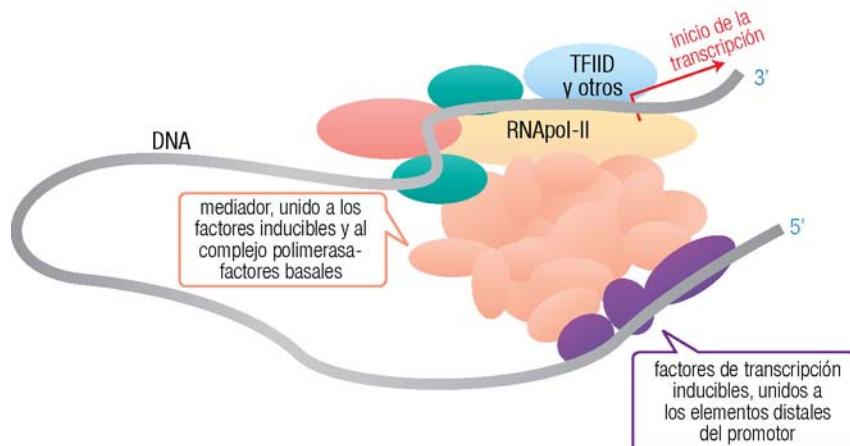
18.15

De esta forma, los factores de transcripción unidos a las regiones basal, proximales y distales pueden contactar entre sí y con la RNAPol-II, por lo que participan al unísono en el control del inicio de la transcripción.

Algunos TF no interactúan de manera directa con la polimerasa, sino que ejercen de intermediarios entre otros TF; esto es particularmente común para los elementos promotores distales. Se han identificado proteínas o complejos de proteínas que cooperan con los factores activadores, conectándolos con la maquinaria basal para regular la transcripción. Por esta razón, se conocen como **coactivadores** o también **cofactores generales**; entre ellos se incluyen algunos TAF (pág. 291). Un coactivador puede, pues, definirse como una proteína que es necesaria para que actúen los activadores que se unen al DNA, pero no para la transcripción basal, y que no se une por sí mismo a una secuencia específica del DNA. Habitualmente los coactivadores poseen múltiples subunidades, lo que les permite recibir señal de distintos potenciadores. A uno de los mejor conocidos, bastante bien conservado entre especies, se le ha puesto el nombre de **mediador**; con 20 subunidades, es más grande que la polimerasa. Mediador interacciona con activadores y con el dominio CTD de la polimerasa, y puede regular la actividad quinasa de TFIIH/K.



18.16

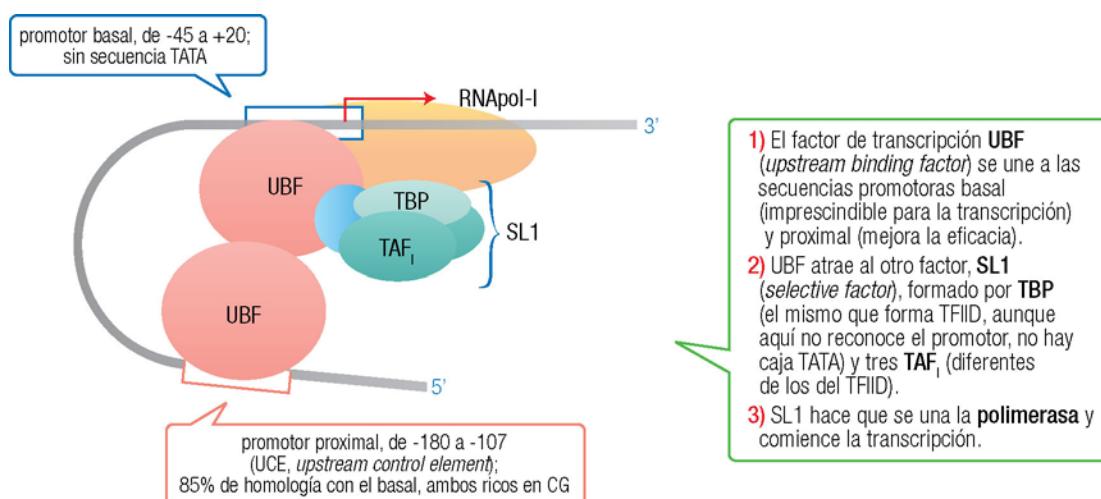


18.17

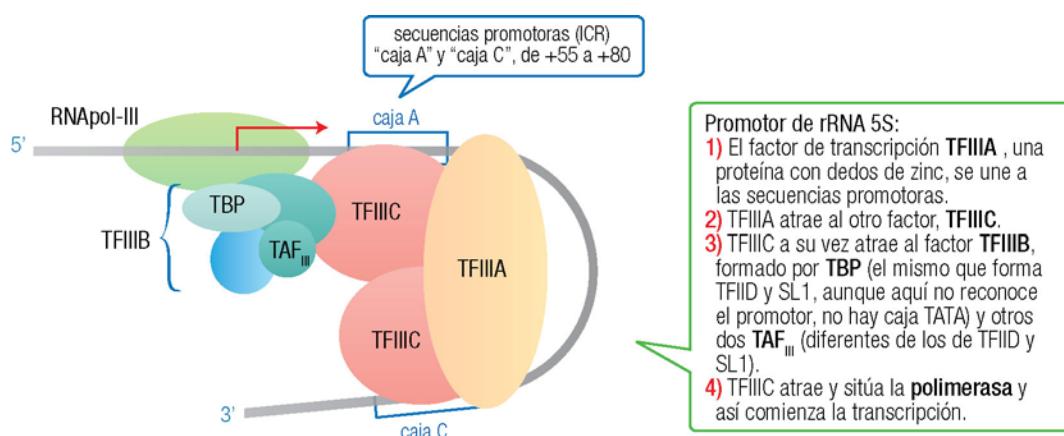
18.4.4 Regulación de la transcripción de genes de las clases I y III

Como ya se ha indicado, las secuencias promotoras y el conjunto de factores de transcripción empleados por cada una de las RNA polimerasas eucarióticas son diferentes (pág. 288). La diversidad de promotores y TF para el inicio por RNAPol-I y III es muy inferior a la correspondiente a la RNAPol-II. Esto puede reflejar simplemente el menor número de genes que transcriben aquellas dos enzimas: la RNAPol-I sintetiza un solo transcripto primario (el precursor grande de RNA ribosómico), mientras que la RNAPol-III transcribe unos pocos genes (el pre-rRNA 5S, los cerca de 40 tRNA y algunos RNA pequeños, pág. 270). Ello contrasta con las decenas de miles de genes proteicos de cuya transcripción es responsable la RNAPol-II, por lo que necesita una *maquinaria* más compleja para su regulación individual, como se acaba de estudiar.

Los promotores de genes de la clase I (para RNAPol-I) se localizan, al igual que los de la clase II estudiados hasta ahora, en dirección 5' del punto de inicio (cadena arriba).



Entre los genes transcritos por la RNAPol-III se encuentran el de rRNA 5S, los de tRNA, el del snRNA U6 (uno de los nucleares pequeños que intervienen en la eliminación de intrones; pág. 304), algunos microRNA y otros RNA pequeños; todos ellos tienen promotores y TF diferentes. Los promotores de rRNA 5S y tRNA aparecen en dirección 3' del inicio (cadena abajo), situándose dentro de la región estructural del gen; por ello se llaman **regiones internas de control (ICR)**. La posición e identidad de las secuencias promotoras y de los correspondientes TF son diferentes para los genes de rRNA 5S, tRNA y snRNA.

**18.19**

Web 18.4. Estructura del factor TFIIIA y su interacción con el DNA.



Web 18.5. Diferentes promotores y factores de transcripción usados por la RNAPol-III.

Cabe resaltar que en la iniciación por las 3 RNA polimerasas intervienen sendos factores de transcripción con varias subunidades (SL1, TFIIID y TFIIB) una de las cuales es la proteína ligante de TATA, TBP. Las otras subunidades son TAF específicos para cada tipo de promotor. Mientras que en los genes de tipo II TBP es el factor inicial que reconoce la caja TATA (pág. 291), esta secuencia no existe en los otros casos (genes de tipo I y III). En ellos, TBP no se une directamente al DNA reconociendo la secuencia promotora, sino que requiere otros factores de transcripción para asociarse. Parece que TBP desempeña una función común a los 3 tipos de promotores: la interacción con la polimerasa respectiva para situarla en el punto de inicio.

Maduración del RNA o procesamiento postranscripcional



19.1 INTRODUCCIÓN	297		
19.1.1 Matizaciones al concepto de gen	298		
19.2 CARACTERÍSTICAS DIFERENCIALES DE LA MADURACIÓN	298		
19.3 PROCESAMIENTO DEL RNA MENSAJERO	300		
19.3.1 Modificación del extremo 5'	300	19.3.3.1 Magnitud de los intrones	303
19.3.1.1 Formación de la caperuza 5'	300	19.3.3.2 Secuencias que delimitan el intrón	303
19.3.1.2 Metilación de la caperuza 5'	301	19.3.3.3 Formación del ayustosoma	304
19.3.2 Modificación del extremo 3'	302	19.3.3.4 Mecanismo de la reacción	305
19.3.2.1 Formación del extremo 3' y su relación con el final de la transcripción	302	19.3.3.5 Otros tipos de intrones	305
19.3.2.2 Poliadenilación del extremo 3'	302	19.3.4 Riboedición	306
19.3.3 Ayuste: eliminación de intrones y empalme de exones	303	19.4 PROCESAMIENTO DE LOS RNA RIBOSÓMICO Y TRANSFERENTE	307
		19.5 FORMACIÓN DE RNA INTERFERENTES	309
		19.6 MADURACIÓN DEL RNA MITOCONDRIAL	309
		19.7 REGULACIÓN POSTRANSCRIPCIONAL DE LA EXPRESIÓN GÉNICA	309
		19.7.1 Ayuste alternativo	310

19.1 INTRODUCCIÓN

La transcripción del DNA es un proceso que no difiere, en esencia, entre procariotas y eucariotas: las RNA polimerasas se unen a secuencias promotoras específicas, y se realiza la síntesis en dirección 5' → 3' a lo largo del gen. Sin embargo, surgen importantes diferencias entre ambos tipos de células cuando se considera la formación del RNA funcional.

Se llama **transcrito primario o precursor del RNA** a la molécula de RNA que resulta directamente del proceso de transcripción. Todos los transcritos primarios poseen la misma secuencia que el fragmento correspondiente de la hebra no transcrita (codificante, pág. 267) del DNA, salvo por la presencia de U en lugar de T. Para todos los tipos de RNA eucariótico y para los tRNA y rRNA procarióticos el transcrito primario es diferente de la molécula funcional de RNA, que se genera a partir de él. En contraste, para el mRNA procariótico el transcrito primario ya es la molécula funcional; esto permite su empleo inmediato para la traducción (que ocurre incluso antes de finalizar su síntesis por transcripción, pues no existe compartimentación subcelular, págs. 262-263).

Los precursores de los diversos tipos de RNA formados en el núcleo eucariótico no pueden desempeñar directamente su función (en el caso del mRNA, su utilización como molde para la traducción), sino que han de convertirse antes en **RNA maduro**, funcional, mediante un proceso que recibe el nombre de **modificación o procesamiento postranscripcional** o simplemente **maduración del RNA**. Ésta tiene lugar en el núcleo, y sólo cuando se ha completado las moléculas de RNA pueden transportarse al citoplasma, donde ejercen su función. Este transporte tiene lugar mediante el paso de las moléculas de RNA maduro a través de los poros de la membrana nuclear (**poros nucleares**).

El procesamiento postranscripcional consiste, de modo general, en una combinación de las siguientes reacciones (v. también pág. 307):

- Modificación covalente de nucleósidos.
- Adición de nucleótidos a los extremos.
- Escisión de nucleótidos de los extremos.
- División de la molécula en varios fragmentos.
- Proceso de ayuste, eliminación de intrones, o corte de intrones y empalme de exones.

Como se verá más adelante, la maduración confiere estabilidad al RNA, básicamente por una mayor resistencia frente a las nucleasas y por un plegamiento tridimensional compacto (salvo en el caso del mensajero). Asimismo, facilita su reconocimiento por otros componentes celulares, lo que mejora la funcionalidad para la traducción.

19.1.1 Matizaciones al concepto de gen

La maduración postranscripcional es uno de los procesos que dan lugar a situaciones donde la definición de gen debe matizarse o forzarse ligeramente (pág. 265). En concreto, pueden citarse dos situaciones:

- Un gen se transcribe dando un pre-RNA que se procesa por escisión generando varias moléculas de RNA. En este caso, el concepto de gen se puede aplicar al gen “global” de partida, puesto que es una unidad de transcripción, con sus regiones estructural y reguladora. Sin embargo, también es aplicable a cada una de las regiones “parciales” de DNA que codifican uno de los RNA finales, pues son éstos los productos génicos funcionales. Ejemplos de esta situación los tenemos en los precursores de RNA ribosómico, tanto procarióticos como eucarióticos (pág. 308), cuyo “gen global” da lugar a varios rRNA maduros, y también en el DNA mitocondrial, cada uno de cuyos transcritos primarios se procesa dando una combinación de varios rRNA, tRNA y mRNA (pág. 309).
- El transcripto primario obtenido de una región de DNA puede en ocasiones sufrir su maduración por dos rutas o variantes distintas, dando lugar por ello a dos productos génicos diferentes. Éste es el caso del **ayuste alternativo** (pág. 310), donde se pueden emplear dos secuencias diferentes de corte y empalme para la eliminación de intrones (pág. 303) dependiendo, por ejemplo, del tejido en el que se exprese el gen o de la llegada a la célula de diversas señales.

19.2 CARACTERÍSTICAS DIFERENCIALES DE LA MADURACIÓN

Las diferencias entre procariotas y eucariotas relativas a la maduración del RNA se centran en el RNA mensajero; los tRNA y rRNA sufren una maduración similar en ambos tipos de organismos. Puesto que, además, afecta a la síntesis de proteínas, nos centraremos especialmente en el análisis de la maduración del mRNA.

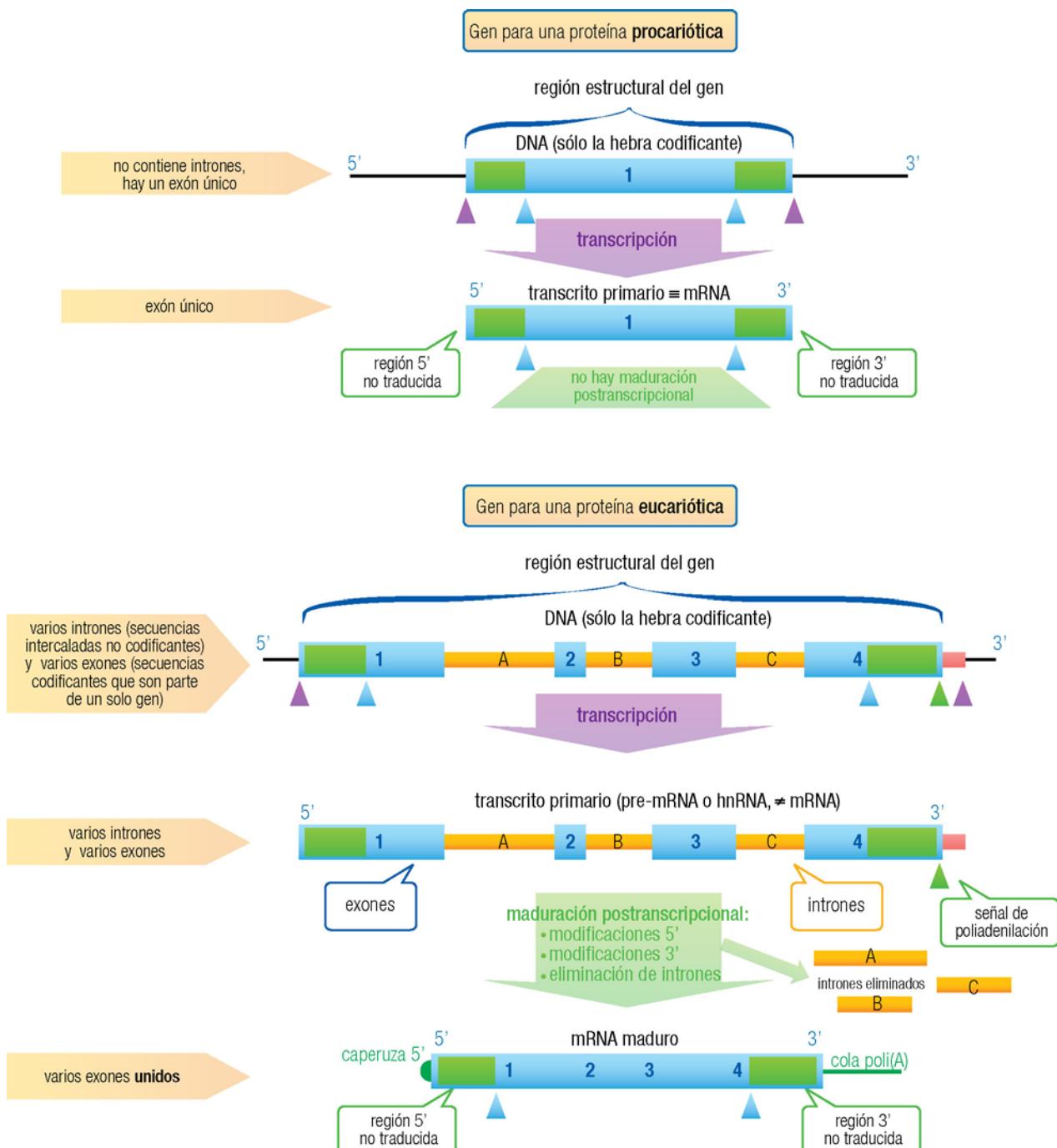
En un gen proteico **procariótico** todo el DNA es codificador (excluyendo la región reguladora, que en este capítulo no nos interesa considerar, pues no se transcribe), de modo que al transcribirse origina un transcripto primario que es ya el mRNA funcional, cuya secuencia corresponde exactamente a la del DNA. Es decir, el mRNA de procariotas es el único RNA que no requiere maduración. Por el contrario, la transcripción de un gen proteico **eucariótico** genera un pre-mRNA que debe sufrir una intensa maduración para transformarse en una molécula funcional de mRNA. El conjunto de los diferentes pre-mRNA se llama **RNA nuclear heterogéneo (hnRNA)**, porque se observa como una población de moléculas de tamaño muy variado.

Dentro de las diferencias entre pre-mRNA y mRNA maduro, la más llamativa es que la secuencia del mRNA no corresponde por completo a la del pre-mRNA, ni a la del DNA del que se ha transcrita (se dice que “se ha perdido la colinealidad entre DNA y RNA”). Esto se debe a que grandes segmentos internos del gen desaparecen durante la maduración. Estas *regiones del DNA que se eliminan del pre-mRNA en su maduración* se denominan **intrones**. Aunque son DNA no codificador (pág. 110), deben contemplarse como parte del gen (pág. 264) pues son esenciales para el desarrollo de su función.

Se pueden definir también los intrones como *regiones no codificantes situadas en el interior de un gen*, lo que da lugar al sinónimo **secuencias intercaladas (IVS, intervening sequences)**. No son intrones aquellas secuencias que forman los extremos 5' y 3' del mRNA maduro, aunque se trate igualmente de DNA no codificador; por ello, una tercera definición de intrón como *región que se transcribe pero no se traduce* no es estrictamente correcta.

Aunque los intrones son característicos de los genes eucarióticos que codifican proteínas, también están presentes en algunos genes de tRNA y rRNA, en especial eucarióticos. En éstos el **intrón** se puede definir únicamente como la región de DNA (o de su producto transcripcional, el pre-RNA) que se elimina durante la maduración.

Finalmente, las regiones que flanquean a los intrones y que se conservan en el RNA maduro se denominan a su vez **exones**. En general, son la parte codificante del gen, pero el primer exón y el último también incluyen, respectivamente, las regiones 5' y 3' no traducidas (5'UTR y 3'UTR), que no son codificantes.



Por tanto, el DNA de los genes **procarióticos** de proteínas se puede considerar (aunque no es terminología frecuente) formado por un exón único, sin intrones, que al transcribirse origina un transcrito primario también sin intrones. Por el contrario, casi todo el DNA génico de **eucariotas** contiene varios intrones y exones, y se transcribe a precursores de RNA con los mismos intrones y exones. Como parte del proceso de modificación postranscripcional se liberan los intrones y se unen los exones.

La formación del mRNA eucariótico maduro a partir de su correspondiente pre-RNA constituye un modelo ideal para el estudio del procesamiento postranscripcional, pues sufre todos los tipos de transformación implicados en dicho proceso: adición de nucleótidos al extremo 5', metilación posterior, alargamiento de la cadena en el extremo 3' y eliminación de intrones con unión de exones ("ayuste", o "corte y empalme"). Una vez completada esta maduración, también se produce el paso del mRNA maduro a través de los poros de la

membrana nuclear hacia el citoplasma, lugar donde se empleará para la síntesis proteica. Se estudiará con detenimiento cada uno de estos pasos.

19.3 PROCESAMIENTO DEL RNA MENSAJERO

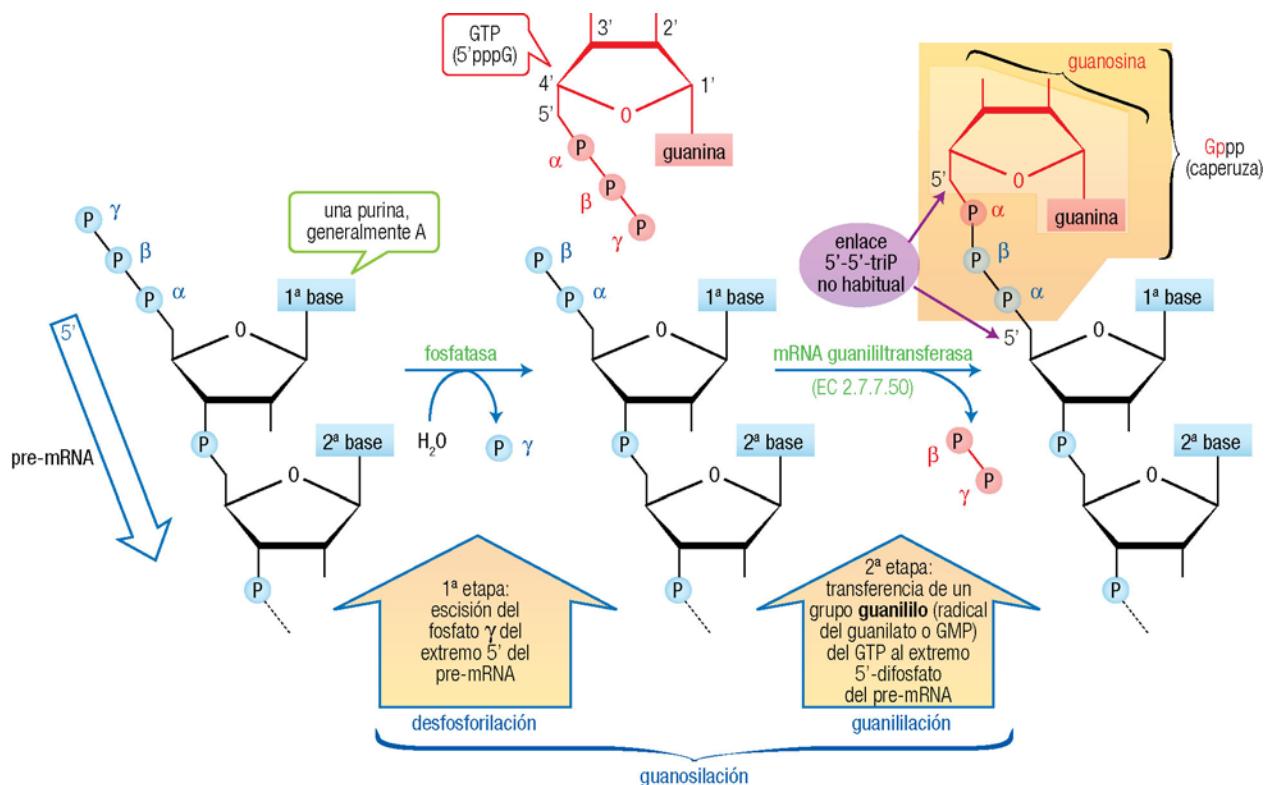
La maduración del mRNA puede describirse mediante tres tipos de modificaciones, que pueden transcurrir de forma sucesiva o simultánea, todas ellas realizadas en el núcleo.

19.3.1 Modificación del extremo 5'

El extremo libre 5'-PPP del mRNA eucariótico comienza a modificarse muy pronto, poco después de iniciada la transcripción. Se trata de una modificación covalente que, globalmente, implica tres tipos de enzimas (fosfataza, guanililtransferasa y metilasas) y que conduce a la incorporación singular de un nucleótido protector sobre el NTP de la posición 5' terminal del RNA.

19.3.1.1 Formación de la caperuza 5'

Se realiza en dos etapas: una desfosforilación previa y una guanililación a costa de GTP.



	base:	nucleósido:	nucleótido:
molécula:	guanina	guanosina	guanilato (GMP)
radical:	guanilo	guanosilo	guanililo

Aunque en la segunda reacción se transfiere un residuo guanilico, el resultado conjunto de las dos reacciones es la incorporación al mRNA de tan sólo un residuo de guanosina, es decir, una **guanosilación**; se dice que el mRNA ha sido guanosilado en 5'. La estructura terminal resultante se conoce como **caperuza de guanina**, **caperuza 5'** o **casquete 5'** (del inglés *cap*). Se trata de una estructura singular, que no se encuentra en ninguna otra parte de nucleótidos ni ácidos nucleicos, con respecto a los siguientes criterios:

- Unión de dos nucleótidos a través de un enlace trifosfato.
- Un nucleótido (guanosina) unido en sentido opuesto al del resto de la cadena polinucleotídica (enlace trifosfato entre dos posiciones 5').

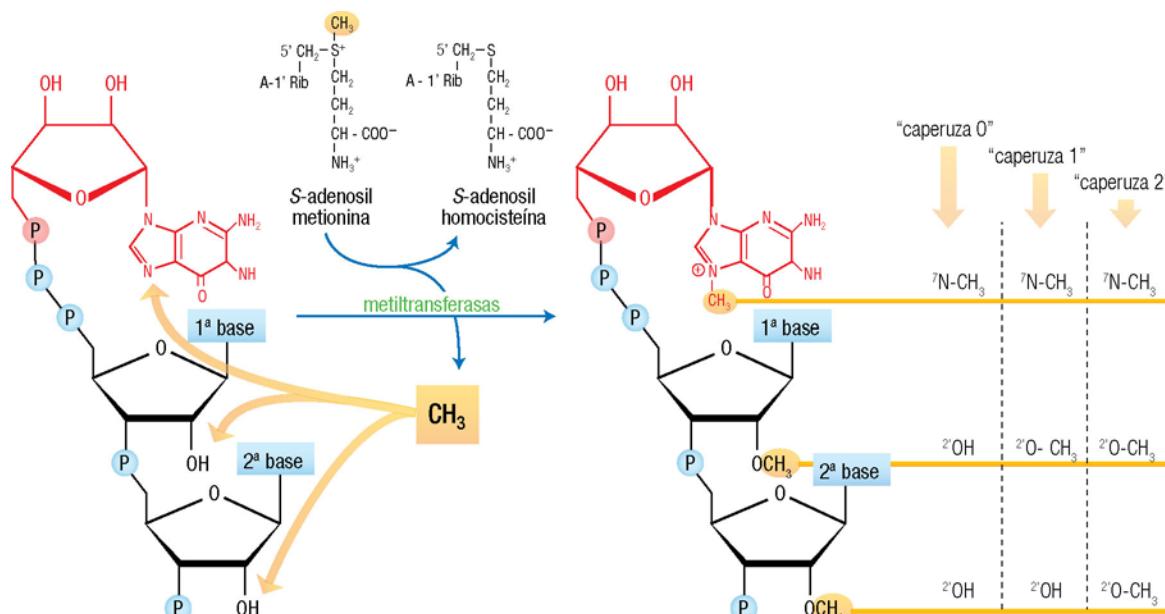
Esta caperuza ejerce un efecto protector para la molécula frente a las exonucleasas (pues éstas sólo actúan sobre enlaces 3'-5'-fosfodiéster), marca el pre-mRNA para su empleo posterior como sustrato en otras reacciones de procesamiento en el núcleo y sirve como punto de unión del mRNA a los ribosomas para iniciar la síntesis proteica.

19.3.1.2 Metilación de la caperuza 5'

Los tres primeros nucleótidos de la estructura 5'-guanosina-trifosfato-mRNA son entonces modificados por metiltransferasas (metilasas), que emplean S-adenosil-L-metionina (pág. 19) como coenzima donadora del grupo metilo (una reacción similar a la de metilación de citosinas como modificación epigenética del DNA, pág. 283).

La metilación puede tener lugar sobre varias posiciones:

- El nitrógeno 7 de la guanosina terminal, para dar la **caperuza de tipo 0**, o caperuza de 7-metilguanosina-trifosfato. La reacción la cataliza la mRNA(guanina-N7)-metiltransferasa, EC 2.1.1.56.
- Además del anterior, el 2'-OH de la ribosa adyacente (primer nucleótido original del pre-mRNA; segundo de la secuencia actual), dando la **caperuza de tipo 1**. Esta reacción la cataliza una metilasa diferente, la mRNA (nucleósido-2'-O)-metiltransferasa, EC 2.1.1.57.
- Además de las dos anteriores, puede metilarse también el 2'-OH de la ribosa siguiente (segundo nucleótido original; tercero de la secuencia actual), para dar la **caperuza de tipo 2**.

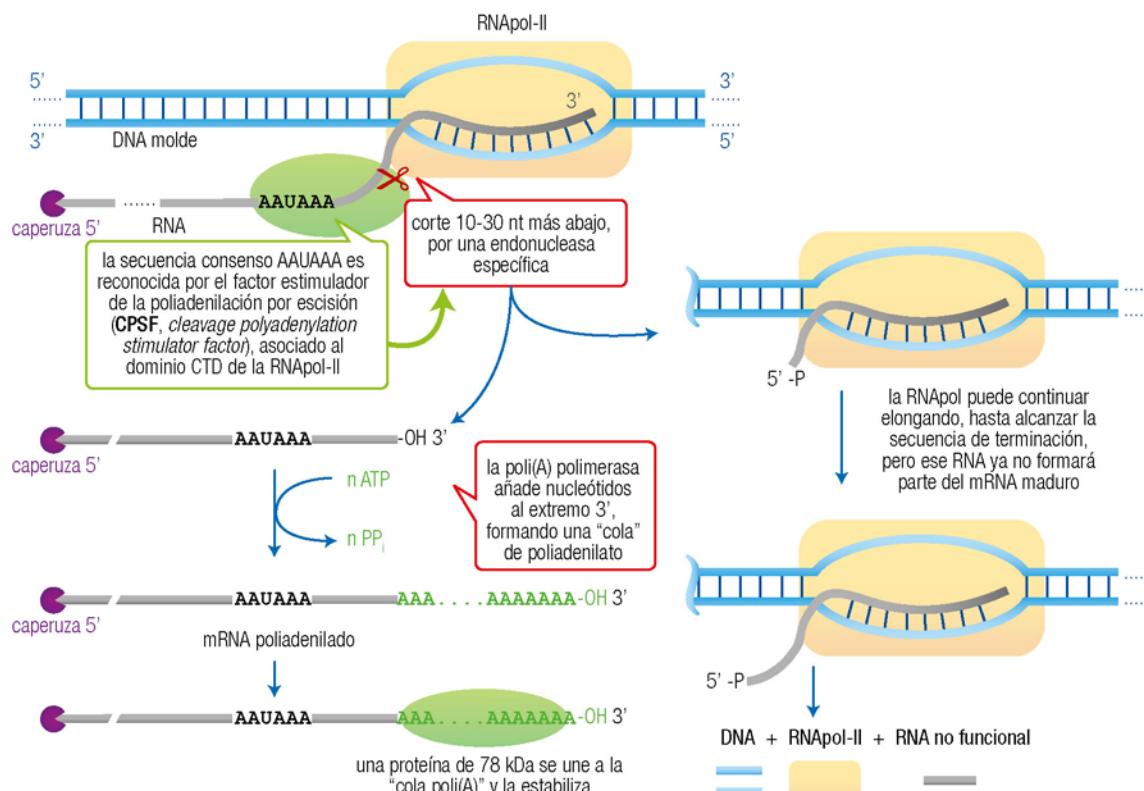


Se sabe que las metilasas y las guanililtransferasas se asocian –gracias al factor de elongación hSPT5– al dominio CTD fosforilado de la RNAPol-II (pág. 271) y llevan a cabo la reacción de procesamiento del RNA antes de que el transcripto haya alcanzado un tamaño de unos 30 nucleótidos. Parece, por tanto, posible que la formación de la caperuza participe en el cambio de iniciación de la transcripción a elongación.

19.3.2 Modificación del extremo 3'

19.3.2.1 Formación del extremo 3' y su relación con el final de la transcripción

Otra peculiaridad del mRNA eucariótico es que el extremo 3' del mRNA no corresponde a la posición donde se termina la transcripción (que se produce gracias a la interacción de las RNAPol con proteínas de pausa o terminadoras, pág. 275), sino que deriva de la escisión de la molécula de pre-mRNA en una posición cadena arriba de su extremo 3'. Como consecuencia, el mRNA es más corto en su extremo 3' que su transcripto primario.



19.3.2.2 Poliadenilación del extremo 3'

Sobre el extremo 3'-OH resultante del corte por la endonucleasa actúa una RNA polimerasa especial, que no usa molde y sólo acepta como sustrato al ATP, denominada poli(A) polimerasa (EC 2.7.7.19). Como consecuencia, se añade al RNA un gran número de residuos de adenilato (AMP), formando una “cola” de poliadenilato o cola de poli(A), de longitud típica entre 40 y 250 nt (como excepción, los mRNA de histonas carecen de esta cola). Esta estructura, que permanece intacta hasta la formación final del mRNA, puede participar en procesos de gran relevancia, como el transporte del mRNA al citoplasma y la determinación, en parte, de la estabilidad y vida media del mRNA (pág. 347).

La presencia de la cola de poli(A) únicamente en los mRNA eucarióticos se aprovecha con fines experimentales, para la purificación del conjunto de mRNA de una célula, empleando técnicas de afinidad sobre soportes que contienen cadenas oligonucleotídicas de poli(dT), poli(T) o poli(U), complementarias a la cola de poli(A) (pág. 139).

19.3.3 Ayuste: eliminación de intrones y empalme de exones

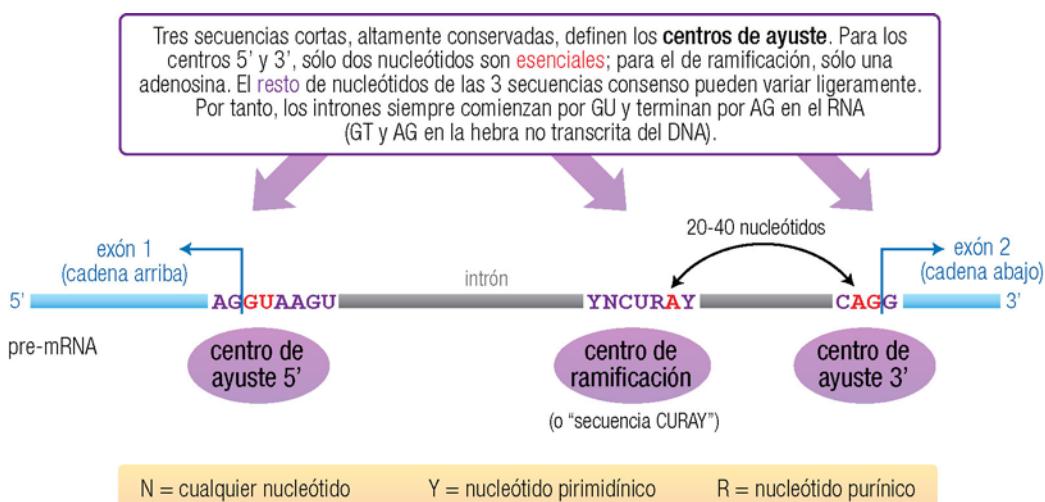
La etapa esencial de la maduración postranscripcional del transcripto primario hasta dar el mRNA eucariótico consiste en la eliminación de los intrones y la unión o empalme de los exones. Este proceso se conoce como **corte y pegado, corte y empalme** o **ayuste** (en inglés, *splicing*).

19.3.3.1 Magnitud de los intrones

Debe recordarse que los genes eucarióticos, en especial en los vertebrados, se encuentran casi siempre interrumpidos por **intrones**, que separan las regiones codificantes o **exones** (págs. 298-299). Como excepción, las histonas son uno de los escasos ejemplos de genes que carecen por completo de intrones. Lo más frecuente es que la longitud total de los intrones supere ampliamente a la suma de los exones, por lo que el tamaño del gen depende de forma notable del número y longitud de sus intrones. Aunque las cifras son variables, se habla de hasta 40 intrones por gen y de una longitud de hasta 20 kb por intrón. Los exones son de tamaño mucho más pequeño, habitualmente inferior a 1 kb.

19.3.3.2 Secuencias que delimitan el intrón

La eliminación de un intrón está determinada por su secuencia, y no por su longitud. Sólo intervienen tres secuencias específicas: dos que definen los extremos del intrón y otra en su interior, conocidas en su conjunto como **sitios o centros de ayuste** (centro 5', centro 3' y centro de ramificación). Todos los sitios de ayuste parecen ser equivalentes, mientras que la secuencia del resto del intrón no tiene trascendencia sobre el proceso.



Un cambio en el sitio de corte (por ejemplo, por mutación de uno de los nucleótidos consenso esenciales) daría lugar a un ayuste erróneo: se utilizaría para el corte y empalme la siguiente secuencia consenso, si la hubiera, con lo que el intrón o parte de él quedaría incluido en el mRNA. Esto conduciría a la síntesis de un polipéptido alterado, con una inserción grande en su secuencia de aminoácidos, o incluso con cambio de toda la secuencia a partir del punto de alteración debido a un desplazamiento del marco de lectura (pág. 391).

19.3.3.3 Formación del ayustosoma

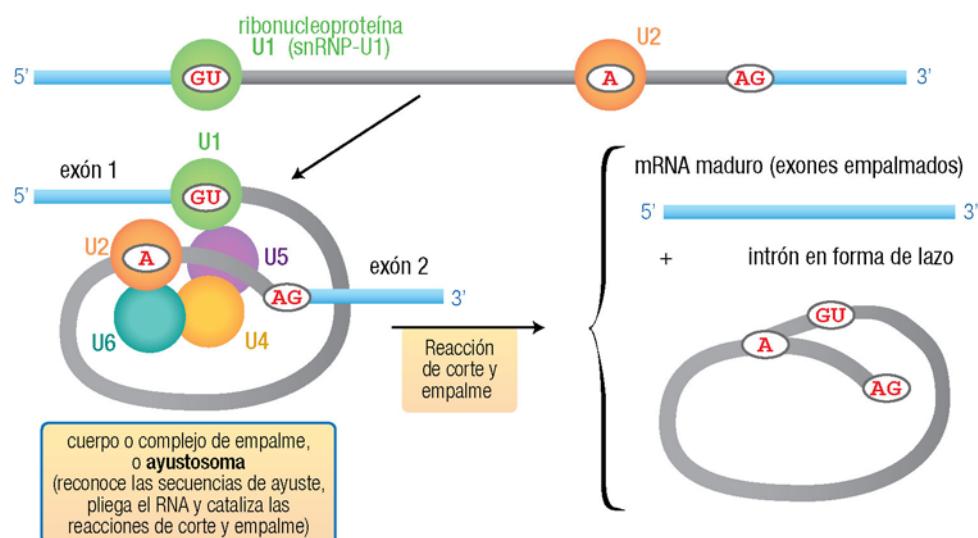
El corte de intrones y empalme de exones está mediado por una asociación macromolecular denominada **ayustosoma** (*spliceosome*; también, cuerpo o complejo de empalme). Éste se forma por asociación del pre-mRNA (en la región de los centros de ayuste) con varias **ribonucleoproteínas nucleares pequeñas** (snRNP, denominadas U1, U2, U4, U5 y U6), constituidas cada una de ellas a su vez por un **RNA nuclear pequeño** (snRNA, del mismo nombre) y cerca de 10 proteínas. El tamaño del ayustosoma es de unos 3.000 kDa, con un coeficiente de sedimentación de 60S, similar a la subunidad mayor del ribosoma eucariótico.

Tipos, tamaño y función de los RNA nucleares pequeños (snRNA)

snRNA	Tamaño (nt)*	Función de la snRNP respectiva (U1, U2, U4, U5 y U6 participan en el ensamblado del ayustosoma por interacciones mutuas)
U1	165	Ayuste: reconoce y se une al extremo 5' del intrón (por emparejamiento de bases)
U2	188	Ayuste: se une al centro de ramificación (empareja sus bases, salvo la A esencial), lo acerca al centro 5'; forma parte del centro catalítico
U3	210	Procesamiento del pre-rRNA
U4	142	Ayuste: inicialmente asociado con U5 y U6, acerca U6 al sitio 5'
U5	116	Ayuste: se asocia inicialmente con U4 y U6; se une a ambos extremos del intrón
U6	107	Ayuste: asociado inicialmente con U4 y U5; junto con U2 forma el centro catalítico
U7	65	Procesamiento del extremo 3' de los pre-mRNA de histonas
U11	131	Terminación y poliadenilación de los pre-mRNA

*Corresponden a los snRNA de mamíferos; en levaduras son de 2 a 10 veces mayores

Este ayustosoma es necesario tanto para el reconocimiento de los puntos de escisión como para la catálisis.



19.3.3.4 Mecanismo de la reacción

La reacción de ayuste supone, aparentemente, la ruptura de dos enlaces fosfodiéster (los que separan el intrón de ambos exones) y la formación de un nuevo enlace fosfodiéster entre los dos exones. Sin embargo, el mecanismo catalizado por el ayustosoma no transcurre de esa forma: no hay acción de nucleasas y ligasa, sino dos reacciones de **transesterificación**:



Web 19.1. Reacción de ayuste del mRNA.

- En primer lugar, se escinde el transcripto primario en el extremo 5' del intrón, mediante un ataque nucleófilo del 2'-OH del adenilato del centro de ramificación. Quedan así unidos ambos por un enlace fosfodiéster atípico 5'-2', y se libera el exón 1 con un extremo libre 3'-OH.
- A continuación, se escinde el centro de ayuste 3', por ataque nucleófilo del extremo 3'-OH libre del exón 1, con lo que quedan ya unidos los dos exones y el intrón se libera con su extremo 3'-OH libre y su extremo 5' formando un lazo (en inglés *lariat*; curiosamente derivado de la palabra española “la reata”). En éste, la adenosina del centro de ramificación está enlazada de manera simultánea a 3 nucleósidos por sendos enlaces fosfodiéster en sus posiciones 2', 3' y 5'. El intrón en lazo se degrada luego rápidamente.

Como se puede observar, en ambas etapas se escinde un enlace fosfodiéster a la vez que se forma otro, en sendas reacciones de transfosforilación (un tipo de transesterificación). Como consecuencia, las reacciones transcurren sin gasto energético apreciable, no hay consumo de ATP. Sin embargo, sí que se requiere hidrólisis de ATP para el ensamblado previo del ayustosoma y para la separación de emparejamientos temporales tanto dentro de las snRNP como entre los snRNA y el mRNA (gracias a actividades similares a helicasa).

Para finalizar, cabe destacar que el centro catalítico del ayustosoma está formado por los snRNA U2 y U6; aunque se requiere la presencia de las proteínas componentes de las snRNP, sólo actúan como soporte para la adecuada conformación del RNA y no intervienen en la catalisis directamente. Se trata, pues, de un ejemplo más de ribozima.

19.3.3.5 Otros tipos de intrones

La descripción anterior corresponde al proceso de corte de los intrones de genes de mRNA eucarióticos. Existen, además, otros tipos de intrones, en función de su ubicación en distintos genes y de características concretas del mecanismo de maduración. Se clasifican en 4 tipos:

Grupo I: incluye los intrones de algunos genes que codifican rRNA (nucleares, mitocondriales y de cloroplastos), RNA de orgánulos, tRNA de bacterias y mRNA de bacteriófagos. Su procesamiento requiere, como cofactor externo, un nucleósido o nucleótido de guanina (guanosina, GMP, GDP o GTP; su grupo 3'-OH actúa como nucleófilo). El proceso de corte y empalme no utiliza ATP como fuente de energía.

Grupo II: aparecen en genes que codifican mRNA de orgánulos, y en algunos mRNA de bacterias y arqueas. No requieren cofactor externo, sino que emplean el 2'-OH de un residuo adenosina del propio intrón, produciendo intermedios en lazo (págs. 304-305). En este sentido, se asemejan más al mecanismo de corte y empalme recién estudiado, aunque difieren por no existir ayustosoma. Al igual que el grupo I, no se necesita ATP como cofactor energético, y experimentan autoayuste, catalizado por el propio RNA en lugar de por enzimas proteicas clásicas. Estos RNA con función autocatalítica constituyen uno de los casos de **ribozimas** (pág. 72).

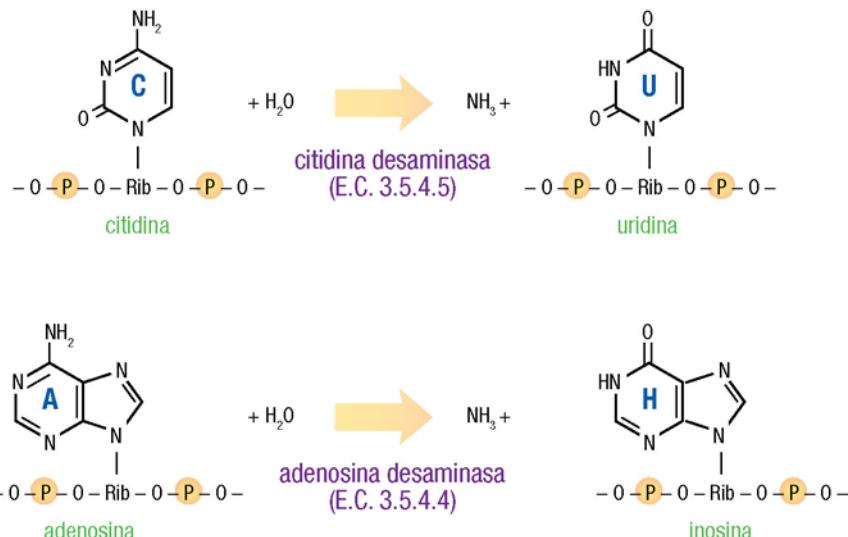
Grupo III: es el más numeroso, incluye los intrones presentes en la mayoría de los genes nucleares eucarióticos que codifican mRNA. Su proceso de ayuste, el que ya se ha estudiado (págs. 303-305), sigue el mismo mecanismo de formación de lazo que los intrones del grupo II, pero catalizado por las ribonucleoproteínas componentes del ayustosoma.

Grupo IV: se encuentran en ciertos genes de tRNA eucarióticos y de arqueas. El corte y empalme requiere ATP, a diferencia de los intrones de los grupos I y II, pues depende de endonucleasas y ligasas (no hay transfosforilación).

19.3.4 Riboedición

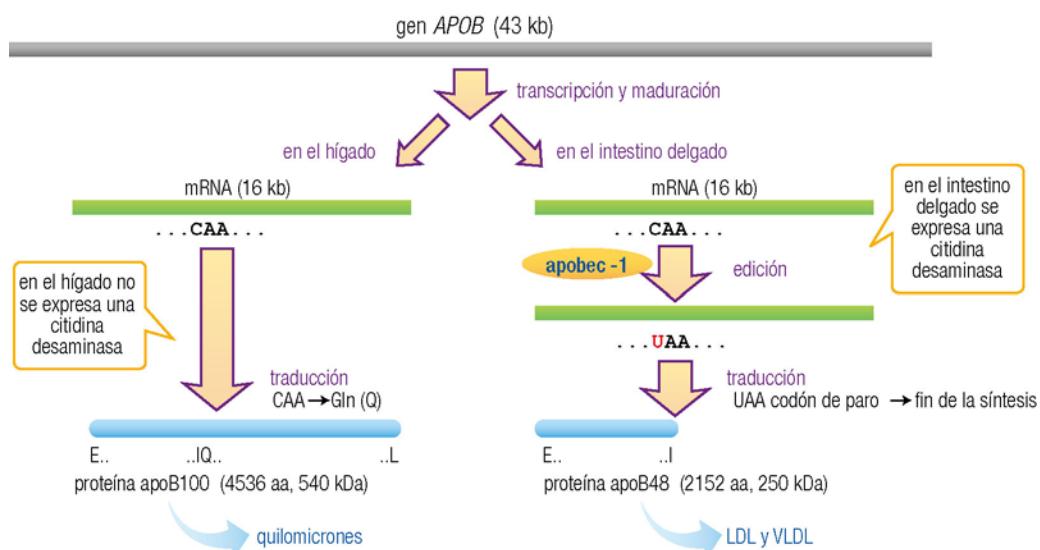
En algunos casos, la secuencia de un RNA puede terminar presentando diferencias con respecto a la del DNA del que procede, aun considerando el resultado de la maduración de extremos y la eliminación de los intrones. Se trata del cambio de uno de los nucleótidos por otro diferente; esta modificación postranscripcional se conoce como riboedición o edición del RNA.

El primer mecanismo que permite este fenómeno es la desaminación específica de una base: la citidina se convierte en uridina, y la adenosina en inosina. Se introduce así un cambio en la secuencia de la proteína traducida a partir del mRNA editado (la inosina emparejará con C en el tRNA, aunque también puede hacerlo con U o A gracias al balanceo, pág. 320).



www Web 19.2. Detalle de las reacciones de desaminación en la edición del mRNA.

Existe un ejemplo paradigmático de riboedición en el ser humano: el gen de la apolipoproteína B. Como consecuencia de una riboedición específica de tejido, en el hígado se sintetiza la proteína completa, apoB100, mientras que en el intestino se sintetiza la forma truncada, apoB48. Las funciones de ambas isoformas en el transporte y metabolismo de lípidos son diferentes: apoB100 participa en el transporte del colesterol y los triacilgliceroles de síntesis endógena, mientras que apoB48 interviene en la distribución de los lípidos de la dieta hacia los tejidos.



19.8

Otras formas de edición incluyen la delección o inserción de nucleótidos (que altera el marco de lectura y, por ello, cambia radicalmente la secuencia de la proteína, pág. 391) y la incorporación de uridinas transferidas desde un pequeño RNA llamado *RNA guía* (gRNA; este mecanismo se ha observado en el genoma mitocondrial del tripanosoma).

19.4 PROCESAMIENTO DE LOS RNA RIBOSÓMICO Y TRANSFERENTE

Los rRNA y tRNA también sufren maduración postranscripcional, tanto en procariotas como en eucariotas. A pesar de su importancia, no se expondrán aquí con detalle sus características; como aproximación didáctica se considera suficiente como modelo la maduración del mRNA. Lo mismo puede decirse de la maduración de RNA pequeños, como los snRNA.

En procariotas, un solo transcripto primario da lugar, mediante varias reacciones de corte catalizadas por endonucleasas (RNAs), a los tres rRNA y uno o varios tRNA, los cuales sufren modificaciones adicionales en sus extremos 3' y 5' y en sus nucleósidos (v. la tabla) hasta dar las formas maduras, funcionales.

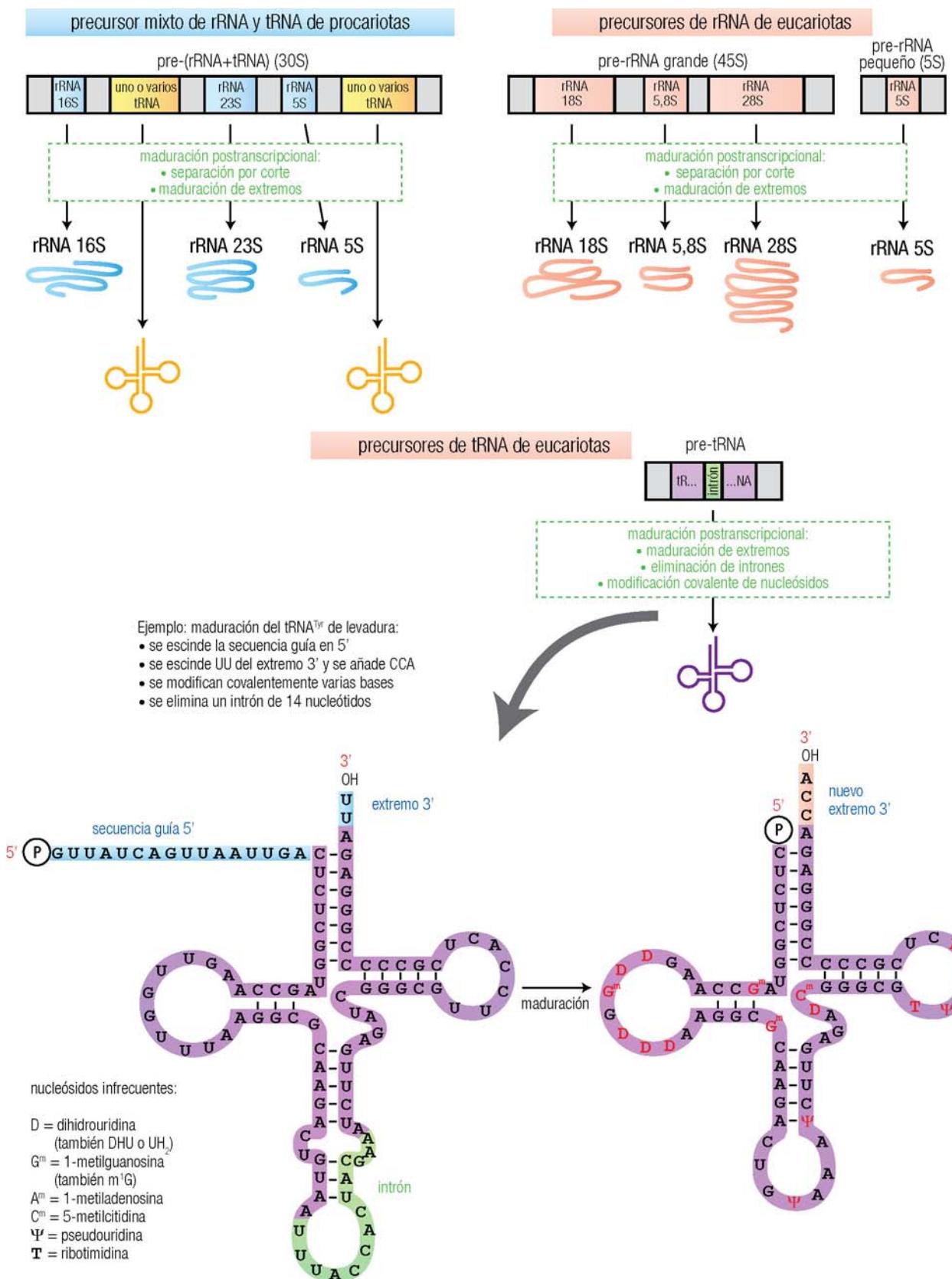
En eucariotas, el proceso es similar, con un transcripto primario de la RNAPol-I que genera 3 de los 4 rRNA y con transcriptos primarios independientes para el otro rRNA y para cada tRNA (todos sintetizados por la RNAPol-III). También en algunos casos hay eliminación de intrones.

En el caso de tRNA se produce, además de un recorte de los extremos 3' y 5', la adición, a las moléculas de tRNA que no la posean, de la secuencia -CCA en el extremo 3'-OH. Los tRNA, en especial, sufren numerosas reacciones de modificación química que determinan que en las moléculas maduras abunden los nucleósidos infrecuentes (págs. 15, 19, 65 y 67), diferentes a los cuatro que participan en la síntesis del RNA (A, G, U, C). Estas modificaciones contribuyen a hacer más resistentes los tRNA a los procesos de degradación celular, aumentando, por tanto, su estabilidad metabólica y su vida media.

Resumen de las modificaciones postranscripcionales experimentadas por los distintos tipos de RNA

	precursor	corte por endonucleasas	ayuste	corte por exonucleasas	adición en 3'	modificación en 5'	modificación de nucleósidos
PROCARIOTAS							
mRNA	- - - - - (no sufren procesamiento postranscripcional) - - - - -						
rRNA	de 30S, da lugar a los tres rRNA (16S, 23S, 5S) y algún tRNA	a ambos lados (RNasa III)	en algunos organismos (autoayuste)	procesamiento en ambos extremos			algunas bases metiladas
tRNA	da lugar a varios tRNA y en algunos casos también a rRNA	a ambos lados (RNasa P* y RNasa III)		extremo 3' recortado por RNasa D			muchas bases modificadas, gran diversidad
EUCARIOTAS							
mRNA	pre-mRNA (da lugar a un solo mRNA)	sólo en el extremo 3'	sí (con autoayuste en el mRNA mitocondrial)		cola de poli(A)	caperuzas 5'	en casos puntuales (ej.: desaminación de C→U)
rRNA	de 35 a 47S, da lugar a tres rRNA (18S, 5,8S, 28S); otro precursor da el rRNA 5S	(similar a procariotas)	en algunos unicelulares (autoayuste)	(similar a procariotas)			metilación de ribosas
tRNA	da lugar a un solo tRNA		en algunos; distinto mecanismo	en ambos extremos	adicción de CCA (tRNA-nucleotidil-transferasa)		muchas bases modificadas, menor diversidad

*La RNasa P es una ribozima (v. apdo. 6.3.1).



19.5 FORMACIÓN DE RNA INTERFERENTES

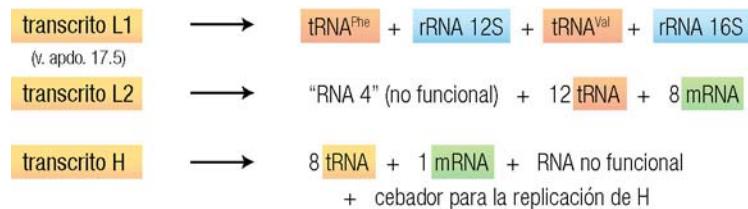
Los microRNA (pág. 69) proceden de genes transcritos por la RNAPol-II o la III, generando un transcripto primario denominado microRNA primario, **pri-miRNA**. Éste posee habitualmente una región con emparejamiento intracatenario imperfecto, que forma una horquilla con bucle. A continuación, aún dentro del núcleo, un complejo proteico que incluye **Drosha**, una endorribonucleasa de la familia de las RNAsas III, elimina la parte de la molécula ajena a la horquilla, generando el precursor de microRNA, **pre-miRNA**. Éste se exporta al citosol, donde continúa su procesamiento a cargo de la enzima **Dicer** (del inglés *dice*, que significa cortar en dados). Finalmente, el miRNA se integra en el complejo **RISC** para ejercer su función.

En cuanto a los siRNA, o RNA interferentes pequeños, se generan a partir de RNA bicatenarios, que pueden ser propios de la célula (generados por copias duplicadas de genes), introducidos en ella por un virus (en células vegetales) o introducidos en la célula con propósitos experimentales de silenciamiento génico. La misma enzima Dicer los fragmenta generando los **siRNA**, que se integran también en un complejo RISC.

Los detalles de estos procesos se describen más adelante (pág. 348), como parte de la regulación de la traducción.

19.6 MADURACIÓN DEL RNA MITOCONDRIAL

La transcripción de la molécula de DNA mitocondrial humano (pág. 276) produce 3 RNA transcritos primarios de gran tamaño, cuyo procesamiento da lugar a todos los mRNA, tRNA y rRNA de la mitocondria. Esta maduración tiene lugar principalmente por escisión del pre-RNA y, en el caso de los mRNA, también por poliadensilación (pero no hay adición de caperuza 5').



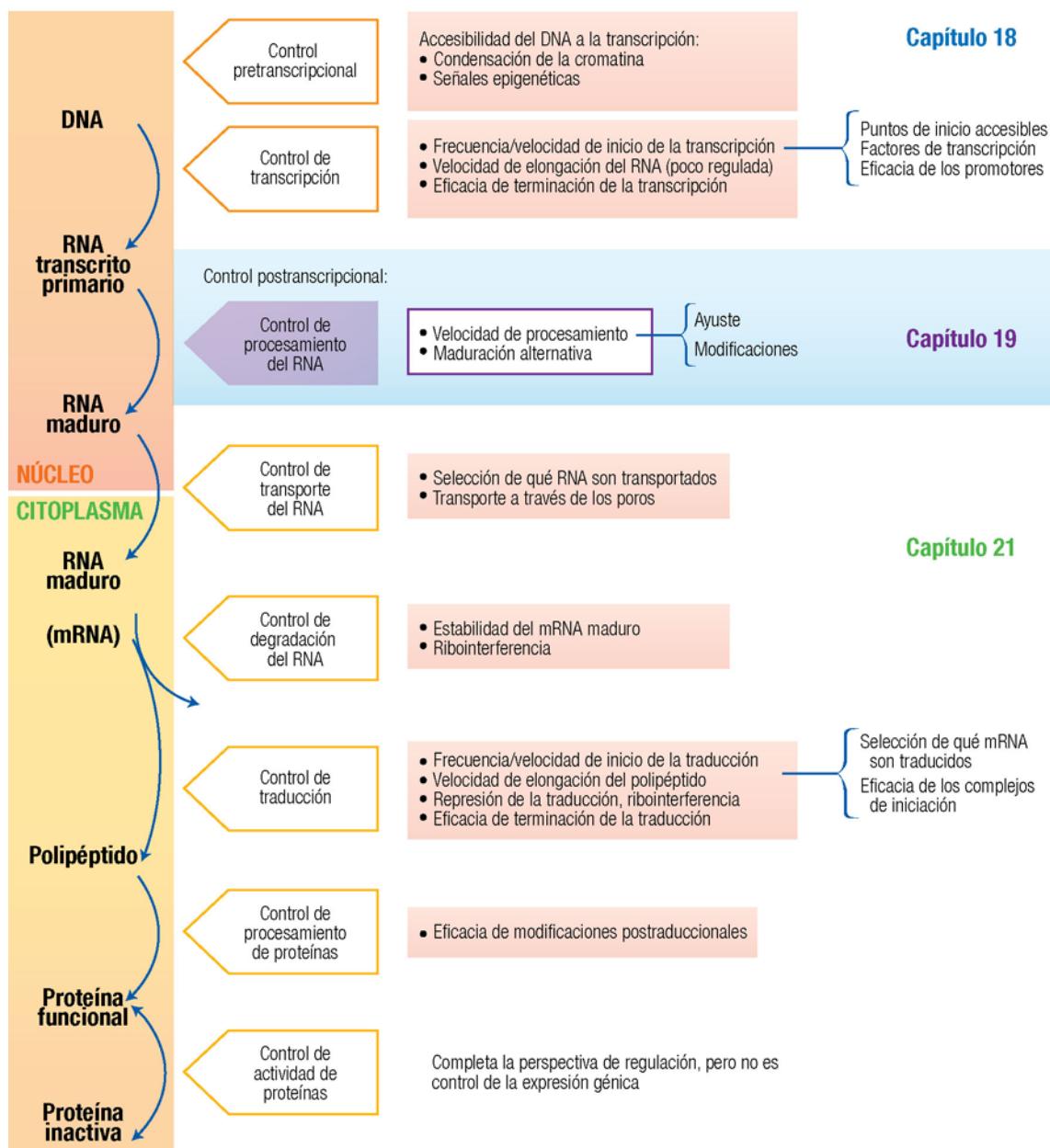
19.11

19.7 REGULACIÓN POSTRANSCRIPCIONAL DE LA EXPRESIÓN GÉNICA

A pesar de que la mayor parte del control de la expresión génica se ejerce previamente y durante el inicio de la transcripción, a través de las señales epigenéticas y de los factores de transcripción (Capítulo 18), existen mecanismos de regulación en todas las etapas de la expresión. En lo que respecta a los procesos posteriores a la transcripción, se conocen ejemplos de control tanto en la maduración del RNA como en su transporte al citoplasma y también a través de la regulación de la vida media del mRNA en el citoplasma (velocidad de su degradación). Se expone a continuación el primero de ellos, dejando los otros dos puntos de control para el apartado dedicado a la regulación de la traducción (Capítulo 21).

Las distintas etapas de procesamiento de los pre-RNA, previamente estudiadas, son lógicamente susceptibles de regulación, lo que afectará a la tasa de expresión de los genes correspondientes. Esta regulación no sólo tiene lugar en cuanto a la velocidad o eficacia de las reacciones de maduración, sino que en algunos casos un mismo pre-RNA madura por vías distintas dependiendo de la célula considerada, formando mRNA adaptados a las necesidades proteicas particulares de un tejido o de una situación fisiológica concreta. La implicación de este tipo de variabilidad en el control del producto final es enorme.

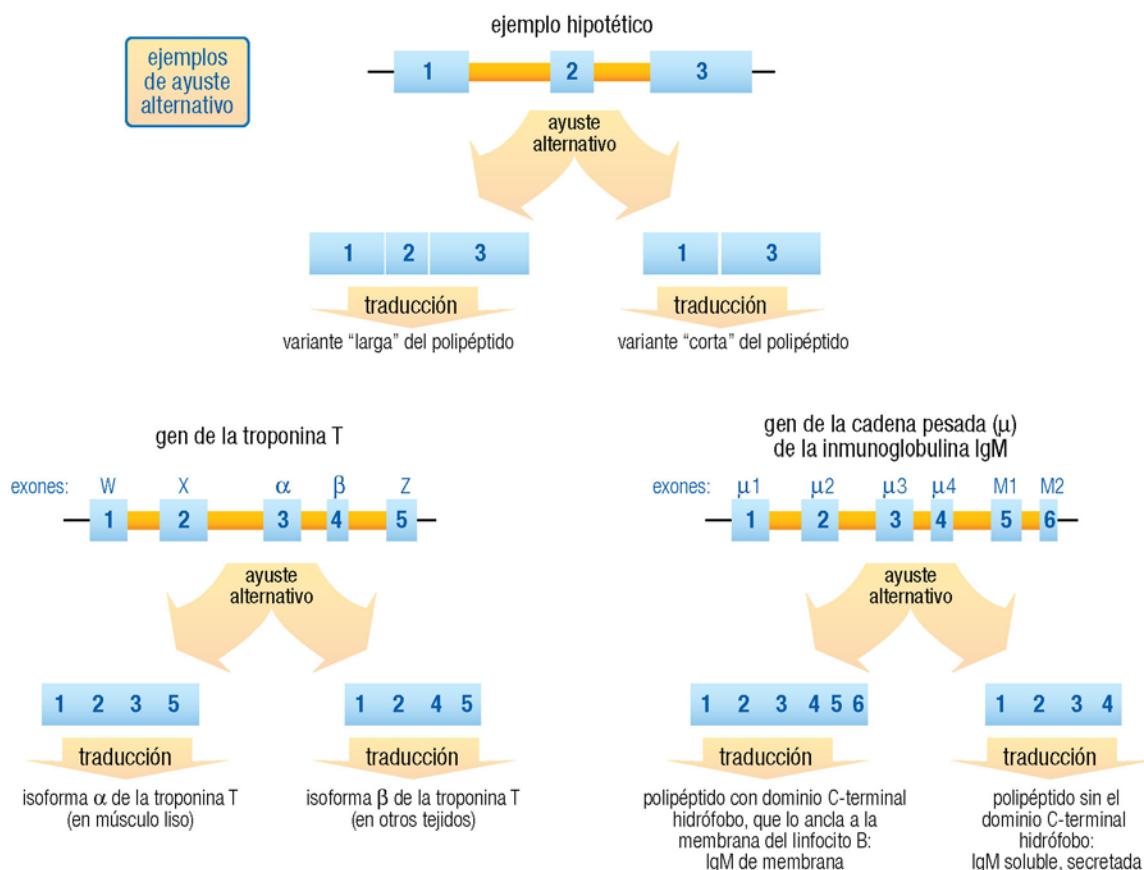
NIVELES DE REGULACIÓN DE LA EXPRESIÓN GÉNICA EN EUCARIOTAS



19.7.1 Ayuste alternativo

El ejemplo más característico de regulación postranscripcional es el **ayuste alternativo**, que comúnmente desempeña un papel en la diferenciación celular y en el control del desarrollo (por ejemplo, en la determinación del sexo en *Drosophila*). Este proceso ha pasado de considerarse algo exótico a convertirse en un mecanismo relevante de control: se ha estimado que el 90% de los genes humanos experimentan ayuste alternativo. Constituye uno más de los retos para una definición completa de gen (pág. 265) y contribuye a la comprensión de cómo el número de genes en el genoma humano ha resultado ser muy inferior al que se esperaba (pág. 6).

Este fenómeno consiste en que algunos genes pueden dar lugar a transcriptos maduros diferentes, que formarán proteínas distintas, debido al empleo alternativo de señales de ayuste. Esta decisión está marcada por secuencias en el mRNA que actúan bien como potenciadores, bien como silenciadores, forzando el salto de un intrón o el salto de un exón. Dichas secuencias son reconocidas por algunas de las proteínas llamadas hnRNP (ribonucleoproteínas nucleares heterogéneas), que enmascaran la secuencia que no se ha de emplear en el ayuste, o bien por proteínas SR, que se unen al CTD de la polimerasa durante la iniciación de la transcripción y actúan potenciando uno de los ayustes alternativos.





El código genético

20.1 ANTECEDENTES Y PROPIEDADES GENERALES DEL CÓDIGO		
20.1.1 Los codones del código son tripletes	313	
20.1.2 Los tripletes no se solapan	314	
20.1.3 Existen tres marcos o pautas de lectura	314	
20.1.4 Papel adaptador del tRNA	315	
20.2 ASIGNACIÓN DE CODONES A AMINOÁCIDOS CONCRETOS		
20.2.1 Traducción de polinucleótidos sintéticos	315	
20.2.2 Evidencia <i>in vivo</i> del código genético	316	
20.3 REPRESENTACIÓN DEL CÓDIGO GENÉTICO		316
20.4 CARACTERÍSTICAS ESPECÍFICAS		317
20.4.1 Codón de inicio	317	
20.4.2 Codones de terminación	317	
20.4.3 Degeneración	318	
20.4.3.1 Hipótesis del balanceo	319	
20.4.3.2 Número de RNA transferentes necesarios	320	
20.4.4 Universalidad	321	

El código genético es el conjunto de pautas que rigen la transferencia de la información contenida en el mRNA (y en definitiva en el DNA) para la síntesis de las proteínas. Describe, tanto en procariotas como en eucariotas, la correlación entre dos lenguajes informativos: la secuencia de nucleótidos en el mRNA (el *lenguaje con cuatro letras* de los ácidos nucleicos) y la secuencia de aminoácidos en el polipéptido (*lenguaje con veinte letras* de las proteínas). Más concretamente, el código genético establece la correspondencia entre los 64 posibles tripletes de bases del mRNA y los 20 aminoácidos de las proteínas. (Nota: Es frecuente, en especial en el código genético y la traducción, utilizar la palabra “base” donde realmente debería decirse “nucleótido”, en lo relativo a la secuencia del mRNA.)

20.1 ANTECEDENTES Y PROPIEDADES GENERALES DEL CÓDIGO

Hoy día es relativamente fácil conocer la correspondencia entre aminoácidos de una proteína y tripletes de nucleótidos del mRNA que los codifican, mediante simple comparación de la secuencia de ambos. Sin embargo, fueron muchas las aproximaciones experimentales realizadas, a partir de la década de 1950, que contribuyeron a descifrar el código genético, a la vez que a entender la transcripción y la traducción. Considerados de forma global, estos estudios han supuesto uno de los grandes retos en la historia de la bioquímica y la genética.

La información previa en la que se basa el conocimiento del código genético es de dos tipos. Por un lado, lo referente al tRNA (pág. 64) cuya intervención, a pesar de no ser un elemento propiamente dicho del código genético, es esencial para entender la correspondencia entre mRNA y proteína. Por otro lado, el ribosoma (págs. 74-78), orgánulo citoplasmático en el que se realiza la síntesis proteica, según se demostró originalmente empleando aminoácidos marcados con isótopos radiactivos.

Se describen a continuación las características generales y particulares del código genético, basadas en algunos de los experimentos más representativos que contribuyeron a descifrar el código genético y su relación con la síntesis proteica.

20.1.1 Los codones del código son tripletes

La “hipótesis de los codones” se demostró al observar que cambios en la secuencia de un mRNA conducían a alteraciones en la secuencia de aminoácidos de la proteína sintetizada (en experimentos con el bacteriófago T4).

El número mínimo de nucleótidos necesario para “codificar un aminoácido” –dicho de otro modo, el tamaño de las unidades de codificación o la relación de codificación– se deduce por un simple cálculo matemático:

Unidad de información	N.º de combinaciones posibles	
1 nucleótido	4	Insuficiente para codificar 20 aminoácidos
2 nucleótidos	$4^2 = 16$	
3 nucleótidos	$4^3 = 64$	Más que suficiente

La unidad de codificación es el triplete de nucleótidos

Por tanto, para poder codificar los 20 aminoácidos naturales se necesita que cada aminoácido venga especificado por una secuencia de 3 nucleótidos, denominada *triplete* o *codón*. El triplete es la característica más sobresaliente del código. Obviamente, las 64 posibles combinaciones exceden las necesidades para 20 aminoácidos; más adelante se analizará cómo se acomoda ese “exceso de información” (pág. 318).

20.1.2 Los tripletes no se solapan

Una de las propuestas iniciales para la distribución de los codones en el mRNA fue que tripletes contiguos compartieran nucleótidos.



20.1

En la hipótesis solapante, la secuencia del primer triplete restringe las posibles secuencias del segundo, éstas a su vez las del tercero, y así sucesivamente, por lo que la secuencia de aminoácidos de las proteínas estaría condicionada, no podría ser cualquiera. Bajo esta hipótesis, la alteración de una base afecta a varios codones, lo que produciría alteraciones en más de un aminoácido. Puesto que estas previsiones no se cumplen en la realidad, se pudo demostrar que la hipótesis solapante es incorrecta, los tripletes nunca comparten nucleótidos, no se solapan. Esto es así para el código genético de todos los seres vivos.

20.1.3 Existen tres marcos o pautas de lectura

El hecho de que los codones contiguos no se solapan no implica que exista una señal de separación entre ellos (es decir, no existen “comas”, “puntos” o interrupciones; tal separación sólo se hace notar sobre el papel, para facilitar la lectura e interpretación de la secuencia). Como consecuencia, en principio la lectura de un mRNA en la traducción se podría iniciar en cualquier nucleótido, lo que definiría a éste como el primer nucleótido del codón de inicio de la traducción. Esto significa la existencia potencial de tres mensajes en cualquier mRNA, dependiendo de en qué nucleótido se inicie la lectura de los codones; se habla de tres **marcos o pautas de lectura**. La traducción del mRNA siguiendo cada uno de ellos formaría tres polipéptidos distintos.

3 formas de "leer" la secuencia del mRNA
3 marcos de lectura

- 1º: ABC, DEF, GHI, JKL, ...
- 2º: BCD, EFG, HIJ, KLM, ...
- 3º: CDE, FGH, IJK, LMN, ...

(el 4º ya coincide con el primero, el mensaje es el mismo aunque con un aminoácido menos)



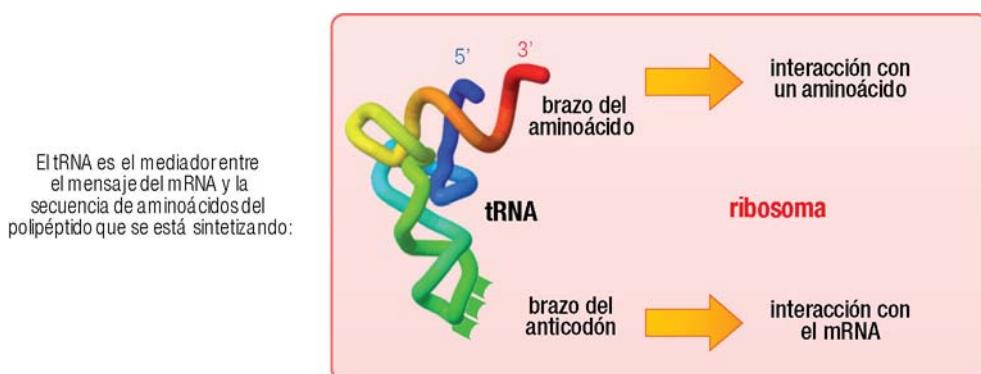
20.2

Sin embargo, en la traducción (tanto *in vivo* como *in vitro*) no se manifiesta esta aparente ambigüedad (por ejemplo, al traducir el mRNA correspondiente a un gen). Una vez iniciada la traducción según un marco de lectura, se puede decir que los otros dos carecen de información. Sin embargo, uno de los otros marcos puede representar un mensaje válido para otro gen distinto, caso de los genes solapantes (pág. 266).

El marco de lectura puede cambiar debido al desplazamiento en 1 o 2 posiciones del punto de inicio o, durante la traducción, por salto accidental de un nucleótido mientras el mRNA es recorrido por el ribosoma. La consecuencia del cambio de marco es que a partir de ese punto la secuencia de la proteína nueva es distinta a la original. Éste es el mecanismo empleado a veces para formar dos o más proteínas a partir de un mismo transcripto. De modo relevante, también es el resultado de algunas mutaciones puntuales, aquellas en las que se insertan o eliminan nucleótidos en número no múltiplo de 3 (pág. 391).

20.1.4 Papel adaptador del tRNA

Debe recordarse que la "lectura" de la secuencia de codones está protagonizada por los tRNA (pág. 64), que ejercen el papel de **adaptadores** entre la información genética y la secuencia de las proteínas. Intervienen para ello como moléculas activas especializadas (formando los aminoacil-tRNA), que dirigen los aminoácidos hacia el ribosoma situado sobre el mRNA. En el ribosoma tiene lugar la interacción del triplete de bases del aminoacil-tRNA (**anticodón**) con el triplete de bases del mRNA (**codón**) determinando, durante la traducción, la incorporación del aminoácido específico frente a cada posición del mRNA. Por tanto, en la expresión habitual se dice simplemente que "un codón codifica un aminoácido".



20.3

20.2 ASIGNACIÓN DE CODONES A AMINOÁCIDOS CONCRETOS

20.2.1 Traducción de polinucleótidos sintéticos

El descifrado del código se pudo realizar principalmente mediante experimentos de traducción *in vitro* a partir de cadenas de polinucleótido sintéticas con secuencias muy simples y repetitivas. El conocimiento de la secuencia suministrada al sistema permitió interpretar los resultados e ir asignando un aminoácido concreto a cada triplete de nucleótidos.

20.2.2 Evidencia *in vivo* del código genético

Son varios los experimentos *in vivo* que confirmaron las conclusiones alcanzadas *in vitro*; en especial, los que analizaban el efecto de las mutaciones sobre la secuencia peptídica (tanto inducidas como espontáneas, pág. 384). Como ejemplos ilustrativos se pueden adelantar aquí algunos correspondientes al gen de la subunidad α de la hemoglobina. En ambos casos, la sustitución de una sola base en el gen normal modifica un codón, conduciendo a la aparición de un aminoácido distinto en la secuencia peptídica.

	<u>normal</u>	<u>sustitución</u>	<u>anormal</u>
	AAA (Lys)	A → G	GAA (Glu)
Sustituciones de nucleótidos en el gen, que causan cambios en un aminoácido del polipéptido, encontradas en la cadena α de hemoglobinas humanas anormales:	GAA (Glu)	A → G	GGA (Gly)
	GAA (Glu)	G → A	AAA (Lys)
	GAA (Glu)	G → C	CAA (Gln)
	GGU (Gly)	G → A	GAU (Asp)
	CAU (His)	C → U	UAU (Tyr)
	AAU (Asn)	U → A	AAA (Lys)

20.4

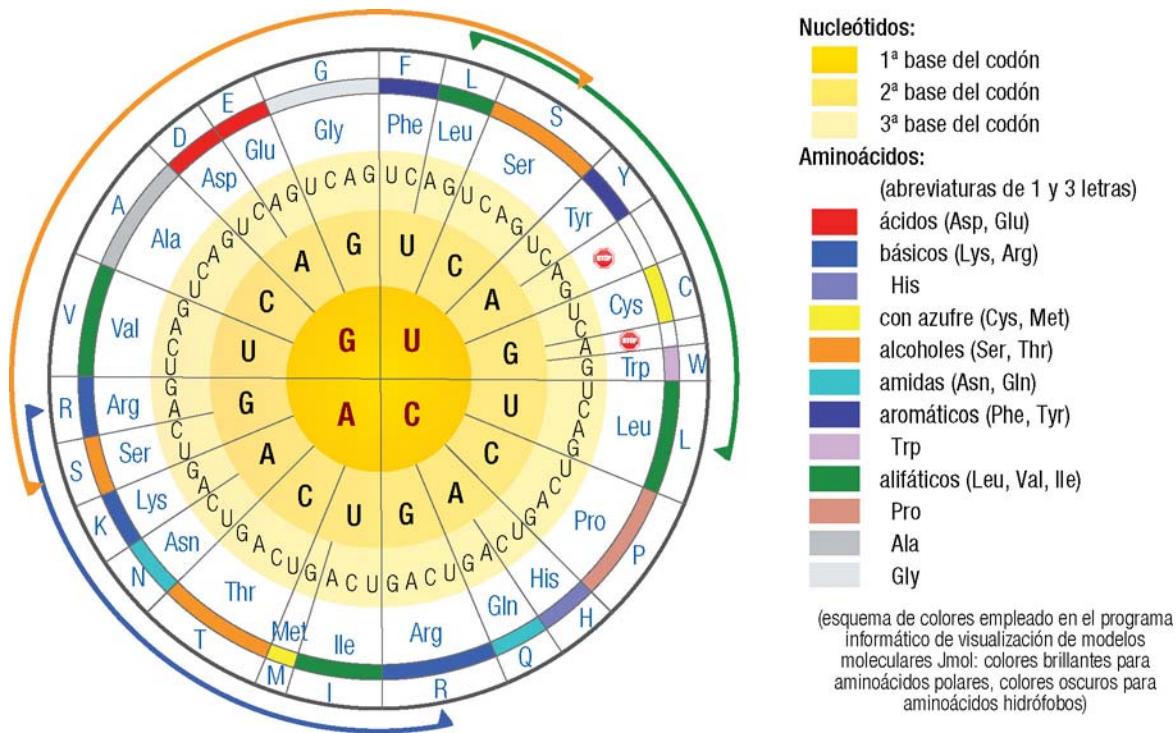
20.3 REPRESENTACIÓN DEL CÓDIGO GENÉTICO

Se han elaborado diversos modelos (tablas, esquemas, etc.) para representar de forma clara y sencilla la correlación existente entre codones del mRNA y aminoácidos del polipéptido traducido. En todos ellos, la secuencia del triplete se refiere al codón del mRNA (de ahí la presencia de U y no de T), nunca se indica como triplete del DNA ni como triplete anticodón del tRNA (aunque se pueden y deben hacer referencias cruzadas entre las secuencias correspondientes). Las bases de cada codón se numeran, como siempre, en la dirección 5' → 3', la misma en la que se sintetiza el mRNA; así, la primera posición es el extremo 5' del codón y la tercera posición es su extremo 3'.



Web 20.2. Representaciones rectangulares del código genético.

Aunque aún es de uso muy común la representación rectangular en forma de tabla, es más recomendable, por su sencillez de empleo, la representación circular. Se utilizan círculos concéntricos en los que se indican, del centro hacia fuera, los 3 nucleótidos del codón y el aminoácido que codifica.



20.5

Representación construida con base en el "Diagrama del sol naciente para el código genético", portada de Biochemical Education, vol.23, nº 3, julio 1995, Pergamon Press

20.4 CARACTERÍSTICAS ESPECÍFICAS

Empleando la representación anterior se pueden analizar con facilidad las características que definen al código genético:

20.4.1 Codón de inicio

El codón AUG, el único que codifica metionina, es responsable como cualquier otro codón de la incorporación del aminoácido correspondiente al polipéptido, pero al mismo tiempo posee la particularidad de actuar de forma específica como **codón de inicio** en la mayoría de los casos (con mucha menos frecuencia se usan como inicio otros codones, como ACG, CUG y GUG). Es decir, la síntesis de la mayoría de las proteínas comienza en un codón AUG; ello explica que Met siempre ocupe la posición 1 (inicial o N-terminal) de cualquier polipéptido recién sintetizado. Como se verá posteriormente (pág. 334), las proteínas eucarióticas poseen una metionina N-terminal, mientras que en procariotas se incorpora como formil-metionina. Para que AUG actúe como codón de inicio en eucariotas se requiere que forme parte de una secuencia específica (secuencia de Kozak, pág. 333).

20.4.2 Codones de terminación

Se puede observar en las representaciones del código genético que los codones UAA, UAG y UGA no codifican aminoácido alguno, sino que causan la finalización de la síntesis proteica (pág. 340). Se denominan **codones de terminación, de paro, codones "stop"**, o también **codones sin sentido** (es decir, sin significado, sin mensaje).

La presencia de 3 codones de terminación en un total de 64 indicaría que, estadísticamente, la síntesis no se alargaría más allá de unos 20 aminoácidos ($3/64 = 1/21 = 4,7\%$). Sin embargo, la región estructural de los genes de proteínas contiene una frecuencia de codones de terminación muy inferior a la que le correspondería de forma aleatoria.

La región del DNA comprendida entre un codón de inicio y uno de terminación se llama “marco de lectura abierto” (*open reading frame, ORF*). Por ejemplo, para una proteína de masa 60 kDa se requeriría un marco de lectura abierto con unos 545 codones (masa media de un residuo aminoácido = 110 Da; 60.000 Da/110 Da = 545).

20.4.3 Degeneración

Ésta es una de las características más llamativas del código. Se entiende por *degeneración* el hecho de que un aminoácido esté codificado por más de un codón; los codones que codifican el mismo aminoácido se denominan **codones sinónimos**. Aunque a primera vista pueda parecerlo, ello no supone indefinición o ambigüedad alguna, pues una secuencia concreta de bases define inequívocamente los aminoácidos que deben incorporarse; si habría ambigüedad si un mismo codón codificase varios aminoácidos alternativos. La degeneración es una consecuencia directa del superávit de codones sobre aminoácidos (debe recordarse que existen 64 tripletes posibles y sólo 20 aminoácidos; pág. 314).

Esta degeneración del código, además, no es uniforme, pues cada aminoácido está codificado por un número diferente de codones sinónimos, desde sólo uno (Met, Trp) hasta un máximo de 6 (Arg, Leu, Ser).

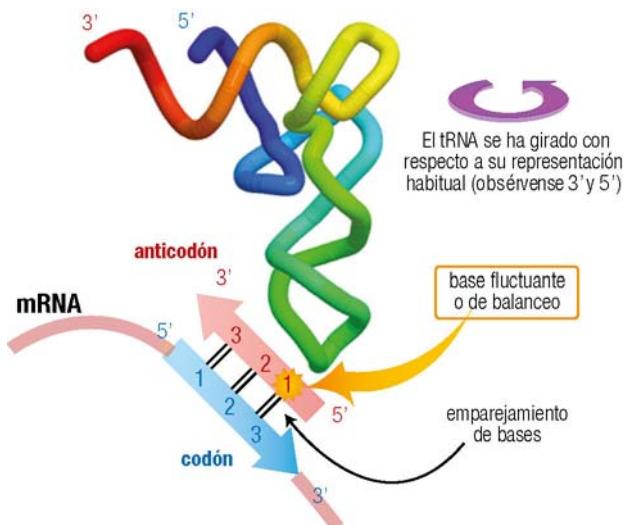
el conjunto de codones sinónimos para cada aminoácido se puede resumir, empleando los símbolos habituales:
 N= cualquier nucleótido
 R= purínico (A o G)
 Y= pirimidínico (C o U)

aminoácidos			codones sinónimos					
M	Met	metionina	AUG					
W	Trp	triptófano	UGG					
D	Asp	aspirótico	GAC	GAU				GAY
N	Asn	asparagina	AAC	AAU				AAY
C	Cys	cisteína	UGC	UGU				UGY
F	Phe	fenilalanina	UUC	UUU				UUY
E	Glu	glutámico	GAA	GAG				GAR
Q	Gln	glutamina	CAA	CAG				CAR
H	His	histidina	CAC	CAU				CAY
K	Lys	lisina	AAA	AAG				AAR
Y	Tyr	tirosina	UAC	UAU				UAY
I	Ile	isoleucina	AUA	AUC	AUU			
A	Ala	alanina	GCA	GCC	GCG	GCU		GCN
G	Gly	glicina	GGA	GGC	GGG	GGU		GGN
P	Pro	prolina	CCA	CCC	CCG	CCU		CCN
T	Thr	treonina	ACA	ACC	ACG	ACU		ACN
V	Val	valina	GUA	GUC	GUG	GUU		GUN
R	Arg	arginina	CGA	CGC	CGG	CGU	AGA	AGG
L	Leu	leucina	CUA	CUC	CUG	CUU	UUA	UUG
S	Ser	serina	UCA	UCC	UCG	UCU	AGC	AGU

Obsérvese que 17 aminoácidos están codificados por codones situados en una misma cuadrícula de la representación del código, y sólo 3 (Ser, Leu, Arg) lo están por codones de dos cuadrículas distintas. Esto indica que la asignación de codones a un aminoácido no es aleatoria, sino que entre los codones sinónimos existe una cierta similitud. Una consecuencia importante de la degeneración del código es que disminuye la frecuencia con que las mutaciones puntuales provocan la incorporación de un aminoácido distinto (pág. 389).

20.4.3.1 Hipótesis del balanceo

El triplete del anticodón de un tRNA (formando parte de un aminoacil-tRNA) puede interaccionar por complementariedad de bases con un codón del mRNA (pág. 315). Esta *lectura o reconocimiento* se realiza por emparejamiento antiparalelo de las bases 1, 2 y 3 del codón con las bases 3, 2 y 1 del anticodón, respectivamente.



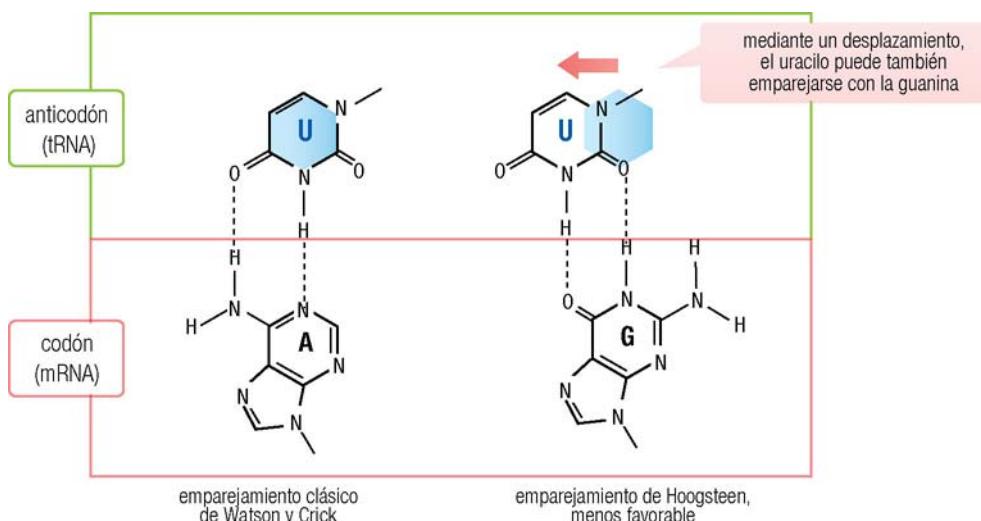
20.7



Web 20.3. Estructura molecular de la interacción entre codones y anticodones.

Si los tres emparejamientos tuviesen lugar de acuerdo con las reglas de Watson y Crick sólo habría una posibilidad de interacción entre anticodón y codón, por lo que existiría igual número de tRNA (anticodones) que de codones, es decir, 61. Sin embargo, se sabe que en las células hay menos moléculas diferentes de tRNA (págs. 15 y 19), lo que quiere decir que un mismo anticodón puede reconocer varios codones. Para poder explicar este hecho se propuso la “hipótesis del balanceo” (también llamada fluctuación, tambaleo o titubeo), que se puede expresar en dos reglas:

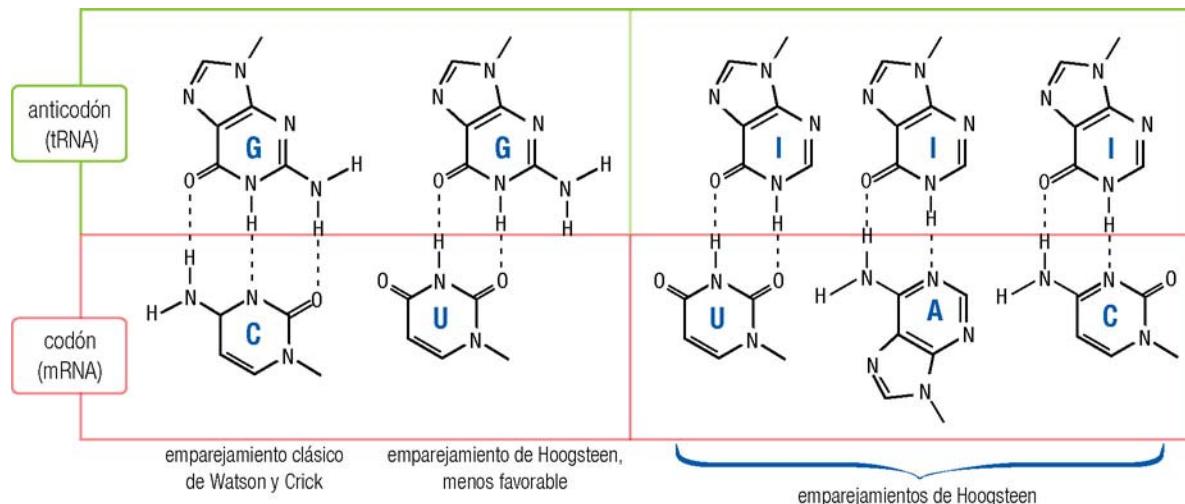
1. Las bases tercera y segunda del anticodón forman puentes de hidrógeno canónicos (de tipo Watson y Crick, bien definidos) con las bases primera y segunda del codón, determinando la especificidad de interacción codón/anticodón.
2. La primera base del anticodón (posición 5' del tRNA) puede orientarse de formas ligeramente distintas (lo que se ha llamado **balanceo** o **fluctuación**) para interaccionar con bases alternativas en la posición tercera del codón. Este emparejamiento tiene lugar mediante enlaces de hidrógeno diferentes de los clásicos de Watson y Crick. Debido a ello, la interacción es más débil (coloquialmente, un emparejamiento “suelto”).



20.8

A esta flexibilidad de interacción contribuyen no sólo los emparejamientos atípicos entre bases normales, sino también la presencia, en algunos anticodones, del nucleósido infrecuente **inosina** (abreviado como **I**, con la base nitrogenada hipoxantina, pág. 15 y 19), que es capaz de emparejarse tanto con A como con U o C. La inosina se genera por la acción de una adenosina desaminasa sobre la A presente en el transcripto primario del tRNA (reacción en págs. 306 y 395).

Los detalles de las interacciones de Watson y Crick (pág. 40) y las menos favorables (denominadas aquí genéricamente enlaces de Hoogsteen, págs. 55 y 67) son los siguientes:



Resumiendo, los emparejamientos posibles mediante “balanceo” son los siguientes:

1 ^{er} nucleótido (5') en el anticodón (tRNA):	3 ^{er} nucleótido (3') en el codón (mRNA):	resultado:
C	G (Watson-Crick)	C reconoce un solo codón
A	U (Watson-Crick)	A reconoce un solo codón
U	A (Watson-Crick) G (más débil)	U reconoce dos codones
G	C (Watson-Crick) U (más débil)	G reconoce dos codones
I	A (más débil) U (más débil) C (más débil)	I reconoce tres codones

20.10

20.4.3.2 Número de RNA transferentes necesarios

Todo lo anterior determina que el anticodón de un tRNA puede leer uno, dos o tres codones de un mRNA, en los cuales las 2 primeras bases son idénticas. Con este fundamento, puede calcularse el número de anticodones, el número de moléculas diferentes de tRNA necesarias para los 61 codones codificadores de los 20 aminoácidos. Aunque en las células ésta es una cifra imprecisa, desde un planteamiento teórico, basado en el aprovechamiento al máximo del balanceo, puede afirmarse que se necesitan *al menos* 31 tRNA para “leer” todo el código.



Los cuatro codones de una cuadrícula precisan siempre dos anticodones, bien a costa de los balanceos de U y G o bien del balanceo de I

	U	C	A	G
U	UUU } AAG trNA Phe UUC } AAU trNA Leu ₁ UUA } UUG trNA Leu ₁	UCU } AGG trNA Ser ₁ UCC } AGI trNA Ser ₂ UCA } AGU trNA Ser ₂	UAU } AUG trNA Tyr UAC } PARO	UGU } ACA trNA Cys UGC } PARO
	UCG }	■ AGC	UAA } UAG	UGG } ACC trNA Trp
C	CUU } GAG trNA Leu ₁ CUC } GAI trNA y CUA } GAU trNA Leu ₁	CCU } GGG trNA Pro ₁ CCC } GGI trNA Pro ₂ CCA } GGU trNA Pro ₂	CAU } GUG trNA His CAC } GGU trNA Gln	CGU } GCG trNA Arg ₁ CGC } GCI trNA Arg ₂
	CGU }	■ GGC	CAG } GUU	CGA } GCU trNA Arg ₂
A	AUU } UAI trNA Ile AUC } UAC trNA Met	ACU } UGG trNA Asn ACC } UGI trNA	AAU } UUG trNA Asn AAC } UUU trNA Lys	AGU } UCG trNA Ser ₃ AGC } UCC trNA
	AUA }	■ UGC	AAA } UUU	AGA } UCC trNA Arg ₃
G	GUU } CAG trNA Val ₁ GUC } CAI trNA y GUA } CAU trNA Val ₂	GCU } CGG trNA Ala ₁ GCC } CGI trNA Ala ₂ GCA } CGU trNA	GAU } CUG trNA Asp GAC } CUC trNA Glu	GGU } CCG trNA Gly ₁ GGC } GUI trNA Gly ₂
	GUG }	■ CGC	GAA } CUU GAG }	GGG } CCU trNA Gly ₂

20.11 Los anticodones con G como 1^a base (5') emparejan por balanceo con 2 codones, los terminados en U y C.
Los anticodones con U en 1^a posición (5') emparejan por balanceo con los 2 codones cuya 3^a base (3') es A y G.

Los anticodones con inosina (I) como 1^a base emparejan por balanceo con 3 codones, los terminados en U, C y A.
Se requiere otro anticodón para el 4^o codón de la cuadrícula, sin balanceo.

nº de codones por aminoácido	nº de anticodones por aminoácido	nº total de aminoácidos	nº total de anticodones	nº total de codones
1	1	2	2	2
2	1	9	9	18
3	1	1	1	3
4	2	5	10	20
6	3	3	9	18
suma:		20	31	61

20.12

20.4.4 Universalidad

Los primeros estudios indicaron que la correspondencia entre codones y aminoácidos era la misma independientemente del organismo observado, lo que condujo a la afirmación de que “el código genético es universal”. Hoy día se sabe que esto no es completamente cierto; se han encontrado excepciones, aunque escasas, en las mitocondrias humanas y de otros mamíferos, y en ciertas bacterias. Debe, pues, decirse que el código genético es *casi universal*.

Diferencias en el código genético de humanos		
	núcleo	mitocondrias
AGA, AGG	Arg	paro
AUA	Ile	Met
UGA	paro	Trp

20.13

En consecuencia, comparten el mismo código genético los procariotas, los núcleos de todos los eucariotas y las mitocondrias de plantas. Parece probable que todas las formas de vida hayan tenido un antepasado evolutivo común, con un solo código genético, que se ha conservado muy estrictamente a todo lo largo de la evolución biológica. Las pocas diferencias existentes pueden haber surgido durante la evolución por transferencia del genoma procariótico original, con el código genético universal, al genoma mitocondrial.



Síntesis de proteínas: traducción

21.1 CARACTERÍSTICAS DE LA TRADUCCIÓN	323
21.1.1 Sentido de avance de la síntesis proteica	324
21.1.2 Carácter monocistrónico o policistrónico del mRNA	324
21.1.3 Fases de la traducción	326
21.2 ACTIVACIÓN DE LOS AMINOÁCIDOS EN FORMA DE AMINOACIL-tRNA	326
21.2.1 Reacciones de aminoacilación	327
21.2.2 Aminocil-tRNA sintetasas	329
21.2.2.1 Tipos de sintetasas	329
21.2.2.2 Especificidad de las aminoacil-tRNA sintetasas	329
21.2.2.3 Mecanismo de reconocimiento de los aminoácidos por las sintetasas	329
21.2.2.4 Mecanismo de reconocimiento de los tRNA por las sintetasas	329
21.2.3 Corrección de errores cometidos por las sintetasas	330
21.2.4 Balance energético de la activación	333
21.3 INICIACIÓN	333
21.3.1 Especificidad del punto de inicio	333
21.3.2 Formación del metionil-tRNA iniciador	334
21.3.3 Pasos sucesivos en la fase de inicio	334
21.3.3.1 Disociación del ribosoma (paso 1)	335
21.3.3.2 Unión a la subunidad menor del ribosoma: complejo de preiniciación (paso 2)	335
21.3.3.3 Unión de la subunidad mayor: complejo de iniciación (paso 3)	336
21.3.4 Reutilización del ribosoma	336
21.3.5 Balance energético de la iniciación	336
21.4 ELONGACIÓN O ALARGAMIENTO DE LA CADENA PEPTÍDICA	337
21.4.1 Ubicación del aminoacil-tRNA en el sitio A (paso 4)	337
21.4.2 Transpeptidación: formación del enlace peptídico (paso 5)	339
21.4.3 Translocación: desplazamiento del ribosoma en un codón (paso 6)	339
21.4.4 Repetición del ciclo de elongación (pasos 4, 5 y 6)	340
21.5 TERMINACIÓN	340
21.5.1 Unión del factor de liberación (paso 7)	341
21.5.2 Hidrólisis del peptidil-tRNA (paso 8)	341
21.5.3 Disociación (paso 9)	342
21.5.4 Reutilización del ribosoma	342
21.6 ENERGÉTICA DE LA SÍNTESIS DE PROTEÍNAS	342
21.7 INHIBIDORES DE LA TRADUCCIÓN	343
21.8 REGULACIÓN DE LA SÍNTESIS PROTEICA	345
21.8.1 Control del transporte de mRNA	346
21.8.2 Vida media y degradación del mRNA	347
21.8.3 Control de la traducción	348
21.8.3.1 Ribointerruptores	348
21.8.3.2 Ribointerferencia	348
21.8.3.3 Control de la accesibilidad del mensajero	350
21.8.3.4 Control de la estabilidad del mensajero	351
21.8.3.5 Regulación del complejo de iniciación	352

21.1 CARACTERÍSTICAS DE LA TRADUCCIÓN

La traducción consiste en la síntesis de las proteínas mediante la unión covalente de aminoácidos según el orden establecido por la secuencia de nucleótidos del mRNA y de acuerdo con el código genético. Se completa así el “dogma central” de transmisión de la información genética a partir del DNA genómico original (pág. 3 y 262).

Puesto que este proceso es esencialmente igual en todo tipo de organismos (salvo pequeñas diferencias en los mecanismos y componentes implicados), nuestro estudio se centra en eucariotas, aludiendo a procariotas cuando lo requieran sus aspectos particulares en favor de una mejor comprensión del proceso en términos moleculares. Se plantea en estrecha relación con el código genético (Capítulo 20) y se completará con las modificaciones postraduccionales (Capítulo 22), necesarias para el ejercicio de la función de las proteínas; finalmente, se analizarán los mecanismos de degradación, de forma análoga al planteamiento tradicional para otras biomoléculas (glúcidos y lípidos, principalmente). Con todo ello se completan las dos grandes áreas –biosíntesis y degradación– del metabolismo proteico. Este enfoque es imprescindible hoy día para conocer el funcionamiento normal y patológico de las proteínas (base de las enfermedades moleculares, Capítulo 25) y para abordar en el futuro los avances en la naciente proteómica (pág. 239) con vistas al diagnóstico y a la terapéutica.

Desde un punto de vista bioquímico, la traducción se caracteriza por la gran variedad de proteínas formadas, su elevado coste energético (consume un 80-90% de la energía que emplea la célula en biosíntesis) y la necesidad de una regulación muy estrecha en respuesta a las necesidades celulares y al ritmo de degradación de las proteínas. Es, posiblemente, el más complejo de los procesos de síntesis, en el que participa un mayor número de macromoléculas diferentes. De modo resumido, las principales son:

- Un tRNA especial para la iniciación ($tRNA_f^{Met}$ en procariotas, $tRNA_i^{Met}$ en eucariotas; pág. 334).
- 31 tipos de tRNA portadores de aminoácidos, en forma de aminoacil-tRNA. En realidad, éste es el número mínimo necesario teóricamente (págs. 320 y 330); en las células suele haber un número superior de tRNA diferentes; por ejemplo, 40 es una cifra frecuente y se han encontrado hasta 84.
- Ribosomas, formados por varias moléculas de rRNA y numerosas proteínas (págs. 75-76).
- Un mRNA, de secuencia distinta para cada polipéptido sintetizado.
- Factores proteicos de inicio, elongación y terminación de la traducción.

En una célula existen millares de estos participantes: unos 20.000 ribosomas, 100.000 moléculas de enzimas y factores proteicos, 200.000 de tRNA, etc. Se requiere, además, un acervo de, al menos, 20 aminoácidos, ATP, enzimas aminoacil-tRNA sintetasas para activar los aminoácidos, GTP para varias etapas, grupos reductores, iones (Mg^{+2} y K^+), enzimas auxiliares, etc. Intervienen, directa o indirectamente, casi todas las estructuras subcelulares (núcleo, citoplasma, orgánulos y membranas). A pesar de ello, el proceso es muy rápido: en bacterias, la velocidad de traducción alcanza 20 aminoácidos por segundo, mientras que en eucariotas es más reducida, pero se incorporan entre 3 y 8 aminoácidos por segundo.

21.1.1 Sentido de avance de la síntesis proteica

El primer aspecto básico, el sentido en el que se realiza la síntesis, puede considerarse desde un doble punto de vista: la secuencia de nucleótidos del mRNA se lee de 5' a 3', mientras que la secuencia de aminoácidos del polipéptido se construye desde el extremo amino hacia el carboxilo.



Web 21.1. Demostración experimental del sentido de síntesis de las proteínas.

21.1.2 Carácter monocistrónico o policistrónico del mRNA

Un aspecto importante en relación con la traducción es que un mRNA puede dar lugar a la síntesis de un polipéptido o de varios. Esta propiedad se define mediante el concepto de **cistrón**, que conviene distinguir del de **gen**, aunque están estrechamente relacionados y ambos pueden considerarse una “unidad” de funcionalidad genética.

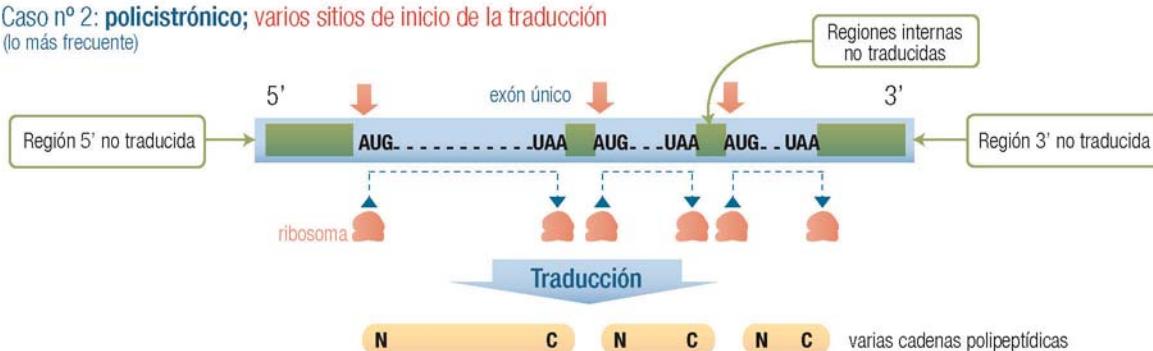
- Los **mRNA monocistrónicos**, presentes en procariotas y típicos en eucariotas, codifican *un solo polipéptido*. Ello implica que la traducción se inicia en un único sitio del mRNA, el codón de inicio, para finalizar en otro sitio, un codón de terminación.
- Los **mRNA policistrónicos**, que sólo aparecen en procariotas y virus, codifican *varios polipéptidos*, por lo que la traducción se inicia (simultáneamente o no) en varios sitios, en varios codones de inicio, para finalizar en sendos codones de terminación.

mRNA procariótico

Caso nº 1: **monocistrónico**; un sitio de inicio de la traducción



Caso nº 2: **policistrónico**; varios sitios de inicio de la traducción (lo más frecuente)



mRNA procariótico

Siempre **monocistrónico**; un sitio de inicio de la traducción



21.1

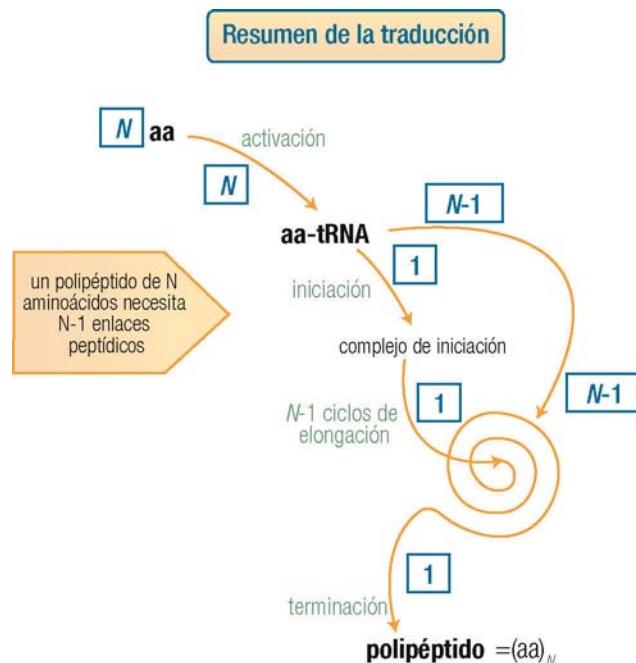
El concepto de **gen** (pág. 263) corresponde a una región codificante del DNA genómico, responsable de un producto génico, sea éste RNA (por transcripción y maduración) o proteína (por transcripción, maduración, traducción y modificación postraduccional). En el caso de mRNA monocistrónicos, el concepto de **cistrón** es prácticamente indistinguible del de gen. La diferencia importante surge en los genes que se transcriben dando un mRNA policistrónico (y, por tanto, sólo afecta a los organismos procarióticos).

El concepto de **cistrón**, en su forma genérica, se diferencia: *a)* por ser más amplio que el de gen, al considerar como región codificante tanto al DNA como al RNA (ambos dan lugar a un producto génico), y *b)* por ser al mismo tiempo más estricto, pues se define como la región del genoma (DNA) o del mRNA que codifica un producto génico único. Habitualmente se emplea sólo para este último caso, por lo que un **cistrón** es la secuencia de mRNA que codifica una molécula de polipéptido. Debe recordarse que la definición más amplia de **gen** obligaba a incluir casos límite como consecuencia de la maduración postranscripcional (pág. 298); de forma similar ocurre en el caso que nos ocupa, el mRNA policistrónico, con una unidad de transcripción que

codifica varios productos (conocida también habitualmente como *operón*). Por lo tanto, en este caso se trata de un *gen policistrónico*.

21.1.3 Fases de la traducción

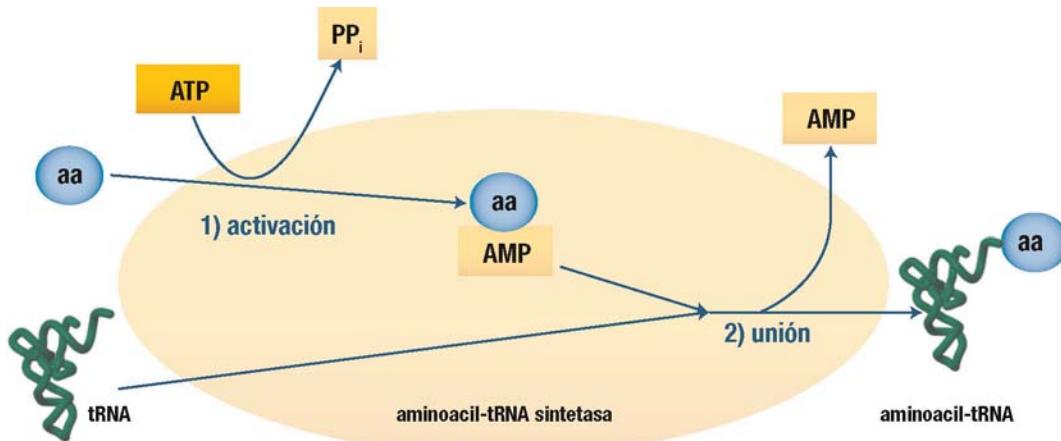
Dentro del proceso de traducción se pueden distinguir varias etapas o fases: una previa o de activación de los aminoácidos precursores y tres bien diferenciadas, **iniciación**, **elongación** y **terminación**, bajo un criterio análogo al seguido para la replicación y la transcripción (Capítulos 11 y 17). La maduración postraducional, considerada en ocasiones como una quinta fase, tiende a tratarse hoy día de modo independiente (Capítulo 22).



21.2

21.2 ACTIVACIÓN DE LOS AMINOÁCIDOS EN FORMA DE AMINOACIL-tRNA

La unión a cada tRNA del aminoácido correcto (el que, según el código genético, está codificado por el codón complementario del anticodón que posee ese tRNA) es posiblemente el paso más crítico de la expresión génica. Si no tiene lugar de forma correcta, carece de sentido todo el código genético. Esta unión tiene lugar en el citosol celular, como el resto del proceso de traducción, y supone la activación de los aminoácidos (en forma de aminoacil-tRNA) para poder formar enlaces peptídicos. De esta forma, la fase de activación da significado al papel conector del tRNA con los aminoácidos en la traducción.



21.2.1 Reacciones de aminoacilación

La unión del aminoácido con el tRNA se verifica mediante dos reacciones sucesivas, ambas catalizadas por una misma enzima (en un centro activo único o en dos centros próximos). Esta enzima es una de las **aminoacil-tRNA sintetasas** (abreviadamente, sintetasas o aaRS).

Resumen de las reacciones

Enzima = aminoacil-tRNA sintetasa
= aminoácido-tRNA ligasa (EC 6.1.1.x)

Primera reacción: activación del aminoácido



el grupo carboxilo del aminoácido se une con el grupo adenilato del ATP, que se hidroliza

intermediario que permanece unido al centro activo

el grupo amino permanece sin sustituir

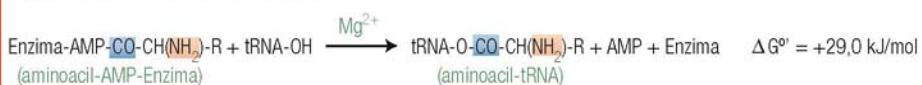
ambas reacciones son exergónicas



$\Delta G^\circ = -33,4 \text{ kJ/mol}$

la hidrólisis del pirofosfato impulsa la reacción

Segunda reacción: unión con el tRNA



el grupo aminoacilo se transfiere desde el intermediario hacia el tRNA

el grupo amino permanece sin sustituir

la reacción es endergónica

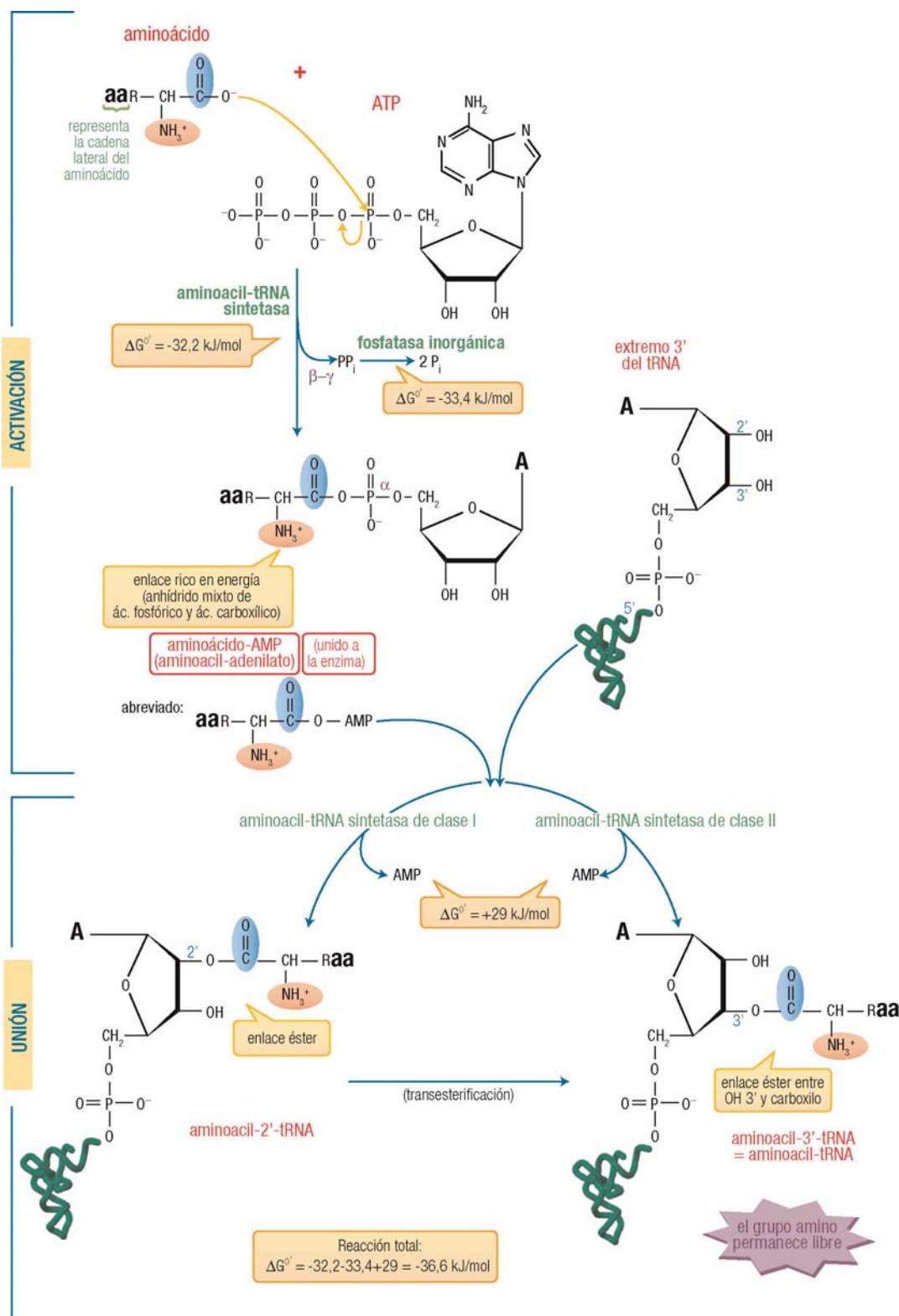
Reacción total:



el aminoacil-tRNA es una forma activada, de alta energía, del aminoácido

la reacción neta es exergónica

El producto final de la doble reacción es un **aminoacil-tRNA**, con un enlace éster entre el grupo OH 3' terminal del tRNA y el grupo carboxilo del aminoácido (véase su nomenclatura en la pág. 64). En el caso de las aminoacil-tRNA sintetasas de clase I (pág. 329), el éster se forma inicialmente sobre el grupo OH 2', pero luego se convierte al isómero en 3'.



21.2.2 Aminoacil-tRNA sintetasas

21.2.2.1 Tipos de sintetasas

Se han aislado multitud de aminoacil-tRNA sintetasas de bacterias, plantas y animales, cuya masa oscila entre 88 y 181 kDa. En algunos casos se han podido secuenciar y cristalizar para estudiar su estructura. Todas las sintetasas catalizan de forma similar la reacción de activación del aminoácido con AMP, pero presentan diferencias en la segunda reacción, la de unión al tRNA. Concretamente, algunas unen el aminoácido al grupo hidroxilo 2' y otras al 3', siempre del último nucleótido del tRNA (pág. 328). De acuerdo con esto, y también con diferencias estructurales, se clasifican las sintetasas en dos grupos denominados, respectivamente, **clase I** y **clase II**.

Tipo de aaRS	Grupo OH del tRNA	Transesterificación posterior	Estructura de la enzima	Aminoácidos que reconocen
De clase I	2'	Es necesaria	La mayoría monoméricas	Arg Cys Gln Glu Ile Leu Met Trp Tyr Val (10, los más grandes y polares)
De clase II	3'	No se requiere	La mayoría diméricas	Ala Asn Asp Gly His Lys Phe Pro Ser Thr (10, los más pequeños e hidrófobos)

21.2.2.2 Especificidad de las aminoacil-tRNA sintetasas

Ya se ha explicado (pág. 319) cómo la hipótesis del balanceo permite que un mínimo de 31 tRNA reconozcan los 61 codones que codifican los 20 aminoácidos. Se deben formar, pues, 32 aminoacil-tRNA diferentes (debe añadirse un tRNA específico para la incorporación de la metionina de iniciación). Cada uno de éstos interaccionará, a través de su anticodón, con el codón del mRNA que codifica el aminoácido correspondiente. En muchos casos, un mismo tRNA puede interaccionar con varios codones sinónimos, gracias al balanceo. Por otra parte, para algunos aminoácidos el grupo de codones (4 o 6) requiere más de un tRNA; éstos se llaman **tRNA sinónimos o isoaceptores** (pág. 64).

Este papel adaptador de los tRNA entre la información genética del mRNA (nucleótidos) y la secuencia del polipéptido (aminoácidos) sólo puede desempeñarse gracias a la especificidad de las sintetasas por ambos sustratos, el tRNA y el aminoácido. Cada aminoacil-tRNA sintetasa reconoce un único aminoácido (del que toma el nombre) y también todos los tRNA cuyos anticodones codifican ese mismo aminoácido (codones sinónimos en el código genético). De acuerdo con ello, existen en la célula 20 aaRS distintas. El papel de las sintetasas es, por tanto, esencial para la fidelidad de la síntesis proteica y constituye la esencia del código genético, el problema central de la traducción; por ello, se suele considerar que las aaRS representan un **segundo código genético**.

21.2.2.3 Mecanismo de reconocimiento de los aminoácidos por las sintetasas

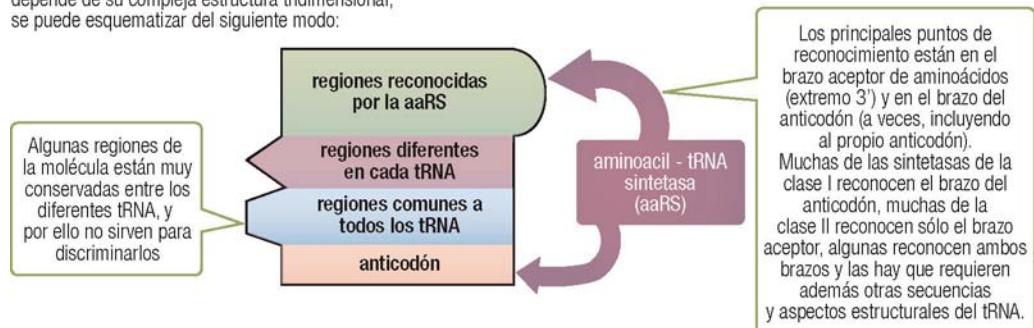
La interacción específica de un solo aminoácido con cada sintetasa es análoga al reconocimiento de cualquier otro sustrato por la enzima correspondiente. Cada aminoácido encaja en un determinado lugar del centro activo de la sintetasa, gracias a interacciones por puentes de hidrógeno, electrostáticas, hidrófobas o de van der Waals. Sólo se unirán aquellos aminoácidos con un número suficiente de interacciones favorables.

21.2.2.4 Mecanismo de reconocimiento de los tRNA por las sintetasas

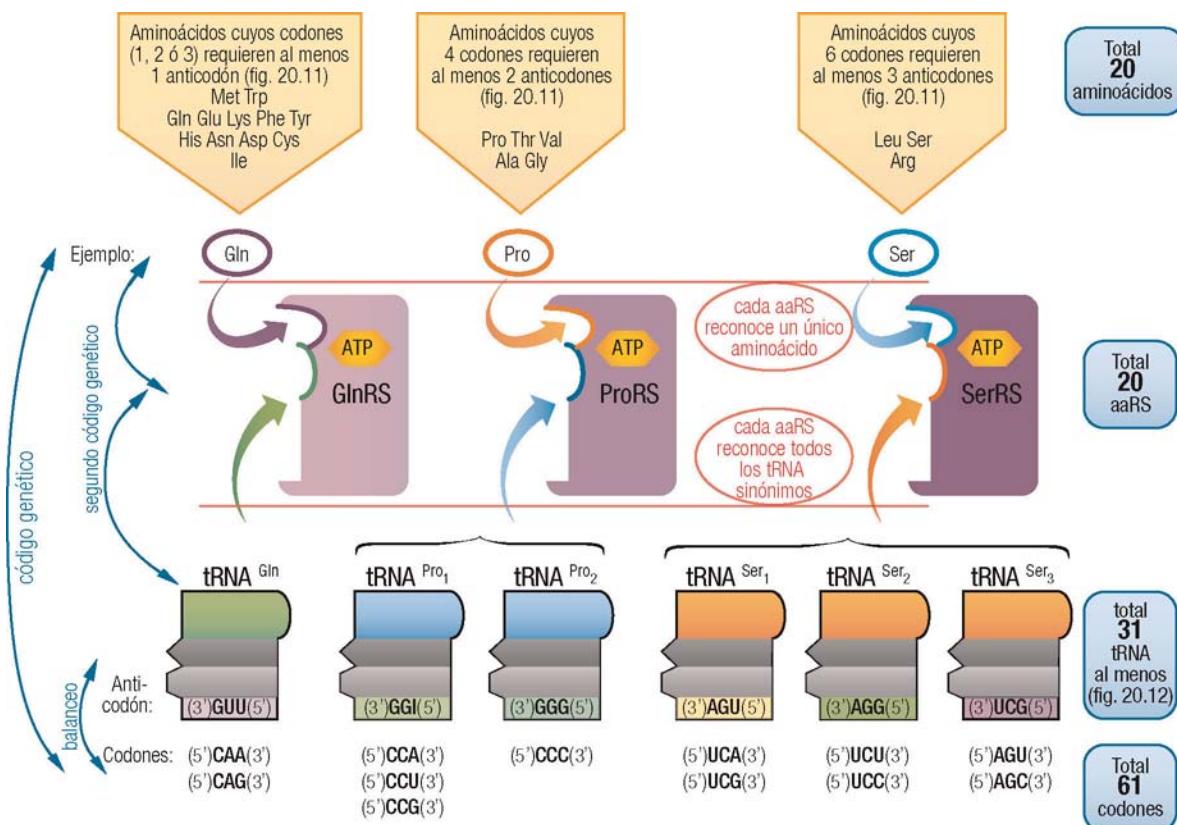
La capacidad de una sintetasa, específica para un único aminoácido, para reconocer varios tRNA, y al mismo tiempo rechazar otros, depende del modo como éstos interaccionan con el centro activo de la aaRS. Este

reconocimiento es complejo y diferente para cada pareja tRNA-sintetasa, e implica no sólo las variaciones estructurales particulares de cada enzima (en especial, según que sean de clase I o II), sino también la interacción con diversas zonas de la molécula de tRNA.

La particularidad de cada molécula de tRNA, que depende de su compleja estructura tridimensional, se puede esquematizar del siguiente modo:



21.6a



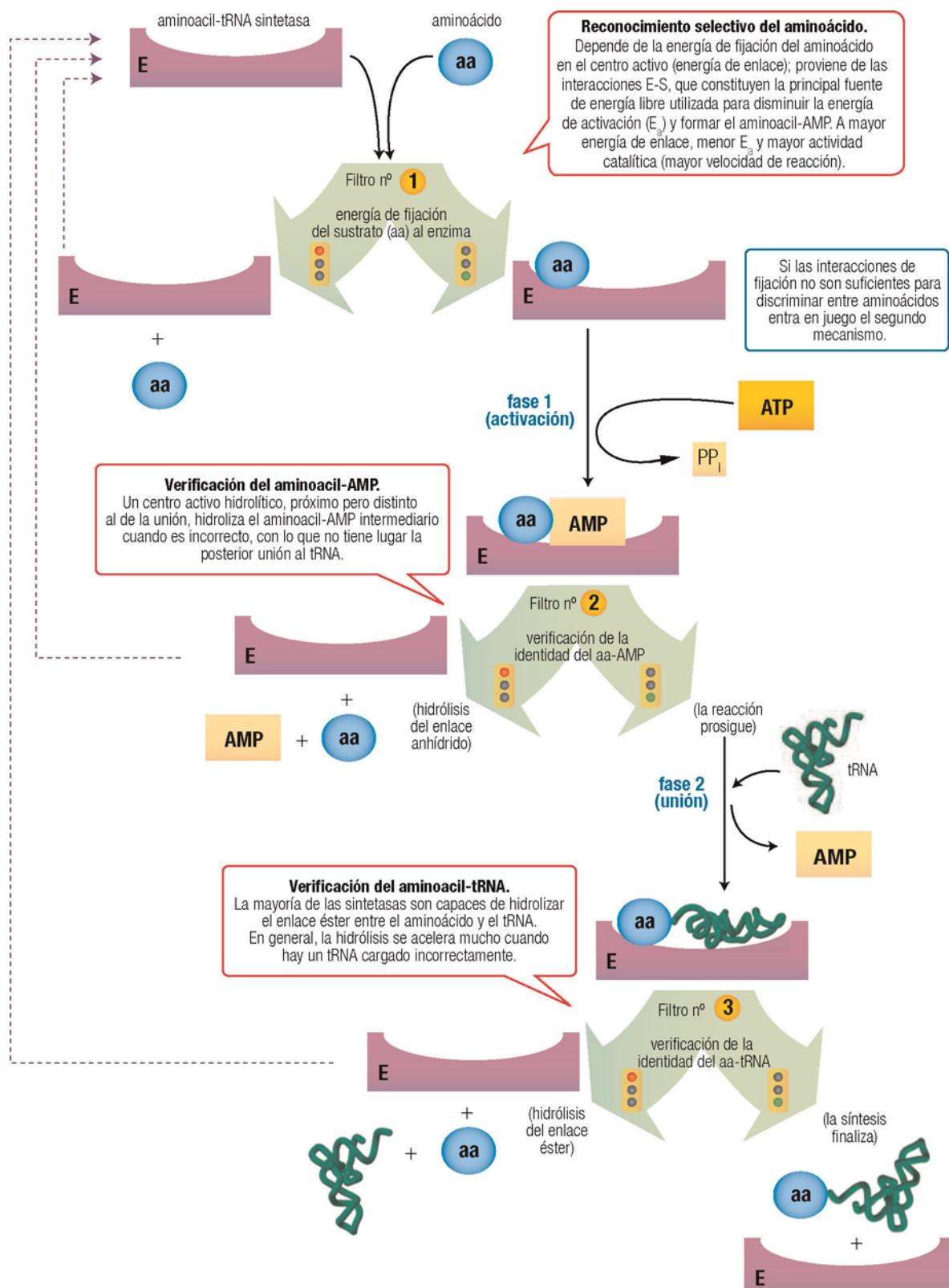
21.6b



Web 21.2. Estructura tridimensional y reconocimiento entre un tRNA y su aminoacil-tRNA sintetasa.

21.2.3 Corrección de errores cometidos por las sintetasas

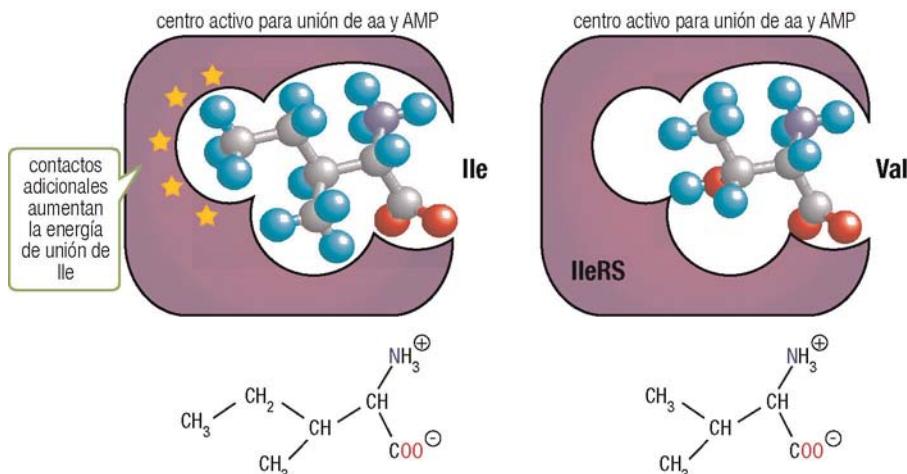
La elevada especificidad de las sintetasas hace que los errores sean infrecuentes. A pesar de ello, se dispone de mecanismos de corrección que actúan de manera sucesiva, en el centro activo de la enzima, antes y después de cada etapa de la reacción de síntesis, y reducen aún más la tasa de error.



21.7

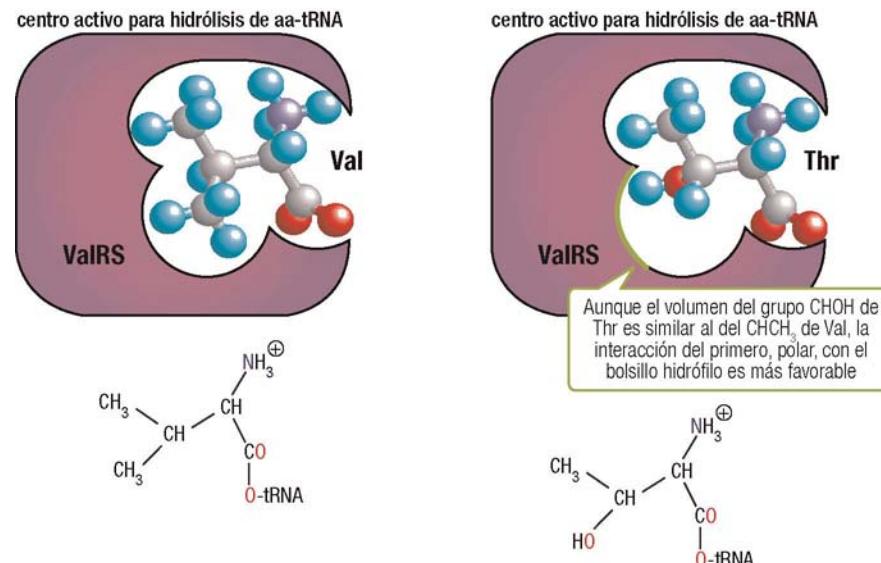
Ejemplo del filtro n.º 1. La enzima isoleucil-tRNA sintetasa (IleRS) no distingue bien entre Ile (su sustrato específico) y Val, aminoácido de estructura similar, aunque de menor tamaño por tener un grupo metileno menos. La activación preferente de Ile respecto a Val en la primera etapa de reacción (formación de Ile-AMP-E frente a Val-AMP-E) está de acuerdo con la interacción más favorable del centro activo con Ile (energía de

fijación superior en 3 kcal/mol a la de Val). A pesar de esta preferencia, dado que *in vivo* la concentración de Val es unas 5 veces mayor que la de Ile, Val debería incorporarse erróneamente con gran frecuencia (en torno al 1%). El hecho de que la tasa de error observada sea muy inferior (0,01%) indica que existen mecanismos adicionales de corrección.



Ejemplo del filtro n.º 2. Con el mismo caso anterior, si se ha llegado a formar erróneamente Val-AMP-E en el sitio activo de IleRS, se percibe el error y se hidroliza el anhídrido entre Val y AMP. Es probable que esta corrección de pruebas tenga lugar porque el sitio hidrolítico de la enzima admite a Val-AMP (incorrecto), pero excluye o rechaza, por impedimento estérico, a Ile-AMP (correcto, de mayor tamaño).

Ejemplo del filtro n.º 3. La enzima valil-tRNA sintetasa (ValRS) sirve para demostrar la capacidad de corregir errores entre dos aminoácidos cuyo tamaño es casi idéntico por tener cadenas laterales muy similares; concretamente, Val y Thr. Sin embargo, ValRS rechaza a Thr por una combinación de dos mecanismos. El centro activo de transferencia es hidrófobo e interacciona de manera más favorable con Val, de cadena lateral más hidrófoba. Por el contrario, el centro activo de hidrólisis es hidrófilo e interacciona mejor con Thr, más polar. Como resultado, la Thr unida por error a tRNA^{Val} se hidrolizará preferentemente, mientras que Val unida correctamente a tRNA^{Val} se conserva.



A pesar de todos estos mecanismos de corrección, la tasa de error de la síntesis proteica es mucho mayor que la de la replicación (respectivamente, uno por cada 10⁴ aminoácidos y uno por cada 10⁶–10⁸ nucleótidos). Esto es permisible puesto que el error en una proteína desaparece al degradarse ésta y no pasa a futuras generaciones, y en la práctica supone un grado de fidelidad suficiente, sin implicar un gasto energético excesivo.

21.2.4 Balance energético de la activación

El *primer gasto energético* que se debe considerar en la traducción es el necesario para la formación de los aminoacil-tRNA para todos los aminoácidos que componen la proteína. Como se ha descrito (pág. 327), en cada reacción de aminoacilación se hidrolizan dos enlaces fosfato de alta energía ($ATP \rightarrow AMP + PP_i$ y $PP_i \rightarrow 2 P_i$) para formar el enlace éster entre aminoácido y tRNA.

No se tendrá en cuenta para el balance global, aunque evidentemente es significativo en la situación real en la célula, el gasto energético desperdiciado cada vez que un aminoácido se incorpora de manera incorrecta (pág. 330); cada error supone dos enlaces fosfato, al haberse sintetizado un aa-tRNA inútil, que se hidroliza durante la corrección por la sintetasa.

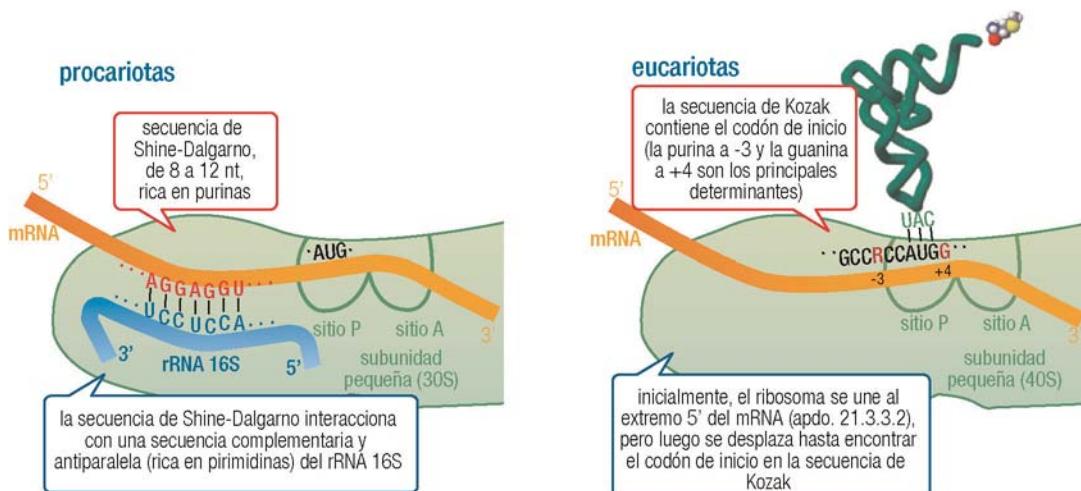
21.3 INICIACIÓN

Esta etapa, previa a la formación del primer enlace peptídico, es habitualmente la etapa limitante para la velocidad global de la síntesis. Transcurre de forma semejante en procariotas y eucariotas, salvo por la identidad de las proteínas implicadas, por lo que la explicación se centrará en los últimos. Tanto en esta fase como en las de elongación y terminación es característica la intervención de complejos formados por las 3 moléculas principales –los tRNA, el ribosoma (completo o una sola subunidad) y el mRNA–, además de otros factores proteicos, coenzimas, iones, etc.

21.3.1 Especificidad del punto de inicio

La traducción nunca comienza exactamente en el extremo 5' del mRNA (por ello, se ha venido representando siempre una “región 5’ no traducida” (5’UTR), por ejemplo en pág. 325). Casi siempre actúa como punto de inicio un codón AUG (pág. 317) situado internamente en la molécula, para lo cual necesita ser reconocido como tal por el ribosoma y por el tRNA^{Met} iniciador.

En procariotas, cada cistrón de un mRNA (pág. 325) posee una secuencia denominada de Shine-Dalgarno, que interacciona específicamente con la subunidad pequeña del ribosoma, definiendo la elección como codón de inicio de un AUG situado unos 10 nt cadena abajo.



21.10

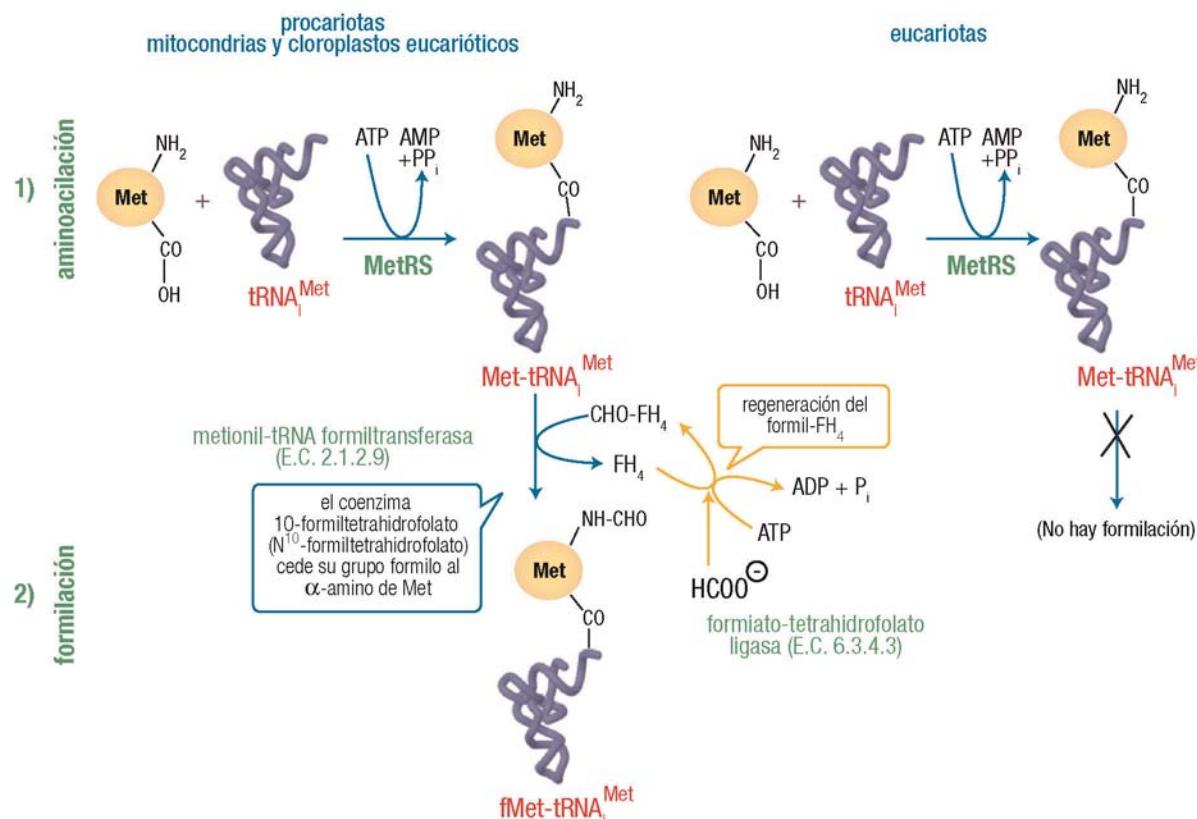
En eucariotas, el mRNA es casi siempre monocistrónico (pág. 325), por lo que tiene un solo codón de inicio. No existe una señal especial que une el mRNA a la subunidad menor del ribosoma, sino que la especificidad se establece a través del tRNA iniciador. Para que éste se una al mRNA, en el entorno proporcionado por la subunidad menor del ribosoma, es preciso que el codón AUG forme parte de la secuencia de Kozak; se forma así un complejo ternario mRNA:ribosoma:tRNA_i^{Met} (como se verá después, el tRNA participa realmente como metionil-tRNA_i^{Met} asociado al factor eIF-2 y a GTP).

21.3.2 Formación del metionil-tRNA iniciador

Los codones AUG en los que se inicia la traducción (codificando el residuo Met del extremo N-terminal del polipéptido) los reconoce un tRNA iniciador, llamado $tRNA_i^{Met}$ (en procariotas suele denominarse $tRNA_f^{Met}$, pero se usará el primer nombre de forma genérica). Los codones AUG que corresponden a posiciones internas en la cadena polipeptídica interactúan con un tRNA distinto, que se llama simplemente $tRNA^{Met}$. En ambos casos el tRNA participa como metionil-tRNA Met (etapa previa de activación del aminoácido, pág. 326).

La incorporación de metionina interna corresponde a la etapa de elongación, y tiene lugar de la misma forma que la de otros aminoácidos; la de metionina N-terminal ocurre, sin embargo, durante la iniciación, y de forma algo diferente en procariotas y eucariotas:

En procariotas (y en mitocondrias y cloroplastos eucarióticos) se forma Met-tRNA Met , que luego se modifica con un grupo formilo en el amino α de la metionina. Es decir, hay **aminoacilación** y **formilación**, para incorporar N-formil-metionina (fMet) al inicio de la traducción. Como consecuencia, todas las proteínas recién sintetizadas comienzan por formilmétionina. Debe resaltarse que no se utiliza fMet como sustrato. Esta etapa supone un pequeño coste energético adicional para la traducción (que no se contemplará al hacer el balance general, pág. 343): aunque la reacción no consume energía directamente, sí lo hace la regeneración del coenzima formiltetrahidrofolato.



En la traducción citoplasmática de eucariotas no hay participación de formilmétionina, sólo se requiere la **aminoacilación** para incorporar metionina al inicio de la síntesis. Por tanto, los polipéptidos formados contienen metionina en su extremo N-terminal.

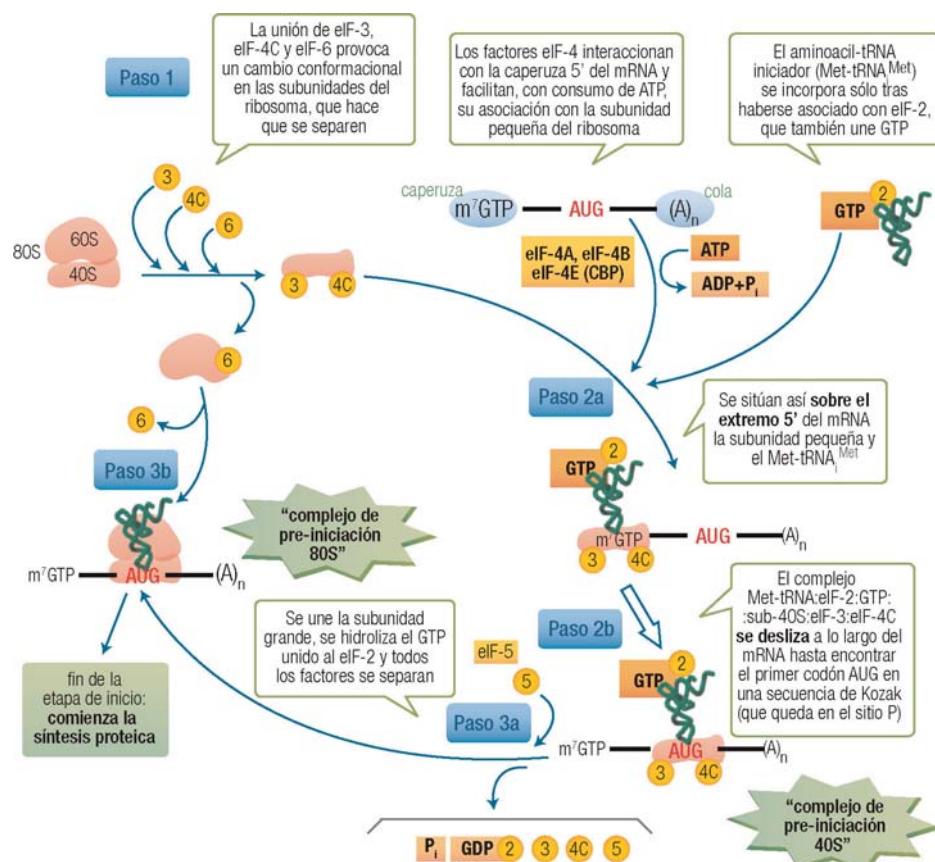
21.3.3 Pasos sucesivos en la fase de inicio

En prácticamente todos estos pasos, así como en las fases de elongación y terminación, intervienen varias proteínas que reciben el nombre genérico de **factores proteicos** de iniciación, elongación o terminación. Los **factores de iniciación** se nombran con la raíz IF en procariotas, eIF en eucariotas. A partir de este momento, sólo se describe la traducción en eucariotas; para facilitar su seguimiento, se establecerá una división en 9 pasos sucesivos: 3 en la iniciación, 3 en la elongación y 3 en la terminación.

En resumen, el inicio de la traducción requiere la unión del tRNA al codón de inicio AUG en el mRNA, en el entorno del sitio P o peptidilo del ribosoma. No intervienen en esta fase el codón anejo (cadena abajo) al AUG ni los sitios A o aminoacilo y E o de salida (*exit*) del ribosoma (pág. 337).

21.3.3.1 Disociación del ribosoma (paso 1)

Debe indicarse que las dos subunidades del ribosoma (40S y 60S en eucariotas), sintetizadas en el núcleo y exportadas a través de los poros nucleares al citosol, se ensamblan en éste, de forma reversible, formando el ribosoma 80S (véase pág. 76 para detalles de estructura y composición). Para su intervención en la traducción, el ribosoma debe disociarse en sus dos subunidades para permitir, en las subetapas siguientes, la unión de tRNA y mRNA y el inicio de la síntesis.



21.3.3.2 Unión a la subunidad menor del ribosoma: complejo de preiniciación (paso 2)

A continuación se incorporan a la subunidad menor del ribosoma el aminoacil-tRNA y el mRNA, en uno u otro orden. Se suele decir que se forma primero un “complejo binario de preiniciación” y después uno “ternario”. En ambos casos intervienen nuevos factores de iniciación.

El mRNA se une a la subunidad menor del ribosoma a través de su caperuza 5', que es reconocida por uno de los factores, eIF-4E, denominado también por ello CBP (*cap-binding protein*, o proteína ligante de la caperuza). En esta unión y en el posterior desplazamiento intervienen, además, los factores eIF-4A y eIF-4B y se produce la hidrólisis de ATP. La subunidad menor del ribosoma, con los factores asociados, se desplaza cadena abajo del mRNA, hasta encontrar un codón AUG. Éste suele ser el más próximo al extremo 5', pero debe estar incluido en una secuencia de Kozak; de no ser así, continuará el avance hasta encontrar otro. El codón de inicio AUG queda ubicado en el sitio P de la subunidad menor. Los factores proteicos implicados en esta etapa posiblemente actúan también desplegando la molécula de mRNA a su forma completamente extendida.

El metionil-tRNA_i^{Met}, por su parte, se incorpora después de formar un complejo con el factor de iniciación eIF-2 y GTP. Este factor reconoce exclusivamente al tRNA iniciador. (Más adelante se verá que todos los aminoacil-tRNA que participan en la elongación lo hacen también asociados con un factor proteico que une GTP, pág. 337.)

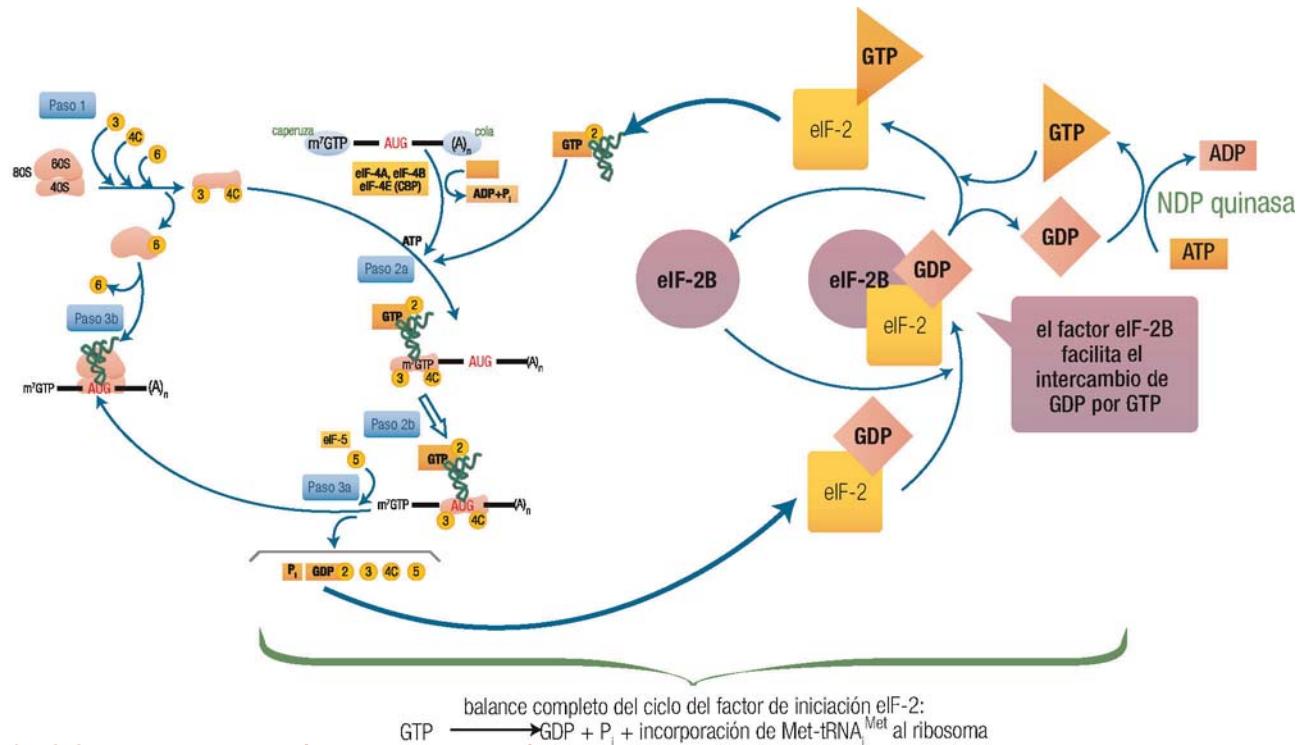
21.3.3.3 Unión de la subunidad mayor: complejo de iniciación (paso 3)

Para que pueda comenzar la síntesis de un polipéptido, el complejo de preiniciación 40S debe unirse con la subunidad mayor del ribosoma, 60S.

En primer lugar, la asociación de un nuevo factor eIF-5 activa la capacidad GTPasa de eIF-2. La forma eIF-2:GDP resultante no retiene la afinidad por el Met-tRNA, con lo que se disocian todos los factores y queda un complejo ternario de iniciación 40S libre de factores proteicos. Éste puede ya reasociarse con la subunidad mayor del ribosoma, liberada del factor eIF-6 que la mantenía separada, formando el definitivo “complejo ternario de iniciación 80S”. Con esta asociación de las dos subunidades del ribosoma, y la presencia del tRNA y el mRNA correctamente ubicados, este complejo queda dispuesto para la elongación.

21.3.3.4 Reutilización de los factores de iniciación

De los factores eIF implicados en el proceso anterior, todos se recuperan intactos y pueden volver a participar en el inicio de una nueva cadena polipeptídica, a excepción del factor eIF-2, que comenzó en su forma GTP y ha terminado con GDP. La regeneración de ese GTP, que evidentemente requiere la energía de hidrólisis de una molécula de ATP, no se hace de forma directa, sino por un intercambio de nucleótidos mediado a su vez por otro factor proteico, eIF-2B. Por su acción y estructura, éste pertenece a la familia proteica de factores intercambiadores de nucleótidos de guanina (GEF, *guanine exchange factors*)



21.3.3.5 Balance energético de la iniciación

Como se acaba de ver, dos pasos de la fase de inicio requieren un consumo energético. En primer lugar (paso 2), el reconocimiento de la caperuza y el desplazamiento hasta encontrar el codón de inicio, y más tarde (paso 3) la regeneración del GTP hidrolizado por eIF-2. Nos referiremos a éstos como *segundo y tercer gastos energéticos* de la traducción, de cara al balance global (pág. 343).

21.4 ELONGACIÓN O ALARGAMIENTO DE LA CADENA PEPTÍDICA

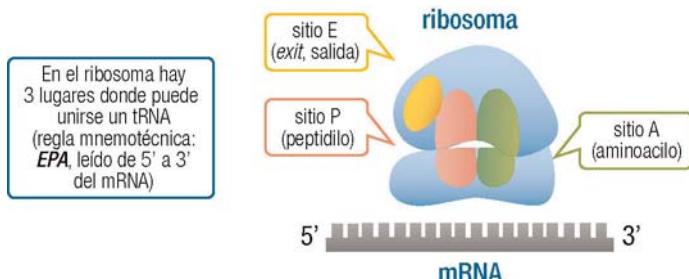
Esta fase incluye desde la adición del segundo aminoácido hasta el último. Se trata de un proceso cíclico que se repite tantas veces como aminoácidos se incorporen. Cada ciclo consta de 3 pasos: ubicación del nuevo aa-tRNA en el sitio A del ribosoma, formación del enlace peptídico y translocación del peptidil-tRNA al sitio P. En ellos intervienen tres factores proteicos citoplasmáticos o factores de elongación: eEF-1A, eEF-1B y eEF-2.

Debe resaltarse que la etapa de elongación no posee la capacidad de diferenciar el aminoácido que se incorpora, sino que es específica únicamente con respecto al anticodón del aa-tRNA; la incorporación de un aminoácido correcto viene determinada por la actuación previa de la aminoacil-tRNA sintetasa durante la etapa de activación.



Web 21.3. Estrategia experimental para discernir la especificidad entre codones, tRNA y aminoácido.

21.4.1 Ubicación del aminoacil-tRNA en el sitio A (paso 4)



21.14



Web 21.4. Asociación de los tRNA en los sitios E, P y A de la estructura tridimensional del ribosoma.

Tras incorporarse el Met-tRNA_i al sitio P del ribosoma en el complejo de iniciación 80S, queda aún vacío el sitio A. En él se producirá la incorporación de cada nuevo aminoácido (en la primera vuelta del ciclo de elongación será el aminoácido n.º 2), por emparejamiento del anticodón de su aa-tRNA con el codón correspondiente del mRNA.

La entrada de todos los sucesivos aa-tRNA en el sitio A requiere que estén unidos al factor de elongación eEF-1A que, de forma similar a lo ya estudiado para el factor de iniciación eIF-2, une GTP y tiene actividad GTPasa latente. (La estructura de los factores eIF-2, eEF-1A y eEF-2 es similar entre sí y a la de otras proteínas G, lo que sugiere su procedencia de una proteína ancestral común.) De hecho, la forma habitual *in vivo* de casi todos los aa-tRNA es como complejo ternario aa-tRNA:eEF-1A:GTP. La demanda de esta preparación de los aminoácidos para la síntesis de todas las proteínas de la célula puede ilustrarse comentando que en *E. coli* hay 10 veces más moléculas de EF-Tu que ribosomas, con lo cual es la proteína más abundante en la célula (EF-Tu es el factor bacteriano homólogo de eEF-1A; EF-Ts y EF-G son homólogos, respectivamente, de eEF-1B y eEF-2).

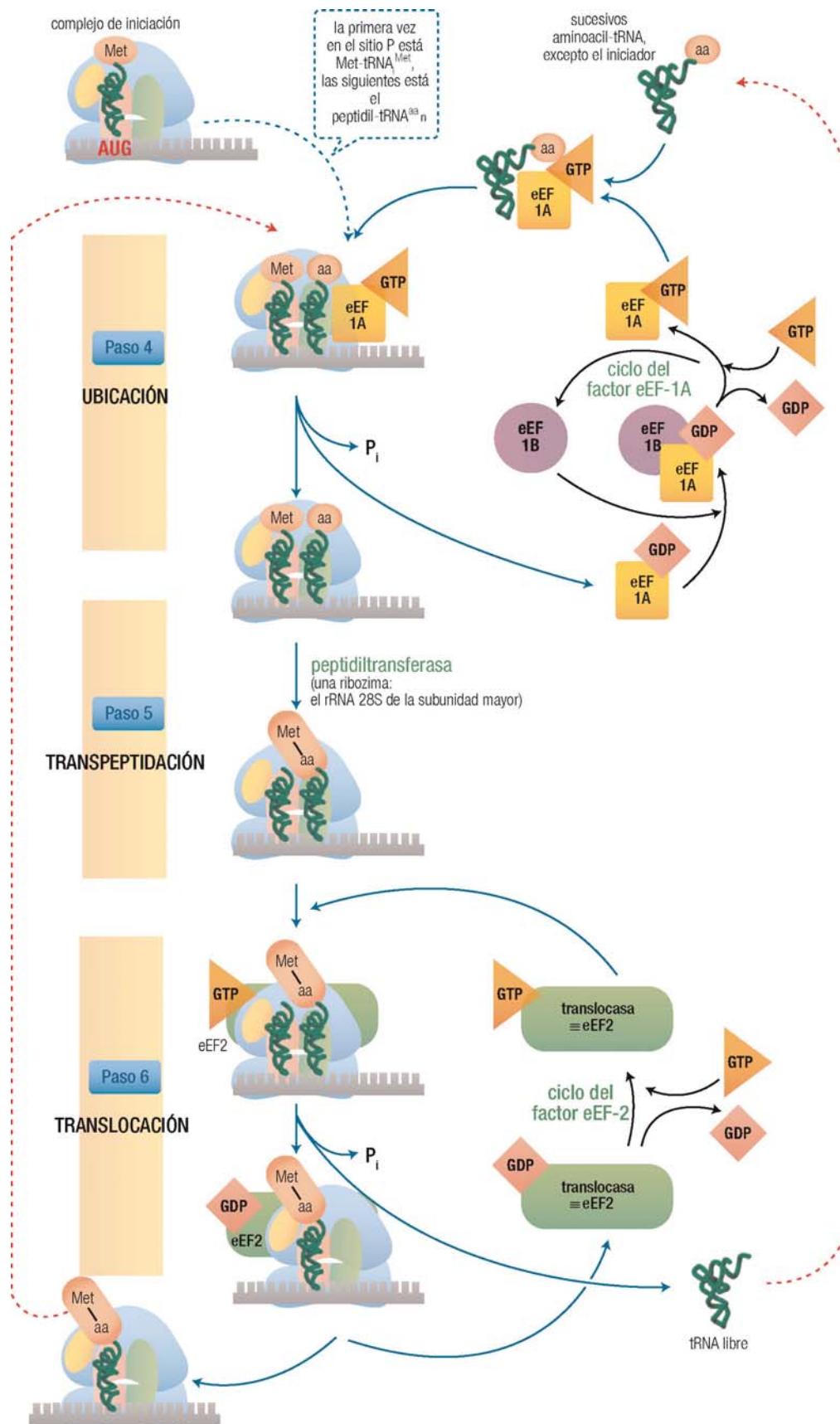
Son los complejos aa-tRNA:eEF-1A:GTP los que difunden al sitio A vacío del ribosoma y reconocen en él sitios de unión adecuados. Si el emparejamiento entre anticodón y codón es correcto, el aminoacil-tRNA se ubica de forma estable en el ribosoma, lo que provoca cambios conformacionales que estimulan la actividad GTPasa de eEF-1A. La forma eEF-1A:GDP no posee afinidad por el ribosoma, liberándose y quedando así el aa-tRNA unido al sitio A, dispuesto para las etapas siguientes.



Web 21.5. Análisis de los complejos portadores de los aminoacil-tRNA hacia el ribosoma y de su reutilización.

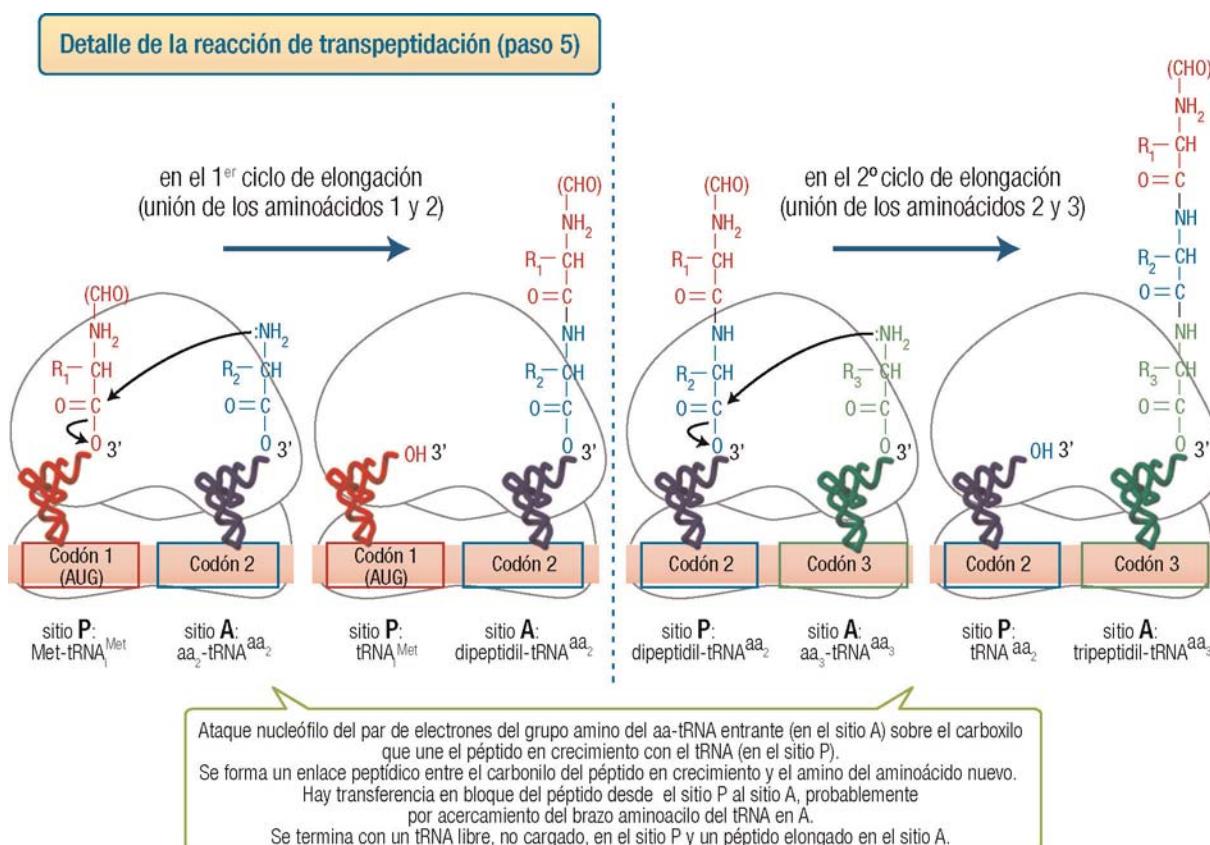
De forma análoga a lo que ocurre con el factor de iniciación eIF-2 (pág. 336), se regenera la forma eEF-1A:GTP por intercambio de nucleótidos mediado por un nuevo factor intercambiador de nucleótidos de guanina, el factor de elongación eEF-1B; se produce, pues, el **cuarto gasto energético** de la traducción (una molécula de ATP por cada aminoácido incorporado). eEF-1A queda así disponible para incorporar un nuevo aminoácido.

Debe resaltarse que eEF-1A interacciona con cualquier aa-tRNA excepto con el iniciador, Met-tRNA_i^{Met}, que se distingue de los restantes tRNA en la estructura secundaria del tallo aceptor. De este modo, durante la elongación, el tRNA iniciador no interviene en la incorporación de residuos internos de metioninas. Esta especificidad, junto con la ya descrita en la fase de inicio, mantiene la dualidad de los tRNA específicos para Met.



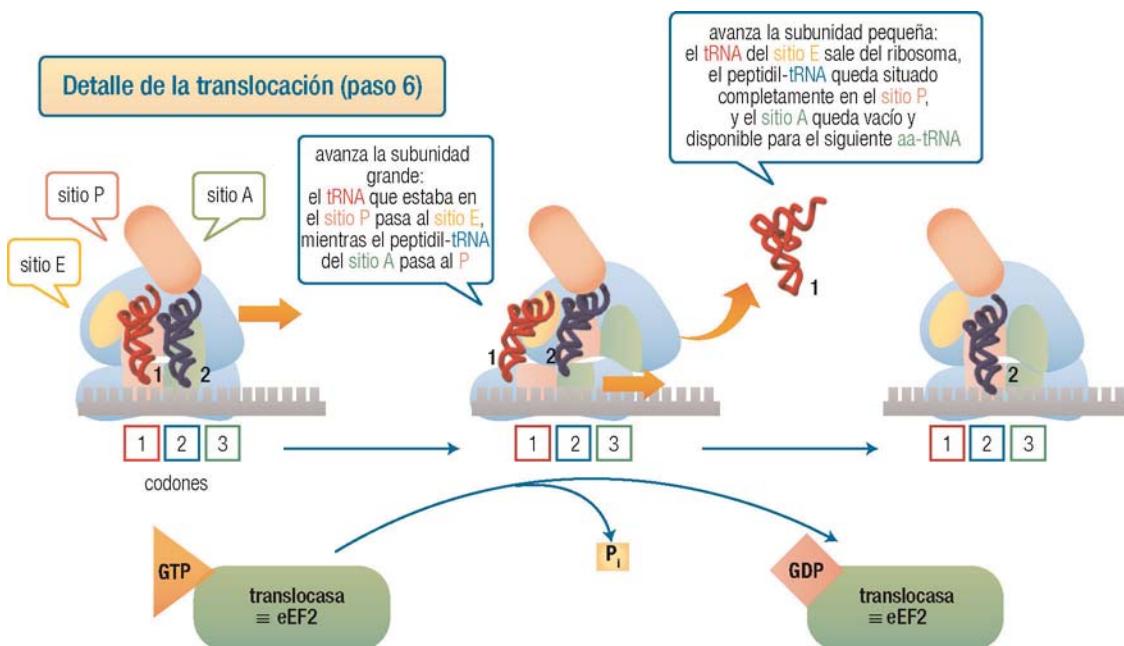
21.4.2 Transpeptidación: formación del enlace peptídico (paso 5)

Una vez situadas las moléculas de aa-tRNA y Met-tRNA_i^{Met} (en ciclos sucesivos, peptidil-tRNA) en los sitios A y P del ribosoma, respectivamente, se forma el enlace peptídico entre ambas. Esta reacción la cataliza una actividad enzimática **peptidiltransferasa** (o **transpeptidasa**) localizada en el rRNA 28S, componente de la subunidad grande 60S; se trata, pues, de una ribozima (pág. 74); aunque las proteínas ribosómicas contribuyen posiblemente a que sea activa, no participan en el centro catalítico.



21.4.3 Translocación: desplazamiento del ribosoma en un codón (paso 6)

A partir de su ubicación inicial sobre AUG el ribosoma completo se desplaza en sentido 5' → 3' del mRNA (cadena abajo), en una distancia de un solo codón cada vez. El responsable de este desplazamiento es el tercer factor proteico de elongación, eEF-2, conocido también como **translocasa**. Este factor interacciona principalmente con la subunidad mayor, originando cambios conformacionales en el ribosoma completo que tienen como resultado el desplazamiento. Al mismo tiempo, impide la unión de un nuevo aa-tRNA en el sitio A, evitando que se pase al siguiente ciclo antes de tiempo.



En esta etapa tiene lugar el *quinto gasto energético*, pues la energía requerida para cada translocación se obtiene de la hidrólisis de GTP gracias a una actividad GTPasa que reside en el propio factor eEF-2. Tanto en este caso como en los anteriores, la hidrólisis de GTP está acoplada a cambios conformativos en la maquinaria de traducción, que conducen al desplazamiento del ribosoma y a la salida del factor eEF-2 (cuya forma eEF-2:GTP deberá regenerarse en un ciclo de intercambio de nucleótidos de guanina, al igual que se ha visto para eIF-2 y eEF-1A).

21.4.4 Repetición del ciclo de elongación (pasos 4, 5 y 6)

Al terminar la primera translocación se ha completado la primera vuelta o ciclo de elongación, llegando a una situación idéntica a la existente al comienzo: sitio A libre y sitio P ocupado por un tRNA. La diferencia estriba en que el Met-tRNA_i del sitio P ha sido sustituido por un dipeptidil-tRNA. La elongación continúa de forma cíclica (figura en pág. 338) hasta que se añade el último aminoácido. La llegada sucesiva al sitio A (situado sobre los codones 2, 3, 4, etc.) de otros aa-tRNA permite la unión de los siguientes aminoácidos en sucesivos ciclos de elongación. Al acabar cada uno de ellos, el péptido en crecimiento siempre queda unido al tRNA del último aminoácido incorporado. Los enlaces peptídicos se forman siempre sobre el grupo NH₂ del último aminoácido incorporado y, por tanto, la cadena polipeptídica crece desde su extremo amino-terminal hacia el carboxilo-terminal.

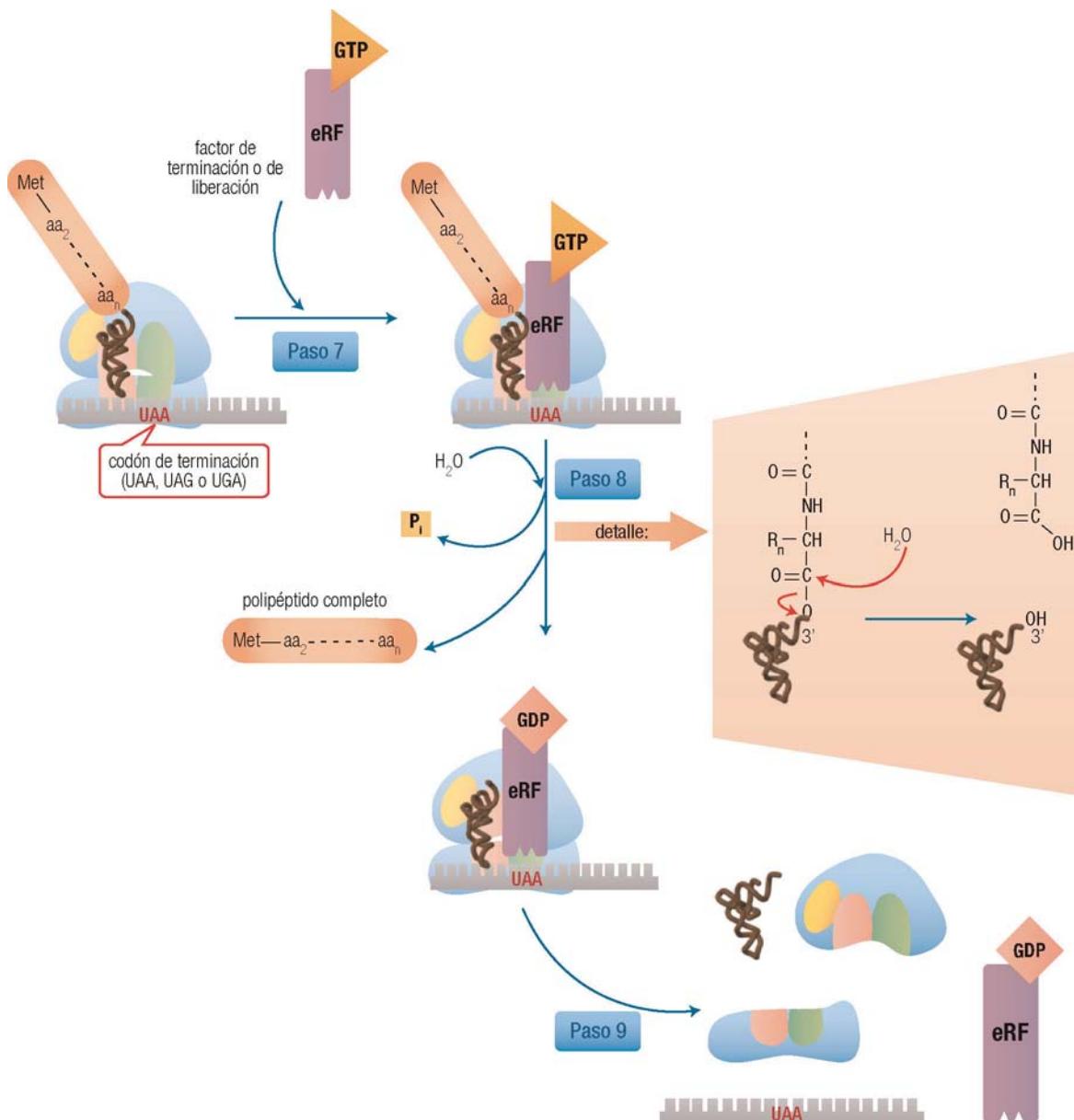
21.5 TERMINACIÓN

Esta etapa comprende los pasos necesarios para liberar el polipéptido ya completo de su unión al tRNA del sitio P y, al mismo tiempo, disociar el ribosoma del mRNA. Esto sólo puede ocurrir en respuesta a la presencia en el sitio A de uno de los 3 codones de terminación (UAA, UAG o UGA en el caso de mRNA codificado por DNA nuclear, y UAA, UAG, AGA o AGG en el caso de mRNA codificado por DNA mitocondrial) (pág. 321).

En eucariotas interviene en el proceso un único factor proteico de terminación o de liberación, eRF (*releasing factor*), en contraste con procariotas donde existen tres (RF1, RF2 y RF3).

21.5.1 Unión del factor de liberación (paso 7)

No existe ningún tRNA que reconozca los codones de paro, por lo que cuando el ribosoma se transloca sobre uno de ellos no puede progresar la elongación. En su lugar, el factor eRF se une al ribosoma en el sitio A, frente al codón de terminación.



21.18



Web 21.6. Estructura del factor de terminación y analogía con la del tRNA.

21.5.2 Hidrólisis del peptidil-tRNA (paso 8)

La unión de eRF al ribosoma altera la actividad peptidiltransferasa de tal forma que ésta emplea como agente nucleófilo el agua, en lugar del grupo amino de un aminoácido (pág. 339). Como consecuencia, se hidroliza el enlace éster del peptidil-tRNA, liberando la cadena polipeptídica. En este proceso, además, se hidroliza una molécula de GTP asociada al factor eRF, lo que supone el **sexto gasto energético** de la traducción.

21.5.3 Disociación (paso 9)

El tRNA no cargado sale del ribosoma. La forma eRF:GDP ya no posee afinidad por el ribosoma y también se libera. Finalmente, se separan las dos subunidades ribosómicas (en eucariotas, 80S → 40S + 60S), liberando el mRNA. Todos los componentes quedan así disponibles para iniciar la síntesis de una nueva proteína (excepto el eRF, que debe recambiar su GDP por GTP, probablemente por un ciclo similar a los estudiados para eIF-2, eEF-1A y eEF-2).

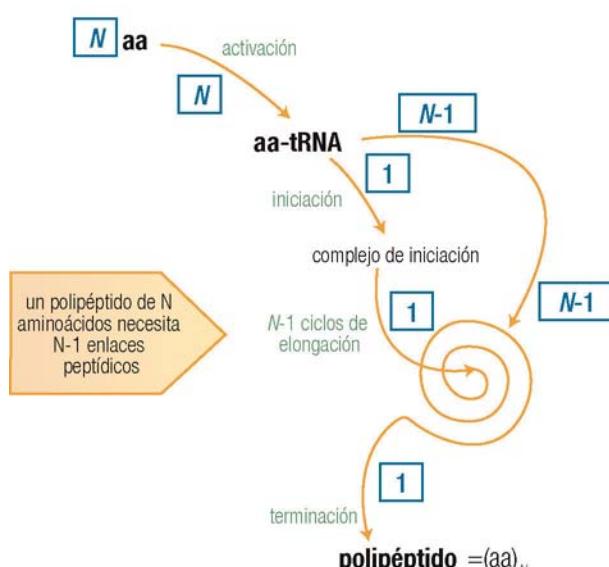
21.5.4 Reutilización del ribosoma

Para completar, debe señalarse que la utilización sucesiva del mismo mRNA para varias rondas de traducción no depende de una nueva asociación –más o menos aleatoria– del ribosoma y todos los factores sobre el codón de inicio. Existen evidencias de que el mRNA forma una estructura circular, cerrada, por aproximación de sus extremos 5' y 3', de modo que es mucho más fácil la reutilización de todos los componentes de la maquinaria de traducción al terminar en el extremo 3', reasociándose rápidamente sobre el extremo 5'. Esto es posible gracias a que hay proteínas que se unen tanto a ambos extremos como entre sí. Por ejemplo, eIF-4E reconoce la caperuza (se la denomina también CBP, *cap binding protein*, proteína ligante de la caperuza), mientras que una proteína PABI [*poly(A)-binding protein*, proteína ligante de poli(A)] se une a la cola de poliadenilato. Ambas proteínas interaccionan con un factor eIF-4G, circularizando así el mRNA. Ello permite que las subunidades ribosómicas disociadas en 3' se reutilicen con eficacia, asociándose rápidamente de nuevo sobre el extremo 5'. Este mecanismo aumenta la tasa de traducción, en tanto no intervenga alguna señal que conduzca, por ejemplo, a la degradación del mensajero.

21.6 ENERGÉTICA DE LA SÍNTESIS DE PROTEÍNAS

La síntesis proteica es un proceso energéticamente muy costoso, que utiliza una gran proporción de todo el equivalente en ATP disponible en la célula. Recapitulando todo lo detallado previamente en cada etapa y teniendo en cuenta la repetición cíclica de la elongación, puede calcularse el requerimiento energético mínimo. Posiblemente la formación o disociación de los numerosos complejos multimoleculares requiera algún aporte energético adicional, así como la corrección de pruebas y, en el caso de procariotas y orgánulos, la formilación.

Resumen de la traducción



Gasto	Fase	Paso	(pág.)	N.º de enlaces fosfato por paso	N.º total de enlaces fosfato por polipéptido de N aminoácidos
N.º 1	Activación	(Aminoacilación del tRNA)	333	2 por cada aa	2N
N.º 2	Iniciación	N.º 2 (incorporación del mRNA)	335	Indeterminado	Al menos 1
N.º 3	Iniciación	N.º 3 (complejo de iniciación)	336	1 por polipéptido	1
N.º 4	Elongación	N.º 4 (ubicación del aa-tRNA)	337	1 por aa añadido	N-1
N.º 5	Elongación	N.º 6 (translocación)	340	1 por aa añadido	N-1
N.º 6	Terminación	N.º 8 (hidrólisis)	341	1 por polipéptido	1
Total:					4N + 1 (al menos)

Esta hidrólisis de cuatro enlaces fosfoanhídrido por cada aminoácido representa un gran cambio en energía libre. Concretamente, $-30,5 \times 4 = -122 \text{ kJ/mol}$ de aminoácido, muy superior a la energía del enlace peptídico (energía libre estándar de hidrólisis de unos -21 kJ/mol). Aunque pueda parecerlo, esto no es un despilfarro energético, sino que la mayor parte de la energía extra (-101 kJ/mol) sirve para compensar la pérdida de entropía que tiene lugar durante la síntesis proteica. Este descenso de entropía se debe principalmente al ordenamiento de los aminoácidos en el polipéptido. También se pierde entropía cuando un aminoácido se une a un tRNA particular y cuando un aa-tRNA se asocia a un codón específico. Así, el aparente exceso energético se emplea en conseguir la casi perfecta fidelidad en la traducción biológica del mensaje genético del mRNA a la secuencia de aminoácidos de las proteínas.

21.7 INHIBIDORES DE LA TRADUCCIÓN

Numerosos y diversos compuestos actúan impidiendo o dificultando la traducción. Su estudio es interesante por varios motivos:

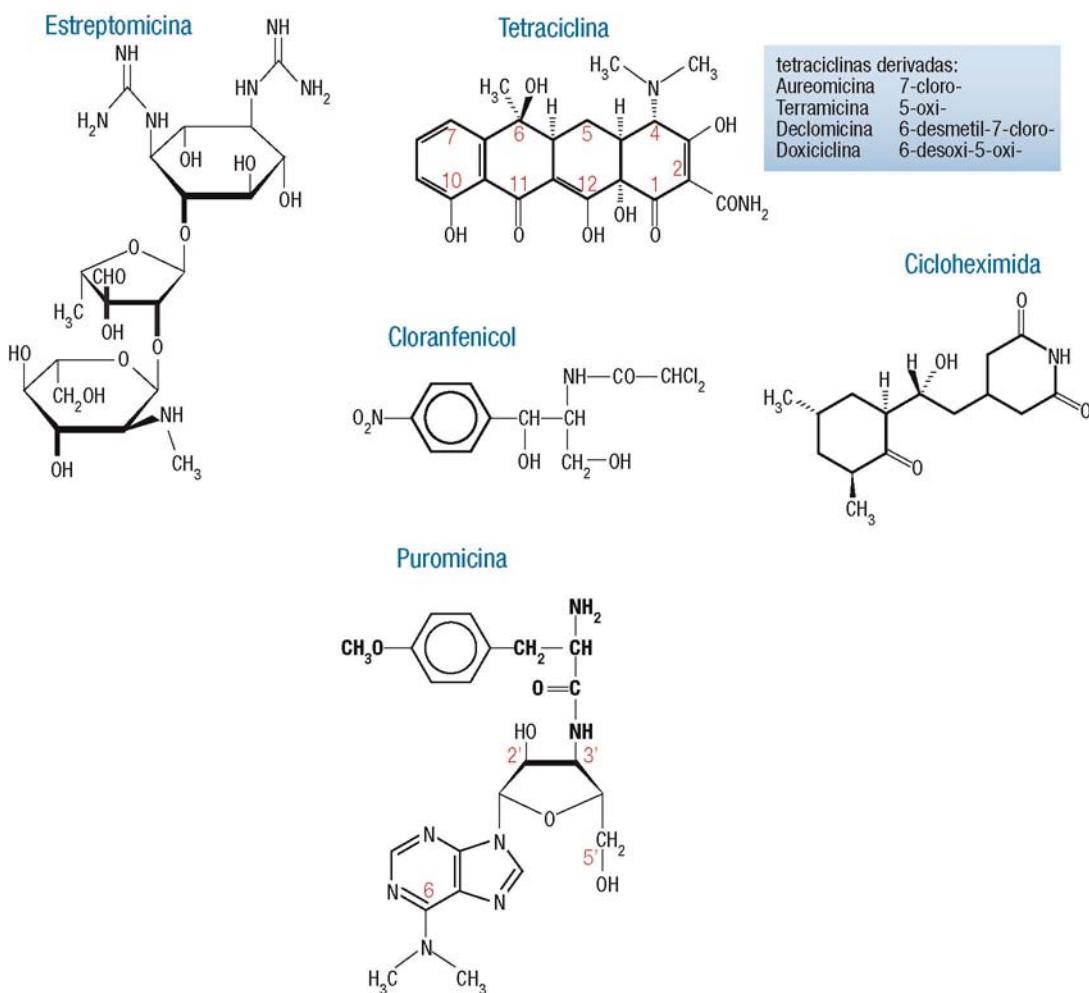
Desde el punto de vista práctico, porque muchos de ellos se pueden emplear como antibióticos, en especial cuando ejercen su efecto selectivamente sobre células procarióticas (bacterias), afectando poco o nada a las eucarióticas. Precisamente el estudio de sus mecanismos de acción ha contribuido a lo largo de muchos años al esclarecimiento de los detalles moleculares del proceso de traducción. A este respecto, debe mencionarse que el uso indiscriminado (sin control médico) de los antibióticos ha creado una resistencia bacteriana de gran trascendencia tanto con respecto a la flora bacteriana intestinal como a la terapia de las infecciones; estos aspectos deben abordarse de forma especializada, en tratados de microbiología.

Desde un punto de vista teórico, es interesante por el papel central de la traducción en la vida celular. Son muchos los antibióticos y toxinas que bloquean o interfieren en ciertos pasos de la síntesis proteica, de forma más o menos específica, en procariotas o en eucariotas. Esta especificidad es la mejor prueba de la diferencia entre la maquinaria de traducción en ambos tipos de organismos.

El efecto inhibidor puede ejercerse en distintas etapas del proceso de traducción:

Etapa afectada	Ejemplos	Efecto sobre procariotas o eucariotas	Detalles de la acción
Activación de los aminoácidos	Mupiroicina (o ácido pseudomónico)	P	Inhibe competitivamente la Ile-tRNA sintetasa, evitando la incorporación de isoleucina y deteniendo la síntesis proteica
Iniciación (paso 2, complejo de preiniciación)	Estreptomicina (un aminoglicósido)	P	Se fija de modo irreversible a la subunidad menor 30S, por interacción con varias de sus proteínas y con el rRNA 16S. Distorsiona la entrada de fMet-tRNA iniciador y también produce errores de lectura del mRNA durante la elongación, al interferir en el emparejamiento codón/anticodón

Etapa afectada	Ejemplos	Efecto sobre procariotas o eucariotas	Detalles de la acción
Iniciación (paso 2, complejo de preiniciación)	Pactamicina	E	Impide la ubicación del Met-tRNA iniciador en la subunidad menor 40S
	Showdomicina	E	Impide la formación del complejo Met-tRNA _i :eIF-2:GTP
	Interferón	E	Induce la expresión de una proteína quinasa que fosforila a eIF-2, inactivándolo (de forma similar al HCl, pág. 352)
Iniciación (paso 3, complejo de iniciación)	Eritromicina	P	Se une a un sitio específico en el rRNA 23S de la subunidad mayor 50S. Impide la asociación con la subunidad menor; otra acción posible es interfiriendo en la translocación
Elongación (paso 4, ubicación)	Tetraciclinas	P > E	Se unen a la subunidad menor, interfiriendo en la fijación del aa-tRNA al sitio A
	Kirromicina	P	Bloquea la desociación de GDP del factor EF-Tu (homólogo bacteriano del eEF-1A eucariótico), lo que evita su salida del ribosoma
	Ricina (glicoproteína de origen vegetal; con 2 cadenas, A y B; la primera es la toxina, 66 kDa)	E	Se une a proteínas de la subunidad mayor 60S, bloqueando la unión de aa-tRNA:eEF-1A:GTP, y posiblemente también de eEF-2 (transpeptidación)
Elongación (paso 5, transpeptidación)	Cloranfenicol	P y mitocondrial	Se une selectivamente a una proteína de la subunidad mayor 50S, interfiriendo en la interacción del aa-tRNA con el centro activo de la peptidiltransferasa, como inhibidor competitivo
	Cicloheximida	E	Similarmente al cloranfenicol, inhibe la actividad peptidiltransferasa, pero sólo en el ribosoma eucariótico (subunidad mayor 60S)
	Puromicina	P y E	Análogo estructural del aa-tRNA, forma enlace con el péptido provocando su terminación prematura
Elongación (paso 6, translocación)	Aminoglicósidos distintos de la estreptomicina (por ejemplo: gentamicina, kanamicina, neomicina)	P	La mayoría interactúan con el rRNA 16S de la subunidad mayor 50S, impidiendo la fijación del factor de elongación G (EF-G, homólogo bacteriano del eEF-2 eucariótico). Los mecanismos de acción son diversos
	Esparsomicina	E	Inhibe la translocación
	Ácido fusídico	P y E	Se une de forma muy estable al complejo EF-G:GDP:ribosoma, impidiendo la liberación de EF-G (homólogo bacteriano del eEF-2). También inhibe a eEF-2
	Toxina diftérica (proenzima proteica de 62 kDa; la toxina es un fragmento de 21 kDa)	E	Cataliza una reacción que inactiva eEF-2 de forma irreversible (ADP-ribosilación a costa de NAD, sobre un residuo His modificado de eEF-2)
Terminación	No se conoce ninguno		



21.20



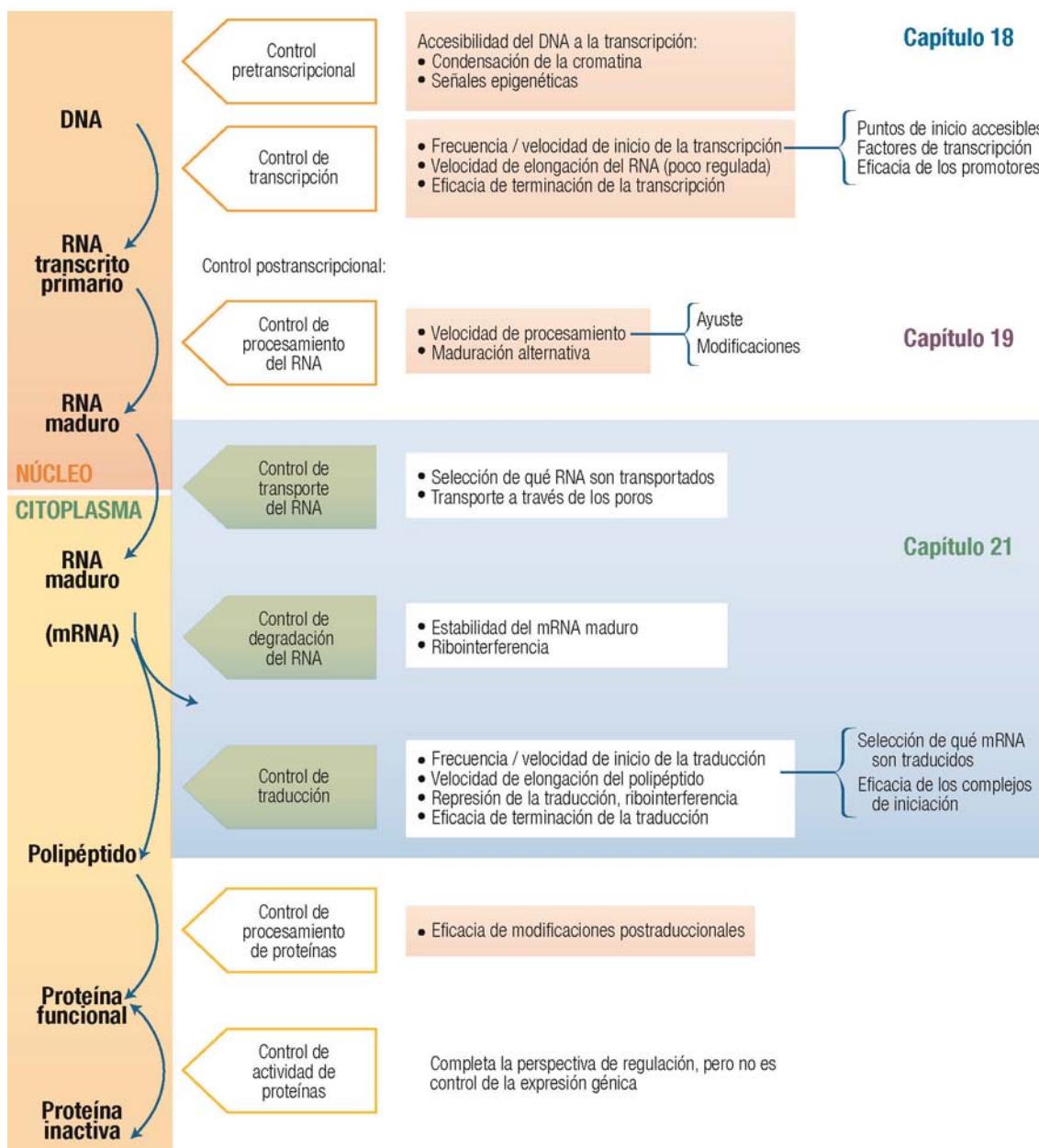
Web 21.7. Estructura tridimensional de antibióticos inhibidores de la traducción y fundamento molecular de su acción.

La selectividad de la mayoría de antibióticos y toxinas sobre organismos procarióticos o eucarióticos es posible gracias a las diferencias en sus ribosomas así como, en algunos casos (por ejemplo, tetraciclinas), a la incapacidad de la molécula para atravesar la membrana celular de organismos superiores.

21.8 REGULACIÓN DE LA SÍNTESIS PROTEICA

La traducción se regula en las células a través de varios mecanismos. Su velocidad depende en cierto modo de la secuencia del mRNA molde; por ejemplo, un mRNA con muchos codones raros se traduce más lentamente (y, por tanto, con menos frecuencia) que otro con codones más comunes. Además, la velocidad de comienzo varía con la secuencia que rodea al codón de inicio. Se sabe que una unión fuerte del mRNA a su sitio en el ribosoma produce un inicio más eficiente (al menos en procariotas).

NIVELES DE REGULACIÓN DE LA EXPRESIÓN GÉNICA EN EUCARIOTAS



21.8.1 Control del transporte de mRNA

Las tres modificaciones esenciales que experimentan los transcritos primarios durante su maduración (formación de la caperuza 5', poliadenilación en 3' y reacciones de ajuste) parecen ser cruciales para que el RNA pueda pasar al citoplasma e intervenir en la traducción a proteínas. Este transporte a través de los poros nucleares está perfectamente regulado, y puede afectar también al control de la expresión.

Es importante resaltar la relevancia (aunque el conocimiento de su mecanismo es aún inicial) del transporte del mRNA maduro desde el núcleo al citoplasma, necesario para que esté disponible para la traducción. Una vez completado el procesamiento postranscripcional del transcríto primario (Capítulo 19) el mRNA maduro se transporta a través de los poros nucleares e inicia la traducción en el ribosoma citoplasmático en un plazo de entre 1 y 5 minutos.

Existen mecanismos celulares de control de la traducción en respuesta a las necesidades bajo diferentes condiciones ambientales. Surge así el término **regulación traduccional** para referirse a estos casos de modulación de la frecuencia de traducción por factores extrínsecos. El estudio de este control tiene un interés especial en relación con el desarrollo animal; en este sentido se ha estudiado intensamente, por ejemplo, en *Drosophila*.

Desde un punto de vista básico, sólo es preciso referirse a una situación general, la de la célula de un individuo adulto. Entre las proteínas específicas que se unen al mRNA recién sintetizado y ayudan a su transporte fuera del núcleo se encuentra un factor proteico multimérico, que incluye al eIF-4E o **proteína ligante de la caperuza (CBP)**, ya estudiada por su papel en el inicio de la traducción (pág. 335).

21.8.2 Vida media y degradación del mRNA

Cada molécula de mRNA se emplea muchas veces como molde para la traducción, dando lugar a numerosas copias del polipéptido que codifica. Como consecuencia, el tiempo de permanencia de una molécula de mRNA en la célula es esencial para determinar la cantidad de proteína producida. Ese tiempo, o vida media de la molécula, depende tanto de la velocidad con que se sintetiza, madura y sale al citoplasma el mRNA, como de la rapidez con que se degrada en este compartimento subcelular.

La degradación intracelular del mRNA es un proceso que tiene lugar continuamente, a mayor o menor velocidad, como consecuencia de la actividad de ribonucleasas específicas (v. web 15.1). Éstas pueden ser tanto endonucleasas como exonucleasas. Frente a estas últimas es importante la acción protectora de la poliadenilación, pues debe digerirse por completo la cola de poli(A) antes de que se alcancen las secuencias codificantes del mRNA. De hecho, las moléculas de mRNA tienen una vida media tanto más prolongada cuanto mayor es la longitud de su cola. De igual manera, la caperuza 5' ejerce un efecto protector esencial, gracias a su estructura singular con enlace 5'-5'-trifosfodiéster (pág. 300). En eucariotas, la vida media del mRNA es unas 10 veces mayor que en procariotas, lo que se debe probablemente al procesamiento postranscripcional, que retarda su degradación; esto es esencial dada la necesidad de que se transporte desde el núcleo al citoplasma (pág. 263) antes de que pueda comenzar a utilizarse en la traducción.

Como consecuencia de estas variables, existen aparentes contrastes entre la abundancia de los distintos tipos de RNA en una célula y la velocidad con la que se sintetizan. Así, por ejemplo, el rRNA es mayoritario y el mRNA minoritario, sin embargo la velocidad de síntesis es superior para el mRNA (y además variable).

Como datos representativos del significado de este control pueden mencionarse los correspondientes, a grandes rasgos, a la diferente estabilidad de las moléculas de RNA (expresada por su vida media). Existe, en general, una relación directa entre este parámetro y el grado de maduración sufrido. La estabilidad también aumenta con el contenido de estructuras secundaria y terciaria de la molécula de RNA.

- Los tRNA y rRNA de eucariotas y procariotas son extremadamente estables (horas o días), pues sufren un gran procesamiento postranscripcional y presentan abundante emparejamiento de bases y plegamiento de la cadena. Los rRNA están, además, fuertemente protegidos por su asociación con las proteínas ribosómicas.
- Los mRNA de eucariotas tienen una estabilidad intermedia (horas), porque experimentan una extensa maduración pero casi no presentan estructura secundaria. Como casos extremos, pueden citarse el mRNA de c-fos, con una vida de 10 a 30 min, y los de globina y ovoalbúmina, que alcanzan 24 h.
- Los mRNA de procariotas son los menos estables (minutos), ya que no sufren maduración alguna y apenas tienen estructura secundaria. Los mRNA de *E. coli* se degradan rápidamente una vez formados (se calcula que las nucleasas degradan en 3 min la mitad del mRNA sintetizado en ese tiempo).

Existen también mecanismos de control de la vida media que dependen de proteínas que se unen al mRNA, protegiéndolo de la degradación, y otros a cargo de complejos ribonucleoproteicos que degradan específicamente un mRNA determinado (**ribointerferencia**). Estos mecanismos se comentan en el apartado siguiente.

21.8.3 Control de la traducción

Una vez que el mRNA ha alcanzado el citoplasma, el control de su traducción se ejerce esencialmente en la iniciación, haciendo de ésta la etapa limitante. Principalmente se regula la formación del complejo de preiniciación 80S y la actividad de los factores de inicio eIF-2 y eIF-4E; también a través de la influencia de estructuras secundarias en el mRNA, por ejemplo regiones intramoleculares bicatenarias, que pueden enmascarar los sitios de unión del ribosoma y el codón de inicio. Se estudian a continuación con cierto detalle algunos mecanismos de regulación de la capacidad y frecuencia de utilización de los mensajeros para la traducción.

21.8.3.1 Ribointerruptores

Se ha denominado **ribointerruptor** o **riborregulador** (en inglés, *riboswitch*; pág. 69) una parte de ciertas moléculas de RNA mensajero que regula su propia expresión, bien en el final de su transcripción o bien en su maduración postranscripcional o en su traducción. Una característica clave de su acción es que reconoce y une específicamente una molécula pequeña, por ejemplo un metabolito. El cambio conformacional inducido en el RNA al unírsele el ligando desencadena una acción sobre la expresión génica de ese mismo RNA. Actúa así como un sensor que sirve a la célula para adaptarse a circunstancias del entorno. La mayoría de ribointerruptores se han detectado en bacterias, pero hay al menos uno que actúa en eucariotas (algunos hongos y plantas).

Cabe destacar que esta regulación, a diferencia de las demás estudiadas, no está ejercida por proteínas. A continuación se estudia otro mecanismo de regulación de la expresión también mediado por moléculas de RNA, la ribointerferencia.

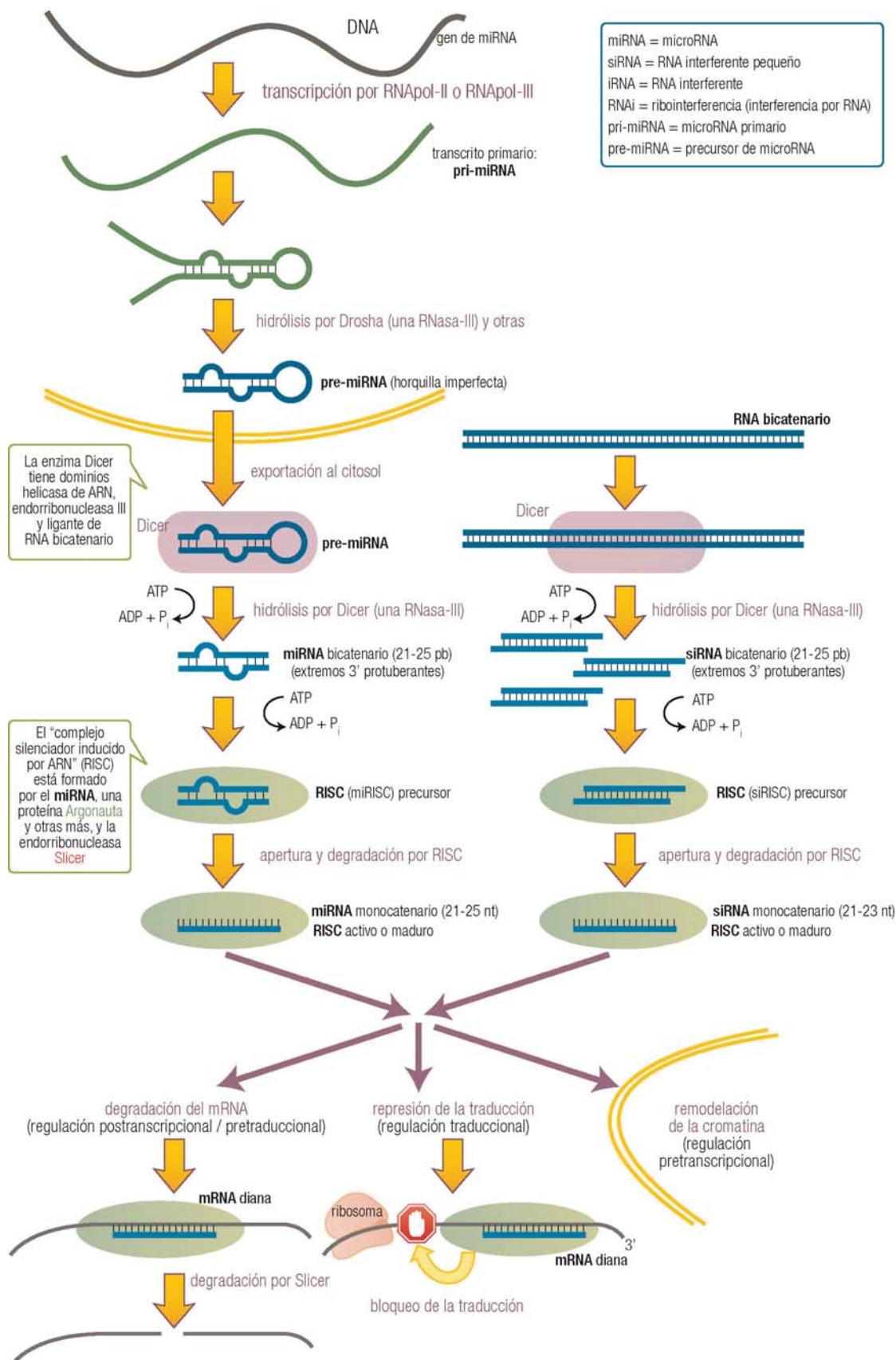
21.8.3.2 Ribointerferencia

Los **miRNA** y los **siRNA** (pág. 69) se descubrieron por separado, en diferentes organismos y con funciones aparentemente distintas, pero cada vez hay más pruebas de que ambos tipos de RNA pequeño pueden actuar en muchas especies y conducir a acciones similares. Todas ellas se engloban bajo el término **ribointerferencia** (o interferencia por RNA), que incluye tanto la acción biológica de regulación como la técnica experimental que aprovecha estos procesos. De forma breve, la función de los RNA interferentes se manifiesta en dos tipos de resultado: el bloqueo de la traducción de RNA mensajeros específicos o bien la degradación selectiva de dichos mensajeros. Ambos casos se pueden denominar también silenciamiento génico postranscripcional. Es importante resaltar que, detalles aparte, la especificidad de su acción sobre la expresión de un único producto génico siempre se basa en el emparejamiento de secuencias complementarias entre el RNA interferente y alguna región del mensajero.

La ruta más común en la que participan los miRNA es su generación endógena por transcripción de un gen propio de la célula, seguida de su maduración en núcleo y citosol (pág. 309) y su integración en el complejo RISC. Se han identificado un gran número de genes de miRNA, tanto en el ser humano como en otras especies, que están implicados como un mecanismo más en la regulación de la expresión génica. En cuanto a los siRNA, éstos proceden de RNA bicatenarios, que, dicho de forma simplificada, la célula percibe como peligrosos (introducidos por un virus, transposones, genes sobreexpresados, etc.). También se pueden generar siRNA a partir de moléculas exógenas de RNA bicatenario introducidas en la célula con propósitos experimentales, para conseguir el silenciamiento de un gen (técnica de ribointerferencia).

La acción del complejo RISC, según los casos, puede ser la represión de la traducción o la degradación del mRNA, pero siempre de forma específica sobre un mensajero que posea secuencia complementaria a la del miRNA o siRNA. La definición de cuál de los dos mecanismos se desencadena es una cuestión aún no del todo resuelta. Uno de los determinantes es el grado de complementariedad del miRNA o siRNA con el mRNA diana: si el emparejamiento es perfecto, es más probable la degradación, mientras que si es parcial se produce con más probabilidad el bloqueo de la traducción.

La **represión de la traducción** tiene lugar mediante la fijación de RISC sobre la región 3' no traducida (es decir, posterior al codón de paro) del mRNA complementario. Dicho de otro modo, esa región 3'UTR contiene la secuencia complementaria al miRNA o siRNA. El resultado, a pesar de que se une en la zona final del mensajero, es el bloqueo de la maquinaria de traducción (ribosoma y factores proteicos asociados), quedando el RNA secuestrado y no disponible para más rondas de traducción.



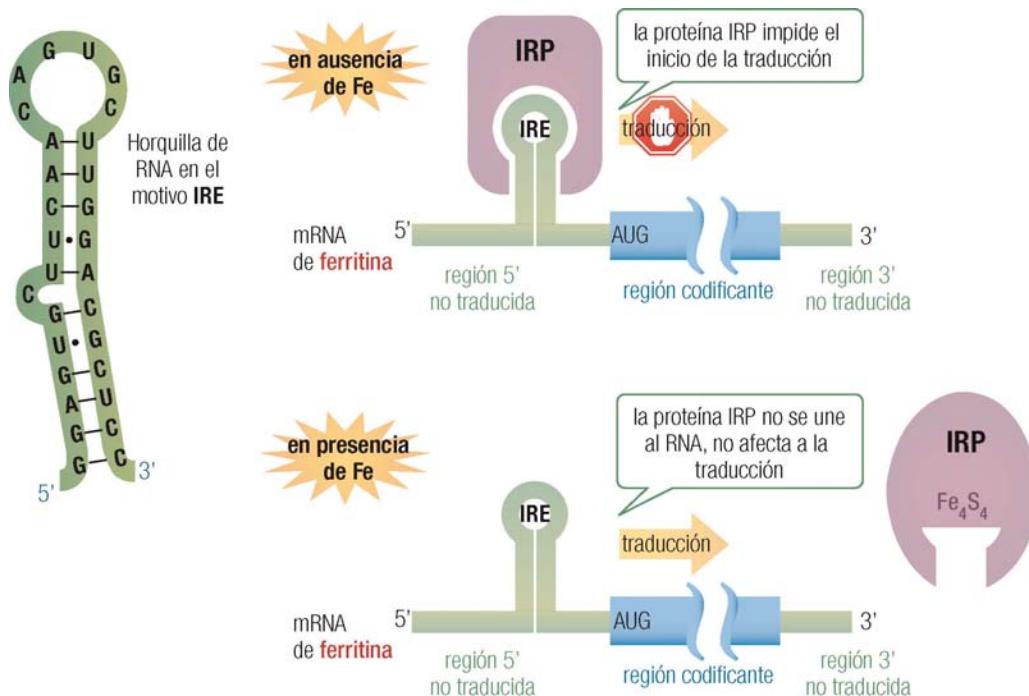
En el caso de la **degradación del mRNA** –sólo si es complementario al RNA pequeño integrado en el complejo RISC–, la responsable de la hidrólisis del mensajero es la enzima Slicer (del inglés *slice*, que significa cortar en rebanadas o rodajas), uno de los componentes proteicos de RISC. Obviamente, mientras este mecanismo esté activo cualquier molécula de ese mRNA concreto que se sintetice será degradada y sus posibilidades de traducirse quedarán reducidas.

Un tercer mecanismo que parece actuar en algunos casos (en este caso, dentro del núcleo) es la interacción del RNA pequeño con la región promotora de un gen, reclutando en esa zona a proteínas que **remodelan la cromatina** y, en consecuencia, reprimiendo la transcripción.

21.8.3.3 Control de la accesibilidad del mensajero

Este ejemplo demuestra un mecanismo de control para la síntesis de ferritina ejercido mediante bloqueo del acceso de la maquinaria de traducción a su mRNA. Ésta es una proteína cuya función es almacenar hierro, como reserva para la biosíntesis de hemoglobina, mioglobina, citocromos y otras muchas proteínas, y al mismo tiempo como protección contra la toxicidad de los iones hierro libres. Su síntesis se controla a través de las proteínas IRP, que se unen a la región 5' no traducida (5'UTR) del mRNA impidiendo el inicio de la traducción.

Las IRP o **proteínas reguladoras del hierro** (*iron-regulatory proteins*) poseen la capacidad de unirse al mRNA, con distinta afinidad en función de la concentración intracelular de ese metal. Por ejemplo, una de ellas, la IRP-1, cuando el hierro es abundante lo incorpora en un complejo hierro-azufre Fe_4S_4 y no interacciona con el mRNA. Por el contrario, cuando el hierro es escaso en la célula la IRP-1 no presenta el complejo Fe-S y adopta una conformación que le permite unirse con elevada afinidad a una horquilla de RNA con bucle (pág. 54). Ésta se conoce como **motivo IRE** o **elemento de respuesta al hierro** (*iron-response element*), por lo que las IRP se llaman también **proteínas ligantes de IRE** (*IRE-BP, IRE-binding proteins*). Existe un motivo IRE en la región 5' no traducida (pág. 264) del mRNA de ferritina, al que se unen las IRP impidiendo el inicio de la traducción y reduciendo así la concentración intracelular de ferritina. La célula se adapta de esta forma a la escasez de hierro, evitando la síntesis de una proteína que no es necesaria. Por el contrario, en condiciones de exceso de hierro, se sintetiza libremente la ferritina, para poder fijarlo.



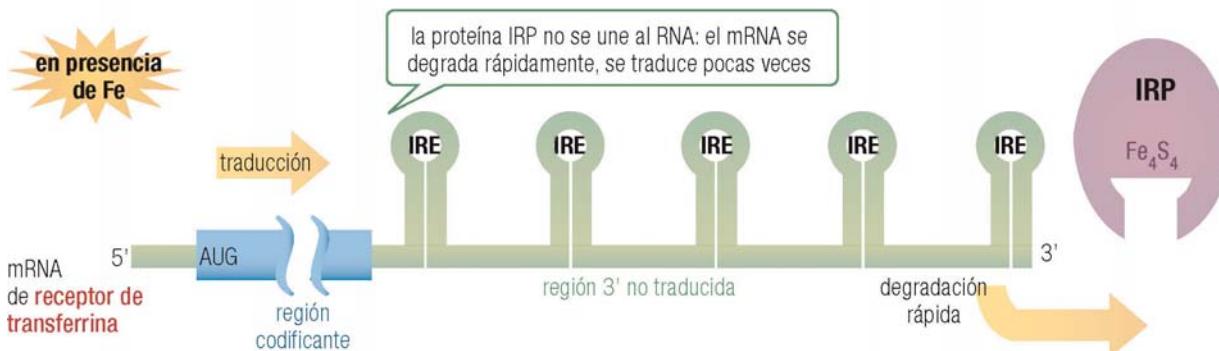
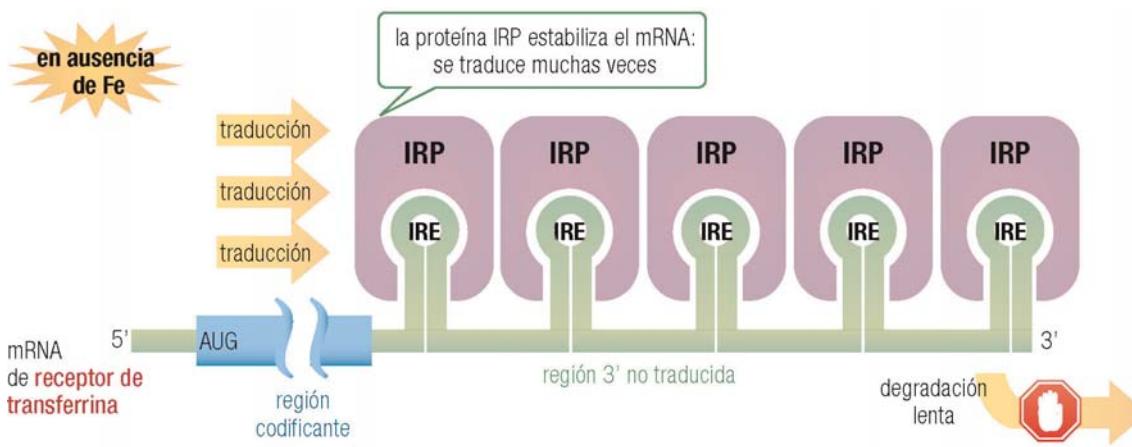
21.23

www

Web 21.8. Estructura tridimensional: reconocimiento del motivo IRE en el RNA por la proteína IRP.

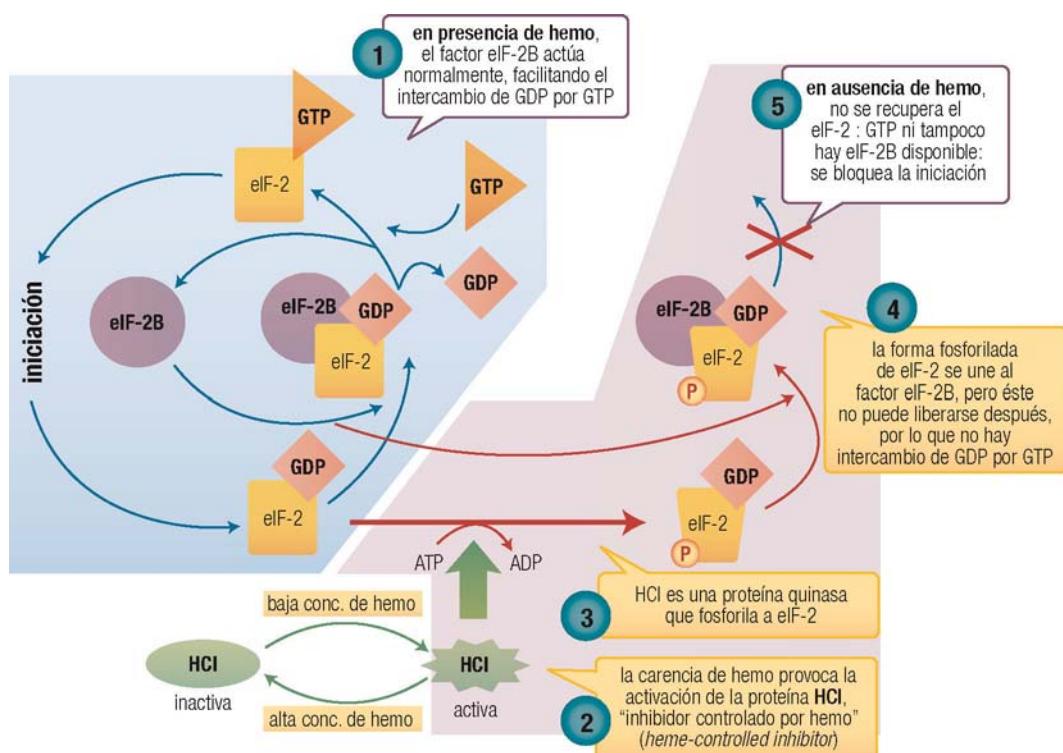
21.8.3.4 Control de la estabilidad del mensajero

Existe un ejemplo muy similar de regulación, coordinada con la anterior, pero en este caso controlando la vida media del mRNA y, por tanto, su disponibilidad para la traducción. Se trata del receptor de transferrina, una proteína de membrana necesaria para que las células captén el hierro. En condiciones de escasez de hierro la proteína IRP se une también a este mRNA, que posee 5 motivos IRE, pero al estar éstos en su región 3' no traducida (3'UTR) el efecto es de protección del mRNA frente a la degradación. En consecuencia, se sintetizan más moléculas del receptor de transferrina a partir de un mismo mRNA, capacitando a la célula para captar hierro con más eficacia. En la situación contraria (abundancia de hierro) IRP no es activa para unirse al mRNA, que se degrada más rápidamente, reduciendo la síntesis del receptor, ya no tan necesario. Ambos mecanismos de regulación, la traducción del mRNA de ferritina y la degradación del mRNA del receptor de transferrina, actúan al unísono con un mismo objetivo, la adaptación de la célula a la disponibilidad de hierro.



21.8.3.5 Regulación del complejo de iniciación

La síntesis de hemoglobina requiere cantidades estequiométricas de hemo y de globinas. La síntesis de las globinas tiene lugar en precursores eritroideos y en reticulocitos (precursores inmediatos de los eritrocitos), donde se expresan de forma activa los genes correspondientes. La velocidad de traducción del mRNA de globina está regulada en función de la concentración de hemo en la célula, de modo que si ésta desciende, disminuye la síntesis. La célula evita así la producción inútil de la proteína para la que no hay grupo prostético (hemo) disponible. La regulación tiene lugar por inhibición del inicio de la traducción, interfiriendo en la reutilización del factor eIF-2 (pág. 336).



21.25

Debe destacarse que resulta inhibida toda nueva incorporación de una Met inicial; puede continuar la síntesis de las cadenas ya iniciadas, pero no puede iniciarse una nueva. Por otra parte, a diferencia de los ejemplos anteriores (ferritina y receptor de transferrina), el control no afecta de manera específica al mRNA de una proteína, sino a la traducción de todos los mRNA de la célula. Este aparente perjuicio potencial para la célula está justificado porque se trata de células muy especializadas, que sintetizan casi exclusivamente globina.



Modificaciones postraduccionales: distribución, maduración, plegamiento y degradación de proteínas

22.1 INTRODUCCIÓN	353
22.2 DISTRIBUCIÓN DE LAS PROTEÍNAS	
A SU DESTINO	354
22.2.1 Compartimentos implicados	354
22.2.2 Tránsito de proteínas en el citosol y entre orgánulos	356
22.2.3 Distribución de proteínas secretadas	357
22.2.3.1 Características del péptido señal	357
22.2.3.2 Mecanismo de entrada al retículo endoplásmico	357
22.3 MADURACIÓN O PROCESAMIENTO DEL POLIPÉPTIDO NACIENTE	358
22.3.1 Maduración por modificación de aminoácidos	358
22.3.1.1 Unión de sustituyentes a los aminoácidos	358
a) Acetilación	359
b) Carboxilación	359
c) Fosforilación	359
d) Hidroxilación	360
e) Metilación	361
22.3.1.2 Modificación con lípidos	361
a) Acilación (ácidos grasos)	361
b) Prenilación (terpenos)	362
22.3.1.3 Incorporación de glúcidos: glicosilación	363
a) Características de las glicoproteínas	363
b) Tipos de glicosilación	364
c) Biosíntesis de las glicoproteínas	367
22.4 PLEGAMIENTO DE PROTEÍNAS	372
22.4.1 Mecanismo del plegamiento de las proteínas	372
22.4.2 Proteínas implicadas en el plegamiento: carabinas moleculares	373
22.4.2.1 Tipos de carabinas	374
a) Carabinas Hsp70 y Hsp40	374
b) Carabinas Hsp60 y Hsp10	375
c) Otras carabinas	375
22.4.2.2 Función de las carabinas	376
22.5 DEGRADACIÓN DE LAS PROTEÍNAS	376
22.5.1 Degradación lisosómica	377
22.5.2 Degradación citosólica	377
22.5.2.1 Ubiquitina	377
22.5.2.2 Señales desencadenantes de la ubiquitinación	378
a) Residuo amino-terminal	378
b) Secuencias PEST	379
22.5.2.3 Degradación por proteasas y proteasoma	379

22.1 INTRODUCCIÓN

La síntesis proteica no termina cuando el polipéptido completado se separa del complejo de traducción, sino que en todos los organismos se requieren procesos adicionales para que el *mensaje* lineal o unidimensional del polipéptido (secuencia de aminoácidos, determinada por la secuencia de nucleótidos del mRNA) se transforme en un mensaje tridimensional, la estructura nativa de la proteína, responsable de su función. Estos procesos, recogidos de forma conjunta bajo el término **maduración o modificación postraduccional**, constituyen el colofón de la síntesis proteica, que a veces se considera la quinta fase de la traducción, después de activación del aminoácido, iniciación, elongación y terminación.

La maduración postraduccional es esencial para la funcionalidad y estabilidad de las proteínas, y más en la célula eucariótica. Cada compartimento subcelular requiere proteínas diferentes, que no se sintetizan en él, para sus variadas funciones: actividad catalítica en todas las vías metabólicas, papel estructural en la membrana celular, transporte o almacenamiento de moléculas e iones, función motora muscular, transmisión de señales intra- e intercelular, regulación de la expresión génica, etc. Incluso las mitocondrias y los cloroplastos, que poseen un genoma propio, necesitan multitud de proteínas que están codificadas en el genoma nuclear, por lo que una vez sintetizadas en el citosol deben llegar hasta estos orgánulos.

Para una descripción representativa de la modificación postraduccional se seguirá un esquema relativamente común, dividido en los siguientes apartados. (Esta separación, sin embargo, es puramente formal y didáctica, pues muchos de estos procesos tienen lugar de forma simultánea en la célula.)

1. Distribución de las proteínas hacia diferentes localizaciones, subcelulares o extracelulares, para el ejercicio de su función.
2. Maduración o procesamiento del polipéptido. Comprende principalmente modificaciones químicas de los residuos de aminoácido y eliminación de fragmentos del polipéptido.
3. Plegamiento correcto del polipéptido hasta alcanzar su conformación nativa. Requiere, en muchos casos, una maduración previa por modificación química, con lo que plegamiento y maduración están íntimamente ligados.
4. Una vez que la proteína ha llegado a su destino y cumplido su misión, debe también describirse su catabolismo, su degradación hasta aminoácidos.

Con todos estos procesos se cierra de manera definitiva el “ciclo biológico de la proteína”, es decir:



22.1

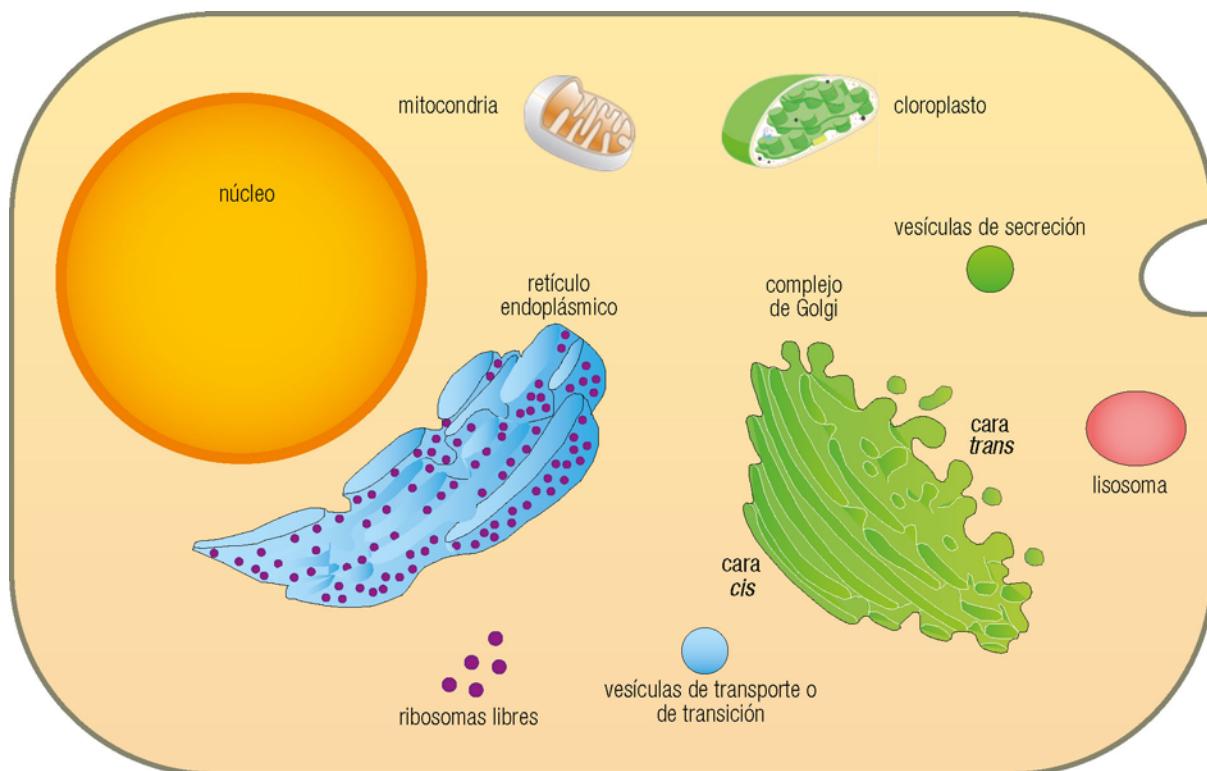
22.2 DISTRIBUCIÓN DE LAS PROTEÍNAS A SU DESTINO

Se conoce como tráfico, tránsito, destino, localización, topogénesis e incluso “clasificación” de proteínas la ruta que siguen estas moléculas en la célula eucariótica hasta alcanzar su localización final (intracelular o extracelular), donde pueden ejercer su función, y los mecanismos moleculares responsables de que sigan esa ruta. Este proceso de distribución se desarrolla de forma estrechamente conectada con la maduración y con el plegamiento del polipéptido.

22.2.1 Compartimentos implicados

El tránsito de proteínas tiene lugar con referencia al citosol, a la membrana plasmática y a los compartimentos cerrados presentes en aquél: los orgánulos delimitados por membranas, constituyentes del denominado “sistema endomembranoso”:

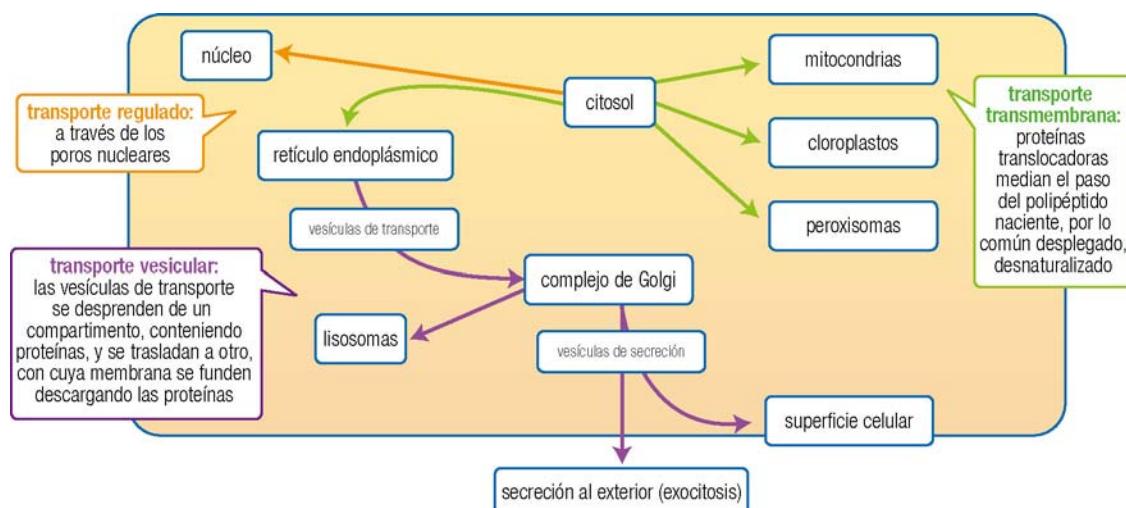
- **Retículo endoplásmico (RE).** Es un sistema intracelular muy ramificado, replegado e interconectado, de estructura variada (canales, esferas, tubos, cisternas aplanadas), que delimita un espacio interior (luz o lumen) aislado del citosol. En su mayor parte, el RE posee ribosomas asociados a su cara citosólica (formando el RE rugoso), donde se realiza la síntesis de las proteínas integrales de membrana (celular, del RE, del Golgi y de lisosomas), de las destinadas al lumen de RE, Golgi o lisosomas y de las destinadas a secreción. Las proteínas con destino y función en el citosol, el núcleo u otros orgánulos (mitocondrias, cloroplastos, peroxisomas, etc.) se sintetizan en ribosomas no asociados al RE. Las proteínas que se sintetizan en el RE se incorporan al lumen de éste de forma cotraduccional (es decir, simultáneamente con su síntesis), para allí sufrir los primeros pasos de la maduración. La salida posterior del RE se hace bajo la forma de pequeñas *vesículas*, que se dirigen al complejo de Golgi, a lisosomas o a la membrana externa.



- **Complejo de Golgi** (o aparato de Golgi). Se trata de un compartimento situado entre el RE y la membrana plasmática en una disposición (cara “cis” frente al RE y cara “trans” frente a la membrana) que facilita la recepción de las vesículas procedentes del RE cargadas de proteínas. Está formado por un apilamiento de sacos o bolsas aplanadas, con forma de disco (*dictiosomas*), no interconectadas entre sí. Los polipéptidos sufren aquí nuevos procesamientos hasta generar la proteína madura, ya dotada de las señales moleculares necesarias para definir su destino final. Las proteínas se reenvían al citosol dentro de nuevas vesículas que se dirigen bien hacia otros orgánulos, bien hacia la membrana externa; (durante este transporte muchas proteínas sufren también su procesamiento por escisión proteolítica limitada. En el caso de las proteínas que la célula secreta al exterior, se habla de *vesículas de secreción*, que finalmente se funden con la membrana celular liberando su contenido al exterior por *exocitosis*. Además de esta función, el complejo de Golgi también genera los lisosomas, cargados de enzimas hidrolíticas que no sólo degradan moléculas intracelulares, sino también partículas extracelulares incorporadas por *endocitosis* (mediante la fusión de endosomas con lisosomas).
- **Mitocondrias.** En ellas tiene lugar específicamente la fosforilación oxidativa, con producción de ATP para todas las necesidades de energía de la célula. Sólo unas pocas de las proteínas necesarias están codificadas en el genoma mitocondrial y se sintetizan en ribosomas mitocondriales, por lo que se requiere la entrada a través de la doble membrana mitocondrial de gran cantidad de proteínas sintetizadas en el citosol.
- **Peroxisomas.** Son orgánulos de tamaño pequeño delimitados por una membrana única, ricos en enzimas que intervienen en reacciones de oxidación de moléculas tóxicas.
- **Otros compartimentos.** Quizá con menor trascendencia en lo que respecta a la distribución y la maduración, pero obviamente con importancia funcional y que también son destino específico de sus proteínas respectivas, deben citarse los cloroplastos, lisosomas, etc.

22.2.2 Tránsito de proteínas en el citosol y entre orgánulos

Todas las proteínas comienzan su síntesis en el citosol (excepto las codificadas en los genomas mitocondrial y cloroplástico). La clave del proceso por el que las proteínas se dirigen a su localización son **señales de clasificación** o **secuencias señal** que forman parte de la proteína naciente y dirigen la distribución de la molécula hacia distintos compartimentos de la célula. Las proteínas que no presentan secuencia señal permanecen en el citosol. Estas señales suelen ser regiones continuas de la secuencia de aminoácidos, entre 15 y 60 residuos, con frecuencia ubicadas cerca de los extremos de la cadena y casi siempre eliminadas por escisión proteolítica (pág. 369-370) una vez han cumplido su función. En otros casos, sin embargo, la señal es una “**región señal**”, una disposición tridimensional específica, dependiente del plegamiento de la proteína, a la que contribuyen residuos aminoácidos no contiguos en la estructura primaria.



22.3

El aspecto más destacable de la distribución intracelular –y el más comúnmente estudiado– es el que corresponde a las proteínas secretadas. Ésta tiene lugar a través de las membranas del RE y del complejo de Golgi. Este sistema de transporte implica la introducción de la cadena polipeptídica (sintetizada en la superficie del RE rugoso) hacia el lumen, a través de la membrana, la salida de la proteína completada en vesículas hacia el complejo de Golgi y, finalmente, su transporte en vesículas de secreción hasta la membrana plasmática y su secreción al exterior. Éste será el único ejemplo que se estudie en detalle, a continuación; no se abordarán otros aspectos concretos, de gran interés pero que escapan al carácter generalista de esta obra. Entre éstos, pueden citarse la integración de proteínas en la membrana de cualquier orgánulo, la distribución a orgánulos especiales como mitocondria, cloroplasto y peroxisomas, el transporte mediante vesículas desde el RE al complejo de Golgi y de éste a la membrana plasmática, la liberación final de su contenido por exocitosis al exterior celular o la importación de proteínas del medio circundante mediante endocitosis facilitada por receptor.

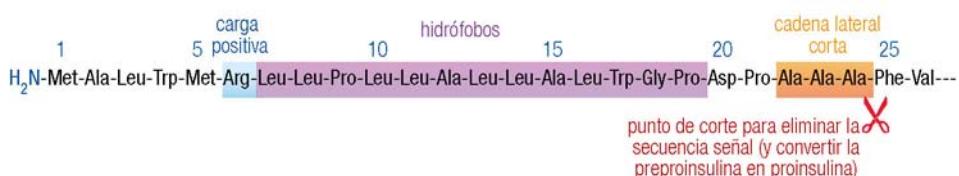
Como ejemplo de otro tipo de destino, cabe mencionar el transporte de proteínas desde el citosol hacia el núcleo para su incorporación a la cromatina o para su intervención en los procesos de replicación, transcripción y maduración postranscripcional. Estas proteínas deben pasar a través de los *poros nucleares* (estructuras muy complejas, formadas por hasta 100 proteínas diferentes, que atraviesan la membrana doble o *envoltura nuclear*). A través de los poros tiene lugar, asimismo, la distribución, en sentido inverso, de las moléculas de RNA y las subunidades ribosomales, respectivamente sintetizadas y ensambladas en el núcleo. Aunque el paso puede realizarse por libre difusión, parece existir una difusión selectiva, sin que se despliegue la estructura tridimensional (a diferencia de los otros tipos de tránsito a través de membranas). Las proteínas que se han de transportar al núcleo parecen poseer una secuencia señal (secuencia de localización nuclear) mediante la cual se unen a otras proteínas, “receptores de importación nuclear”, que dirigen la entrada a través de los poros, a costa de la energía liberada por la hidrólisis del GTP.

22.2.3 Distribución de proteínas secretadas

22.2.3.1 Características del péptido señal

La señal de localización para las proteínas que deben secretarse es habitualmente una secuencia peptídica corta en el extremo N-terminal del polipéptido recién sintetizado, que se denomina **secuencia señal o de etiquetado, etiqueta, péptido señal** o secuencia líder (por su posición al comienzo de la cadena). Consta de uno o más residuos con carga positiva (Lys o Arg), seguidos de entre 5 y 15 aminoácidos hidrófobos y de unos pocos residuos relativamente polares y de cadena corta, en especial Gly y Ala.

Ejemplo de péptido señal en la preproinsulina humana



22.4

Como se verá a continuación, es frecuente que el péptido señal se elimine, por la acción de una **proteasa de la señal** (o proteasa señal). En estos casos, las proteínas recién sintetizadas, cuya secuencia señal aún no se ha eliminado, se denominan **formas pre-proteína**, para diferenciarlas de la molécula que ya ha sufrido la proteólisis del péptido señal; en algunos casos ésta a su vez es una forma inactiva que precisa un procesamiento adicional por proteólisis, por lo que recibe el nombre de **pro-proteína**, y su precursor inicial es, pues, una **pre-pre-proteína** (pág. 370).

22.2.3.2 Mecanismo de entrada al retículo endoplasmático

Como se ha indicado, todas las proteínas cuyo destino es la membrana celular, la membrana de RE o Golgi, el interior de estos orgánulos o la secreción al exterior se sintetizan en ribosomas asociados al RE. Además, a poco de comenzar su síntesis la cadena polipeptídica se introduce al interior del RE, se **transloca** al lumen. La condición única e imprescindible para esta translocación es la presencia en la cadena polipeptídica de la secuencia señal. A continuación se estudia con detalle este proceso, concretándolo para el caso de proteínas secretadas.



Web 22.1. Entrada de la proteína al lumen del RE.

1. Primer proceso de unión: la secuencia señal se sintetiza pronto tras el comienzo de la traducción, por estar situada en el extremo amino del polipéptido naciente, y emerge del ribosoma cuando el polipéptido alcanza unos 70 aminoácidos de longitud. Entonces rápidamente la reconoce un complejo proteico citosólico llamado **partícula de reconocimiento de la señal (SRP)**, con una masa de 325 kDa).

La SRP es una ribonucleoproteína formada por 6 polipéptidos y una molécula de RNA. Es ésta, llamada RNA 7SL, la responsable directa del reconocimiento de la secuencia señal. Pertenece al grupo de los RNA citoplásicos pequeños (scRNA) y su secuencia está relacionada con la familia *Alu* de DNA repetitivo (pág. 118).

2. Segundo proceso de unión: la unión de la SRP provoca una detención temporal de la elongación del polipéptido en el ribosoma, hasta que el complejo anterior (mRNA:ribosoma:aa-tRNA:polipéptido naciente:SRP) contacta con la superficie citosólica del RE, donde la SRP es reconocida y se une a una proteína integral de la membrana del RE, denominada **proteína receptora de SRP (SRP-R)** o **proteína de anclaje**. Como resultado, todo el complejo queda fijado fuertemente al RE.

El receptor de SRP es una proteína heterodimérica con la capacidad de unir GTP y de catalizar su hidrólisis (actividad GTPasa). La forma con GTP unido es la que interacciona con la SRP, mientras que al hidrolizarlo a GDP pierde afinidad y se libera la SRP.

3. Tercer proceso de unión: el complejo ha quedado dispuesto sobre el RE de tal forma que se favorece su interacción con otra proteína integral de la membrana, situada en una posición vecinal a SRP-R. Esta nueva proteína recibe el nombre de **proteína receptora del ribosoma** o **riboforina** porque su misión es, precisamente, la de reconocer la subunidad mayor del ribosoma que forma parte del gran complejo y conducir el polipéptido naciente al interior del RE. Como consecuencia de este “anclaje” ocurre lo siguiente:
 - a) La subunidad α de SRP-R hidroliza el GTP unido, rindiendo GDP y P_i (gasto energético de un GTP por polipéptido).
 - b) Se disocia la SRP, quedando disponible en el citosol para volverse a emplear (“ciclo de la SRP”).
 - c) Al liberarse la SRP se elimina la inhibición de la traducción, es decir, se reanuda la síntesis proteica, ahora ya dirigida hacia el lumen del RE.
 - d) En la riboforina se abre un “canal de translocación” que permite la entrada en el lumen de la cadena polipeptídica creciente. De ahí los nombres alternativos de **translocasa** para la riboforina y de **complejo de translocación** para el complejo que se formó.
4. En la cara luminal del RE, asociada a la porción interior de las riboforinas del poro, se encuentra una nueva proteína integral de membrana llamada **peptidasa de la señal**, pues escinde el péptido señal, separándolo del resto del polipéptido. Una vez liberado, aquél se degrada hasta aminoácidos.
5. La finalización de la síntesis en el lumen del RE, con liberación de la proteína, tendrá lugar cuando el extremo carboxilo del polipéptido haya pasado a través de la membrana del RE. El ribosoma se disocia entonces del RE y se desorganiza todo el complejo de traducción para volver a reciclarse (“ciclo del ribosoma”).

22.3 MADURACIÓN O PROCESAMIENTO DEL POLIPÉPTIDO NACIENTE

Conceptualmente, éste es el componente más sencillo de la modificación postraduccional. Puede tener lugar tras haber finalizado la síntesis, una vez liberado el polipéptido del ribosoma o, más comúnmente, de forma simultánea con la traducción, es decir, la región de la molécula que se ha sintetizado antes (N-terminal) va siendo modificada mientras aún se está elongando el polipéptido por una región más próxima al extremo C-terminal. Esta segunda posibilidad recibe el nombre de **modificación cotraduccional**, aunque los procesos sufridos son los mismos en ambos casos. Se utilizan enzimas y cofactores específicos para una multitud de reacciones de modificación.



Web 22.2. Modificaciones cotraduccionales.

La maduración implica dos grandes tipos de modificaciones del polipéptido: la transformación de las cadenas laterales de algunos aminoácidos y la escisión proteolítica.

22.3. I Maduración por modificación de aminoácidos

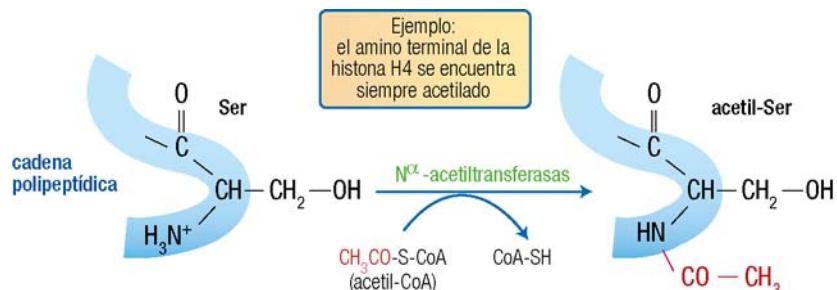
Diversos aminoácidos del polipéptido, siempre en posiciones específicas, sufren reacciones de modificación **química**, aunque catalizada enzimáticamente. Se describen a continuación algunas de las más comunes.

22.3. I. I Unión de sustituyentes a los aminoácidos

Es frecuente la adición de determinados grupos (como acetilo, carboxilo, fosfato, hidroxilo o metilo) a la cadena lateral de aminoácidos o a los extremos amino- o carboxi-terminales. Se han descrito más de un centenar de modificaciones en residuos aminoácidos; veamos algunos de los ejemplos más comunes.

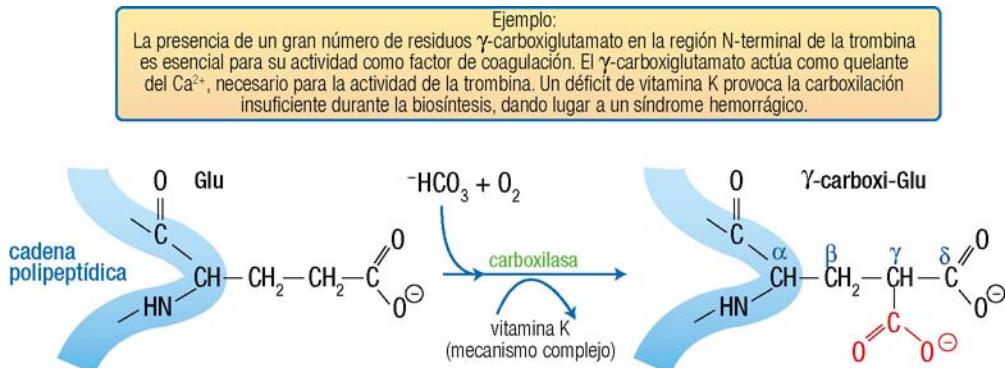
a) Acetilación

La adición de grupos acetilo tiene lugar en un 50% de las proteínas eucarióticas. Una posición frecuente es el grupo **amino terminal**, previa eliminación del residuo de metionina N-terminal (pág. 369). Uno de los efectos de esta modificación es una mayor resistencia a la degradación, al proteger el grupo amino. También se acetila el grupo amino en la **cadena lateral** de residuos de lisina en las histonas; esta modificación, que es reversible, contribuye a la regulación de la condensación de la cromatina (pág. 284).



b) Carboxilación

En algunas proteínas se añade un grupo carboxilo a la cadena lateral de un aminoácido; en concreto, en los factores de coagulación y en proteínas estructurales del hueso se carboxilan residuos de glutamato produciendo γ -carboxiglutamato (Gla). También sufre carboxilación el aspartato dando β -carboxiaspartato (Asa), encontrado en proteínas ribosomales de *E. coli*.



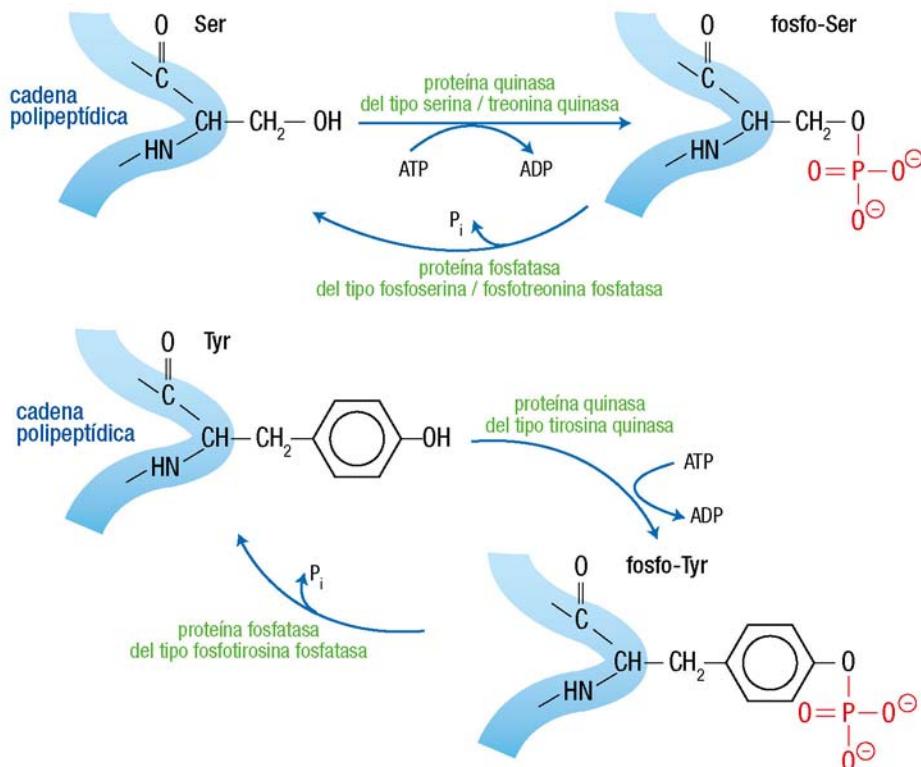
Web 22.3. Carboxilación del glutamato mediada por la vitamina K.

c) Fosforilación

Es quizás la modificación más frecuente, y actúa casi siempre de forma reversible (fosforilación y desfosforilación) (v. web 22.4). Ambas reacciones constituyen un mecanismo para la regulación de la actividad de multitud de proteínas, en particular las implicadas en rutas de transducción de señal y algunas enzimas (v. por ejemplo, pág. 464). Pertenece este mecanismo a la categoría de **regulación por modificación covalente**. En ambos casos, se trata de una modificación estrictamente postraduccional, que tiene lugar en el citosol después de que se han completado la síntesis y el plegamiento de la proteína.

La modificación afecta a los aminoácidos alcoholes, serina, treonina y tirosina, en su grupo -OH, y está catalizada por proteína quinasas, que hidrolizan el ATP transfiriendo su grupo fosfato γ al residuo aminoácido. Una de las consecuencias es un incremento de carga negativa en la proteína. La reacción opuesta, de desfosforilación, la catalizan proteína fosfatases.

Ejemplos:
La **glucógeno fosforilasa b** (forma inactiva) es fosforilada en el residuo Ser-14 por la **fosforilasa quinasa** dando la **glucógeno fosforilasa a** (activa), que libera glucosa-1-P del glucógeno. La forma **a** se transforma de nuevo en la **b** por la acción de la **fosforilasa fosfatasa**. La fosforilación de varios residuos de Ser en la **caseína** permite la fijación del Ca^{2+} , actuando así la leche como nutriente de P y Ca para el lactante.



22.7

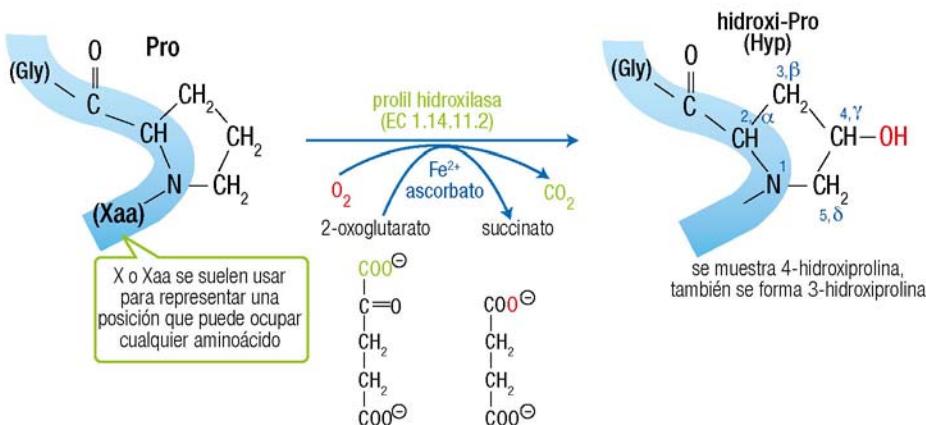


Web 22.4. Quinasas y fosfatases de proteínas.

d) Hidroxilación

Varias hidroxilasas presentes en el RE catalizan la incorporación de grupos -OH a algunas proteínas. Esta modificación se observa, por ejemplo, en residuos de prolina y lisina del colágeno, y resulta esencial para las correctas propiedades de esta proteína fibrosa.

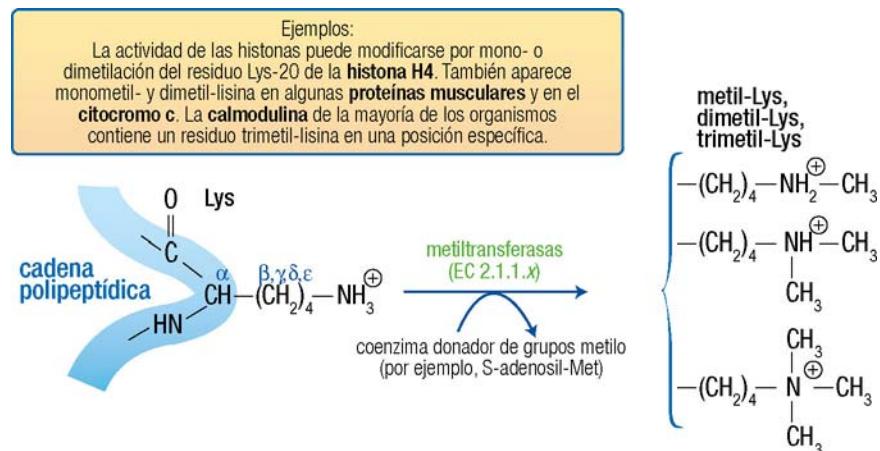
Ejemplo:
El **procolágeno** se modifica (como parte de su transformación en colágeno) por hidroxilación de alrededor de un 50% de sus residuos Pro, en la secuencia X-Pro-Gly. La enzima responsable es la **prolil hidroxilasa**, una dioxigenasa que utiliza oxígeno molecular y 2-oxoglutarato como cosustratos, y ascorbato y Fe^{2+} como cofactores. El escorbuto, enfermedad derivada de una insuficiente hidroxilación del colágeno, se produce por un déficit de vitamina C (ácido ascórbico).



22.8

e) Metilación

Consiste en la incorporación de metilo al grupo ε -amino de la cadena lateral de lisina, o bien al grupo γ -carboxilo del glutamato. La metilación de lisinas en las histonas es un importante mecanismo de regulación de la expresión génica, una de las marcas epigenéticas (pág. 285).



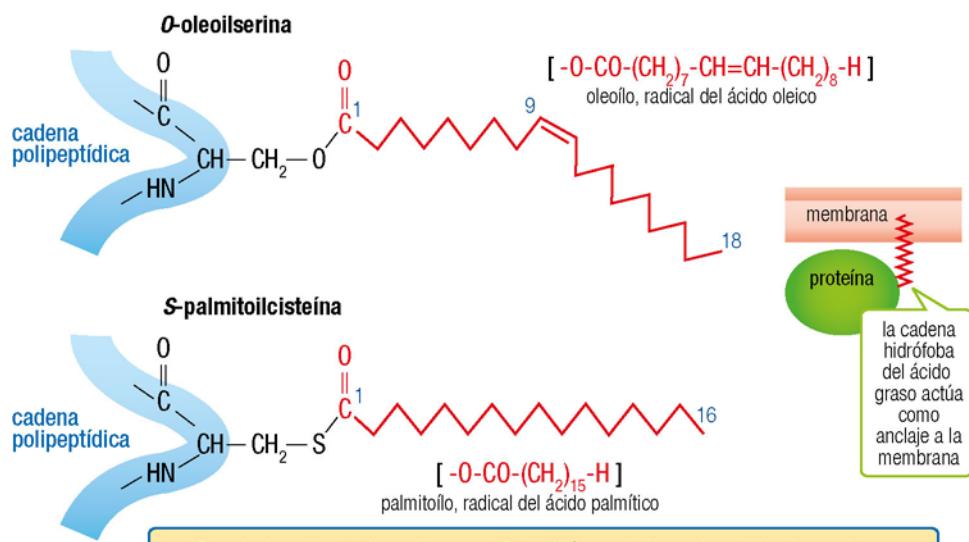
22.9

22.3.1.2 Modificación con lípidos

a) Acilación (ácidos grasos)

La acilación es una modificación relevante en algunas proteínas, que en general aumenta su hidrofobia. Más concretamente, el lípido supone a menudo un punto de anclaje a la membrana (comúnmente en la cara interna). Puede afectar a las cadenas laterales o a los extremos del polipéptido. En general esta reacción tiene lugar sobre proteínas citosólicas solubles, sintetizadas en ribosomas libres, y es cotraduccional.

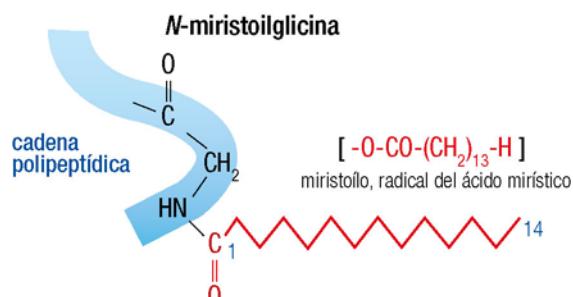
- Acilación de cadenas laterales de serina y treonina (enlace éster con el grupo OH) o de cisteína (enlace tioéster con el SH), con ácidos grasos de 14C (miristato), 16C (palmitato) y 18C (estearato y oleato).



22.10

- Acilación de los extremos. El grupo amino terminal puede formar enlace amida con miristato (14C) o palmitato (16C), lo que tiene lugar cotraduccionalmente (antes de completar la síntesis del polipéptido). La reacción parte de miristoil-coenzima A, catalizada por una miristoil-CoA:proteína N-miristoiltransferasa.

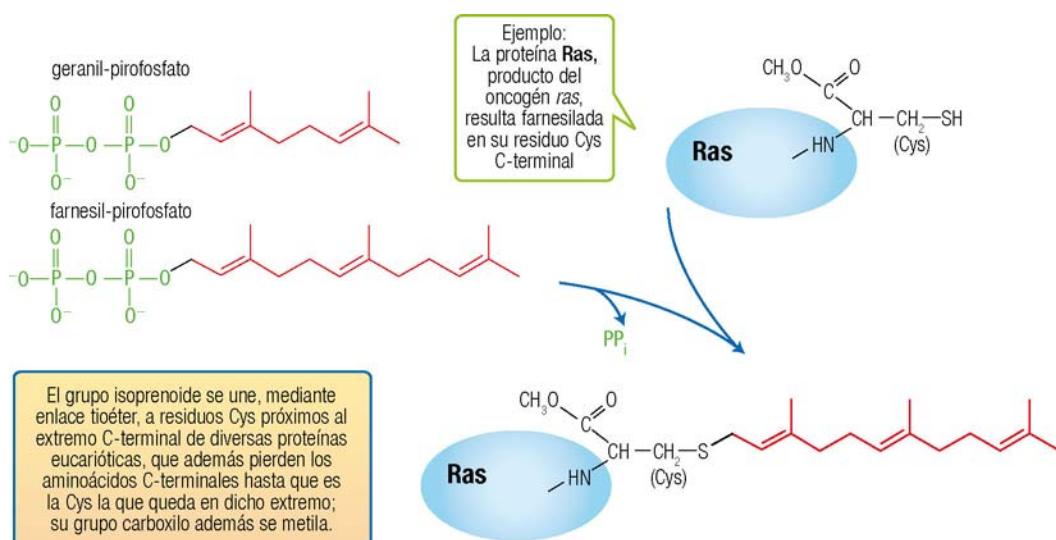
Ejemplo: Muchas proteínas implicadas en la transducción de señales, tales como Ser/Thr-quinasas, Tyr-quinasas, proteínas G y mediadores del transporte de proteínas en vesículas, están miristoiladas. La miristoilación ocurre siempre sobre secuencias $\text{H}_2\text{N-Gly-X-X-Ser-Y-Y-Thr}$ donde Y son aminoácidos básicos.



22.11

b) Prenilación (terpenos)

Principalmente se unen tres tipos de radical terpenoide (formados por unidades de isopreno, de ahí el nombre “prenilación”): geranilo (10C), farnesilo (15C) y geranylgeranilo (20C); los dos primeros son metabolitos intermedios de la ruta biosintética del colesterol. Al igual que la acilación, sufren esta modificación proteínas sintetizadas en ribosomas citosólicos.



22.12

Entre las proteínas modificadas con farnesilo se encuentran la proteína Ras, proteínas G y proteínas de la matriz nuclear. La oncoproteína Ras utiliza el farnesilo para anclarse a la superficie interna de la membrana plasmática y poder activar las señales intracelulares que desembocan en el fenotipo canceroso. El bloqueo de la prenilación conlleva la pérdida de la actividad transformante (carcinogénica), por lo que se han desarrollado fármacos anticancerosos basados en la inhibición de la farnesiltransferasa.

22.3.1.3 Incorporación de glúcidos: glicosilación

Se llama glicosilación a la unión covalente de cadenas de **oligosacáridos** (o **glicanos**) a la cadena lateral de algunos aminoácidos en las proteínas (preferimos el prefijo *glico-* al alternativo *gluco-*, que queda reservado para lo específico de glucosa). Esta modificación es muy frecuente en proteínas de membrana y en las secretadas al exterior (entre las que se incluyen las proteínas del plasma sanguíneo), y menos común en proteínas intracelulares. Por otra parte, es característica de eucariotas y virus, pero inexistente en procariotas.

a) Características de las glicoproteínas

Las glicoproteínas en su conjunto no tienen una función común definida por la glicosilación. Sin embargo, la adición de azúcares posee un papel esencial en la estructura y en la función de cada una de ellas. Por ejemplo, la glicosilación:

- Contribuye a establecer la conformación de la proteína.
- Aumenta su estabilidad (por ejemplo frente al calentamiento) y su resistencia a la digestión por proteasas. Como consecuencia, aumenta la vida media en plasma (disminuye el aclaramiento).
- Hace las proteínas más hidrófilas, aumentando su solubilidad y, por tanto, facilitando su interacción con las células del entorno.
- Aporta estructuras individualizadas que median la interacción con receptores específicos y actúan como determinantes antigenicos.

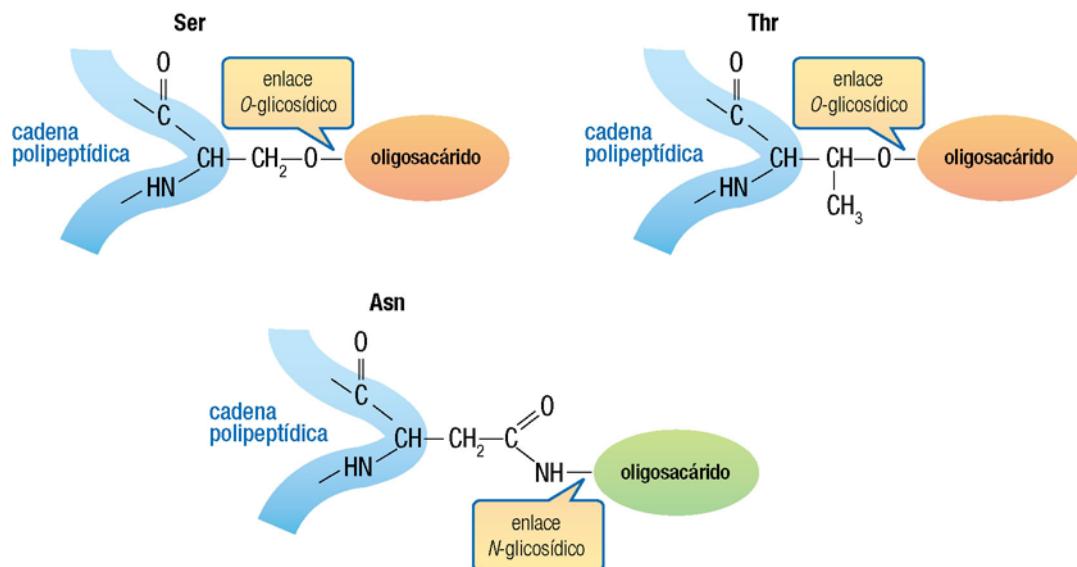
Son innumerables las proteínas que están glicosiladas (más que las no glicosiladas), así como sus funciones, y no se discutirán en este texto. Como meros ejemplos, que no representan el conjunto de funciones, cabe citar su papel como base de los distintos tipos de **grupos sanguíneos**, entre ellos el grupo AB0 (pág. 410), cuyos antígenos A, B y 0 son oligosacáridos constituyentes de glicoproteínas y glicolípidos presentes en la superficie de los eritrocitos y de otras células, y en glicoproteínas secretadas. Por otro lado, un método rutinario de vigilancia de la diabetes es la medida de **hemoglobina glucosilada**, cuya concentración depende de la de glucosa en sangre (la hemoglobina une covalentemente la glucosa que atraviesa de forma libre la membrana eritrocitaria). Asimismo, existen abundantes ejemplos de patologías provocadas por la alteración del patrón de glicosilación.

La unión de los oligosacáridos, que tiene lugar tanto de forma cotraduccional como postraduccional, da lugar a una gran variedad de glicoproteínas en las que la parte glucídica representa una proporción también variable de la masa total. La clasificación de los compuestos cuya molécula posee tanto glúcidos como polipéptidos no es sencilla, debido a la gran diversidad existente. En cualquier caso, existen diferencias tanto en lo relativo a la proporción de las partes glucídica y peptídica como al tipo de azúcares que participan y a la forma como se unen entre sí. A partir de aquí, la descripción se limitará a lo referente a las glicoproteínas propiamente dichas.



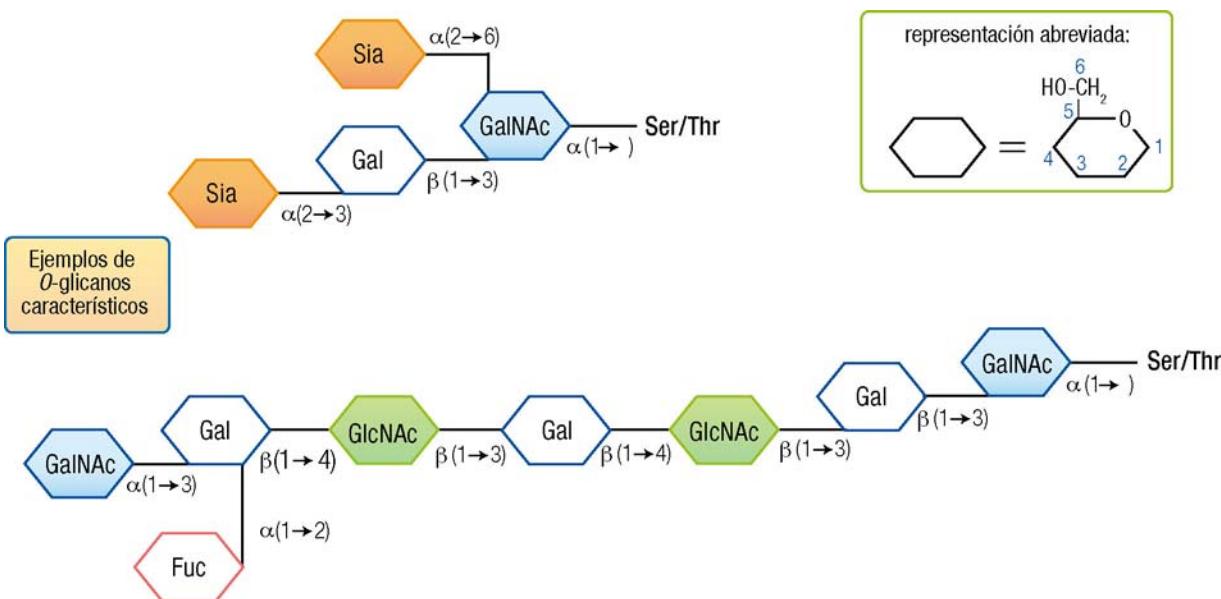
b) Tipos de glicosilación

Las cadenas glucídicas de las glicoproteínas se clasifican en dos grandes familias, en función de cuál sea el enlace entre el oligosacárido y el polipéptido. Esta clasificación puede aplicarse en cierta medida a las propias glicoproteínas, pero algunas presentan de forma simultánea azúcares unidos mediante ambos tipos de enlace, en puntos distintos de su molécula.



22.14

- **Oligosacáridos unidos por oxígeno (O-glicoproteínas):** el azúcar se une al átomo de oxígeno de cadenas laterales de serina o treonina. En general, los O-glicanos u oligosacáridos unidos por O son estructuras sencillas, aunque variadas. Los más frecuentes tienen como núcleo un residuo de N-acetilgalactosamina unido a la serina o treonina, pero existen también uniones xilosa-serina, galactosa-hidroxilisina, arabinosa-hidroxiprolina, N-acetilglucosamina-serina/treonina e incluso galactosa-cisteína (con un enlace S-glicosídico).

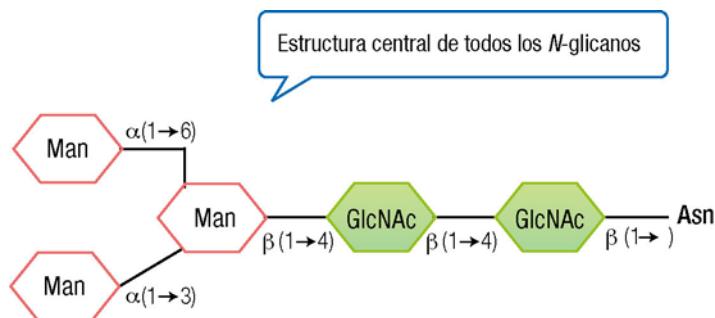


22.15

Monosacárido:	α -D-glucosa	N-acetil- β -D-glucosamina	α -D-manosa (α Man)	α -D-fucosa (6-desoxi-D-Gal)
Abreviatura:	Glc	GlcNAc	Man	Fuc
Monosacárido:	β -D-galactosa	N-acetil- α -D-galactosamina	ácido siálico (N-acetilneuramínico)	
Abreviatura:	Gal	GalNAc	Sia (NeuNAc)	

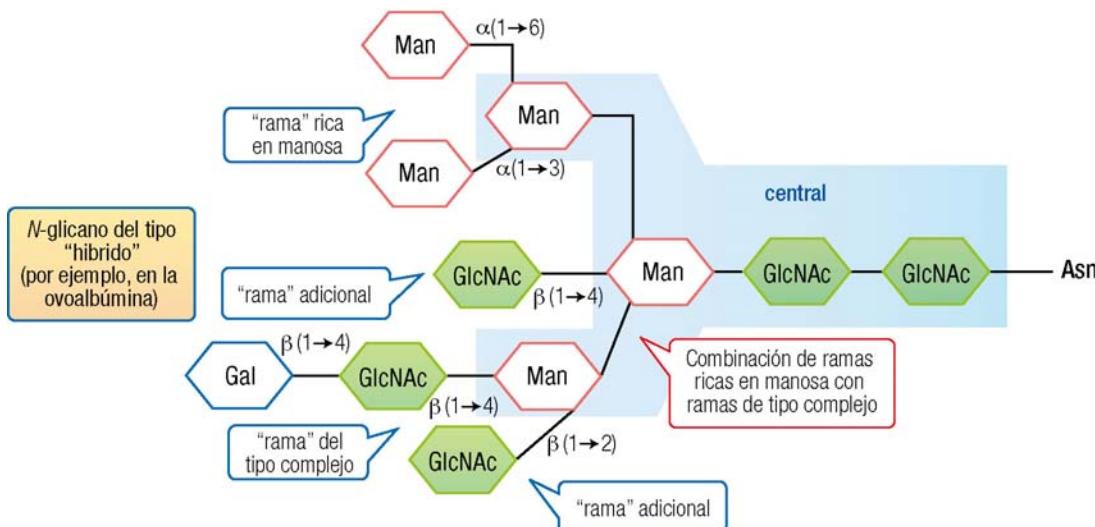
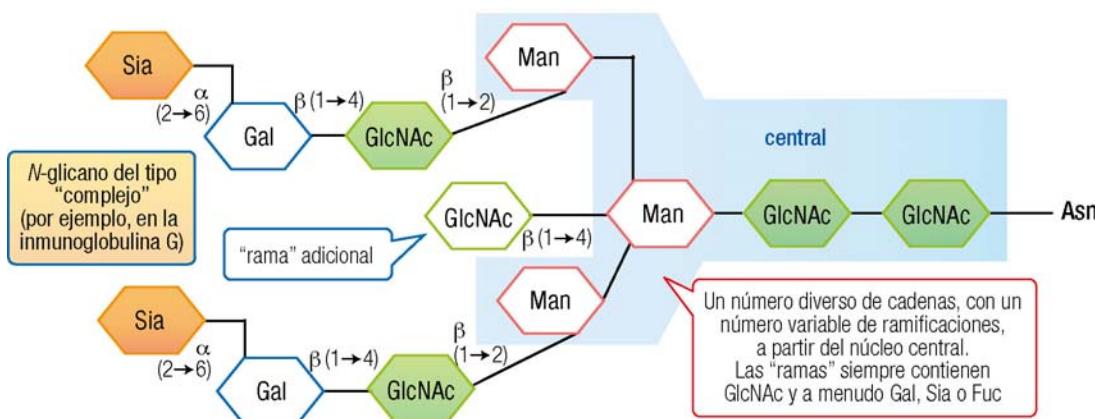
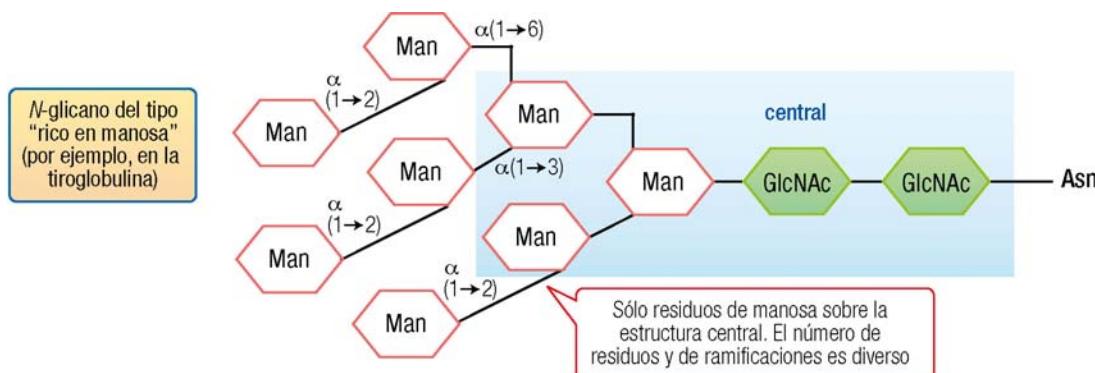
Nota: Estrictamente, las abreviaturas sólo están recomendadas (IUPAC) para usarse cuando estos monosacáridos están formando parte de oligosacáridos o polisacáridos

- **Oligosacáridos unidos por nitrógeno (N-glicoproteínas):** el azúcar se une al átomo de nitrógeno amídico de la cadena lateral de Asn. Ésta debe formar parte de la secuencia Asn-X-Ser-Y, donde la posición de Ser puede estar ocupada también por Thr o Cys, y X e Y son cualquier aminoácido excepto Pro. Las estructuras de los N-glicanos u oligosacáridos unidos por N son mucho más complejas y diversas. El tipo de monosacáridos constituyentes es generalmente diferente del de los O-glicanos.

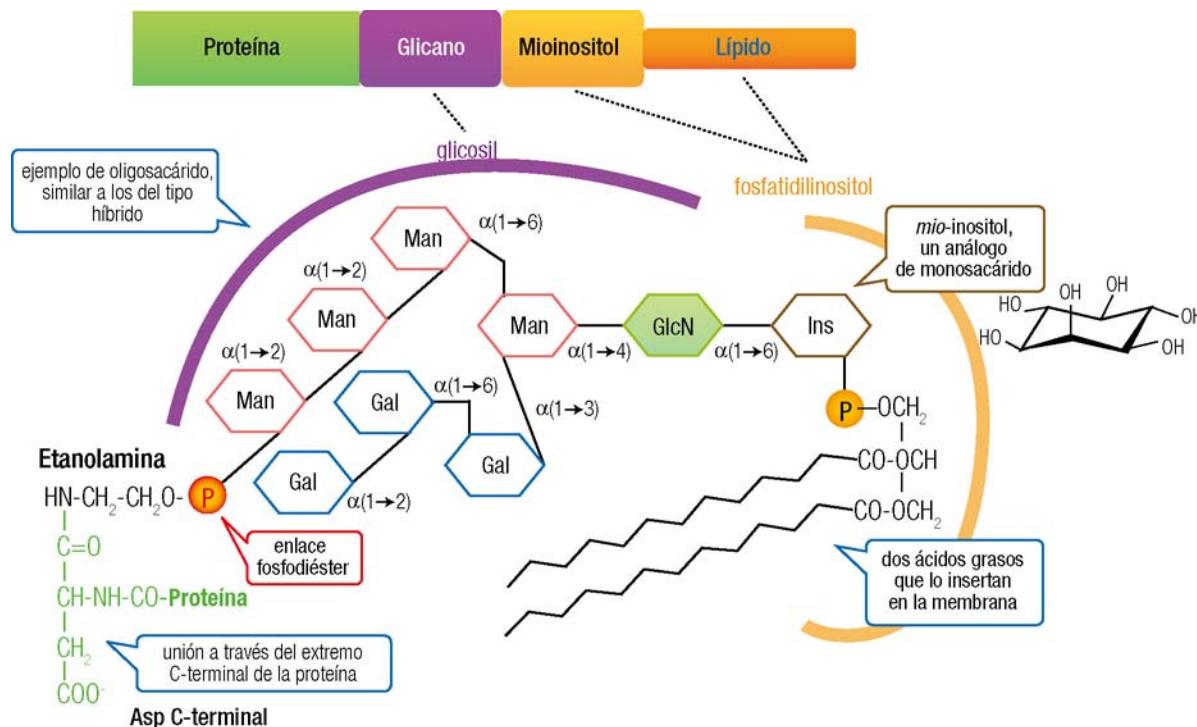


22.16

Alrededor de una estructura central común se unen monosacáridos muy diversos (ácido siálico, galactosa, N-vacetilglucosamina, etc.), formando cadenas con distintas disposiciones. De acuerdo con éstas, los N-glicanos se clasifican en 3 grupos: ricos en manosa, complejos e híbridos.



Anclaje GPI (glicosilfosfatidilinositol): este tercer tipo de glicosilación posee importantes diferencias con los anteriores. Se trata de una estructura particular que permite a ciertas proteínas fijarse sobre una membrana (habitualmente, en la cara extracelular), a través de un lípido que se inserta en ésta. La unión de la proteína al lípido no es directa (como lo son la acilación y la prenilación presentadas anteriormente), sino mediante un oligosacárido conector. Se puede contemplar como una proteína modificada por unión de un glicolípido.



22.18



Web 22.5. Estructura tridimensional de algunos oligosacáridos de glicoproteínas.

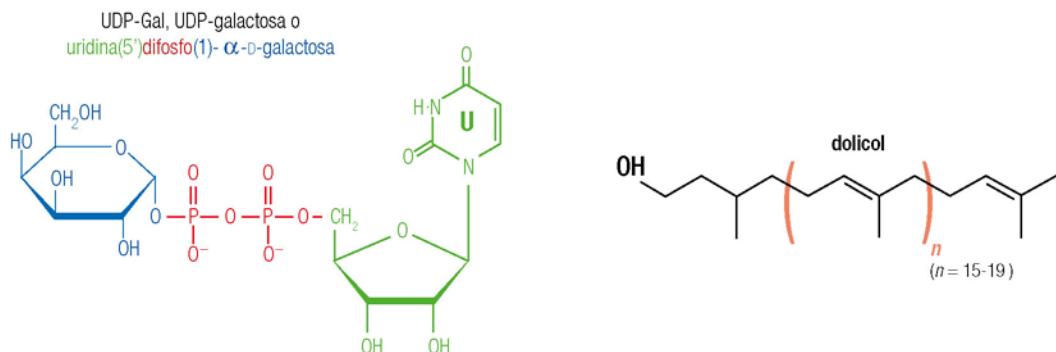
c) Biosíntesis de las glicoproteínas

La incorporación de las cadenas oligosacáridicas al polipéptido es uno de los principales procesos metabólicos que tienen lugar en el lumen del RE y del complejo de Golgi. Las proteínas, por tanto, se mantienen aisladas del citoplasma durante su glicosilación, que tiene lugar mientras la proteína va viajando a lo largo de la ruta RE → vesículas de transferencia → complejo de Golgi → vesículas → secreción u otro destino final de la glicoproteína.



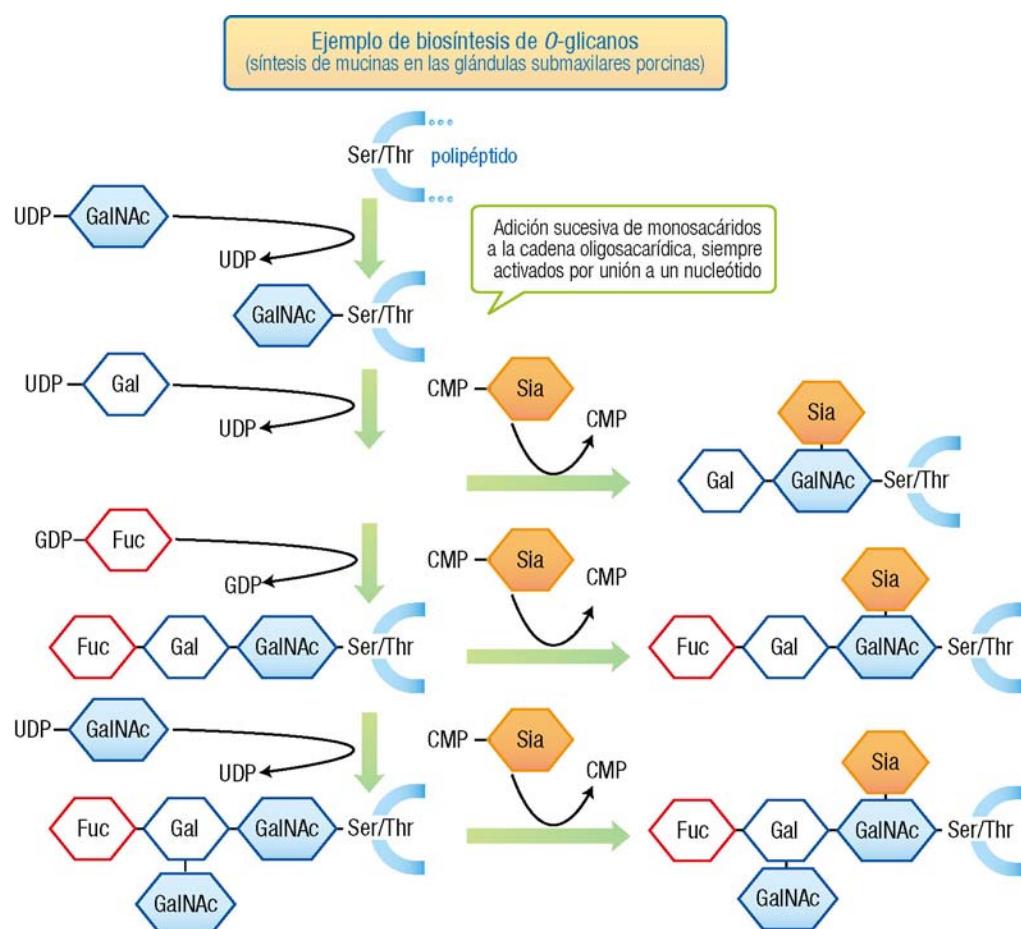
Web 22.6. Rutas de procesamiento y distribución de las glicoproteínas.

Los O-glicanos y N-glicanos se sintetizan por rutas metabólicas claramente diferentes, pero en todos los casos la unión de los monosacáridos está catalizada por **glicosiltransferasas**, específicas no sólo para reconocer el monosacárido que se incorpora y la cadena oligosacáridica que actúa de acceptor, sino también para formar el enlace en posiciones muy definidas del anillo de cada uno de los azúcares. Estas enzimas se encuentran en la cara interior de las membranas del RE y del complejo de Golgi. Para poderse emplear como sustratos, los monosacáridos deben estar activados, siempre por unión a un nucleótido; en concreto, en forma de UDP-Glc, UDP-GlcNAc, UDP-Gal, UDP-GalNAc, GDP-Man, GDP-Fuc y CMP-Sia. En otros casos, el precursor es el azúcar unido a una molécula de dolicol (un lípido terpenoide).



22.19

- La adición de los O-glicanos al residuo de Ser o Thr se realiza de manera directa, por unión sucesiva de los monosacáridos, catalizada por las respectivas glicosiltransferasas. Todas las etapas tienen lugar en el complejo de Golgi, de forma postraduccional (pues la síntesis de la cadena polipeptídica se ha completado en el RE).



22.20

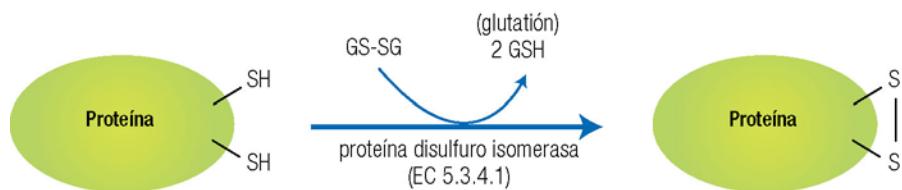
- La síntesis de N-glicoproteínas es un proceso más complejo, que comienza por la síntesis de un oligosacárido precursor unido al dolic平, seguida de la transferencia *en bloque* de ese glicano al residuo de asparragina del polipéptido y, finalmente, de nuevas reacciones de adición de monosacáridos y también de eliminación de algunos de los previamente incorporados. El proceso se desarrolla de manera secuencial, parte en el RE (cotraduccional) y parte en el complejo de Golgi (postraduccional).



Web 22.7. Biosíntesis de N-glicanos para las glicoproteínas.

22.3.1.4 Formación de puentes disulfuro

Son numerosas las proteínas que forman entrecruzamientos covalentes mediante puentes disulfuro entre residuos de cisteína de la misma o de distinta cadena (intracatenarios o intercatenarios, respectivamente). Los enlaces disulfuro son más frecuentes en las proteínas secretadas que en las proteínas intracelulares, lo cual está relacionado con el ambiente más reductor en el citosol. Estos enlaces participan en la estabilización del plegamiento correcto (pág. 372) y protegen la conformación nativa de la proteína frente a la desnaturización en un ambiente extracelular que puede ser más agresivo. La formación de enlaces disulfuro (una reacción redox) está asistida por la enzima **proteína-disulfuro isomerasa** y por moléculas pequeñas con grupos SH (tiol), la principal de las cuales es el glutatión (γ -glutamilcisteinilglicina).



22.21



Web 22.8. Estructura del glutatión.

22.3.1.5 Otras modificaciones

- Para ejercer su actividad muchas proteínas procarióticas y eucarióticas emplean **grupos prostéticos**, algunos de ellos unidos covalentemente una vez que el polipéptido ha abandonado el ribosoma. Entre estos grupos pueden citarse la biotina en la acetil-CoA carboxilasa y el grupo hemo en el citocromo c.
- La actividad de algunas enzimas se regula añadiendo **mononucleótidos**. Por ejemplo, la glutamina sintetasa (del sistema fijador de nitrógeno en bacterias) en su forma inactiva está sustituida con un grupo adenililo sobre un residuo de tirosina. El grado de adenililación lo controla una proteína reguladora.
- La **ADP-ribosilación** es otra reacción de modificación que actúa de forma reversible regulando la función de proteínas citoplásmicas y nucleares. Transferasas específicas actúan sobre His, Arg, Asn, Lys o Glu, empleando como cosustrato el NAD⁺. La misma actividad la ejercen las toxinas diftérica, colérica y pertusica, perturbando la regulación fisiológica.
- En algunas proteínas se produce la **sulfatación** de tirosina con 3'-fosfoadenosina-5'-fosfatosulfato; esta reacción tiene lugar en el complejo de Golgi.
- La **ubiquitinación** es la unión de la ubiquitina (una proteína pequeña, con 76 aminoácidos) a residuos de lisina. Es una señal o *etiqueta* para la degradación de aquellas proteínas que se han ubiquitinado varias veces (pág. 377-378). La **sumoilación** es la unión –también sobre la cadena lateral de lisina– de una molécula de SUMO, otra proteína pequeña. Como resultado, se puede modular la interacción con otras proteínas, alterar la localización en la célula o neutralizar el efecto de la ubiquitina.

22.3.2 Maduración por escisión proteolítica

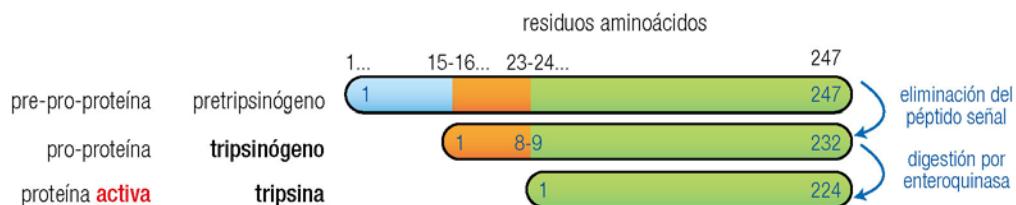
Este segundo tipo de maduración proteica consiste en la ruptura específica del polipéptido (**proteólisis**) en polipéptidos más pequeños, cada uno con distinta actividad, o en la eliminación de una porción sin función (en general pequeña) para dar un solo polipéptido activo. En cualquier caso, se trata de un proceso de **activación**, de precursores inactivos a proteínas activas. Se conoce como procesamiento proteolítico o **proteólisis limitada**. Es una hidrólisis absolutamente específica y de acción limitada a ciertos enlaces peptídicos, y no debe confundirse con la responsable de la degradación proteica (pág. 377).

Además de los procesos de activación por proteólisis, que se describen a continuación, se deben considerar dentro de este apartado la eliminación de residuos amino y carboxi-terminales –esencialmente la liberación enzimática de la metionina N-terminal (previa desformilación en proteínas procarióticas)– y la eliminación del péptido o secuencia señal (responsable de la distribución de las proteínas maduras hacia su destino, pág. 357).

22.3.2.1 Activación proteolítica de enzimas

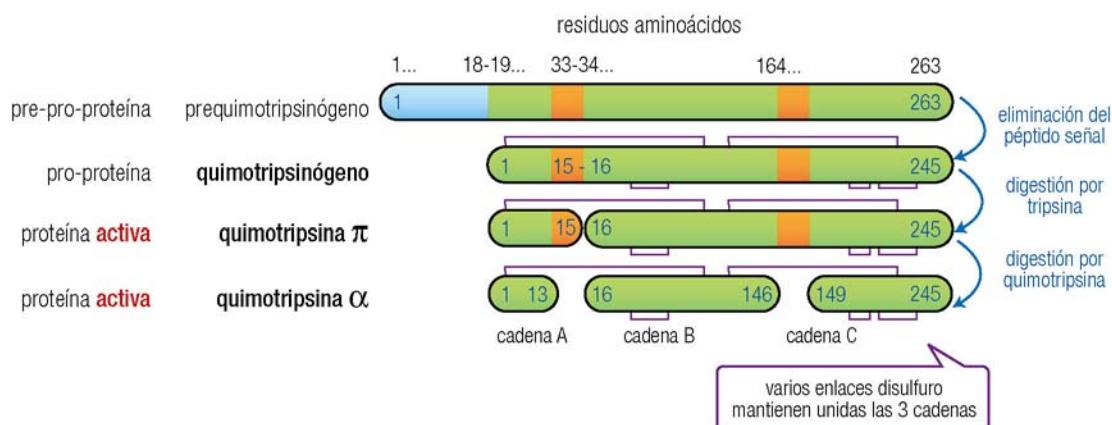
Las proteasas digestivas tripsina y quimotripsina se sintetizan en forma inactiva, como precursores o proproteínas, denominadas proenzimas o zimógenos. De ese modo se protegen las células donde se sintetizan y se reserva la actuación para el tubo digestivo. (En realidad, recién sintetizadas son *pre-pro*-proteínas, pues poseen un *péptido señal* que las marca como proteínas que deben secretarse y luego se elimina.) Una situación similar afecta a la mayoría de factores de coagulación, que son inactivos hasta que una proteólisis limitada los activa como proteasas, fenómeno responsable de la cascada de eventos que desencadena la coagulación.

La **tripsina**, producida en el páncreas exocrino, se activa mediante la eliminación de su octapeptido N-terminal, por acción de la **enteroquinasa** producida por las células de la pared del intestino delgado:



22.22

La tripsina actúa, a su vez, activando otras proenzimas o zimógenos, entre ellos el **quimotripsinógeno**, precursor de otra enzima digestiva pancreática, la quimotripsina; esta activación es algo más compleja:



22.23

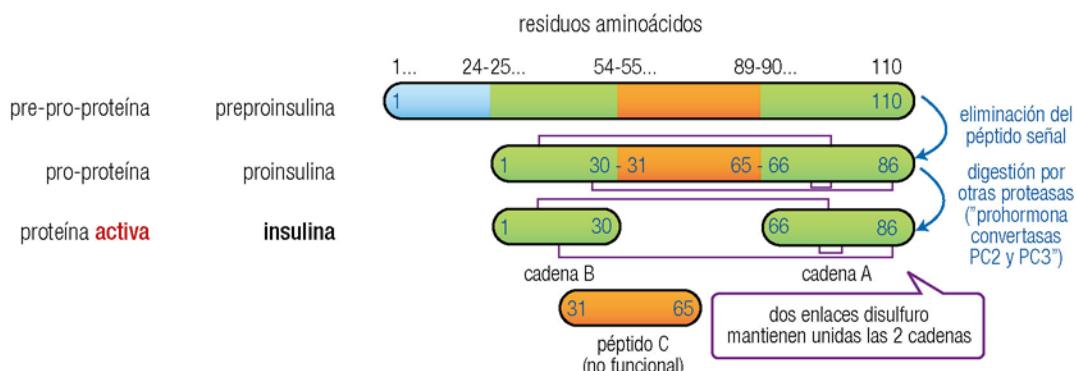


Web 22.9. Activación proteolítica de proteínas.

22.3.2.2 Activación proteolítica de hormonas

Como ejemplo característico, veamos la síntesis de la **insulina**. Además de demostrar la maduración por escisión proteolítica, constituye un ejemplo típico de modificación química por formación de enlaces disulfuro (pág. 369).

1. La insulina se sintetiza en las células B (o β) de los islotes de Langerhans del páncreas endocrino bajo la forma de **preproinsulina**, un precursor polipeptídico inactivo. La presencia en su extremo amino de una secuencia señal determina, como en toda proteína destinada a la secreción, el posterior transporte de esta molécula mediante vesículas hacia el exterior celular.



22.24

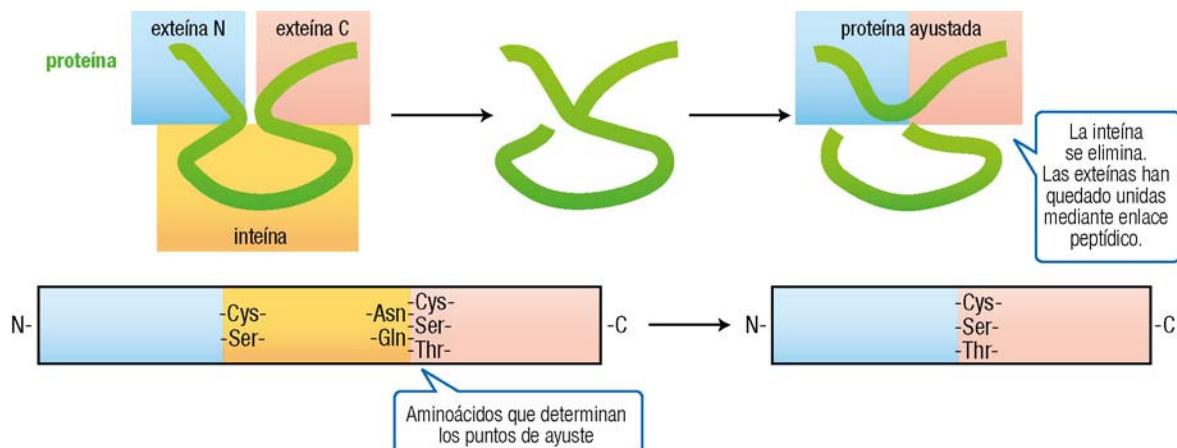


Web 22.10. Estructura tridimensional de proinsulina e insulina.

2. En el RE se elimina la secuencia señal dando la **proinsulina**, el verdadero precursor aún inactivo. Éste se pliega en una conformación tridimensional específica, estable, que facilita la formación de tres enlaces disulfuro.
3. La proinsulina se transporta hasta el complejo de Golgi, donde comienza su conversión en **insulina activa** mediante proteólisis. Las moléculas se acumulan en vesículas denominadas **gránulos de secreción**, en los que se completa su proteólisis y se almacenan hasta que la célula recibe las señales que determinan la secreción de la insulina al medio extracelular, e inmediatamente su paso al los capilares sanguíneos.

22.3.3 Ayuste de proteínas

En algunos casos, el polipéptido sintetizado pierde de forma específica un fragmento interno y los dos segmentos que lo flanquean se unen para formar la proteína funcional. Este proceso es formalmente análogo al de ayuste del mRNA (corte de intrones y empalme de exones, pág. 303), por lo que se denomina ayuste de proteínas. El segmento de polipéptido eliminado recibe el nombre de **inteína** y los dos segmentos que se unen y forman parte de la proteína final, **exteínas**. El proceso está catalizado por la propia inteína. Se ha encontrado este fenómeno en diversas bacterias, arqueas y eucariotas unicelulares.



22.25

El proceso está estrictamente definido y depende de la presencia de ciertos aminoácidos en las fronteras de la inteína (región de ayuste del lado amino y región de ayuste del lado carboxilo). En concreto, en la mayoría de los casos la inteína comienza con un residuo de cisteína o serina y termina con asparragina o glutamina, mientras que la exteína del lado C comienza por cisteína, serina o treonina. Gracias a esos residuos en particular se produce una serie de reorganizaciones de enlaces que terminan en la escisión de la inteína y la conexión de ambas exteínas mediante un enlace peptídico completamente normal.



Web 22.11. Detalles del mecanismo de ayuste de proteínas.

La capacidad de las inteínas para catalizar su propia eliminación conectando los dos polipéptidos adyacentes se ha aprovechado en el laboratorio para conectar proteínas no relacionadas o unir proteínas a etiquetas marcadoras.

22.4 PLEGAMIENTO DE PROTEÍNAS

Uno de los problemas importantes en biología molecular es comprender cómo la información contenida en la secuencia de aminoácidos de una proteína da lugar a su estructura tridimensional distintiva, responsable de su función o actividad biológica. Este **plegamiento** de la cadena polipeptídica, que tiene lugar en unos segundos, se realiza de forma simultánea a la distribución o tránsito y a una parte de la maduración del polipéptido, por lo que maduración y plegamiento están íntimamente ligados.

Los defectos en las rutas de plegamiento y distribución de las proteínas son la causa molecular de varias enfermedades.

- La **fibrosis quística** se origina por una mutación en el gen de la proteína transportadora de cloruro, que suprime uno de sus aminoácidos (pág. 392). Este cambio impide la separación del complejo proteína-carabina (pág. 373), lo que a su vez impide que la proteína salga del RE y continúe su maduración y tránsito hacia la membrana citoplasmática.
- La **enfermedad de Alzheimer**, que afecta al 10% de las personas mayores de 65 años, se origina en parte por la acumulación de agregados de un péptido derivado de la proteína β -amiloide; esos agregados parecen estar causados por un defecto en el plegamiento.
- Una forma hereditaria de **enfisema** (enfermedad pulmonar degenerativa) se produce por mutaciones en el gen de la antitripsina α_1 (pág. 414) que conducen a la agregación de la proteína, impidiendo su secreción; la deficiencia de la proteína en plasma produce susceptibilidad al desarrollo de la enfermedad.
- Por último, la conocida enfermedad de las vacas locas, el *scrapie* en ovejas y la **enfermedad de Creutzfeld-Jacob** en los seres humanos las causa una proteína con plegamiento anómalo, la proteína priónica PrP. Aunque la descripción de este problema supera las posibilidades de este libro, baste decir que una molécula de proteína plegada de forma anómala induce el mismo plegamiento en otras moléculas de la misma proteína, conduciendo así, por un lado, al desarrollo de la enfermedad y, por otro, a su transmisión a otros individuos.

22.4.1 Mecanismo del plegamiento de las proteínas

Algunos polipéptidos, en especial los más pequeños, pueden adoptar su conformación tridimensional nativa de forma espontánea, pero en la mayor parte de los casos el plegamiento correcto requiere la ayuda de proteínas especiales (distintas de las que se están plegando); la traducción del término inglés *chaperones* ha llevado a llamarlas chaperonas, chaperoninas, celadoras o **carabinas**; emplearemos este último nombre, más fiel al significado original (persona de edad acompañante de una joven, que evita “interacciones incorrectas”). Estas carabinas no sólo están implicadas en el plegamiento, sino también en la translocación a través de membranas subcelulares y en la degradación de las proteínas.

Además de las carabinas, en las células contribuyen al plegamiento dos tipos de isomerasas que, al catalizar ruptura y reconstrucción de enlaces covalentes, aceleran el plegamiento sin afectar a su resultado final. Entre ellas cabe citar la **proteína-disulfuro isomerasa**, que intercambia enlaces disulfuro entre dos moléculas (pág. 369), y otras enzimas que catalizan la interconversión de las conformaciones *cis* y *trans* de la prolina formando enlace peptídico, contribuyendo ambas a que la cadena polipeptídica adopte distintas conformaciones en la búsqueda de la de mínima energía.

El conocimiento de los mecanismos e interacciones que estabilizan la estructura tridimensional surge de experimentos pioneros de desnaturalización y renaturalización proteica (con la ribonucleasa A, de pequeño tamaño y un solo dominio) y posteriormente de la aplicación de métodos biofísicos de análisis y tecnología del DNA para producir polipéptidos alterados. Dada la variedad de proteínas y la complejidad de los mecanismos implicados, no se conocen aún las reglas que rigen el plegamiento *in vivo*: aún no somos capaces de predecir la estructura tridimensional de una proteína a partir de su secuencia. Se están haciendo importantes avances, no obstante, mediante el uso de técnicas computacionales que simulan el proceso de plegamiento a partir de los parámetros fisicoquímicos de las interacciones entre átomos.

Sin embargo, está claro que es la propia secuencia de aminoácidos la que determina por sí misma la estructura tridimensional activa. La estructura primaria o secuencia lineal resultante de la asociación de los

aminoácidos mediante enlaces peptídicos (determinada a su vez por la secuencia lineal de bases del DNA a través del mRNA) define la disposición de la cadena polipeptídica para formar puentes de hidrógeno, la capacidad de sus residuos hidrófobos para acomodarse en el interior de la molécula y los requerimientos estéricos que definen las posiciones más idóneas de la conformación nativa. Este punto tiene implicaciones de índole práctica, biotecnológica, porque cualquier modificación de la secuencia altera la estructura final de la proteína.

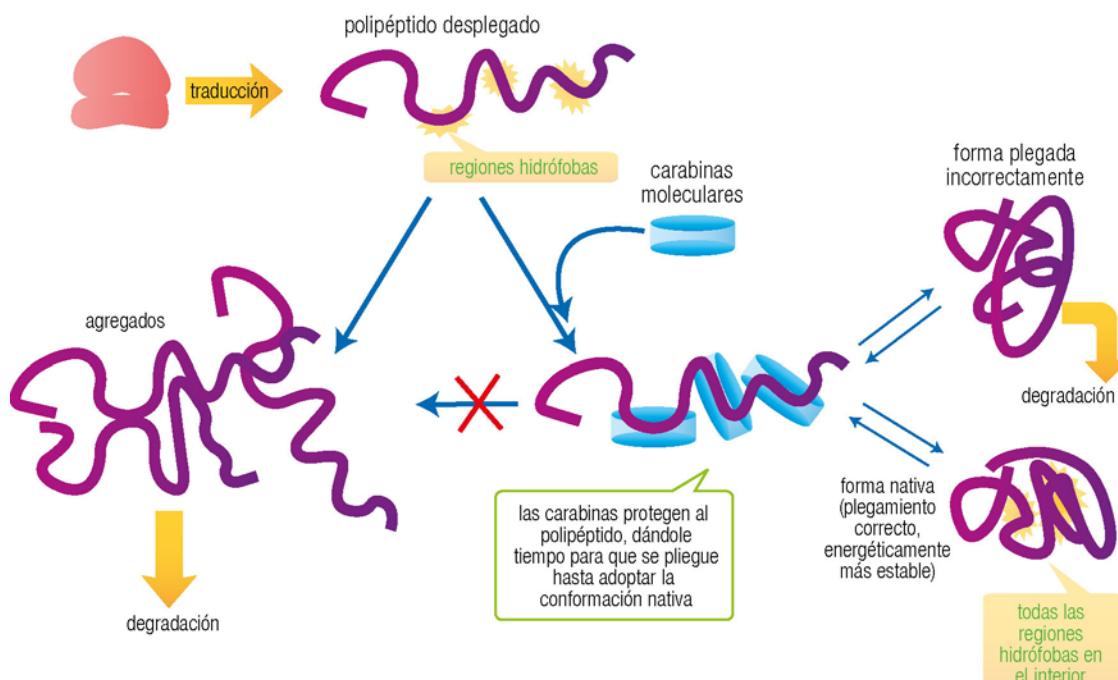
Las etapas del proceso de plegamiento pueden analizarse, de forma aproximada, siguiendo los niveles de organización de las proteínas globulares. En general, el plegamiento viene determinado por consideraciones energéticas correspondientes a las estructuras secundaria y terciaria, las que permiten el máximo número de puentes de hidrógeno e interacciones de Van der Waals, iónicas e hidrófobas, es decir, la conformación de menor energía (termodinámicamente, la más estable).

Muchas proteínas se pliegan a través de estados globulares intermedios, conocidos como “glóbulo fundido” (*molten globule*), estructuras compactas, aunque más abiertas y flexibles que la conformación nativa, con un contenido considerable de estructura secundaria, más o menos similar a la de aquélla, pero que carecen aún de su estructura terciaria definida. Las cadenas laterales de los aminoácidos en el glóbulo fundido no han establecido todavía sus interacciones mutuas, presentan libertad de movimiento e interacción con el disolvente, incluso las hidrófobas. Este primer plegamiento sería como un armazón o esqueleto sobre el que se elaboraría la estructura tridimensional definitiva, algo así como un pliegue de la estructura final por etapas.

El principio básico del plegamiento, en su conjunto, es la presencia de residuos hidrófobos de la cadena en un ambiente acuoso altamente polar: el “efecto hidrófobo”, el mismo que determina la asociación de los lípidos en las membranas celulares, es la principal fuente de estabilidad de la conformación plegada de una proteína.

22.4.2 Proteínas implicadas en el plegamiento: carabinas moleculares

Los polipéptidos en estado de glóbulo fundido, los que están plegados parcial o incorrectamente y en especial los que aún están saliendo del ribosoma (cadenas nacientes) exponen en su superficie residuos hidrófobos que en la proteína nativa estarán siempre dispuestos en el interior de la molécula (“escondidos” o “enterrados” por la envoltura de residuos hidrófilos que forma la superficie de la proteína). Estas regiones hidrófobas, para evitar la interacción con el medio acuoso polar, tienden a asociarse entre sí, conduciendo a la formación de agregados de varias moléculas (que en algunos casos constituyen los denominados “cuerpos de inclusión”). El plegamiento correcto, por tanto, requiere una protección frente a esta posibilidad, hasta que la proteína tenga la ocasión de encontrar su conformación nativa, la más estable; éste es uno de los papeles de las **carabinas moleculares**.

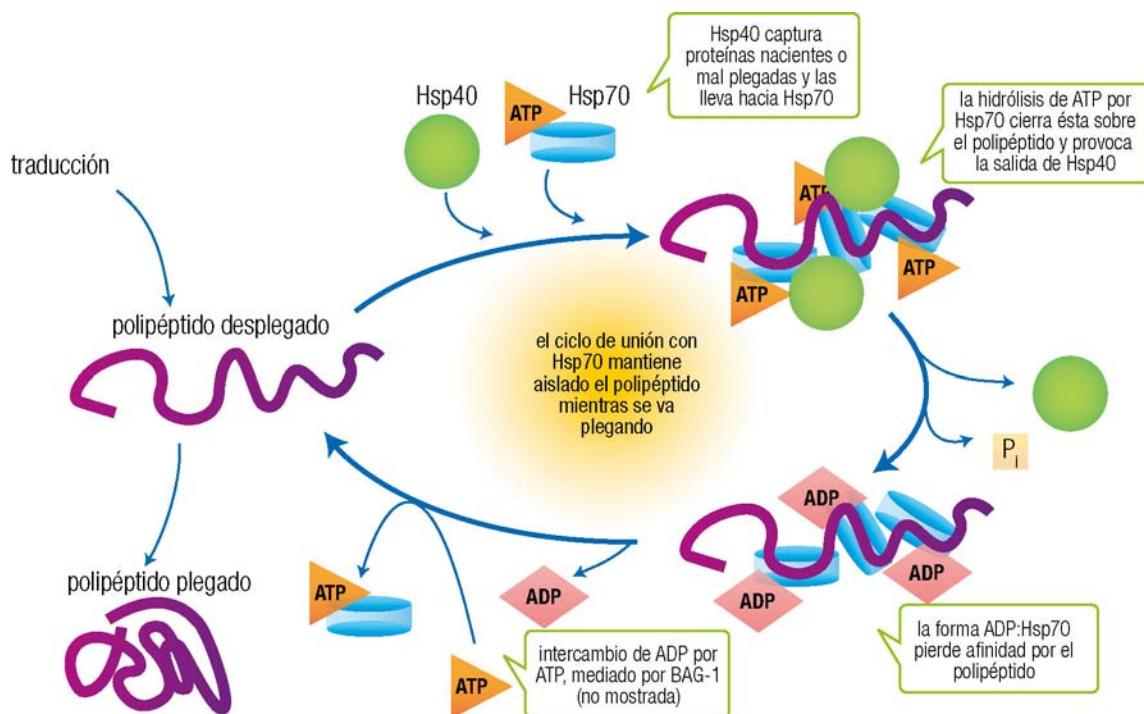


22.4.2.1 Tipos de carabinas

Las primeras proteínas identificadas como accesorias, auxiliares o catalizadoras del plegamiento se llamaron **proteínas de choque térmico (Hsp, por heat shock protein)**, porque su síntesis se estimula al someter células en cultivo a una elevación transitoria de la temperatura. Esto es así porque en esas condiciones aumenta la cantidad de moléculas de proteína parcialmente desplegadas, que es necesario replegar. Hoy día, el grupo de carabinas incluye otras muchas proteínas, no relacionadas entre sí, importantes para la viabilidad celular a todas las temperaturas y bajo todas las circunstancias. De acuerdo con su modo de actuación, se pueden distinguir dos grandes grupos de carabinas moleculares:

a) Carabinas Hsp70 y Hsp40

La familia Hsp70 (cuyo nombre procede de su masa molecular de 70 kDa) comprende proteínas que se asocian a segmentos peptídicos cortos, en conformación extendida o desplegada y ricos en aminoácidos hidrófobos. Por ello, reconocen las cadenas polipeptídicas nacientes según están saliendo del ribosoma. Su acción principal es evitar que se agreguen por interacciones hidrófobas intermoleculares (recordando la metáfora de la joven acompañada). Existen carabinas de esta familia en procariotas, eucariotas, mitocondrias, cloroplastos y en el interior del RE. Su unión y separación, reversibles, con la proteína desplegada dependen de la hidrólisis de ATP por parte de una actividad ATPasa propia de la Hsp70. Todo el proceso está modulado por otro tipo de proteínas, la familia Hsp40, que posee capacidad activadora de la Hsp70 (co-carabinas) y también de carabinas moleculares por sí mismas.

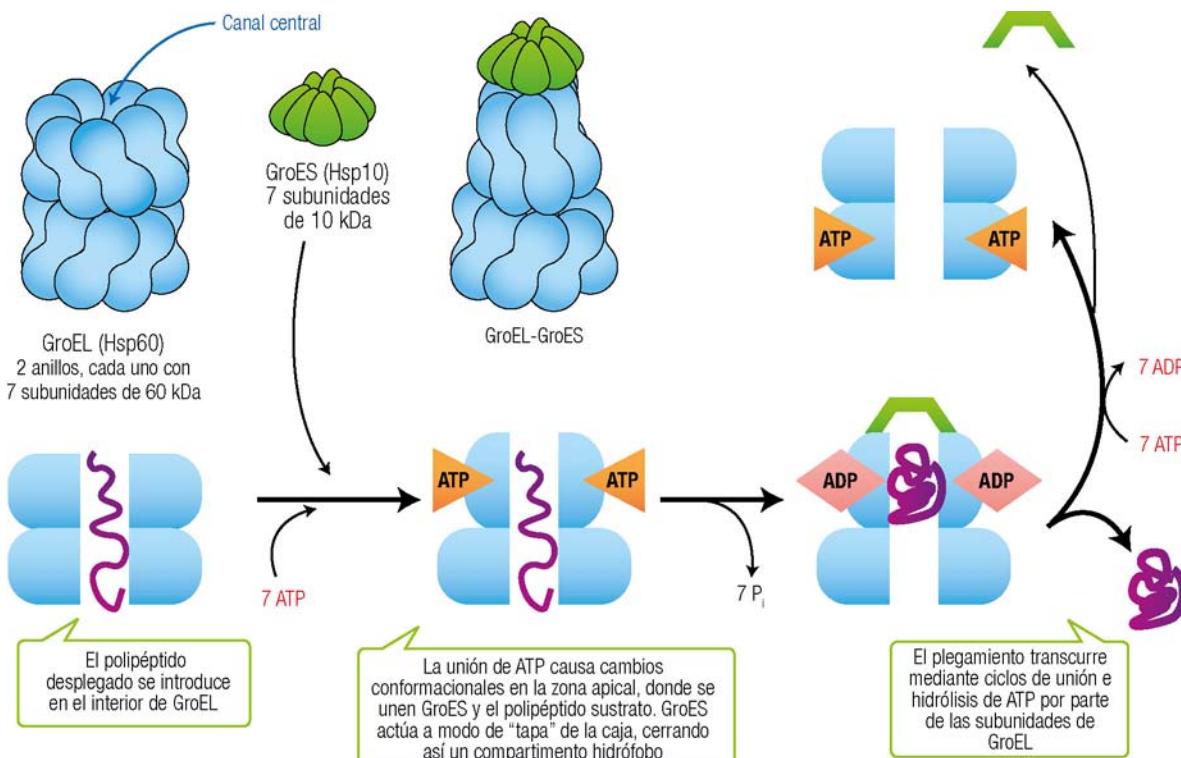


22.27

Las carabinas Hsp70 estabilizan, por tanto, las cadenas polipeptídicas nacientes, protegiéndolas de la agregación, y mediante los ciclos controlados de unión y separación proporcionan un entorno “seguro” para que la proteína pueda plegarse de manera adecuada. Sin embargo, no contribuyen a definir el resultado del plegamiento, que sólo depende de la secuencia del polipéptido.

b) Carabinas Hsp60 y Hsp10

Estas proteínas forman una estructura multimérica muy característica, en forma de caja cilíndrica, en cuyo interior se alojan los polipéptidos y donde experimentan su plegamiento, aislados del entorno y de nuevo evitando la interacción con otras moléculas desplegadas o parcialmente plegadas. El ejemplo mejor conocido es la proteína GroEL de *E. coli*, cuya estructura se ha determinado por cristalografía de rayos X a alta resolución. Con la colaboración de GroES, proteína de la familia Hsp10, canaliza el plegamiento del polipéptido. Si éste no se pliega por completo en un ciclo, puede volver a entrar en la cavidad para continuar plegándose, hasta alcanzar la estructura nativa.



22.28

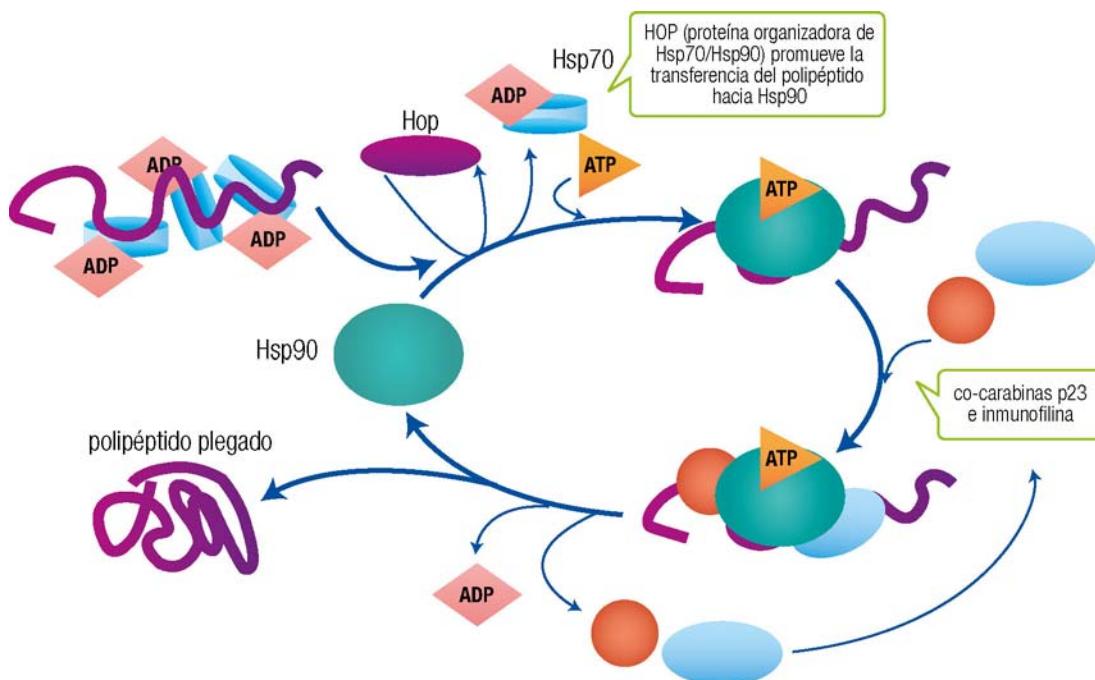


Web 22.12. Estructura tridimensional de las carabinas GroEL y GroES.

Se han encontrado homólogos del sistema GroEL-GroES en mitocondrias y cloroplastos, y proteínas menos estrechamente relacionadas en el citosol eucariótico. Las Hsp60/Hsp10 del “grupo I”, en las mitocondrias, captan los polipéptidos semiplegados cedidos por Hsp70:ADP una vez que han atravesado la membrana mitocondrial. Las del “grupo II”, en el citosol, reciben también los polipéptidos desde Hsp70:ADP. La prefoldina (homóloga a Hsp10) actúa de co-carabina transfiriendo el polipéptido hacia TRiC/CCT, análoga a Hsp60.

c) Otras carabinas

La Hsp70:ADP puede ceder el polipéptido que protege a otro grupo de carabinas, las Hsp90; éstas son menos abundantes que las anteriores, pero esenciales para que el plegamiento de algunas proteínas se complete correctamente.



22.29

Se han encontrado también unas carabinas “Hsp pequeñas”, que forman complejos oligoméricos. Entre ellas se encuentran las Hsp26, que se activan cuando la célula se ve expuesta a alta temperatura, y las Hsp27, que se activan en respuesta a señales mediante quinasas de proteínas. La activación disocia sus oligómeros en dímeros que pueden asociarse al polipéptido desplegado, evitando su agregación hasta que sufra un plegamiento correcto o bien sea degradado.

En general, se ha comprobado que las proteínas van pasando de uno a otro tipo de carabinas de forma ordenada, no se liberan para encontrar al azar la siguiente carabina necesaria.

22.4.2.2 Función de las carabinas

Las carabinas moleculares desempeñan un papel esencial en varias circunstancias:

- En el citosol celular, dada la gran densidad de macromoléculas, la variedad y elevada concentración de proteínas allí presentes en fase de plegamiento aumenta la probabilidad de interacciones hidrófobas y de agregación de las proteínas que, como se ha comentado, evitan las carabinas.
- En la mitocondria y el cloroplasto intervienen en el desplegamiento necesario para el tránsito de proteínas de origen citosólico a través de las membranas externa e interna del orgánulo, así como en el plegamiento en su interior, para adquirir su funcionalidad.
- En el lumen del RE interaccionan con las proteínas plegadas de forma incorrecta (o las proteínas diméricas o multiméricas no ensambladas de forma adecuada), con lo que éstas quedan retenidas en el RE hasta que tiene lugar el plegamiento correcto. Sólo entonces (al igual que las proteínas bien plegadas desde el principio) se transportan en vesículas hacia el complejo de Golgi. En caso contrario, las proteínas no plegadas se degradan.

22.5 DEGRADACIÓN DE LAS PROTEÍNAS

En paralelo con el desempeño de su función, las proteínas se degradan constantemente en la célula. Este proceso continuo, unido a la biosíntesis, constituye el llamado **recambio proteico** (en inglés, *turnover*), que tiene lugar a lo largo de toda la vida celular y es necesario para mantener en todo momento un nivel adecuado de proteínas.

(estado estacionario), indispensable para la actividad celular, para impedir la acumulación de proteínas anómalas o no necesarias y para facilitar el reciclado de los aminoácidos. De forma genérica, se puede decir que en una célula una proteína “sobrevive” en promedio unos dos días antes de degradarse, lo que ofrece una idea de la intensa actividad de los procesos de biosíntesis y de degradación.

Existen proteínas que se sintetizan y degradan (se recambian) de forma muy rápida respecto a la vida media de la célula en la que se hallan. Entre ellas se encuentran las proteínas defectuosas, por ejemplo por haber incorporado aminoácidos incorrectos en la síntesis, las que no desempeñan de manera correcta su función; a veces un simple cambio químico convierte a una proteína en blanco de una acción proteolítica. Por el contrario, otras proteínas, las necesarias en todo momento, son muy estables y por ello no requieren una síntesis continua a elevada velocidad. Así ocurre, por ejemplo, con la hemoglobina, que puede durar todo el tiempo de vida media del eritrocito (120 días en humanos).

En eucariotas existen dos sistemas principales de degradación de las proteínas, uno mediante los lisosomas y el otro en el citosol.

22.5.1 Degradación lisosómica

El primer sistema de degradación descrito se realiza en los lisosomas, liberados inicialmente por gemación a partir del complejo de Golgi (lisosomas primarios) y que después reciben enzimas hidrolíticas mediante vesículas procedentes también del Golgi. Estos orgánulos delimitados por una membrana sencilla acumulan más de 50 enzimas hidrolíticas diversas, capaces de actuar sobre proteínas, ácidos nucleicos, lípidos, etc., con una gran actividad al pH ácido característico del lisosoma (alrededor de 5). Las enzimas lisosomales que hidrolizan proteínas (proteasas) se denominan **catepsinas**.

Las enzimas hidrolíticas concretas presentes en los lisosomas dependen de cada tipo de célula; como ejemplo, las células pancreáticas contienen lisosomas con enzimas capaces de intervenir en las siguientes rutas de degradación:

- Para la exocitosis, o transporte de su contenido enzimático al exterior de la célula (secreción).
- Para la degradación de orgánulos por autofagia, éstos se incorporan en una vesícula que luego se fusiona con un lisosoma.
- Para hidrolizar nutrientes o destruir agentes patógenos procedentes del exterior de la célula: la vacuola fagocítica o fagosoma se fusiona con un lisosoma formando un fagolisosoma, donde se degrada el material.
- En ocasiones, para romper intracelularmente los lisosomas y liberar sus enzimas, conduciendo a la autólisis de la célula (una de las formas de muerte celular programada, que interviene, por ejemplo, en el desarrollo embrionario).

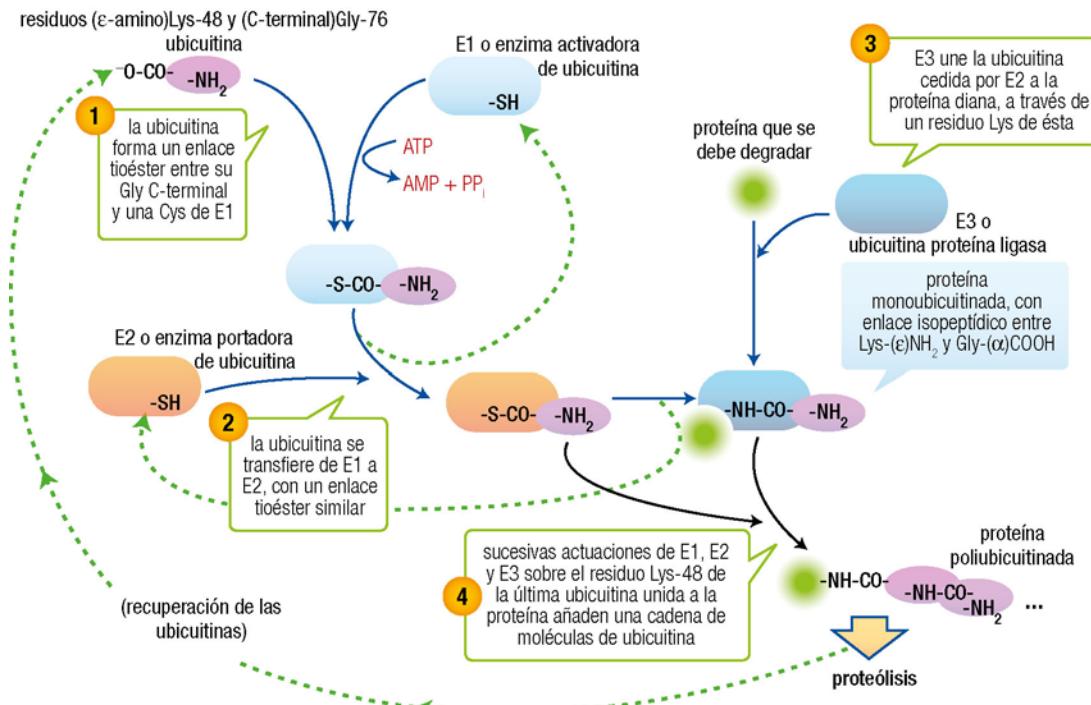
22.5.2 Degradación citosólica

La mayoría de las proteínas presentes en forma libre en el citosol se pueden degradar por esta ruta. El proceso es muy selectivo, está bajo un estricto control y es la ruta de degradación que requiere mayor gasto de ATP. Al igual que en la distribución de proteínas, intervienen como mediadores en la degradación señales moleculares específicas, en concreto aminoácidos N-terminales y la proteína **ubiquitina**.

22.5.2.1 Ubiquitina

La ubiquitina es una proteína pequeña, de 76 aminoácidos, llamada así por encontrarse en todas las células eucarióticas (es ubicua). Es posiblemente una de las proteínas más conservadas entre especies (las de levadura y humanos sólo se diferencian en tres aminoácidos). De carácter globular, rígida y muy estable, interviene como marcador selectivo para la degradación de otras proteínas, al unirse covalentemente a ellas, de forma que las proteínas *ubiquitinadas* quedan dispuestas para ser digeridas por los sistemas proteolíticos.

La **ubiquitinación** –o unión de la ubiquitina a cualquier proteína destinada a la destrucción– se realiza en un proceso de tres reacciones, catalizadas por 3 enzimas diferentes:



22.30

El enlace amida de la ubiquitina con la proteína (denominado isopeptídico porque no tiene lugar con el amino α , sino con el ϵ de la cadena lateral de Lys) constituye la señal para la destrucción, el marcaje para la proteólisis (se le ha llamado “el beso de la muerte”). La especificidad de qué proteínas deben marcarse para su degradación corre a cargo de las ubiquitina proteína ligasas (E3), pues en cada célula hay cientos de variantes. La unión de una sola molécula de ubiquitina no es suficiente para la degradación; sin embargo, puede marcar para otros procesos. A partir de 4 moléculas de ubiquitina encadenadas ya se produce una degradación eficaz de la proteína marcada (se han encontrado hasta 50 ubiquitinas unidas en cadena sobre una misma proteína). Las ubiquitinas que ejercen de señal se reciclan después, una vez liberadas en el proteasoma.

La degradación no es la única función desempeñada por las etiquetas de ubiquitina. Tanto la longitud de las cadenas de poliubiquitina como la posición de los enlaces entre ellas suponen marcas diferentes para procesos diversos. Por ejemplo, la monoubiquitinación de las histonas es una señal reguladora de la transcripción (pág. 286). Otros procesos regulados mediante la unión de ubiquitina incluyen procesamiento de proteínas, activación de una quinasa, compactación de la cromatina, interacciones intermoleculares entre proteínas, y control del transporte de proteínas hacia dentro o hacia fuera de la célula.



Web 22.13. Estructura de las cadenas de poliubiquitina.

22.5.2.2 Señales desencadenantes de la ubiquitinación

a) Residuo amino-terminal

La identidad del aminoácido N-terminal parece determinar que una proteína resulte marcada con ubiquitina con mayor o menor facilidad. Debe recordarse que en eucariotas este residuo resulta de la eliminación de Met y de cualquier otro proceso de maduración de dicho extremo; las señales amino-terminales parecen haberse conservado durante miles de millones de años en la evolución.

Aminoácido en posición N-terminal	Vida media de la proteína
Grupo I, básicos:	Arg, Lys, His
Grupo II, hidrófobos voluminosos:	Leu, Ile, Trp, Phe, Tyr
Grupo III, neutros:	Gly, Ala, Val, Ser, Thr, Cys, Met, Pro

b) Secuencias PEST

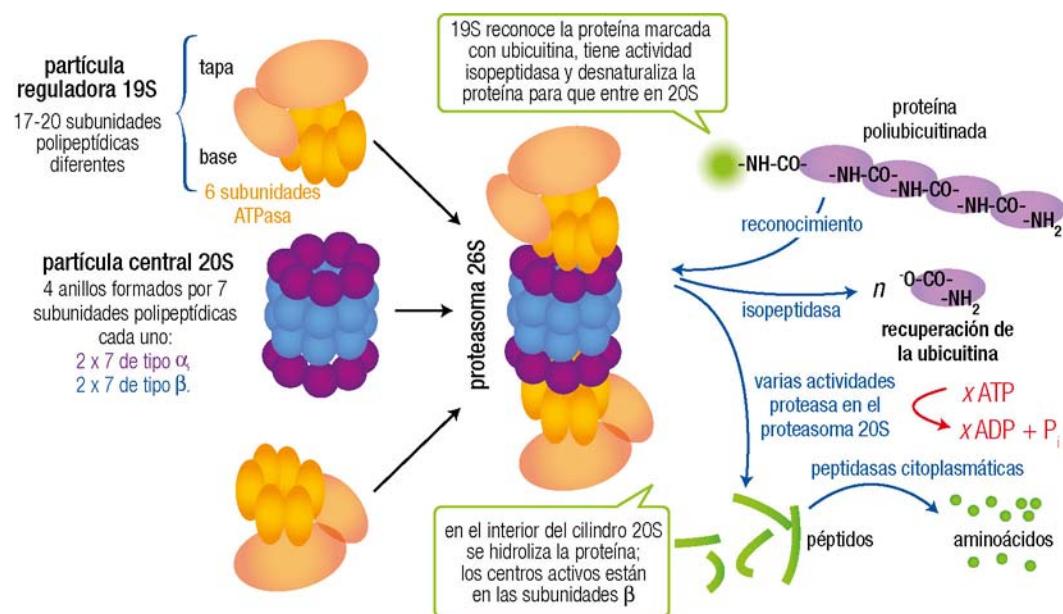
Por otra parte, también se asocia una vida media corta con la presencia en la proteína de regiones de 12 a 60 aminoácidos, ricas en Pro, Glu, Ser y Thr (P, E, S, T, con la abreviatura de una letra). Estas regiones actuarían a modo de señal de reconocimiento para los sistemas enzimáticos que degradan las proteínas de vida corta.

22.5.2.3 Degradación por proteasas y proteasoma

La proteólisis de las proteínas marcadas con ubiquitina pueden ejercerla distintos sistemas de proteasas. Se han descrito **calpaínas**, proteasas que requieren calcio para su actuación, pero el sistema más importante es probablemente el del proteasoma.

Se ha denominado **proteasoma** a un gran complejo proteico (2.000 kDa) que aparece en muchas células eucarióticas, formado por distintas subunidades proteolíticas. El conjunto, **proteasoma 26S o complejo proteína multicalitico**, contiene dos componentes asociados entre sí:

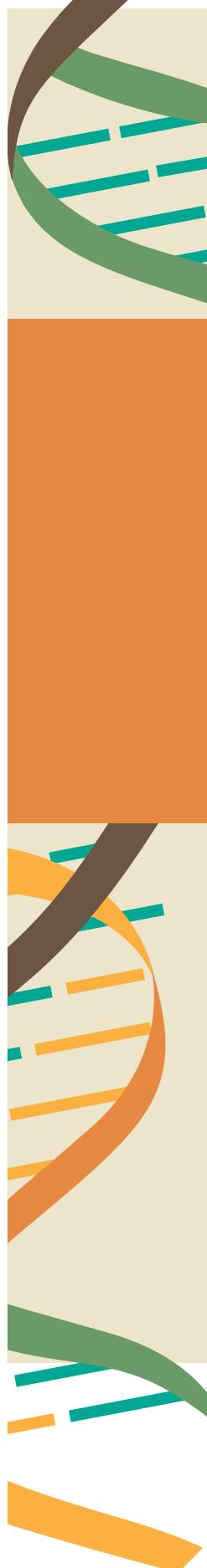
- Una **partícula central 20S**, formada por 28 subunidades en una disposición cilíndrica, que es el núcleo catalítico responsable de la proteólisis secuencial, con liberación de fragmentos peptídicos. (En cierto modo, recuerda a la estructura de las carabinas Hsp60, una caja cilíndrica.) La cara interior posee 6 centros activos con actividad proteasa.
- Uno o dos componentes que se unen a los extremos del cilindro: **partícula reguladora 19S**, a su vez compuesta por una **base** –6 moléculas con actividad ATPasa formando un anillo– y una **tapa**. Esta partícula 19S posee, además, la capacidad de reconocer específicamente los conjugados proteína-poliubiquitina y es responsable (gracias a una actividad isopeptidasa) de la separación de las ubiquitinas intactas, que pueden así reutilizarse.



Web 22.14. Estructura del proteasoma.

Sección IV

Aspectos aplicados





Bases moleculares de la mutación y la reparación del DNA

23.1 CONCEPTO DE MUTACIÓN	383
23.2 CLASIFICACIÓN DE LAS MUTACIONES	384
23.3 TIPO DE CÉLULA QUE SUFRE LA MUTACIÓN	384
23.3.1 Mutación en células de la línea germinal	384
23.3.2 Mutación en células somáticas	385
23.4 MUTACIONES A PEQUEÑA ESCALA O PUNTUALES	386
23.4.1 Cambios en la secuencia del DNA	387
23.4.1.1 Sustitución	387
a) Transiciones	388
b) Transversiones	388
23.4.1.2 Deleción, pérdida o eliminación	388
23.4.1.3 Inserción	389
23.4.2 Efecto sobre la secuencia de la proteína sintetizada	389
23.4.2.1 Mutaciones silenciosas	389
23.4.2.2 Mutaciones no silenciosas	390
a) Mutaciones de aminoácido	390
b) Mutaciones que cambian el marco de lectura	391
c) Mutaciones que no cambian el marco	392
d) Mutaciones con terminación prematura de la proteína	392
e) Mutaciones con terminación retrasada	394
23.5 CAUSAS Y MECANISMOS BÁSICOS DE LAS MUTACIONES	394
23.5.1 Mutaciones endógenas	394
23.5.1.1 Incorporación de nucleótidos erróneos durante la replicación	394
23.5.1.2 Otras mutaciones espontáneas	394
a) Sustituciones por desaminación oxidativa	394
b) Pérdida de bases por inestabilidad del enlace N-glicosídico	395
c) Modificación de las bases por mutágenos endógenos	396
23.5.2 Mutaciones inducidas o exógenas: mutagénesis	396
23.5.2.1 Lesiones y mutaciones inducidas por agentes químicos	396
23.5.2.2 Lesiones y mutaciones inducidas por agentes físicos	398
23.6 REPARACIÓN DEL DNA	398
23.6.1 Eliminación de agentes mutágenos	399
23.6.2 Reversión directa de la lesión	399
23.6.2.1 Reparación de fotodímeros	399
23.6.2.2 Reversión de bases modificadas con grupos alquilo	400
23.6.3 Reparación por escisión	400
23.6.3.1 Mecanismo general de reparación por escisión de nucleótidos	400
23.6.3.2 Mecanismos específicos de reparación por escisión de bases	402
23.6.4 Reparación de emparejamientos incorrectos	403
23.7 ENFERMEDADES HUMANAS ASOCIADAS A LA REPARACIÓN	404

23.1 CONCEPTO DE MUTACIÓN

Clásicamente, se define la **mutación** como una alteración en la secuencia del DNA de un individuo que se transmite por herencia a sus descendientes. La forma inalterada del DNA, llamada tipo *silvestre* (en inglés, *wild type*), se convierte así en otra forma, portadora de la mutación o tipo *mutante*. Las mutaciones se producen por errores en la replicación, por la alteración espontánea de nucleótidos o debido a la acción de agentes físicos o químicos (**mutágenos**).

El carácter hereditario incluido en la definición supone que la mutación debe presentarse de modo estable en las células germinales. Ahora bien, ello no es estrictamente así. Aunque con frecuencia se asume que la mutación sólo corresponde a secuencias codificantes (genes), siendo las alteraciones en los productos génicos (generalmente proteínas) las responsables de las enfermedades *hereditarias*, el concepto de mutación es hoy día más amplio.

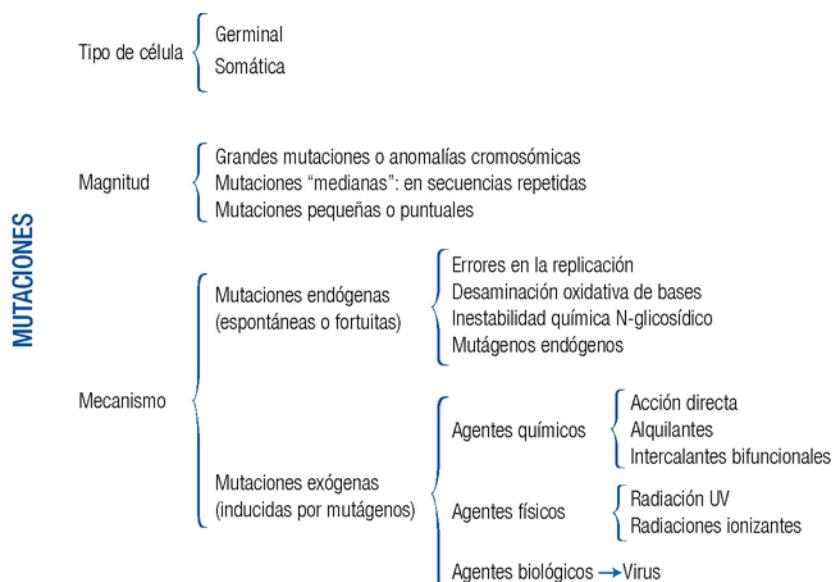
Las mutaciones tienen lugar en todo el genoma, incluyendo las secuencias codificantes o no, del genoma nuclear o mitocondrial, en células germinales (heredable) o somáticas (no heredable). Por tanto, la mutación es, junto con la recombinación meiótica (pág. 108), la principal fuente de variabilidad genética en todo tipo de organismos, y su estudio tiene interés básico y aplicado por su relación con procesos aparentemente tan alejados como la evolución de las especies, la identidad genética de un individuo, las alteraciones patológicas de

su genoma responsables de enfermedades, el diagnóstico prenatal o el tratamiento de enfermedades tanto somáticas (por ejemplo, el cáncer) como de células germinales.

Aunque muchas mutaciones se generan al azar, y existe la misma susceptibilidad de mutación en todas las regiones del genoma, las que ocurren sobre el DNA codificante tienen peores consecuencias. Dentro de esta **mutación génica** se debe distinguir entre la que afecta a la región del gen directamente responsable de la información (*región estructural*), que altera el RNA o la proteína sintetizados, y la que ocurre en las secuencias no codificantes del gen implicadas en el control de la expresión (*región reguladora*), que no altera la proteína sino su síntesis (aumento o descenso en concentración).

23.2 CLASIFICACIÓN DE LAS MUTACIONES

Dada la variedad de factores implicados, existen diversos planteamientos para clasificar las mutaciones. Se estudiarán bajo tres criterios: tipo de célula, magnitud de la mutación y mecanismos causantes de la mutación.



23.3 TIPO DE CÉLULA QUE SUFRE LA MUTACIÓN

23.3.1 Mutación en células de la línea germinal

Esta alteración se origina en alguna de las divisiones mitóticas o meióticas de la gametogénesis (pág. 102), a partir de células progenitoras diploides de las gónadas. La mutación puede no influir en el fenotipo de la célula o individuo que la sufre, ni en la capacidad reproductora de éste, pero, puesto que el genoma se transmite a la descendencia, la mutación se **hereda**, transmitiéndose a **todas las células** del organismo hijo y generando, en su caso, la alteración o enfermedad hereditaria. Obviamente, dicha transmisión depende de que en la fecundación participe un gameto portador de la mutación. A su vez, las células germinales del hijo también la poseerán (al igual que todas sus células somáticas), por lo que se volverá a transmitir a la siguiente generación. En consecuencia, la mutación se comporta como una variación heredable, que estará sujeta a la selección natural; se puede decir que son la *materia prima* para los cambios evolutivos.

La importancia de estas mutaciones depende directamente de cuál sea la función afectada: si ésta es esencial, se originan organismos inviables (abortos tempranos, a veces no detectados), mientras que si la función no es esencial, pero sí importante, se origina una enfermedad congénita o hereditaria, o una subnormalidad. Sin embargo, no debe olvidarse que algunas mutaciones resultan beneficiosas y son la base de la evolución.

Las mutaciones en la línea germinal son habitualmente acontecimientos esporádicos, minoritarios y relativamente infrecuentes en la población humana. Su tasa es de unas 10^{-6} mutaciones por locus y división celular. A este respecto, se puede calcular, como ejemplo, el número de mutaciones que tienen lugar durante las divisiones

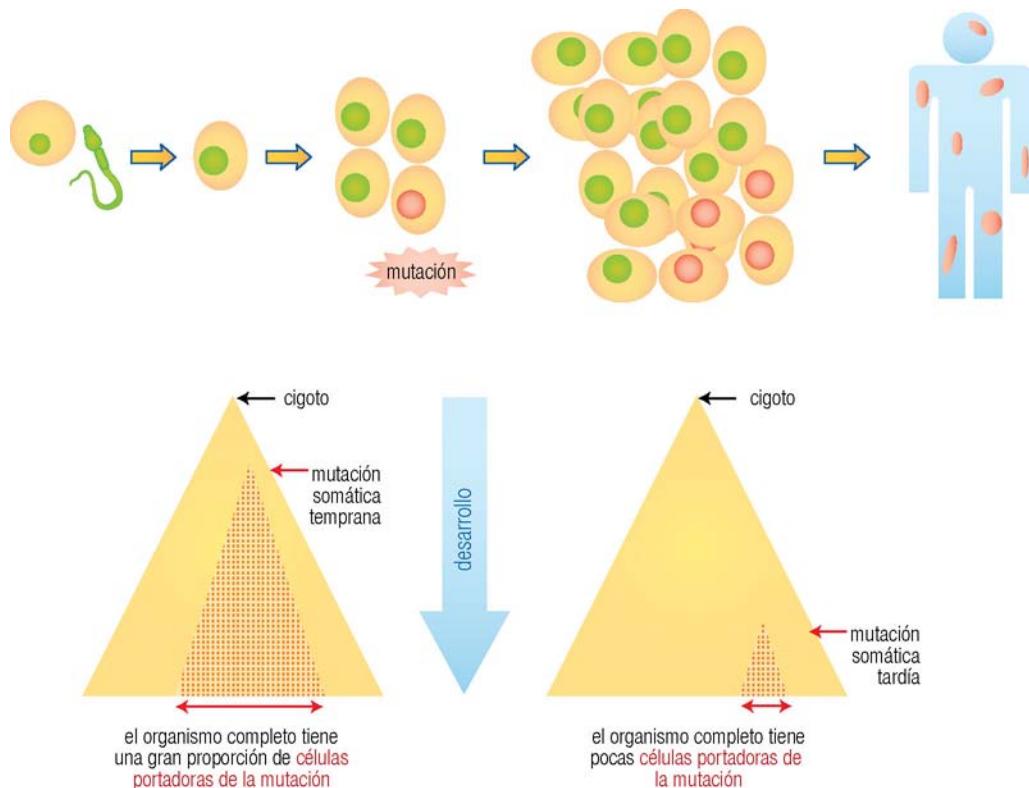
mitóticas y meióticas de la espermatogénesis. En la formación de un espermatozoide humano aparecen alrededor de 100 mutaciones debidas a errores en la replicación, lo que, teniendo en cuenta el tamaño de su genoma haploide, supone una mutación por cada 32 Mb. La mayoría de ellas son recesivas o, por el contrario, tan letales que no son fenotípicamente evidentes, pues no alcanzan la concepción o el nacimiento. De ahí que la tasa de mutaciones de la línea germinal con consecuencias patológicas sea muy pequeña.

23.3.2 Mutación en células somáticas

A diferencia de las anteriores, las mutaciones somáticas (en células no reproductoras) sólo se transmiten a las células hijas (en la mitosis), y no a un organismo completo. Pueden afectar a cualquiera de los 46 cromosomas de cualquier célula somática, diploide, en cualquier órgano, tejido o célula aislada, durante el desarrollo precoz o final del organismo, o en su fase adulta. Estas mutaciones determinan las características normales o patológicas del individuo, pero los cambios **nunca son heredables**.

Estas son mutaciones **mayoritarias** porque las células somáticas son las que más se dividen (unas 10^{13} divisiones en humanos desde el cigoto hasta el individuo adulto). La tasa de mutación, del orden de 10^{-10} por pb y división celular, es inferior a la de células germinales. Sin embargo, se incrementa como consecuencia de la exposición a agentes mutágenos (algunos de ellos cancerígenos), de la edad y de la alteración de los mecanismos de reparación del DNA, llegando a provocar miles de nuevas mutaciones durante la vida de un adulto, prácticamente en cada posición en el genoma y en cualquier parte del organismo.

El individuo portador de una mutación somática que se ha producido después del cigoto, pero antes de alcanzar su desarrollo completo, puede considerarse un *mosaico*, término genético que define la presencia de dos o más líneas celulares genéticamente distintas derivadas de un mismo cigoto. A través de la mitosis, las células somáticas del individuo adulto pertenecerán a dos líneas o clones, la *propia* o normal y la *anómala*, derivada de la célula que sufrió la mutación. Los individuos con una réplica fiel y completa de ambos tipos de células muestran *mosaicismo somático*. Este fenómeno se puede observar, por ejemplo, en ciertos trastornos citogenéticos con alteración en el número de cromosomas (pág. 447), como la trisomía 21. Lógicamente, cuanto más temprano tiene lugar la mutación en el desarrollo del organismo, mayor será la abundancia de células de la línea anómala sobre las de la línea normal.



Las mutaciones que sean compatibles con la vida de la célula pueden, sin embargo, afectar a su crecimiento. El efecto de la mutación sobre la célula somática depende en parte de su compatibilidad con el ciclo de división celular, pues la célula afectada se divide de forma activa y conducirá a la proliferación de la mutación, pudiendo occasionar problemas clínicos. Así ocurre, por ejemplo, en procesos de envejecimiento celular y transformación neoplásica. El **cáncer** se puede iniciar por la mutación de genes que controlan la división celular (Capítulo 26).

23.4 MUTACIONES A PEQUEÑA ESCALA O PUNTUALES

El estudio de las mutaciones en función de su magnitud, longitud o grado de la lesión permite clasificarlas en pequeña y gran escala (y, a veces, mediana). Las **anomalías a gran escala** ocurren por duplicaciones, pérdidas o reagrupamientos que alteran una región de millones de pares de bases y se reflejan en la estructura del cromosoma. Se llaman también **mutaciones cromosómicas o voluminosas**. Su estudio se reserva para el Capítulo 25, como anomalías cromosómicas. En éste nos limitaremos a las mutaciones a pequeña escala.

Las **mutaciones a pequeña escala, sencillas, puntuales, mínimas o micromodificaciones** implican habitualmente a un solo nucleótido (y, como consecuencia, a su complementario en la otra hebra del DNA). Sin embargo, también se incluyen aquí aquellas otras lesiones a pequeña o mediana escala que afectan a varios pares de bases. Nuestro estudio pretende cubrir, esencialmente, los tipos de mutación puntual en cuanto a la modificación en la secuencia del DNA, como punto de partida para su influencia sobre la estructura y función proteicas.

Clases y tipos de mutaciones puntuales y medianas, y su grado de incidencia

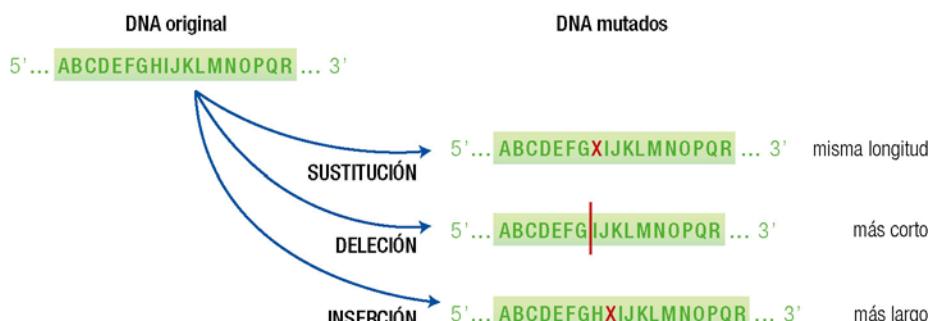
Clase	Tipo	Incidencia
Sustitución		Tipo de mutación comparativamente común en DNA codificante y no codificante
	Transiciones y transversiones	Más comunes las transiciones que las transversiones, sobre todo en DNA mitocondrial
	Sustituciones sinónimas y no sinónimas	Las sinónimas son mucho más comunes que las no sinónimas en DNA codificante; las conservadoras son más comunes que las no conservadoras
	Sustitución de múltiples bases	Es infrecuente, excepto en ciertos loci repetidos en tandem o repeticiones agrupadas
Inserción	De uno o unos pocos nucleótidos	Muy común en DNA no codificante pero infrecuente en DNA codificante
	Expansiones repetidas de tripletes	Infrecuente, pero puede contribuir a varios trastornos, especialmente neurológicos
	Otras inserciones grandes	Infrecuente; produce ocasionalmente duplicaciones en tandem e inserción de elementos transponibles
Eliminación o delección	De uno o unos pocos nucleótidos	Muy común en DNA no codificante pero infrecuente en DNA codificante
	Grandes delecciones	Infrecuente; ocurre a menudo en regiones que contienen repeticiones en tandem o entre repeticiones dispersas

Las mutaciones puntuales son tan representativas que, a veces, se asumen como el único tipo de mutación existente. Es importante resaltar que es la **estabilidad** del cambio lo que hace que la mutación se perpetúe en la progenie celular, al pasar por replicación a cada nueva copia y permanecer en las secuencias de DNA de cada individuo. Estas variaciones genéticas heredadas contribuyen a las diferencias entre individuos dentro de una población. En algunos casos, cuando la modificación se produce en una región génica (estructural o reguladora) puede ser suficiente para alterar su mensaje, afectando a la funcionalidad de su producto génico y conduciendo a una enfermedad hereditaria. Al depender de la alteración de un único gen, estas enfermedades se denominan **monogénicas**, y se caracterizan por transmitirse a la descendencia según las leyes de Mendel.

Como excepción, las mutaciones en el genoma mitocondrial se transmiten por herencia materna exclusivamente, debido a que el número de mitocondrias que aporta el espermatozoide en la formación del cigoto es muy inferior al procedente del óvulo.

23.4. I Cambios en la secuencia del DNA

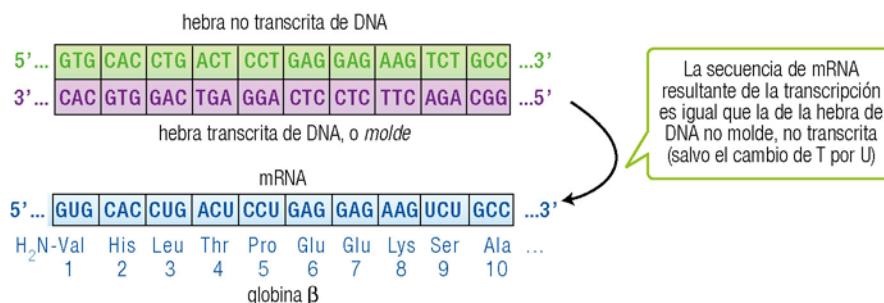
La descripción de los distintos tipos de mutaciones se hará atendiendo a su causa molecular en el DNA: mutaciones por sustitución, por inserción y por delección.



23.3

Este estudio se acompaña, en cada uno de sus apartados, de ejemplos específicos indicativos de las consecuencias patológicas de la mutación.

Para la descripción de las mutaciones puntuales y sus efectos sobre la proteína codificada se hará uso, siempre que sea posible, de la secuencia de 30 pb de DNA que codifica los primeros 10 aminoácidos de la proteína globina β humana (con un total de 147 aminoácidos):

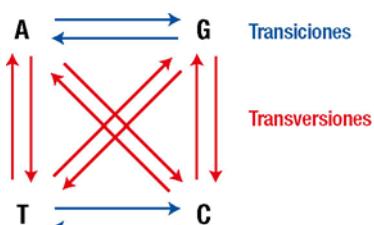


23.4

23.4. I. I Sustitución

Como su nombre indica, esta mutación consiste en un **cambio** de un nucleótido por otro. Es relativamente común en el DNA, tanto codificador como no codificador. En casos raros, pueden reemplazarse de forma simultánea varias bases agrupadas.

Según la naturaleza purínica o pirimidínica de las dos bases implicadas, se distingue entre transición y transversión.



23.5

a) Transiciones

Son sustituciones de una pirimidina por otra ($C \rightarrow T$; $T \rightarrow C$) o de una purina por otra ($A \rightarrow G$ o $G \rightarrow A$). Existen, pues, 4 posibilidades, una para cada base original:

23.6

transición	par inicial	par final
$A \rightarrow G$	A/T	G/C
$G \rightarrow A$	G/C	A/T
$T \rightarrow C$	T/A	C/G
$C \rightarrow T$	C/G	T/A

b) Transversiones

Son sustituciones de una purina por una pirimidina (A o $G \rightarrow T$ o C), o viceversa (T o $C \rightarrow A$ o G). Existen 8 posibilidades, dos por cada base original:

23.7

transversión	par inicial	par final
$A \rightarrow T$	A/T	T/A
$A \rightarrow C$	A/T	C/G
$G \rightarrow T$	G/C	T/A
$G \rightarrow C$	G/C	C/G
$T \rightarrow A$	T/A	A/T
$T \rightarrow G$	T/A	G/C
$C \rightarrow A$	C/G	A/T
$C \rightarrow G$	C/G	G/C

Teóricamente, si la sustitución de las bases fuera aleatoria, la frecuencia de una transversión en el genoma debería ser mayor que la de una transición, pues hay doble posibilidad de que ocurra. Sin embargo, se observa en diferentes procesos mutagénicos que las transiciones son más comunes que las transversiones, en especial en el DNA no funcional.

Como ejemplo específico puede citarse la transición de una base ($G \rightarrow A$) en la región 5' no traducida del gen del factor IX de la coagulación (región que contiene islotes CpG, págs. 167 y 282). El cambio de secuencia en esa región reguladora impide la unión de un factor de transcripción clave que, en condiciones normales, activa el promotor del gen. Como consecuencia, disminuye la expresión del gen y se sintetiza menos factor IX, lo que produce hemofilia de tipo B, un trastorno de la coagulación con una actividad coagulante disminuida en dos tercios.

23.4.1.2 Deleción, pérdida o eliminación

Consiste en la pérdida de uno o más nucleótidos de una secuencia. Es muy común en el DNA no codificante, pero infrecuente en el DNA codificante. Suelen ser mutaciones pequeñas, y sólo en raras ocasiones afectan a zonas lo suficientemente grandes como para ser visibles a escala citogenética (en regiones con repeticiones en tandem o entre repeticiones dispersas). Si se dan en una región codificante, normalmente producen un cambio en el marco de lectura (pág. 391), razón por la cual se comprende que sean tan infrecuentes: el efecto sobre la proteína es dramático.

23.4.1.3 Inserción

Se trata de la situación contraria a la anterior, la aparición de uno o varios nucleótidos adicionales en una secuencia. Es muy común en el DNA no codificador, pero rara en el DNA codificador (en el que, al igual que la delección, introduce un cambio de marco de lectura). Un ejemplo es la “repetición de trinucleótidos” o “expansión de tripletes”, un aumento en el número de copias de un triplete de nucleótidos observado en un gen o en sus proximidades.

23.4.2 Efecto sobre la secuencia de la proteína sintetizada

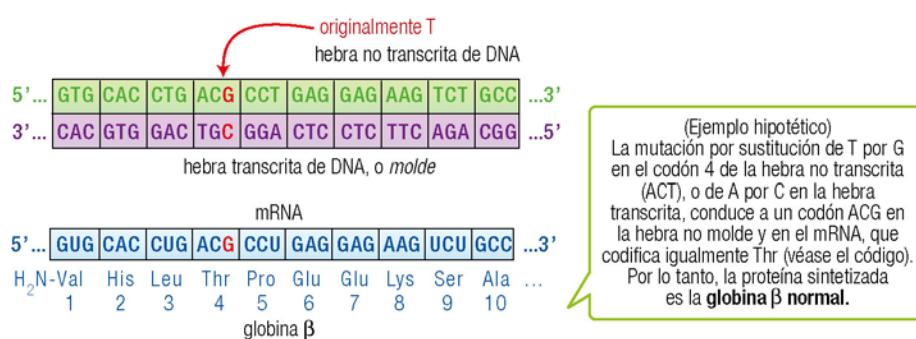
El cambio de secuencia del DNA generado por la mutación puede tener distintos efectos sobre la secuencia de aminoácidos codificada y, en consecuencia, sobre la proteína que se sintetice y su función. Se distingue por ello entre mutaciones *silenciosas* y *no silenciosas*.

Tipo de mutación	Causa posible		
	Sustitución	Delección	Inserción
Mutaciones silenciosas	✓		
Mutaciones no silenciosas	✓	✓	✓
De aminoácido	✓		
Conservadoras			
No conservadoras			
Desplazamiento del marco		✓	✓
Sin desplazamiento del marco		✓ (triple)	✓ (triple)
Terminación prematura	✓	✓	✓
Terminación retrasada	✓	✓	✓

(Cuando no se indica lo contrario, el cambio afecta a una sola base)

23.4.2.1 Mutaciones silenciosas

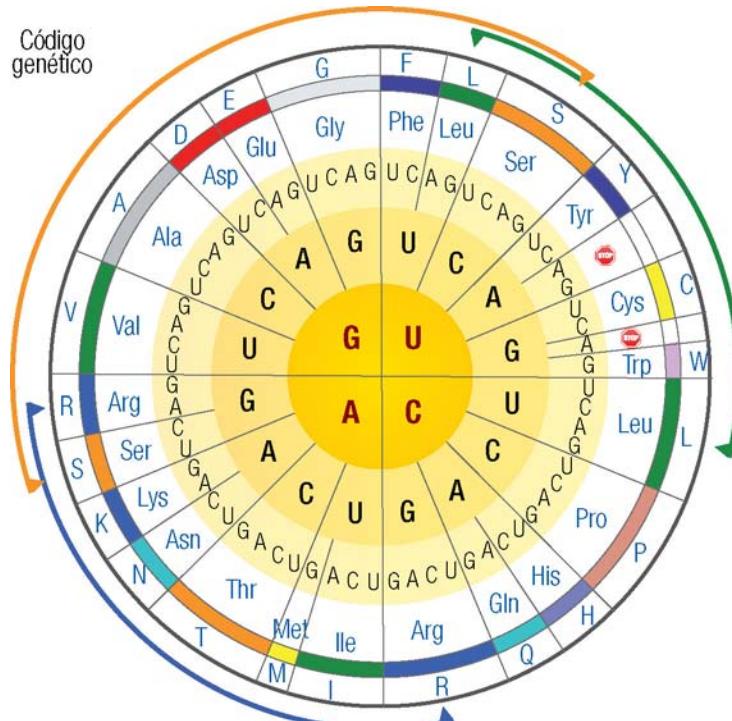
Se denominan también **sinónimas**, **neutrales** o **asintomáticas** (término no recomendable: mutaciones *con sentido*, calco del inglés *sense mutations*), y son muy numerosas (23-25% de todas las mutaciones en DNA codificador). Se caracterizan porque, a pesar de haber un cambio en la secuencia del DNA, no se altera la secuencia de su producto proteico. Esto es posible debido a la degeneración del código genético (varios tripletes o codones codifican un mismo aminoácido, pág. 318). Habitualmente, la mutación es silenciosa porque afecta a la tercera base de un codón de tal forma que el nuevo triplete es sinónimo del original, es decir, codifica el mismo aminoácido.



23.8

Igual ocurre si la sustitución de la base ocurre en la primera posición de determinados tripletes; concretamente, CUA \leftrightarrow UUA y CUG \leftrightarrow UUG, que codifican Leu (véase el código), y AGA \leftrightarrow CGA y AGG \leftrightarrow CGG, para Arg.

Obviamente, las mutaciones silenciosas pasan inadvertidas, pues no se detectan en el fenotipo ni confieren ventaja o desventaja alguna al organismo en cuyo genoma surgen. En consecuencia, no están sometidas a la presión de selección evolutiva ni influyen en la susceptibilidad frente a enfermedades. Sin embargo, pueden originar una enorme diversidad genética en una población, conocida como **polimorfismo genético**, en este caso no génico (Capítulo 24).



23.9

(se incluye aquí una copia del código genético para facilitar el análisis de los ejemplos)

23.4.2.2 Mutaciones no silenciosas

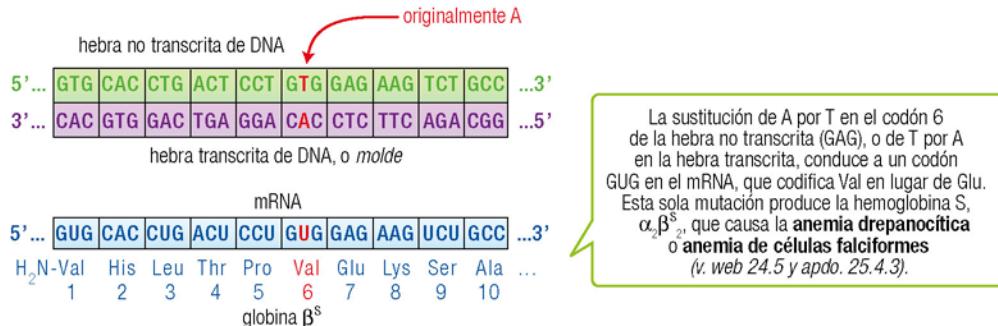
En este caso, la alteración en la secuencia de nucleótidos afecta a la secuencia proteica, pues codifica uno o varios aminoácidos diferentes de la secuencia original. Dependiendo del efecto que tenga la alteración de la secuencia proteica sobre su función, se puede distinguir entre mutaciones con efecto beneficioso y con efecto perjudicial. Aunque unas pocas mutaciones (menos del 1%) pueden tener un **efecto beneficioso o positivo**, por ejemplo, favoreciendo la función de la proteína codificada o la interacción entre genes, son tan poco frecuentes que, en la práctica, es común que no se consideren. En cualquier caso, los mecanismos que las generan son los mismos que si el efecto es perjudicial. Por el contrario, interesa estudiar con todo detalle las mutaciones no silenciosas que se manifiestan con efecto negativo o deletéreo.

A continuación, se comentan los efectos más frecuentes de las mutaciones no silenciosas sobre los codones y, en consecuencia, sobre la secuencia del producto proteico.

a) Mutaciones de aminoácido

Contrariamente a las mutaciones silenciosas, en las que la sustitución de una base origina un nuevo triplete que codifica el mismo aminoácido, es más común que dicho cambio produzca un codón nuevo que codifica un aminoácido diferente. Por ello se llaman **mutaciones de aminoácido o no sinónimas** (términos no recomendables: *de sentido equivocado, falso o alterado*, por calco del inglés *missense mutation*). La mayoría de las mutaciones en regiones codificantes son de este tipo (73%), por lo que son causa de muchos trastornos observados. Además, la formación de más de una versión proteica supone que el **polimorfismo génico** generado por estas mutaciones da lugar también a un **polimorfismo proteico** (Capítulo 24).

Lógicamente, estas mutaciones son particularmente relevantes cuando tienen lugar en **regiones génicas estructurales**. Por ejemplo, son de este tipo la mayoría de las mutaciones que producen hemoglobinopatías.



23.10

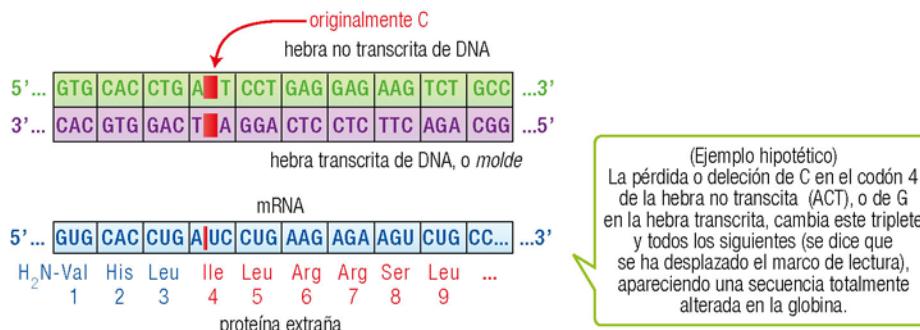
Dentro de las mutaciones de aminoácido se distinguen dos subtipos:

- **Sustitución conservadora.** El aminoácido se reemplaza por otro con una cadena lateral químicamente similar, por lo que el efecto sobre la estructura y función de la proteína puede ser mínimo. Esta situación se aproxima, pues, a las mutaciones silenciosas. En cierto modo, parece que el código genético se ha adaptado para reducir el efecto de las mutaciones pues, además de la degeneración, los codones que especifican aminoácidos similares son a su vez parecidos. Por ejemplo, los codones para Asp (GAC, GAU) y Glu (GAA, GAG) son muy parecidos, de modo que el cambio de la tercera base en un codón GAN (donde N es cualquier nucleótido) o bien es silencioso o tiene un efecto mínimo, el cambio de un aminoácido ácido por el otro. También pueden ser conservadores algunos cambios en la primera posición del codón, por ejemplo, CUN (Leu) \leftrightarrow GUN (Val), donde N representa cualquier nucleótido.
- **Sustitución no conservadora.** El aminoácido se sustituye por otro no similar, diferente en carga, polaridad, tamaño de la cadena lateral, etc. Son no conservadoras casi todas las sustituciones en la primera o segunda posición de un codón; por ejemplo, CGN (Arg) \Rightarrow GGN (Gly), CCN (Pro), CUN (Leu) o CAN (Gln/His), etc.

b) Mutaciones que cambian el marco de lectura

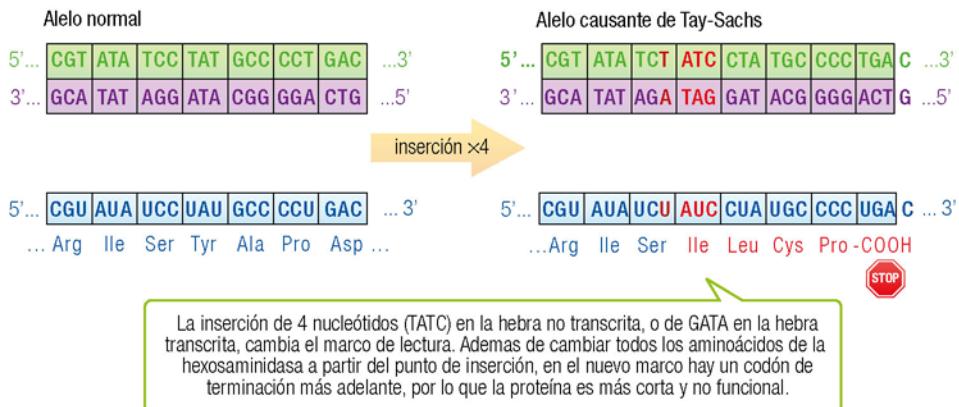
Éste es un concepto importante, por la frecuencia con la que afecta a la secuencia del producto proteico. El desplazamiento, desfase o cambio del marco de lectura (*frameshift mutation*) consiste en que la inserción o delección de nucleótidos, aun sin afectar en gran medida a la secuencia de bases, cambia la forma como se leen los tripletes (pág. 315), por lo que produce una alteración radical en la secuencia de la proteína.

Se produce un cambio de marco cuando hay inserciones o delecciones de uno o más pares de bases, en número no múltiplo de 3 (es decir, un número no entero de codones). El cambio en la secuencia de aminoácidos se produce a partir de la primera inserción o delección, hasta el extremo carboxilo-terminal de la proteína. Por ello, en general la proteína resulta muy alterada en su estructura.



23.11

Como ejemplo específico puede citarse la mutación más común encontrada en judíos asquenazíes (es decir, con antepasados de Europa central u oriental) que padecen la enfermedad de Tay-Sachs, un trastorno autosómico recesivo con acumulación de gangliósidos en los lisosomas neuronales. El gen afectado es el que codifica la subunidad α de la enzima lisosómica hexosaminidasa A (o isoenzima A de la β -N-acetilhexosaminidasa). La deficiencia enzimática total es responsable del fenotipo grave observado clásicamente en la forma infantil de la enfermedad.



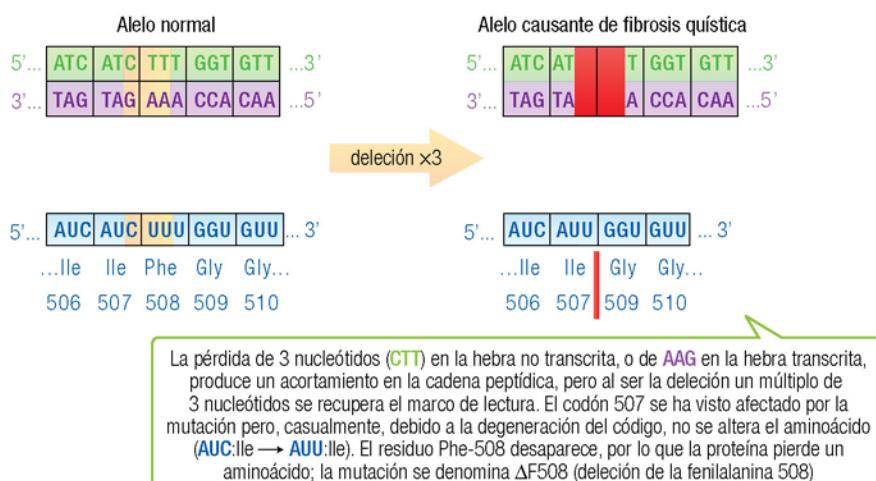
23.12

c) Mutaciones que no cambian el marco

La inserción o delección de 3 pb en el DNA (o múltiplos de tres), aunque añade o elimina algún aminoácido a la proteína, no cambia el marco de lectura, por lo que el resto de la secuencia de aminoácidos es normal. El efecto de estas mutaciones es, por tanto, mucho menor y puede no afectar a la función de la proteína.

Como ejemplo específico, en el 70% de todos los alelos asociados con la **mucoviscidosis**, o **fibrosis quística**, se observa una delección triple, en el cromosoma 7q. La mutación afecta al gen de una proteína de membrana, transportadora de aniones cloruro. La proteína mutada se pliega de manera incorrecta, por lo que durante su procesamiento postraduccional queda retenida en el lumen del retículo endoplásmico, unida a proteínas carabina (págs. 372-373); se impide así su llegada a la membrana plasmática, con consecuencias que afectan a varios tejidos y glándulas endocrinas (insuficiencia pancreática y afecciones respiratorias). La enfermedad se diagnostica por un exceso de cloruro en el sudor.

23.13

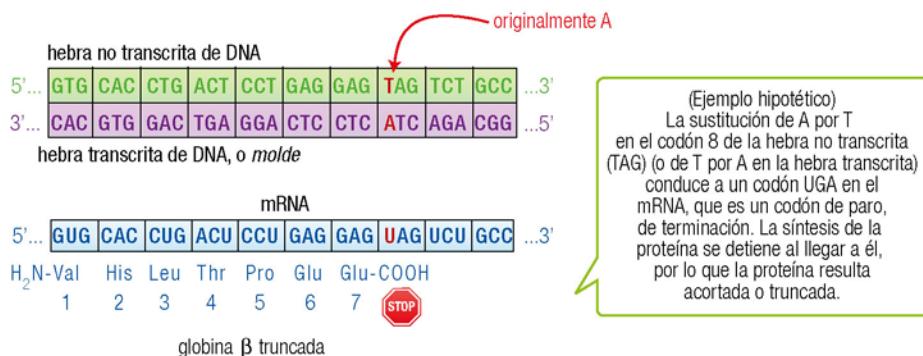


d) Mutaciones con terminación prematura de la proteína

Algunas mutaciones no afectan a la secuencia de aminoácidos codificados, sino que la sustitución, la inserción o la delección de uno o varios nucleótidos tiene como efecto la aparición de un **codón de terminación** o de **paro** (UAA, UAG o UGA en el mRNA citoplásmico de eucariotas superiores, pág. 317). Como consecuencia, se provoca el cese prematuro de la traducción del mRNA y se obtiene una proteína acortada, truncada (con carencia de un segmento C-terminal).

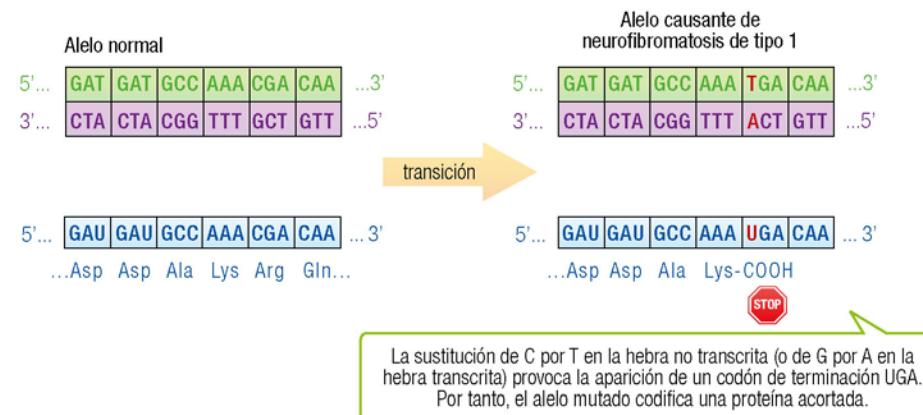
Este tipo de **mutaciones terminadoras** o finalizadoras es poco frecuente; por ejemplo, suponen cerca del 4% de las sustituciones en el DNA codificante. (No es recomendable el término *mutación sin sentido*, calco de

nonsense mutation, pues sí hay un significado: la terminación de la elongación.) El acortamiento de la proteína es a menudo considerable, por lo que se ven afectadas dramáticamente su estructura tridimensional, su estabilidad y su función. Con frecuencia esta alteración acelera la degradación intracelular de la proteína no funcional, ocasionando generalmente un efecto fenotípico.



23.14

Como ejemplo específico cabe citar la mutación causante de la neurofibromatosis de tipo 1 (NF1) o enfermedad de Recklinghausen, un trastorno autosómico dominante (pág. 437) localizado en el cromosoma 17q.



23.15

La mutación puede producir de manera directa la aparición del codón de terminación, como en los ejemplos mostrados, pero también puede hacerlo de forma indirecta, mediante un cambio en el marco de lectura (véase, como ejemplo, el polimorfismo de grupo sanguíneo AB0, pág. 411).

Otro ejemplo similar, aunque no se trata de una mutación sino del resultado de un mecanismo biológico regulado, es la edición del mRNA de la apolipoproteína B (pág. 306), que tiene como consecuencia la conversión de un codón CAA en uno UAA y la formación de la proteína truncada apoB48 en lugar de la apoB100.

e) Mutaciones con terminación retrasada

De la misma forma que una mutación puntual puede provocar la aparición de un codón de terminación, también puede eliminar uno existente, sustituyéndolo por otro que codifique un aminoácido. En ese caso, la traducción del mRNA continúa más allá del extremo 3' codificante original, hasta que se alcance un nuevo codón de terminación. Por tanto, se sintetiza una proteína de mayor longitud, cuya secuencia de aminoácidos en la región C-terminal no tiene significado estructural ni funcional. Las consecuencias fenotípicas son, en general, similares a las de la terminación prematura.

23.5 CAUSAS Y MECANISMOS BÁSICOS DE LAS MUTACIONES

El DNA está sometido a daños, lesiones y modificaciones de diversos tipos, provocados por distintas causas. Éstas pueden ser fortuitas o endógenas y provocadas, inducidas o exógenas.

23.5.1 Mutaciones endógenas

Las mutaciones más frecuentes se generan debido a situaciones o agentes propios del ambiente intracelular; bajo condiciones normales. Pueden originarse por errores en la replicación, o bien por reacciones que ocurren de forma espontánea sobre la molécula de DNA, o por la acción de subproductos del metabolismo celular (mutágenos endógenos).

23.5.1.1 Incorporación de nucleótidos erróneos durante la replicación

La replicación por la DNA polimerasa, moviéndose a lo largo del DNA, combina la incorporación de nucleótidos a la hebra nueva con la corrección de pruebas, de forma que duplica la doble hélice con una gran fidelidad. Más del 99,9% de los errores de replicación se corrigen rápidamente mediante enzimas reparadoras que reconocen el nucleótido incorrecto en la hebra que se está sintetizando y lo reemplazan por uno correcto, complementario al de la hebra molde. Como consecuencia, sólo se introduce una variación por cada 10^9 a 10^{11} nucleótidos incorporados en cada ciclo de división celular, lo que equivale a una mutación cada 10.000 Mb. Puesto que el genoma diploide humano tiene 6.400 Mb, esto se traduce en **menos de una nueva mutación por cada división celular**. A pesar de ello, el gran número de divisiones celulares durante la vida de un individuo hace que el número de mutaciones sea significativo.

Una de las causas de incorporación errónea de nucleótidos se encuentra en la tautomería de las bases nitrogenadas (pág. 16); las formas tautoméricas menos abundantes (imina y lactima) establecen puentes de hidrógeno diferentes, lo que permite el emparejamiento de un nucleótido incorrecto en la hebra de nueva síntesis. Además de los errores en la identidad del nucleótido incorporado, durante la replicación también se introducen mutaciones debido a deslizamientos de la DNAPol sobre la hebra molde, que producen delecciones o inserciones por repetición de una pequeña secuencia.



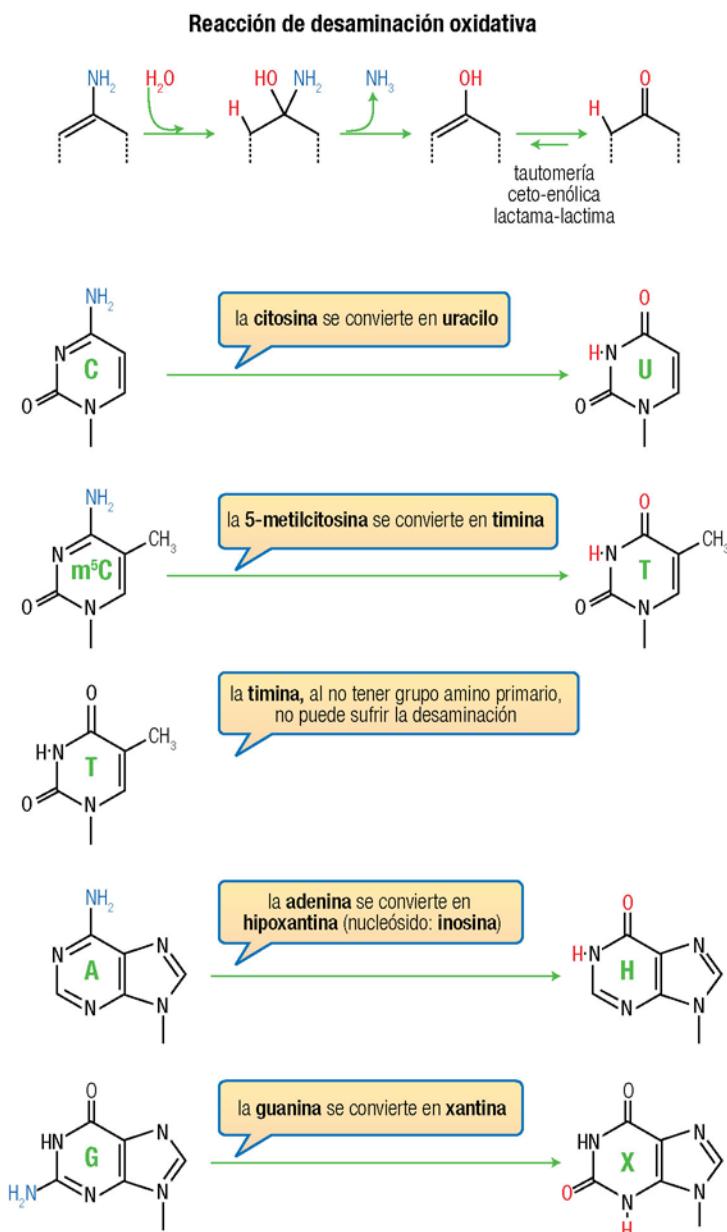
Web 23.1. Mutaciones en la secuencia del DNA inducidas por alteraciones de las bases.

23.5.1.2 Otras mutaciones espontáneas

Las bases en el DNA pueden experimentar espontáneamente reacciones (sin intervención de enzima alguna) que transforman su identidad, de modo que si no son corregidas por los sistemas de reparación se convierten en una mutación.

a) Sustituciones por desaminación oxidativa

Consisten en la pérdida de un grupo amino exocíclico, con aparición de un grupo carbonilo. Ello da lugar a que se altere la formación de puentes de hidrógeno entre las bases. (La desaminación interviene también en la riboedición, pero en ese caso está catalizada enzimáticamente y cumple una función, pág. 306.)



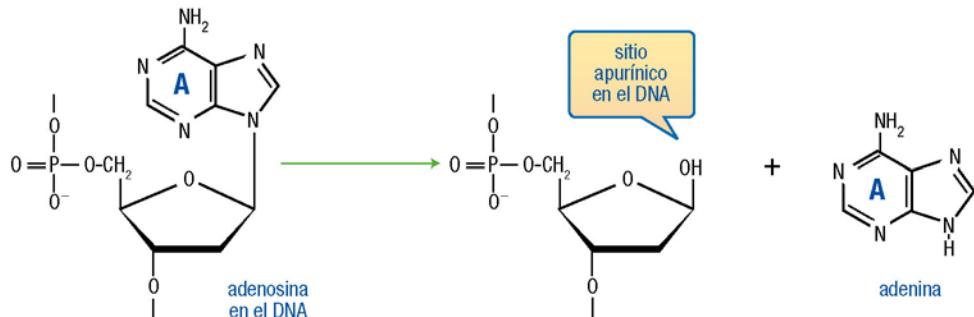
23.16

Las transformaciones $\text{C} \rightarrow \text{U}$ y $5\text{-metil-C} \rightarrow \text{T}$ se producen lentamente, pero unas 100 veces más rápidamente que las $\text{A} \rightarrow \text{H}$ y $\text{G} \rightarrow \text{X}$. Por ejemplo, habrá una desaminación de citosina al día por cada 10^7 C en el DNA de una célula típica de mamífero (lo que equivale a unas 100 mutaciones por día).

Debe observarse que la única base habitual ($\text{A}, \text{G}, \text{C}, \text{T}$) que se transforma por desaminación en otra base igualmente habitual es la citosina (v. web 23.1). Sin embargo, su producto (U) puede reconocerse como ajeno al DNA y corregirse la mutación. Ésta es una ventaja de la presencia en el DNA de T : si el DNA contuviese uracilo, no se podrían distinguir los residuos de uracilo “correctos” de los generados por la desaminación de citosina, y el mensaje genético sería por ello menos estable.

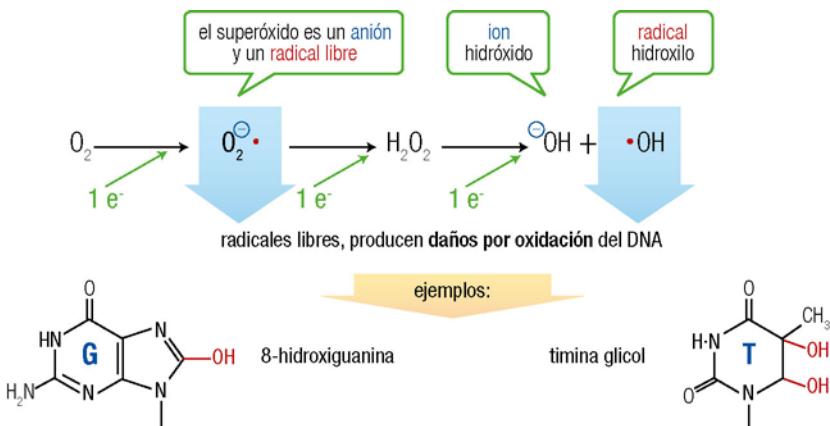
b) Pérdida de bases por inestabilidad del enlace N-glicosídico

Como ejemplo, pueden citarse las despurinaciones espontáneas, en las que se pierde una purina por hidrólisis del enlace N-glicosídico, dando lugar a un sitio apurínico (v. web 23.1). Análogamente, pueden generarse sitios apirimidínicos, aunque con menos frecuencia. En la posición afectada, el nucleósido queda reducido únicamente al esqueleto, es decir, un residuo de desoxirribosa unido por ambos enlaces fosfodiéster.



c) Modificación de las bases por mutágenos endógenos

El propio metabolismo de la célula puede generar compuestos químicos que reaccionan con las bases, o en general con los nucleósidos, modificándolos. El caso más frecuente son las especies reactivas de oxígeno, radicales libres producidos por el metabolismo oxidativo de la célula, en especial en la mitocondria. Las principales son el anión superóxido y el radical hidroxilo, que dañan el DNA rompiendo una hebra u oxidando las bases nitrogenadas.



23.5.2 Mutaciones inducidas o exógenas: mutagénesis

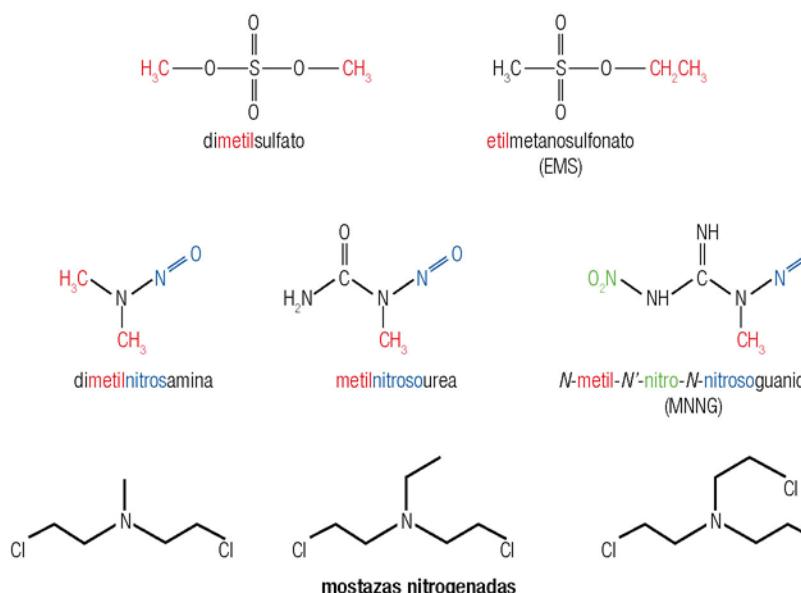
Además de las transformaciones espontáneas de nucleótidos y de los errores de replicación, se producen también mutaciones debidas a agentes fisicoquímicos ajenos a la célula, denominados **mutágenos exógenos**.

23.5.2.1 Lesiones y mutaciones inducidas por agentes químicos

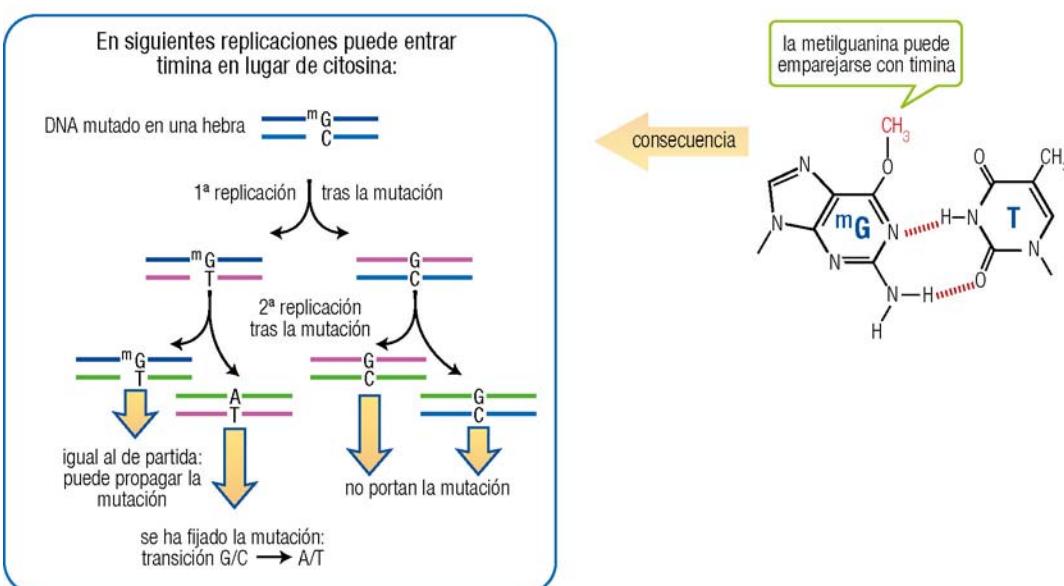
A diferencia de la inestabilidad del DNA ya discutida, producida por cambios químicos espontáneos en los nucleótidos (esencialmente por desaminación o despurinación), también pueden modificarse químicamente las bases por la acción directa de diversas sustancias químicas del ambiente (naturales o no). La lista de estos agentes mutágenos es muy amplia, y los mecanismos y efectos que producen muy diversos. Muchos de ellos se identificaron hace más de dos siglos como responsables de cánceres en obreros en contacto con el humo y hollín. Algunos actúan como mutágenos directos y otros requieren activarse a carcinógenos activos mediante la acción de ciertas enzimas.

- **Agentes químicos alquilantes**

Transfieren grupos, generalmente metilo o etilo, a un átomo nucleófilo (O o N) de las bases nitrogenadas. Entre los mejor estudiados se encuentran:



23.19



23.20

• Agentes intercalantes

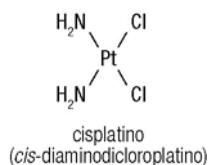
Estos compuestos, con estructura plana e hidrófoba (pág. 141), se insertan entre los pares de bases apilados del DNA y distorsionan su estructura helicoidal. Algunos provocan el desplazamiento del marco de lectura.

- Naranja de acridina, proflavina, etc.
- Benzo[a]pireno, uno de los productos que se encuentran en el humo.
- Otros hidrocarburos aromáticos policíclicos: 2-acetamidofluoreno, *N*-metil-4-antraceno.
- Compuestos naturales: actinomicina D, aflatoxina.

• Agentes entrecruzantes

Reaccionan con el DNA y establecen enlaces cruzados entre bases de la misma hebra o de ambas hebras.

- Compuestos de *cis*platino



23.21

• Otros agentes químicos

- 5-bromouracilo: se incorpora en lugar de T durante la replicación, gracias a su analogía estructural (v. web 10.9). BrU adopta con mayor facilidad que T la forma tautomérica enol (pág. 16), que empareja con G en lugar de A; como consecuencia, la incorporación de BrU provoca mutaciones de par T=A a par C≡G (v. web 23.1).
- Ácido nitroso (HNO_2): acelera la desaminación oxidativa $\text{C} \rightarrow \text{U}$.
- Compuestos que pueden metabolizarse a ácido nitroso, como nitritos o nitratos.
- Aminas aromáticas: *N*-metil-4-aminoazobenceno (MAB).
- Varios fármacos: ciclofosfamida, dietilestilbestrol.
- Compuestos inorgánicos: arsénico, asbestos, berilio, cadmio, cromo.

23.5.2.2 Lesiones y mutaciones inducidas por agentes físicos

Son varios los agentes que originan lesiones por exposición, con efectos mutágenos o cancerígenos.

La **radiación ultravioleta**, componente de la luz solar, provoca la reacción entre dos residuos de timidina adyacentes en la misma hebra, formando a partir de sus dobles enlaces anulares un sistema tricíclico con un anillo de ciclobutano, que se conoce como **dímero de timina**. Esta asociación covalente de dos T impide que puedan emparejarse con sus A complementarias y origina una estructura tridimensional localmente incompatible con el giro de la doble hélice, que fuerza una flexión o “codo” en la molécula de DNA. Además, se paraliza la replicación, pues la DNA polimerasa no puede reconocer los dímeros y se detiene. La irradiación UV también origina otra reacción entre T adyacentes, con formación de dímeros diferentes llamados **fotoproductos**, en los que el enlace aparece entre el C-6 de una T y el C-4 de la adyacente.

www

Web 23.2. Estructura de los dímeros de timina.

Las **radiaciones ionizantes** (rayos X y rayos γ) pueden provocar apertura de los anillos, fragmentación de las bases, así como rotura del esqueleto covalente del DNA. También generan (por fotólisis del agua) radicales hidroxilo, que producen 8-hidroxiguanina, entre otros. Debido a su mayor poder de penetración, estas radiaciones afectan a todo tipo de tejidos, a diferencia de la luz UV, que provoca lesiones solamente en la piel.

23.6 REPARACIÓN DEL DNA

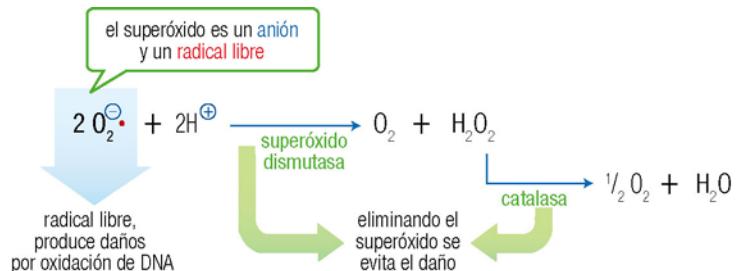
La capacidad de una célula para detectar y reparar daños en el DNA es esencial para el mantenimiento de la integridad de su genoma. Existen, con este objeto, diversos sistemas enzimáticos *in vivo*, que pueden llegar a ser tan complejos como el propio aparato de replicación, lo que pone de manifiesto su gran importancia para la supervivencia celular. De hecho, la velocidad de mutación refleja el balance entre el número de errores ocurridos y los que se corrigen: una baja tasa de mutación indica una gran eficiencia de los sistemas de reparación, mientras que un fallo de éstos conduce a un aumento del número de mutaciones.

En general, no puede establecerse una estrategia única de reparación por parte de la célula, sino que existen distintas vías.

Tipo de reparación	Ejemplo	Tipo de lesión reparada
Eliminación de mutágenos (destoxicificación)	Superóxido dismutasa	Se evita la formación de la lesión oxidativa
Reversión directa	Fotoliasa	Fotodímeros de pirimidina
	Alquiltransferasas	Derivados alquilo de las bases
Escisión de nucleótidos (general)	Sistema de la escinucleasa	Las que causan distorsión en la doble hélice (fotoproductos y sustituyentes voluminosos)
Escisión de bases (específica)	N-glicosilasas del DNA y endonucleasas AP	Bases desaminadas, modificadas, oxidadas, etc., y sitios abásicos
Reparación de emparejamientos incorrectos		Nucleótidos mal emparejados por errores en la replicación o bases modificadas

23.6.1 Eliminación de agentes mutágenos

El primer mecanismo para reducir el daño al DNA no es propiamente una vía de reparación, sino que consiste en evitar las lesiones antes de que se produzcan, mediante la neutralización de compuestos mutágenos. Uno de tales mecanismos de destoxicificación lo ejerce la superóxido dismutasa, que evita la acumulación de radicales superóxido, agentes causantes de lesiones oxidativas en el DNA.

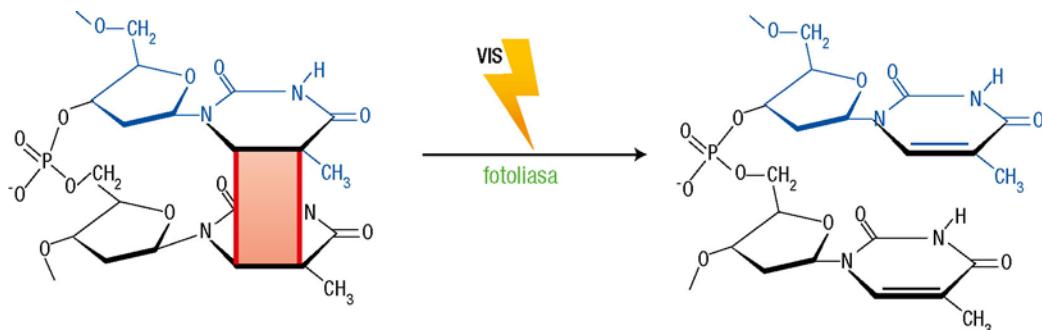


23.6.2 Reversión directa de la lesión

La forma más inmediata de reparar una mutación es la regeneración de la base dañada revirtiendo la transformación química que ha sufrido. Este mecanismo no es factible en la mayoría de los casos, a excepción de los ejemplos siguientes.

23.6.2.1 Reparación de fotodímeros

Ya se ha explicado que la luz ultravioleta provoca la formación de anillos de ciclobutano entre dos pirimidinas adyacentes (pág. 398), lo que supone una distorsión estructural que impide físicamente la replicación y la transcripción. El mecanismo de reparación se basa en el reconocimiento del fotodímero por la enzima fotoliasa, que lo escinde, mediante una reacción de transferencia electrónica iniciada por absorción de luz visible, regenerando sus dos bases originales. La fotoliasa tiene como coenzimas FADH₂ y un cromóforo que absorbe la luz e inicia la transferencia de electrones. Este requerimiento de luz (azul o violeta) hace que también reciba el nombre de **enzima fotorreactivadora**. En condiciones de oscuridad se requieren otros sistemas para reparar estas lesiones. Este sistema se ha encontrado en algunas especies, pero está ausente en otras; así, existe en bacterias, hongos, la levadura *Saccharomyces cerevisiae*, peces, reptiles, anfibios y mamíferos marsupiales, pero no aparece en la levadura *Schizosaccharomyces pombe* ni en mamíferos placentarios, incluyendo a los humanos. En algunos organismos, incluida *Drosophila*, se ha detectado otra fotoliasa que repara los fotoproductos 6-4, ausente a su vez en *E. coli*, levaduras y humanos.

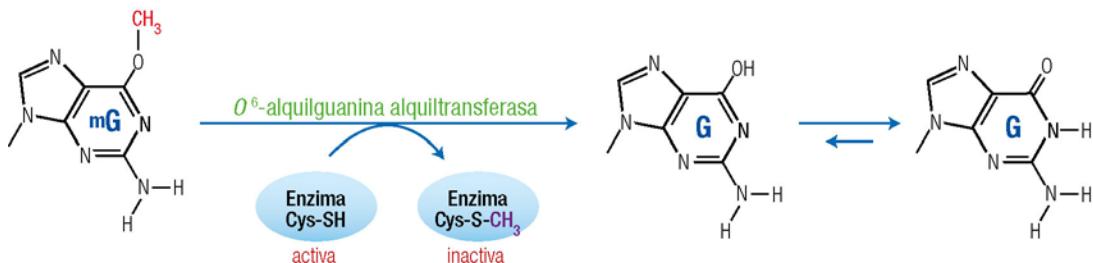


23.23

23.6.2.2 Reversión de bases modificadas con grupos alquilo

El caso más conocido, muy bien caracterizado en *E. coli* e igualmente presente en todas las especies estudiadas (incluida la humana), es la reparación de O⁶-metilguanina y O⁶-etilguanina, productos de la modificación del DNA por agentes mutágenos alquilantes como nitrosoguanidinas, metilnitrosourea y etilmetasulfonato (pág. 397).

El grupo alquilo de la base lo elimina una alquiltransferasa, que lo transfiere a un residuo cisteína de su propio centro activo. Éste no puede regenerarse, por lo que la “catálisis” inactiva la propia enzima (con lo que estrictamente no es enzima ni catalizador, pues no se cumple la condición de que no se altere como consecuencia de la reacción que acelera). Dada esta inactivación, la célula debe sintetizar nuevas moléculas de alquiltransferasa para mantener el proceso reparador, por lo cual el sistema enzimático de reparación puede saturarse si existe un número excesivo de mutaciones alquilantes.



23.24

23.6.3 Reparación por escisión

Se conoce bajo este nombre genérico a la reparación de lesiones basada en la eliminación de la base nitrogenada, nucleótido o segmento de DNA defectuosos y su reemplazo por otro nuevo, con secuencia correcta. A diferencia de la reparación por reversión, éste es un proceso complejo, que tiene lugar en varias fases, que utiliza varias proteínas (sólo algunas específicamente reparadoras) y que lleva siempre la síntesis de un fragmento de DNA (lo que se denomina *síntesis fuera de programa*). El proceso se inicia a través del reconocimiento molecular de la lesión (base alterada o cambio en la estructura tridimensional del DNA).

De forma global, en el mecanismo de reparación se pueden distinguir 3 etapas, la primera distinta para los mecanismos “general” y “específico”, y las dos restantes comunes.



Web 23.3. Reparación por escisión.

23.6.3.1 Mecanismo general de reparación por escisión de nucleótidos

En este mecanismo (NER, *nucleotide excision repair*) intervienen proteínas que reconocen distintos tipos de alteraciones que causan distorsión en la doble hélice, como los dímeros de timina o una base con un sustituyente voluminoso (psoraleno, cisplatino, acetamidofluoreno, etc.). En humanos, es el sistema de defensa más importante contra el daño al DNA causado por el humo del tabaco, agente responsable de más del 30% de las muertes por cáncer en el mundo. El sistema NER actúa no sólo en la reparación global del genoma, sino también en una actividad de reparación acoplada a la transcripción, que revisa el DNA justo antes de ser transcrita.



Web 23.4. Reparación por escisión de nucleótidos.

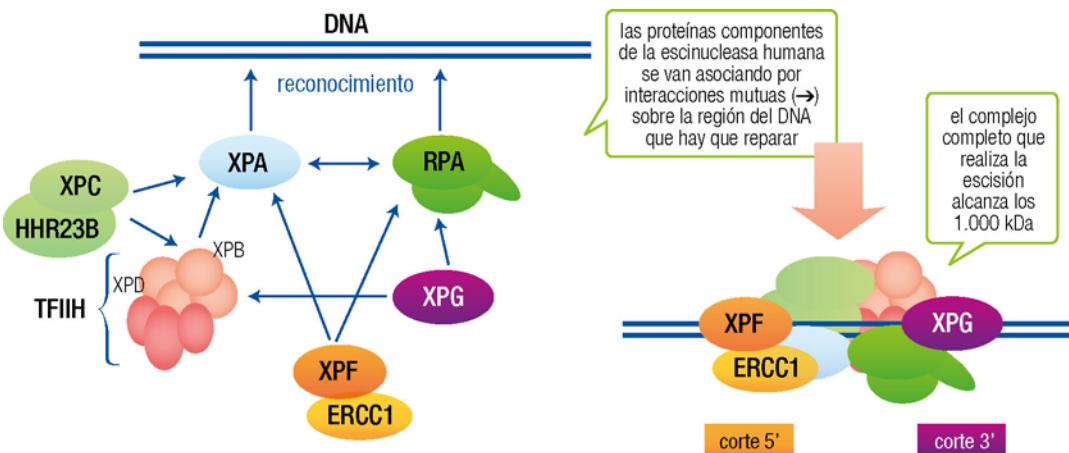
- El proceso comienza con el **reconocimiento de la lesión** por unión de proteínas componentes del complejo de reparación, con gasto de ATP.
- En la **fase de escisión** un complejo proteico que incluye endonucleasas hidroliza sendos enlaces fosfodiéster de la hebra lesionada a ambos lados del punto de mutación (pero no junto a él; concretamente, a -22 y +6 en humanos). La peculiaridad de cortar sólo una de las hebras y en dos puntos separados la distingue de otras nucleasas, por lo que se denomina **escinucleasa**. A continuación, una **helicasa** favorece la separación del oligonucleótido liberado (de 12 o 13 nt en procariotas y unos 28 nt en eucariotas). Algunas de las subunidades de la escinucleasa permanecen unidas al tramo monocatenario hasta que sean desplazadas por la maquinaria de replicación.
- A continuación, en la **fase de “síntesis reparadora”**, una **DNA polimerasa** rellena el hueco dejado por la escinucleasa, mediante la elongación del extremo 3'-OH formado en el corte anterior. El uso de la otra hebra como molde asegura que se introduce la secuencia correcta. En eucariotas, realizan esta síntesis las polimerasas δ y ε, junto con sus factores accesorios PCNA y RFC (web 11.8).
- Finalmente, una vez que la polimerasa ha llenado todo el hueco y sólo queda una mella, una **DNA ligasa** une el extremo 3'-OH recién sintetizado con el extremo 5'-P formado en el corte de la hebra original, cerrando la mella (**fase de sellado o ligamiento**).

Este mecanismo de reparación posee la particularidad de estar relacionado con la transcripción, gracias a que participan proteínas comunes (principalmente, el factor de transcripción TFIIH, pág. 304). La RNAPol II detenida en el punto de lesión puede actuar de señal para el complejo de reparación, y además se observa una reparación preferente de las regiones de DNA que se transcriben y, dentro de ellas, de la hebra molde.

La escinucleasa humana es un gran complejo proteico (cerca de 1.000 kDa en total) que se puede considerar formado por 6 subunidades funcionales que se van asociando y separando durante el proceso, constituidas por un total de 16 cadenas polipeptídicas. Si se suman las proteínas responsables de la replicación y sellado, en total participan en la reparación por escisión de nucleótidos unos 30 polipéptidos diferentes.

Subunidad	N.º de cadenas polipeptídicas	Función
XPA	1	Reconoce la lesión en el DNA y se une a él. Posee un dedo de zinc del tipo C4 (pág. 58).
XPC+HHR23B	2	Reconocen la lesión en el DNA y se unen a él.
RPA	3	Reconoce la lesión en el DNA y se une a él. Es la misma proteína ligante de DNA monocatenario que interviene en la replicación.
TFIIH	7	Es el mismo factor de transcripción general de los genes de clase II. Actúa mediante su actividad helicasa (TFIIH), separando las hebras del DNA en la región a reparar. Su actividad proteína quinasa (TFIIK) no es necesaria aquí.
Helicasa:		
XPB		
XPD		
p62		
p44		
Proteína quinasa:		
p41		
Quinasa dependiente de ciclinas CDK7		
p38		
Ciclina H		
p34		
XPG	1	Endonucleasa que corta el extremo 3' de la región escindida. Presenta homología de secuencia con la endonucleasa FEN1 (web 11.9).
XPF+ERCC1	2	Endonucleasa que corta el extremo 5' de la región escindida.

Nota: los nombres XPA-XPG proceden de xeroderma pigmentosa (pág. 404).



23.25

23.6.3.2 Mecanismos específicos de reparación por escisión de bases

Este sistema (**BER**, *base excision repair*) se encarga de reparar bases específicas; por tanto, reconoce y actúa en sitios concretos, reparando lesiones que son demasiado sutiles para distorsionar gravemente el DNA, es decir, que no reconoce por el sistema general de escisión de nucleótidos. Por ejemplo, puede corregir daños por hidrólisis (desaminación o pérdida de base), oxidación por radicales libres de oxígeno, alquilación, y derivados no voluminosos como timina glicol, N³-metiladenina y 8-hidroxiguanina.



Web 23.5. Reparación por escisión de bases.

- La fase de escisión se diferencia de la reparación general porque, en primer lugar, se elimina la base alterada o incorrecta y no su nucleótido completo. Una **N-glicosilasa de DNA** reconoce un solo tipo de base, anómala o emparejada de manera incorrecta, e hidroliza su enlace N-glicosídico con la desoxirribosa, dando lugar a un **sitio abásico** (*apurimidínico* o *apurínico*, según sea el caso). En todas las células existen varias N-glicosilasas, específicas para bases distintas y que, por tanto, reparan lesiones diferentes.
- Una vez generado el sitio abásico, en una **segunda subetapa de la fase de escisión** se produce el corte de la hebra de DNA junto a esa desoxirribosa que no está unida a una base nitrogenada. Entran en juego para ello las **endonucleasas AP** (de *apurínico* y *apurimidínico*), que hidrolizan un enlace fosfodiéster, originando una mella; en eucariotas (levaduras y mamíferos) cortan siempre al lado 5' del sitio abásico, generando un extremo nucleótido-3'-OH y otro 5'-fosfodesoxirribosa.
- A partir de este momento, la reparación continúa de forma similar al mecanismo general, gracias a la intervención de una **DNA polimerasa** que elonga sustituyendo el nucleótido abásico y, en ocasiones, varios más (e introduce la secuencia correcta por complementariedad con la otra hebra). En eucariotas se han encontrado dos mecanismos:
 - Reparación de un solo nucleótido, mediante la DNAPol β, que posee dos actividades enzimáticas: **desoxiribofosfodiesterasa**, que elimina el nucleótido abásico, y **polimerasa**, que lo sustituye por uno completo. Finalmente, una **ligasa** completa la reparación cerrando la mella.
 - Reparación de un tramo más largo, realizada por las **polimerasas δ o ε**, junto con sus factores accesorios PCNA y RFC, y la **endonucleasa FEN1** (web 11.8 y web 11.9). Finaliza, igualmente, con el sellado de la mella por una ligasa.

Como ejemplo de N-glicosilasa, la **uracilo-DNA glicosilasa** es una enzima muy específica que no actúa sobre residuos U del RNA, pero que recorre el DNA reconociendo como extraño el uracilo (que sólo se diferencia de la timina en la carencia de grupo metilo). En las células se reduce la posibilidad de que U se

incorpore en la síntesis del DNA gracias a una **dUTP pirofosfatasa** que elimina el nucleótido precursor ($\text{UTP} \rightarrow \text{UMP} + \text{PP}_i$). Sin embargo, en el DNA ya sintetizado se forma U con gran frecuencia por la desaminación espontánea de C (pág. 394). Si el DNA estuviese formado por U (en lugar de T) sería imposible diferenciar los residuos U formados a partir de C, y la mutación U → C no podría repararse. Gracias a que el componente del DNA es T, y no U, es posible la reparación de la mutación frecuente U → C y el mantenimiento del par A=T. Este mecanismo se considera una de las ventajas por las que el DNA (con A, G, T, C) ha podido evolucionar como un mejor depositario de la información genética. Sin embargo, no ocurre lo mismo con las citosinas metiladas, relativamente comunes en el DNA eucariótico (pág. 282), cuya desaminación oxidativa produce T, que no puede ser distinguida de las T propias del DNA.

Otras glicosilasas de DNA reparan de forma análoga diversas bases modificadas:

Enzima	Actúa sobre	Lesión que repara
Uracilo-DNA glicosilasa	U	Desaminación de C
3-metiladenina-DNA glicosilasas	3-Me-A, 3-Me-G, 7-Me-G, O ² -Me-C, O ² -Me-T, hipoxantina	Metilación de A, G, C o T, Desaminación de A
Formamidopirimidina-DNA glicosilasa	8-hidroxiguanina, formamidopirimidinas	Oxidación de G
Hidrato de pirimidina-DNA glicosilasa	timina glicol, pirimidinas oxidadas o fragmentadas	Oxidación de pirimidinas

(*) Me = metil.

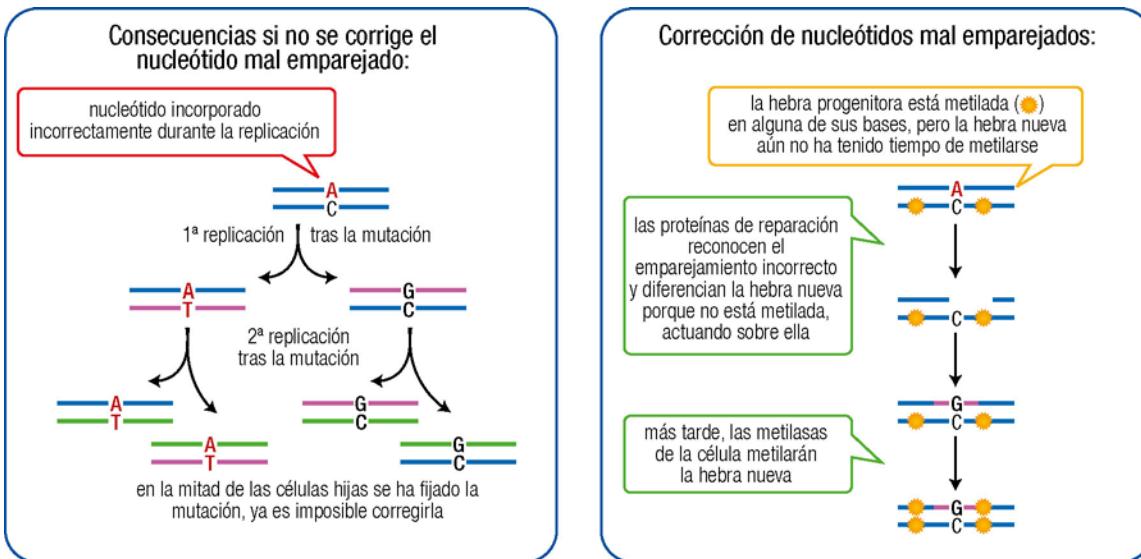
Finalmente, debe mencionarse que el reconocimiento de los sitios AP es vital no sólo para completar la acción previa de las N-glicosilasas, sino también para corregir la despurinación espontánea, que es un suceso relativamente frecuente (pág. 395).

23.6.4 Reparación de emparejamientos incorrectos

Habitualmente, los mecanismos de reparación anteriores actúan al poco tiempo de que se ha producido la mutación, antes de que ésta se perpetúe debido a la replicación. Sin embargo, existen además otros procesos, denominados de **reparación posreplicación**, que corrigen errores cometidos en la replicación o errores que la impiden (actúan, por tanto, una vez por cada división celular). Entre ellos se encuentran la **reparación por recombinación** y la **reparación de emparejamientos incorrectos**. Comentaremos solamente esta última.

Los emparejamientos incorrectos se producen por errores de la replicación o como consecuencia de recombinación; el mismo sistema es también capaz de corregir pequeñas inserciones de 1 a 4 nucleótidos en una de las hebras. Normalmente, la mayoría de los errores cometidos inicialmente por la DNA polimerasa se rectifican gracias a su propia actividad correctora de pruebas (exonucleasa 3', pág. 150), pero unos pocos pueden permanecer sin corregir. Es entonces cuando entra en juego este mecanismo, formado por proteínas que reconocen la presencia de un par de bases incorrectamente emparejadas y eliminan la base errónea o bien un tramo de la hebra de DNA que la contiene. Éstas son luego **repuestas** por los mecanismos de reparación ya estudiados, con intervención de las DNA polimerasas δ y ε y sus cofactores RFC y PCNA. En humanos se han podido identificar 5 genes que codifican proteínas de este sistema de reparación, llamados *hMSH2*, *hMLH1*, *hPMS1*, *hPMS2* y *GTBP/hMSH6*.

La cuestión importante que surge es: ¿cómo determina la maquinaria de reparación cuál de las dos bases de un par incorrecto es la errónea, la que hay que eliminar? Parecen existir varios mecanismos para esto, pero uno de ellos se basa en la **metilación** del DNA (pág. 282). Las enzimas encargadas de la metilación de bases tardan un tiempo en actuar, por lo que en el DNA bicatenario recién replicado la hebra nueva aún no está metilada, mientras que la hebra progenitora sí lo está. Esto permite a las proteínas reparadoras diferenciar la hebra nueva, que ha de ser la que contiene la mutación, y corregir ésta. Este mecanismo se conoce bien en *E. coli*, y posiblemente actúe bajo principios similares en eucariotas.



23.7 ENFERMEDADES HUMANAS ASOCIADAS A LA REPARACIÓN

Se conocen varias enfermedades atribuidas a defectos en genes que codifican proteínas de reparación del DNA. Estos defectos conducen con frecuencia a una mayor incidencia del cáncer, pues se produce una acumulación de mutaciones no corregidas.

El cáncer de piel denominado **xerodermia pigmentosa, o pigmentaria (XP; también, xerodermia de Kaposi o melanosis lenticular progresiva)** es una enfermedad hereditaria causada por una deficiencia en alguno de los genes responsables de la reparación por escisión de nucleótidos, conduciendo principalmente a la acumulación de dímeros de timina. La piel de estos enfermos es extremadamente sensible a la exposición a los rayos ultravioleta de la luz solar, ocasionándoles cánceres de piel con una frecuencia 1.000 o 2.000 veces superior a la de personas normales. También sufren una incidencia entre 10 y 20 veces mayor de cánceres en órganos internos, y síntomas neurológicos como retraso mental, ataxia progresiva y sordera. Se han identificado defectos causantes de XP en 7 genes del sistema de reparación, que codifican componentes de la escinucleasa: XPA, XPB, XPC, XPD, XPF, ERCC1 y XPG (págs. 401-402).

El **síndrome de Cockayne (CS)**, caracterizado por enanismo, retraso mental y síntomas neurológicos progresivos causados por desmielinización, también con una sensibilidad moderada a la radiación ultravioleta, se origina por mutaciones en los genes CSA o CSB, relacionados con la reparación acoplada a la transcripción. Algunas mutaciones de XPB, XPD y XPG dan lugar a síndromes solapantes XP/CS.

Otra enfermedad genética relacionada con la reparación del DNA es la **tricotiodistrofia (TTD)**, debida a mutaciones en los genes TTD-A, XPB o XPD. Se trata más bien de un defecto de la transcripción, que se puede corregir *in vitro* mediante microinyección de TFIIH a las células. Sus síntomas (en los dos últimos casos, solapantes con XP) incluyen caries dental, ictiosis, anomalías esqueléticas y retraso mental progresivo por desmielinización.

El **cáncer colorrectal hereditario no asociado a poliposis (HNPCC, hereditary nonpolyposis colorectal cancer, o síndrome de Lynch)**, uno de los cánceres hereditarios más comunes, está causado por defectos en los genes *hMSH2* o *hMLH1*, ambos responsables de la reparación de emparejamientos incorrectos. Se pueden detectar mediante VNTR, pues el sistema de reparación defectuoso no elimina pequeños bucles monocatenarios de DNA, lo que conduce a una mayor longitud de algunos microsatélites de DNA (pág. 421).

Otras enfermedades relacionadas con defectos en los sistemas de reparación son la **ataxia-telangiectasia**, asociada con una mayor sensibilidad al daño del DNA por rayos X, y la **anemia de Falconi**, que parece deberse a un defecto en la reparación del entrecruzamiento químico de bases.



Diversidad del genoma: polimorfismos

24.1 INTRODUCCIÓN	405
24.2 DIVERSIDAD GENÉTICA DENTRO DE UNA ESPECIE	406
24.3 MECANISMOS IMPLICADOS EN LA VARIABILIDAD GENÉTICA	407
24.3.1 Duplicación del material genético	407
24.3.2 Transposición	407
24.4 CONSECUENCIAS FUNCIONALES DEL POLIMORFISMO	409
24.4.1 Polimorfismo en regiones codificantes, o polimorfismo génico	409
24.4.1.1 Sin efecto fenotípico	409
24.4.1.2 Con alteración del fenotipo	409
24.4.2 Ejemplos representativos de polimorfismo génico	409
24.4.2.1 Polimorfismo de grupo sanguíneo AB0	410
24.4.2.2 Polimorfismo en el locus de la galactosemia	413
24.4.2.3 Polimorfismo de la antitripsina α_1	414
24.4.3 Polimorfismo en regiones génicas no codificantes	415
24.4.4 Polimorfismo en regiones no génicas	415
24.4.5 Polimorfismo en el genoma mitocondrial	416
24.5 ANÁLISIS DE GENES	417
24.5.1 Búsqueda de genes por análisis de ligamiento	417
24.5.2 Empleo de métodos físicos para la identificación de genes	417
24.5.3 Identificación de genes por secuenciación	418
24.5.4 Clonación funcional	418
24.5.5 Identificación de genes y bioinformática	418
24.6 DETECCIÓN DEL POLIMORFISMO DE SECUENCIA DE DNA	418
24.6.1. Polimorfismo de un solo nucleótido	418
24.6.1.1 RFLP: detección de polimorfismo en la longitud de los fragmentos de restricción	418
24.6.1.2 Otras formas de detectar polimorfismo SNP	420
24.6.2. Polimorfismo de repetición	421
24.6.2.1 VNTR: detección de polimorfismo en el número de repeticiones en tandem	421
a) Detección por hibridación	421
b) Detección por PCR	422
24.6.2.2 Polimorfismo STR	423
24.7 APLICACIONES DEL POLIMORFISMO GENÉTICO	424
24.7.1 Mapa de polimorfismos SNP	424
24.7.2 Análisis del DNA en estudios familiares	424
24.7.3 Identificación de individuos: huella dactilar de DNA	425
24.7.4 Análisis de polimorfismos con fines diagnósticos	426
24.7.4.1 Diagnóstico de errores congénitos del metabolismo	427
a) Valoraciones preliminares	427
b) Ensayos genéticos	428
c) Incidencia de las enfermedades congénitas	428
24.7.4.2 Diagnóstico de enfermedades poligénicas	428

24.1 INTRODUCCIÓN

El DNA genómico nuclear es una entidad extremadamente variable, no sólo entre especies diferentes, sino también entre individuos de una misma especie. Esta diversidad es responsable de fenómenos biológicos de tanta trascendencia como la evolución de las especies y, dentro de cada especie, de la presencia de características diferenciales en cada individuo, únicas e irrepetibles. La *diversidad o variabilidad genética* se debe a variaciones en la secuencia del genoma; por tanto, en un sentido amplio el concepto de *diversidad* se hace sinónimo de *polimorfismo* (en su significado literal, *muchas formas*) o, más concretamente, polimorfismo genético, de secuencia o de DNA. Los proyectos de secuenciación del genoma ya tienen en cuenta esta diversidad; una de las iniciativas recientes del proyecto genoma humano es precisamente la secuenciación del genoma de individuos concretos, de distinto origen étnico (pág. 258) para realizar un análisis comparativo de los puntos de variabilidad del genoma.

El polimorfismo afecta tanto a regiones codificantes del genoma (*polimorfismo génico*) como a las no codificantes (*polimorfismo genético* en general). En ambos casos puede consistir en la variación de un solo par de bases del DNA o, menos frecuentemente, de millones de pb (macromutación). El primer caso se conoce como polimorfismo de un solo nucleótido o **SNP** (acrónimo inglés de *single nucleotide polymorphism*; coloquialmente, “*snip*”). Cualquier variación puede tener lugar en células germinales o reproductoras, con lo que se transmite a la descendencia (polimorfismo hereditario, por ejemplo en las enfermedades congénitas), o bien en células somáticas, no reproductoras, en cuyo caso no se transmite (polimorfismo no hereditario, por ejemplo en la mayoría de los cánceres).

En un principio, la definición de polimorfismo se refería a las proteínas, y más tarde a sus genes. Hoy día debe definirse como la “existencia simultánea en una población de genomas con distintos alelos para un locus determinado”, sea éste génico o no. Los *alelos* son variaciones de la secuencia de DNA presente en un *locus* o posición definida en un cromosoma (pág. 97); en consecuencia, en una célula diploide cada locus autosómico está ocupado por dos alelos, uno de origen paterno y otro materno, situados en sendos cromosomas homólogos. El **grado de polimorfismo** de una población es tanto mayor cuantos más individuos contenga, viene determinado por el número de alelos distintos existentes para un locus concreto y se refleja en el grado de **heterocigosis**, o proporción de individuos heterocigóticos en el total de la población (pág. 99).



24.1

Lógicamente, la causa última de la existencia de polimorfismo es la mutación del DNA. La distinción entre el uso de los términos **mutación** y **polimorfismo** no es clara ni unánime, pero en general se asocia el primero con situaciones excepcionales, en especial si son patológicas, mientras que se habla de polimorfismo cuando la presencia de la variación genética es razonablemente común en la población y, por tanto, estable y nada o poco perjudicial. Como consecuencia, el polimorfismo ha de ser heredado, mientras que la mutación puede o no serlo. Por convenio, se dice que un locus es polimórfico cuando la variabilidad afecta a más del 1% de la población. Dicho de otro modo, el segundo alelo más frecuente se presenta con frecuencias superiores a 0,01.

Ejemplos hipotéticos de la frecuencia de alelos en 5 loci de una población

Alelo	Locus A	Locus B	Locus C	Locus D	Locus E
1	99,5% normal	98% polimórfico	99,5% normal	93,5% polimórfico	95,7% polimórfico
2	0,5% mutado	2% polimórfico	0,3% mutado	5% polimórfico	4% polimórfico
3			0,2% mutado	1,5% polimórfico	0,3% mutado

Todo polimorfismo es, pues, reflejo del genotipo de los individuos (constitución alélica para un locus determinado o para un conjunto de ellos). Surge, por tanto, el concepto de **individualidad genética**, entendida como la presencia en cada individuo de un genoma con una secuencia característica, única para ese individuo.

Se presentan, en primer lugar, los fundamentos moleculares del polimorfismo y sus consecuencias funcionales, para continuar con la descripción de las técnicas empleadas para su detección.

24.2 DIVERSIDAD GENÉTICA DENTRO DE UNA ESPECIE

Nuestro objetivo es estudiar la individualidad genética en especies con reproducción sexual, caso de los animales superiores, incluyendo los seres humanos, en los que el polimorfismo se presenta en dos versiones.

- Polimorfismo *fisiológico* o *normal*, de interés en estudios familiares, en la identificación de individuos (indicios criminales o investigación de la paternidad), expresión de proteínas (describiendo, por ejemplo, el color de los ojos o la capacidad atlética), análisis de ligamiento que conduce a la cartografía genética, etc.
- Polimorfismo *patológico*, de interés en el diagnóstico presintomático y prenatal de enfermedades genéticas, en la detección de individuos portadores, la determinación de compatibilidad para trasplantes, etc. Asimismo, puede definir el riesgo para padecer enfermedades (Alzheimer, diabetes, etc.) y condicionar la respuesta a fármacos (existe un gran número de personas con graves problemas de salud debido a reacciones adversas a fármacos).

Curiosamente, los primeros indicios de la variabilidad genética en humanos surgieron, a principios del siglo xx, a partir de la observación de proteínas distintas de las normales en patologías como la alcaptonuria. Hoy sabemos que esta “diversidad proteica” es el resultado de que distintos individuos poseen alelos diferentes en el locus del gen que codifica la proteína (polimorfismo génico).

24.3 MECANISMOS IMPLICADOS EN LA VARIABILIDAD GENÉTICA

Son varios los mecanismos que –actuando en combinación– dan cuenta de la variabilidad genética durante el complejo proceso de la morfogénesis (desde el cigoto a las primeras capas de tejidos embrionarios y su transformación en órganos) y a todo lo largo de la vida del individuo adulto. De estos mecanismos se han considerado ya la recombinación homóloga (pág. 108), la segregación de cromosomas (pág. 106) y la mutación (capítulo 23). Se comentarán a continuación los efectos de la duplicación y la transposición.

24.3.1 Duplicación del material genético

Se habla de **duplicación a gran escala** cuando afecta a grandes porciones de un cromosoma. Sin entrar a describir los aspectos concretos del proceso (propios de una genética molecular), debe indicarse que ha actuado progresivamente durante la evolución, mediante diversos mecanismos complejos.

La **duplicación a pequeña escala** es de interés cuando tiene lugar en regiones codificantes del genoma, afectando a genes específicos, por lo que también se denomina duplicación génica. La aparición de múltiples copias del gen, inicialmente idénticas, suele evolucionar posteriormente hacia la formación de copias ligeramente diferentes. De esta forma, una copia ancestral de un gen puede originar todos los miembros de una **familia multigénica** (genes que codifican polipéptidos con secuencias relacionadas, pág. 113). Estas familias codifican diversas proteínas típicas de eucariotas; por ejemplo, globinas (proteínas de transporte y almacen de oxígeno en sangre y músculo), actinas (parte del aparato contráctil en células musculares y del citoesqueleto en células no musculares), opsinas (que detectan la luz a diferentes longitudes de onda en células de la retina), colágeno (en diferentes tipos de tejido conjuntivo y óseo), albúminas e inmunoglobulinas (en plasma sanguíneo), etc.

El principal mecanismo responsable de la duplicación de genes completos a lo largo de la evolución y, en definitiva, de la generación de tan gran variedad y cantidad de DNA repetitivo en las regiones no codificantes y codificantes del genoma eucariótico, es la **recombinación no homóloga o desigual**. Ésta ocurre durante la meiosis, entre cromosomas homólogos alineados de manera incorrecta, en los que no coinciden las regiones equivalentes (como sí ocurre en la recombinación homóloga), pero sí se enfrentan regiones con similitud de secuencia. Como consecuencia del sobrecrezamiento de cromátidas (pág. 108), un homólogo puede recibir mayor cantidad de material genético que el otro (lo que supone en ambos, respectivamente, una inserción o duplicación y una delección).



Web 24.1. Duplicación génica resultante de la recombinación desigual.

También han surgido por duplicación génica los genes que codifican, en tejidos diferentes, las *isoformas* (formas alternativas de una proteína) o, en su caso, las *isoenzimas* o *isozimas* (formas alternativas de una enzima). Sirve como ejemplo la **fosfatasa alcalina**, codificada por al menos 4 genes diferentes, que se expresan de forma diferencial en cada tejido. Tres de ellos están agrupados cerca del telómero del cromosoma 2q: ALPI de intestino, ALPP de placenta (con un 87% de similitud entre sí) y ALPPL, similar a la de placenta. El gen de la cuarta isozima, ALPL, de hígado, hueso, riñón y otros tejidos, está localizado cerca del telómero de 1p y muestra una similitud de secuencia de 57% con ALPI y 55% con ALPP. Tan elevada homología indica que estas formas específicas de la misma proteína enzimática se originaron por una duplicación génica ancestral.

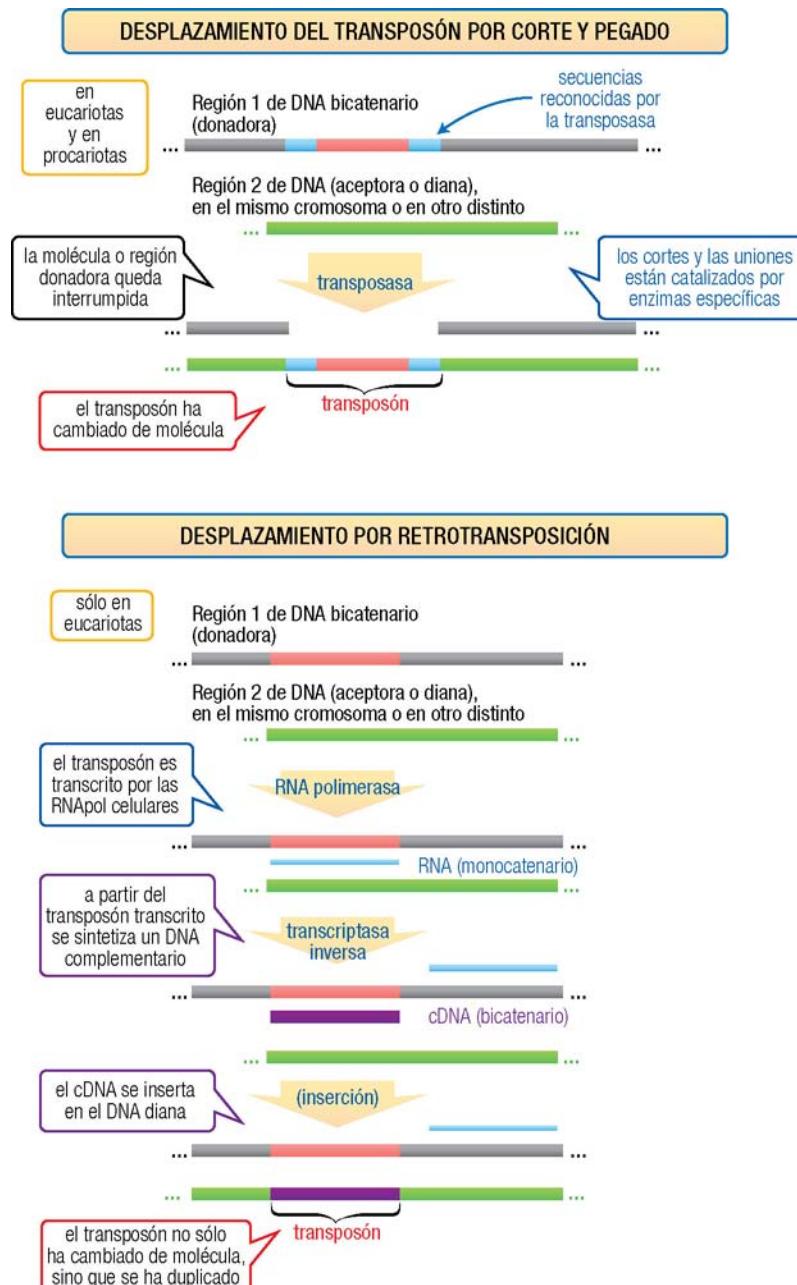
24.3.2 Transposición

Otra fuente importante de variabilidad genética es el proceso de **transposición**, bien caracterizado en bacterias pero también presente en el genoma eucariótico. Consiste en el desplazamiento, de un sitio a otro del cromosoma, de fragmentos de DNA llamados *elementos móviles*, *elementos transponibles* o *transposones*; su

longitud es variada, desde unos cuantos cientos de nucleótidos hasta decenas de miles. La transposición es responsable en eucariotas de una gran proporción del DNA repetitivo no codificante, destacando las secuencias *Alu* y *LINE-1* (pág. 118), que por sí solas suponen el 10% del genoma humano.

Existen dos mecanismos de transposición en eucariotas:

En el **mecanismo de corte y pegado** el *transposón* se escinde del DNA donador y se inserta en el DNA diana. En algunos casos, la enzima responsable, transposasa, está codificada por una parte de la propia secuencia del transposón.



24.2

Mediante el **mecanismo de retrotransposición** se sintetiza un cDNA (págs. 206 y 222-223) a partir del transcripto del transposón, y finalmente se inserta aquél en el DNA diana. El nombre se debe a la similitud con el mecanismo de replicación en retrovirus. Algunos retrotransposones codifican en su propia secuencia la transcriptasa inversa necesaria para su propagación (por ejemplo, *LINE-1*).

24.4 CONSECUENCIAS FUNCIONALES DEL POLIMORFISMO

El polimorfismo puede tener distinta trascendencia desde el punto de vista funcional, dependiendo de si afecta a regiones no codificantes, reguladoras o codificantes del genoma y también del modo como afecte al mensaje genético de estas últimas.

24.4.1 Polimorfismo en regiones codificantes, o polimorfismo génico

De la parte codificante del genoma humano se cree que unas tres cuartas partes son **monomorfos**, genes únicos compartidos sin variación por todos los individuos. Éstos constituyen la parte del genoma común o “de referencia”, que define la especie y determina sus características morfológicas. Dichos genes son casi idénticos a los homólogos en las especies más cercanas (en nuestro caso, el gorila). El resto, una cuarta parte del total, son **polimorfos**; sus variaciones, en muchos casos mínimas, determinan el polimorfismo génico.

Dentro del polimorfismo génico se pueden distinguir dos situaciones de acuerdo con su trascendencia funcional, es decir, su efecto sobre el fenotipo.

24.4.1.1 Sin efecto fenotípico

Aun estando la variabilidad localizada en una secuencia codificante, el polimorfismo puede no tener consecuencias sobre el fenotipo (es decir, sobre el producto génico), principalmente por dos motivos: que no se altere la secuencia proteica codificada (como consecuencia de la degeneración del código genético, caso de las *mutaciones silenciosas*, pág. 389) o que la variación ocurra en una región de la proteína que no es esencial para su función (entre otras, las *mutaciones conservadoras*, pág. 391). Podrían incluirse también en este último caso los polimorfismos que afectan a los intrones, regiones génicas pero no codificantes (que, por tanto, pueden considerarse también polimorfismos no génicos, pág. 415).

El polimorfismo sin efecto fenotípico es el más común, responsable de una diversidad genética enteramente normal. Las proteínas formadas por los distintos alelos de un locus de este tipo son variantes relativamente comunes, con secuencias y estructuras idénticas o similares y manteniendo una funcionalidad similar. Como ejemplos pueden citarse proteínas plasmáticas como inmunoglobulinas, antitripsina α_1 , haptoglobina, transferrina y ceruloplasmina, así como otras proteínas tanto intracelulares como extracelulares. Durante años, estas *formas múltiples* o *variantes* de proteínas fueron el principal producto *apreciable* del polimorfismo, gracias a la posibilidad de separarlas, sobre todo mediante electroforesis.

24.4.1.2 Con alteración del fenotipo

En otros casos, la variación en la secuencia del gen modifica la secuencia de la proteína de tal forma que afecta a su funcionalidad. Aunque esta modificación puede ser beneficiosa (base de la evolución), lo más común es encontrar un polimorfismo patológico o deletéreo (perjudicial). La alteración puede modificar las características bioquímicas, fisiológicas o incluso morfológicas de la célula, y puede llegar a producir patologías, dependiendo de la gravedad de dicho efecto fenotípico. Se puede distinguir entre:

- Variaciones fenotípicas que no influyen en la susceptibilidad a enfermedades, caso de los polimorfismos responsables de la estatura, el color de pelo, el grupo sanguíneo, etc.
- Variaciones fenotípicas que influyen de forma leve o mínima en la susceptibilidad a ciertas patologías, tales como los polimorfismos responsables de rasgos morfológicos complejos.
- Variaciones fenotípicas que desempeñan un papel relevante en la aparición de una patología. Esto es lo que ocurre, por ejemplo, con las hemoglobinas anormales, que se manifiestan en forma de enfermedades genéticas. Precisamente, el concepto de *enfermedad genética* constituye la manifestación más obvia y a veces más extrema de un cambio genético.

24.4.2 Ejemplos representativos de polimorfismo génico

A continuación se van a desarrollar en detalle tres ejemplos de polimorfismo génico en regiones codificantes –y, en consecuencia, polimorfismo proteico– bien conocidos, dos de ellos con implicaciones patológicas.

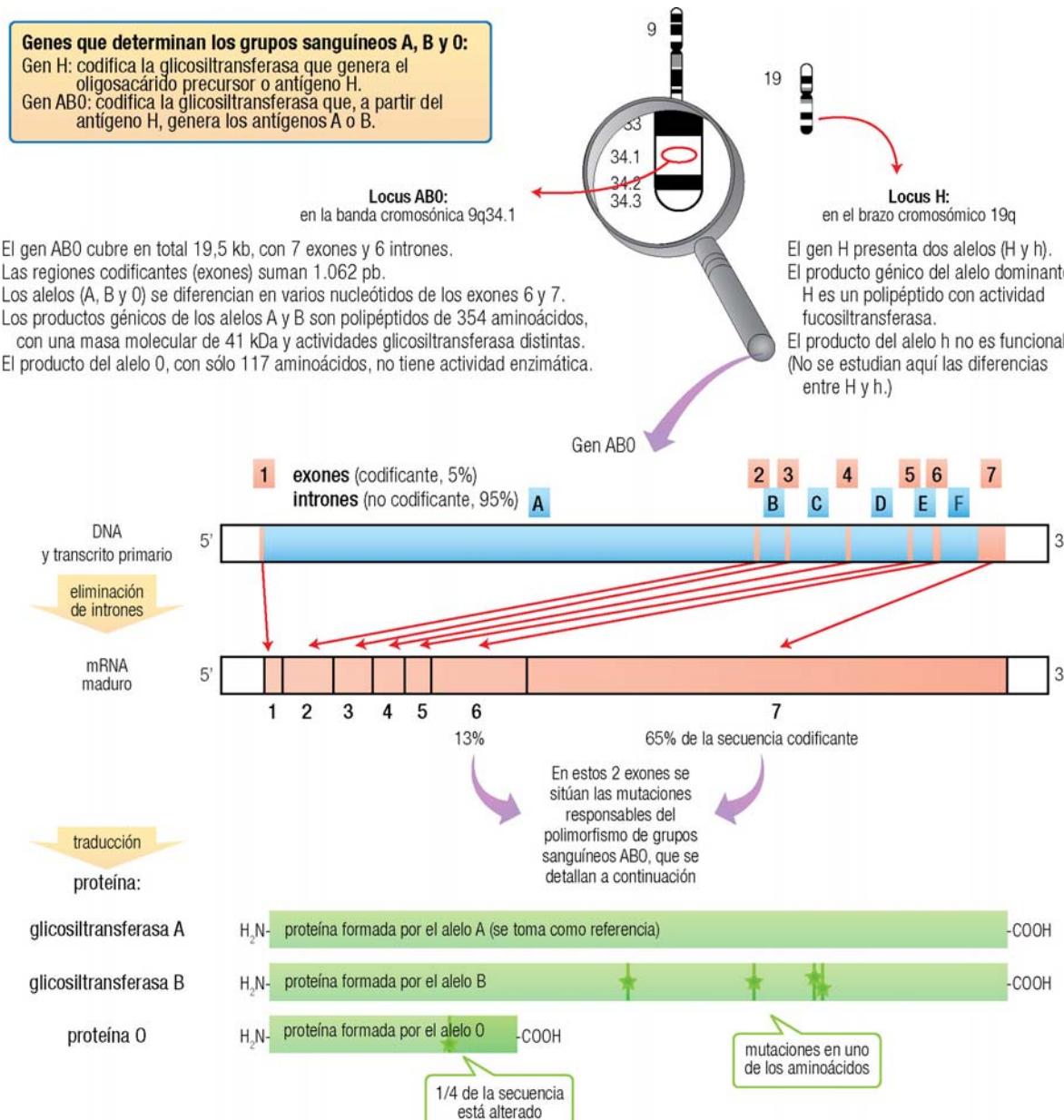
24.4.2.1 Polimorfismo de grupo sanguíneo AB0

La existencia de los grupos sanguíneos humanos del sistema AB0 fue el primer caso descrito (1990) de diversidad proteica determinada genéticamente. La existencia de tres alelos determina cuatro fenotipos en la sangre humana, denominados grupos sanguíneos 0, A, B y AB.

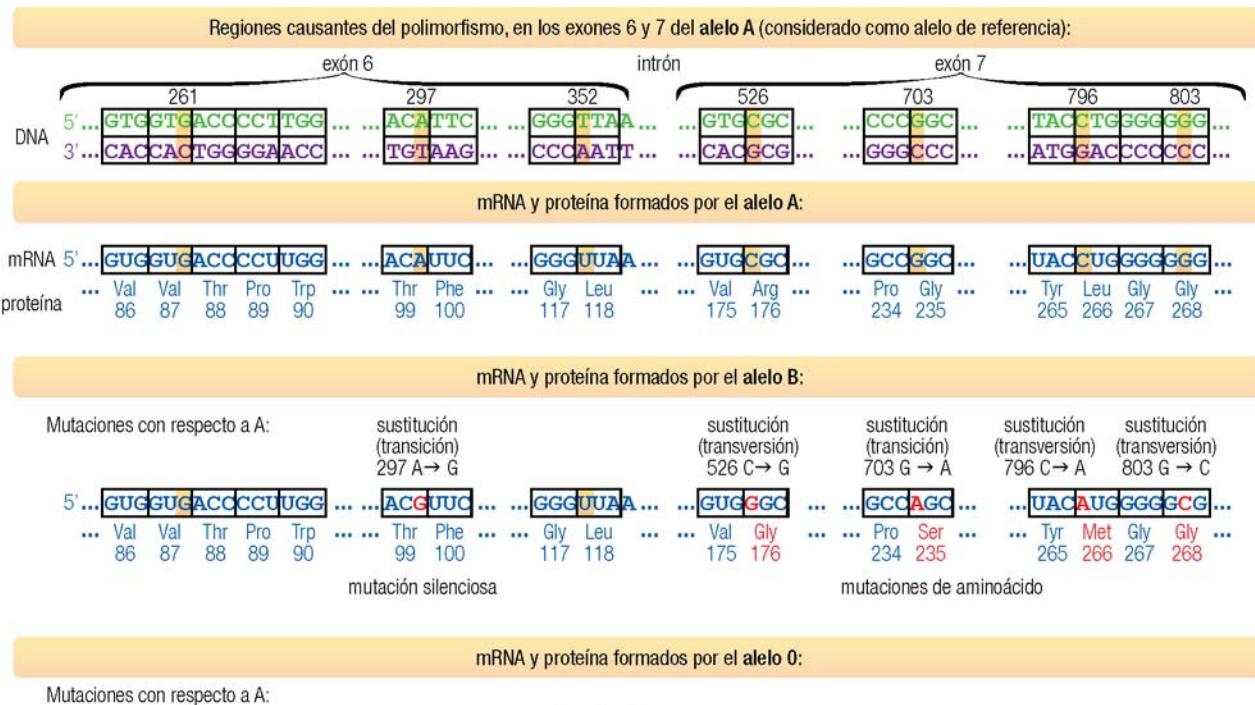


Web 24.2. Identidad de los grupos sanguíneos y aspectos clínicos.

Desde el punto de vista de la biología molecular, es de sumo interés considerar las implicaciones genéticas y bioquímicas de un sistema polimórfico tan representativo como el de los grupos sanguíneos AB0. Éste consiste en un **polimorfismo de oligosacáridos** (carbohidratos), que es el resultado de un **polimorfismo proteico** en las glicosiltransferasas responsables de la síntesis de esos oligosacáridos. Éste se origina, a su vez, en un **polimorfismo génico**, la existencia de 3 alelos para el gen que codifica la glicosiltransferasa. Por otra parte, la expresión de los antígenos A y B está condicionada, además, por la de otro gen polimórfico, el gen H, cuyo producto proteico es otra glicosiltransferasa que sintetiza el carbohidrato precursor (antígeno H) sobre el que actúan las transferasas codificadas por el gen AB0. Los aspectos particulares de todo el sistema se muestran comparativamente a continuación.



Para una mejor comprensión de las diferencias causantes del polimorfismo, se detallan los puntos polimórficos (en los exones 6 y 7) de las secuencias de mRNA y proteína sintetizados por cada alelo, considerando el A como referencia y el B y 0 como mutaciones de aquél.



Mutaciones con respecto a A:

⇒ cambio del marco de lectura

261 G → cambio del marco de lectura → prematura

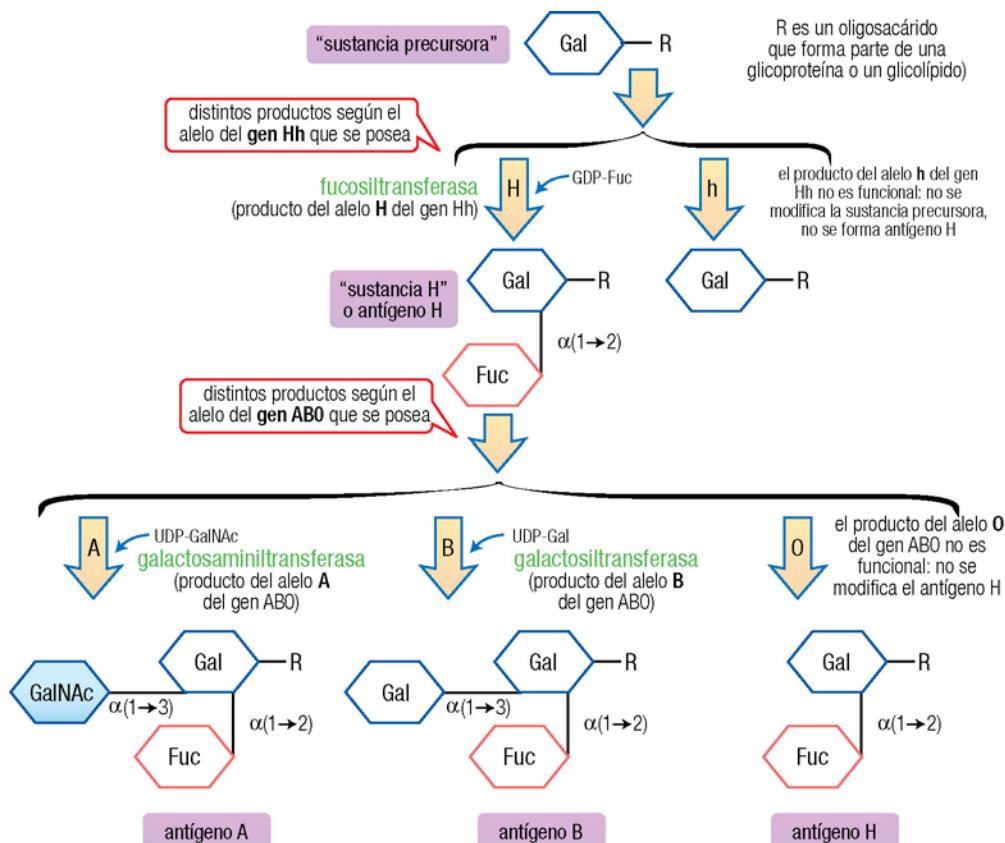
5' ... **GUGGUU**ACCCUCUUGG CAUUC **G**GGGUAA UGGCGC GCGGC ACCUGGGGGGGG ...

... Val Val Pro Leu Gly ... His Ser Gly O

86 87 88 89 90 99 100 117

La región C-terminal de la "proteína O" (residuos 88 a 117) tiene una secuencia completamente diferente a las glicosiltransferasas A y B, y el tamaño de la proteína completa es 2/3 inferior (117 frente a 354 residuos), como consecuencia de todo ello, la proteína O no tiene actividad glicosiltransferasa.

En cuanto al aspecto bioquímico de este polimorfismo, los antígenos A y B se sintetizan mediante la acción enzimática de dos glicosiltransferasas distintas, codificadas por los alelos A y B. Ambas actúan sobre el antígeno H, un oligosacárido componente de glicoproteínas y glicolípidos que, a su vez, se genera previamente mediante otra glicosiltransferasa codificada por el gen independiente del locus H. La especificidad antigenica de A, B y H, que define el grupo sanguíneo y determina la producción de los anticuerpos, viene definida por los azúcares terminales de los oligosacáridos.



24.5

Los aspectos considerados se pueden resumir mostrando las relaciones entre los dos loci génicos, sus alelos, los productos génicos de éstos y los antígenos resultantes de la acción de esas enzimas:

Gen (o locus)	Alelo	Producto genérico (proteína, glicosiltransferasa) (se indican varios nombres y su número de clasificación EC)	Antígeno (oligosacárido) generado por la glicosiltransferasa
Hh	H	Fucosiltransferasa Galactósido 2-l-fucosiltransferasa EC 2.4.1.69	H
2 alelos: <i>H</i> dominante, <i>h</i> recesivo	h	No funcional	Ninguno
ABO	A	Galactosaminiltransferasa Transferasa del grupo histo-sanguíneo A Glicoproteína-fucosilgalactósido α -N-acetilgalactosaminiltransferasa EC 2.4.1.40	A
3 alelos: <i>A</i> y <i>B</i> codominantes, <i>O</i> recesivo	B	Galactosiltransferasa Transferasa del grupo histo-sanguíneo B Glicoproteína-fucosilgalactósido α -galactosiltransferasa EC 2.4.1.37	B
	O	No funcional	Ninguno

El gen AB0 constituye un ejemplo clásico de multialelismo: los alelos *A* y *B* son codominantes, mientras que *O* es recesivo, por lo que los 6 genotipos posibles determinan 4 fenotipos. En cuanto al gen H, posee sólo dos alelos, uno dominante (*H*) y el otro recesivo (*h*), que producen 3 genotipos y 2 fenotipos.

Grupos sanguíneos: genotipo, fenotipo, antígenos y anticuerpos presentes

Genotipo	Fenotipo	Glicosiltransferasas sintetizadas	Antígenos presentes	Anticuerpos en el plasma
HH o Hh	AA	Grupo A	H y A	H (poco) y A
	A0			Anti-B
	BB	Grupo B	H y B	H (poco) y B
	B0			Anti-A
	AB	Grupo AB	H, A y B	H (poco), A y B
	00	Grupo 0	H	Sólo H (mucho)
hh ^(*)	AA	Grupo 0	A	Ninguno
	A0			Anti-H, anti-A y anti-B
	BB	Grupo 0	B	Ninguno
	B0			Anti-H, anti-A y anti-B
	AB	Grupo 0	A y B	Ninguno
	00	Grupo 0	Ninguna	Anti-H, anti-A y anti-B

^(*)En el caso de genotipo hh no se sintetiza el antígeno H por lo que no se pueden sintetizar los antígenos A ni B, aunque existan glicosiltransferasas de tipo A o B. Por tanto, el grupo sanguíneo es aparentemente 0; se le llama fenotipo 0_h Bombay.

24.4.2.2 Polimorfismo en el locus de la galactosemia

La galactosemia es una enfermedad autosómica recesiva causada por la imposibilidad de metabolizar la galactosa, que como consecuencia se acumula en sangre, síntoma que da nombre a la enfermedad. En teoría (v. web 24.4) la galactosemia puede deberse a mutaciones en distintas enzimas; de hecho, así se ha demostrado, aunque con escasa incidencia. En la práctica, la causa más frecuente son mutaciones en el gen que codifica la enzima galactosa-1-fosfato uridiltransferasa o GALT.



Web 24.3. Descripción y bioquímica de la galactosemia.

Los recién nacidos con galactosemia son homocigóticos para un alelo mutante en el locus GALT. Se han encontrado al menos 150 alelos diferentes en este locus (situado en la banda cromosómica 9p13), con influencia variable sobre la actividad enzimática y, en consecuencia, sobre la posibilidad de desarrollar la enfermedad. Las frecuencias de estos alelos varían mucho dependiendo de la población que se considere, como suele ocurrir con todos los polimorfismos. Los alelos más frecuentes, N (normal), G (“galactosemia clásica”), D (“Duarte” o “Duarte-2”) y LA (“Los Ángeles” o “Duarte-1”) dan lugar a las siguientes situaciones:

Genotipo	Actividad enzimática	Fenotipo bioquímico
NN	100%	Normal
GG	< 5%	Galactosemia clásica
DD	50%	Normal, “Duarte” (actividad reducida, sin trascendencia) ^(*)
LA LA	140%	Normal, “Los Ángeles” (actividad excesiva, sin trascendencia) ^(*)
NG	50%	Normal (actividad reducida, sin trascendencia)
ND	75%	Normal (actividad reducida, sin trascendencia)
NLA	120%	Normal (actividad excesiva, sin trascendencia)
GD	25%	Situación límite
GLA	70%	Normal (actividad reducida, sin trascendencia)
D LA	95%	Normal

^(*)El fenotipo bioquímico “Duarte” se identifica porque la enzima GALT se presenta en una isoforma más ácida (coherente con la mutación de asparragina a aspártico que se explica después), de punto isoelectrónico menor, que avanza más rápidamente en electroforesis y en isoelectroenfoque se sitúa más cerca del ánodo. Las variantes Duarte 1 y 2 se diferencian en su actividad enzimática.

Todos los fenotipos resumidos anteriormente son a su vez polimórficos. La reducción de la actividad enzimática varía en distinto grado según la combinación de alelos que presente el individuo.

Detalle de las mutaciones causantes del polimorfismo en el locus GALT

Se muestra la secuencia de la proteína GALT normal (locus genético en la banda cromosómica 9p13). Los aminoácidos mutados que originan las variantes aparecen indicados en azul.

```

1  MSRSGTDPQQ RQQASEADAA AATFRANDHQ HIRYNPLQDE WVLVSAHRMK RPWQQGVEPQ
61  LLKTVPRHDP LNPLCPGAIR ANGEVNQYD STFLFDNDFP ALQPDAPS PG PSDHPLFQAK
121 SARGVCKVMC FHPWS DVTLP LMSVPEIRAV VDAWASVTEE LGAQYPWVQI FENKGAMMGC
181 SNP HPHCQVW ASSFLPDIAQ REERSQQAYK SQHGEPLIME YSRQELLRKE RLVLTSEHWL
241 VLVPFWATWP YQTLLPDRRH VRRLPELTPA ERDDLASIMK KLLTKYDNLF ETSFPYSMGW
301 HGAPTGEAG ANW NHWQLHA HYYPPLLRSA TVRKFMVGYE MLAQAQRDLT PEQAAERLRA
361 LPEVHYHLGQ KDRETATIA

```

La secuencia se indica mediante las abreviaturas de una letra para los aminoácidos (v. apdo. 20.3). Cada bloque contiene 10 aminoácidos (numeración del primero a la izquierda).

HPH (His-Pro-His) son aminoácidos del centro activo ("tríada" característica; los dos residuos de histidina participan en la catálisis). Se ha identificado una mutación, **H184Q** (His→Gln, no descrita aquí), que afecta directamente al centro activo.

La mutación **Q188R** (Gln→Arg, debida a una transición CAG→CGG) es la más frecuente en individuos con galactosemia clásica, especialmente en la raza blanca europea y norteamericana. Conduce a una proteína con actividad enzimática casi nula (Gln-188 es un aminoácido altamente conservado, por tanto esencial para la función, probablemente por su cercanía al centro activo) y un fenotipo galactosémico.

La mutación **K285N** (Lys→Asn, debida a una transversión AAG→AAT) es la segunda más frecuente en la población europea. Tiene también efectos fenotípicos graves (galactosemia).

La mutación **S135L** (Ser→Leu, debida a una transición TCG→TTG) es especialmente abundante en los casos de galactosemia entre la población norteamericana de raza negra (fenotipo "Negro"). La actividad enzimática se ve reducida en distinta medida según el tejido considerado. Se puede detectar este alelo mediante RFLP (v. apdo. 24.6.1.1), aprovechando que la mutación hace desaparecer una diana para la enzima de restricción *Taq I*.

La mutación **N314D** (Asn→Asp, debida a una transición AAC→GAC) es característica de los fenotipos "Duarte" (o "Duarte-2") y "Los Ángeles" (o "Duarte-1"). El primero se caracteriza por una actividad enzimática reducida (cerca del 50%), mientras el segundo muestra actividad superior a la de individuos normales (140%). El fenotipo "Los Ángeles" suele presentar, además de la anterior, una mutación silenciosa **L218L** (Leu) por transición CTA→TTA, que parece ser responsable de una traducción más rápida y, por ello, de una mayor concentración de enzima que explicaría el aumento en actividad enzimática. Al igual que el caso anterior, se pueden detectar ambos alelos mediante RFLP, puesto que la mutación A→G hace aparecer una diana para la enzima de restricción *Ava II*.

24.6

24.4.2.3 Polimorfismo de la antitripsina α_1

La antitripsina α_1 (α_1 -AT o A1AT) es una proteína plasmática que inhibe la actividad de varias proteasas, como tripsina, quimotripsina y elastasa. Actúa principalmente regulando la elastasa leucocitaria, para evitar que ésta destruya la elastina y otras proteínas del tejido conjuntivo pulmonar y se dañe la pared alveolar. La deficiencia en α_1 -AT se asocia con enfisema (enfermedad pulmonar degenerativa) y con enfermedades hepáticas (cirrosis y cáncer). Se ha podido determinar que dicha deficiencia y sus síntomas clínicos corresponden a la presencia de algunos de los múltiples alelos existentes para el locus de la α_1 -AT, denominado *locus PI* (de inhibidor de proteasas) y situado en la banda cromosómica 14q32.1. Estos alelos (se han identificado más de 75) codifican variantes o isoformas de la proteína, algunas de las cuales se pueden distinguir por su migración en electroforesis o isoelectroenfoque.

Alelos más comunes	Posición en la secuencia proteica						Fenotipo (proteína)
	101	160	213 ^(*)	264	342	376	
M1A (el más antiguo evolutivamente)	Arg	Tyr	Ala	Glu	Glu	Glu	Isoformas ligeramente diferentes, pero todas funcionalmente normales
M1V (el más frecuente en la actualidad)	Arg	Tyr	Val	Glu	Glu	Glu	
M2	His	Tyr	Val	Glu	Glu	Asp	
M3	Arg	Tyr	Val	Glu	Glu	Asp	
M4	His	Tyr	Val	Glu	Glu	Glu	
Z (el más frecuente de los patológicos)	Arg	Tyr	Ala	Glu	Lys	Glu	Concentración en plasma reducida (<15%)
S	Arg	Tyr	Val	Val	Glu	Glu	Casi normal
Nulo (variante “Granite Falls”)	Arg	stop					Forma truncada, no funcional

(*) Las variantes M1A y Z pueden distinguirse del resto por RFLP (pág. 418), pues la mutación GCG (Ala) → GTG (Val) hace aparecer una diana para *Bst* EII.

Se ha podido establecer la base molecular de la alteración funcional que sufre la isoforma Z. En la estructura tridimensional de la proteína normal, el residuo Glu-342 (con carga negativa) se sitúa próximo al Lys-290 (con carga positiva) y ambos forman un puente salino. La mutación Glu(342) → Lys, característica del alelo Z, impide este puente salino, provocando la agregación de la proteína dentro de la célula, lo que impide su secreción al plasma, disminuyendo la actividad y provocando la susceptibilidad a la enfermedad. El efecto se manifiesta especialmente, como es lógico, en los homocigóticos ZZ.

24.4.3 Polimorfismo en regiones génicas no codificantes

Teniendo en cuenta la definición más completa de gen, la que se establece bajo un criterio funcional (pág. 263), existen regiones del DNA que forman parte del gen, pero no codifican su producto génico. Tal es el caso de las regiones reguladoras (promotores), los intrones y las regiones 5'UTR y 3'UTR que, aunque se conservan en el mRNA, no se traducen. Por ello se debe considerar la posibilidad de polimorfismos en estas regiones génicas no codificantes. Evidentemente, aunque tal polimorfismo no afectará a la secuencia de la proteína, puede afectar a su expresión, dependiendo de la función que ejerza la secuencia no codificante que lo sufre. Muchos polimorfismos en los intrones no tendrán efecto alguno, y serán equivalentes al polimorfismo en regiones no génicas, estudiado a continuación. Otros podrán afectar al ayuste (pág. 303), lo que da como resultado un polimorfismo proteico. Las variaciones en secuencias promotoras probablemente perturben la expresión del gen. Por tanto, el polimorfismo en regiones génicas no codificantes puede manifestarse o no en un efecto fenotípico.

No parece conveniente complicar la exposición profundizando en este aspecto, pero no debe dejarse de tener en cuenta. Como ejemplo similar, se ha mencionado una mutación en secuencias reguladoras del gen del factor IX de la coagulación, causante de hemofilia B (pág. 388); aunque por su baja frecuencia no entra en la categoría de polimorfismo, como ya se ha indicado la mutación y el polimorfismo comparten el mismo principio molecular.

24.4.4 Polimorfismo en regiones no génicas

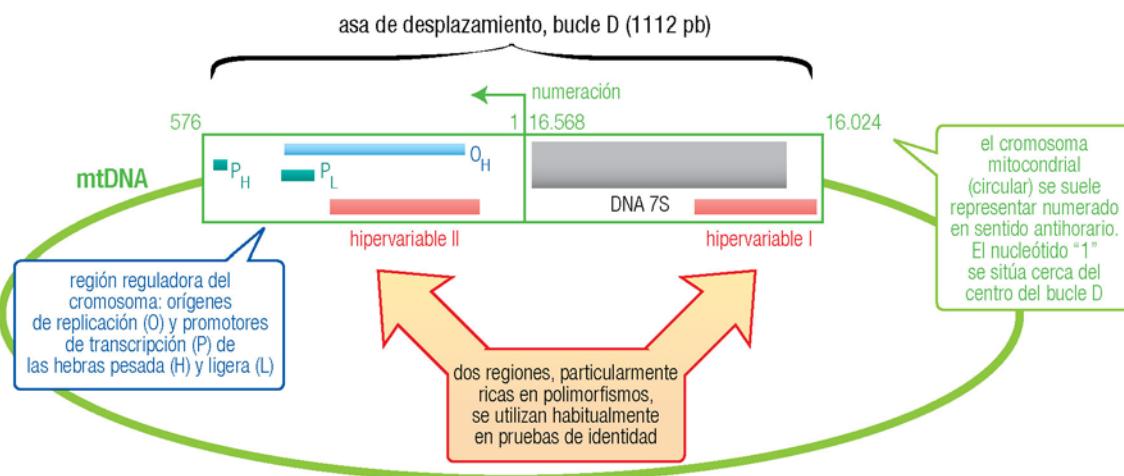
La mayor parte de las mutaciones del genoma ocurren en regiones del DNA que no codifican ningún producto, debido a su mayor abundancia en el genoma humano y de otros eucariotas (Capítulo 9). Por tanto, los cambios de bases en este DNA no tienen efecto fenotípico alguno: no se altera la secuencia, la estructura o la función proteicas, de ahí que tampoco tengan un interés patológico directo. Son, sin embargo, polimorfismos de gran interés aplicado; principalmente, para la cartografía del genoma, la búsqueda de genes relacionados con enfermedades y la identificación genética de individuos.

La elevada cantidad de loci polimórficos que se encuentran en DNA no codificante, junto a la falta de trascendencia funcional de sus alteraciones (que permite que se perpetúen libremente en sucesivas generaciones), hacen que sea éste el DNA que más difiere de unos a otros individuos. Como consecuencia, una de las propiedades más notables del genoma de un individuo es la exclusividad de su perfil genético (salvo para los hermanos gemelos univitelinos, individuos clónicos que tendrán los mismos alelos en cada uno de sus loci, aunque incluso en este caso pueden existir diferencias debidas a mutaciones somáticas). Se puede, por tanto, utilizar el análisis de estos polimorfismos para la identificación de una persona, con una fiabilidad igual o superior a la de las huellas dactilares, lo que ha dado lugar al término **huella genética** o **huella dactilar genética** (pág. 425). Con este objeto se emplean principalmente los polimorfismos en las regiones de DNA minisatélite y microsatélite (págs. 117 y 243). Por otra parte, puesto que las secuencias no codificantes no proporcionan información sobre características individuales de las personas (posibles enfermedades, cualidades físicas o psíquicas, etc.), el análisis basado en polimorfismos de regiones no codificantes posee la ventaja de evitar cualquier conflicto ético.

24.4.5 Polimorfismo en el genoma mitocondrial

En el genoma mitocondrial pueden darse también mutaciones que generen un polimorfismo. Su DNA experimenta mayor proporción de mutaciones que el nuclear, debido a la exposición a especies reactivas de oxígeno, resultantes del metabolismo oxidativo mitocondrial, a encontrarse más desprotegido por carecer de histonas y a la menor eficacia de los sistemas de reparación mitocondriales. De esas mutaciones, las que afectan a regiones codificantes (más del 93% de la molécula de mtDNA) suelen tener consecuencias perjudiciales para la funcionalidad de las proteínas que codifican y para la vida de la célula. En contraste, las mutaciones en la región no codificante se conservan con más facilidad, al no poseer efecto fenotípico. Las mutaciones que no son perjudiciales dan lugar a polimorfismos genéticos; cada vez parece más claro que éstos tampoco son completamente inocuos para las diferentes funciones celulares. Por ejemplo, una combinación particular de alelos en varios loci polimórficos del cromosoma mitocondrial humano (poco frecuente y denominada “haplogrupo T”) da lugar a espermatozoides con motilidad reducida.

Como consecuencia, los polimorfismos más comunes son los que afectan a la única región no codificante del DNA mitocondrial, conocida como **bucle D** (pág. 161). En este corto tramo de DNA (1.112 pb en humanos, menos del 7% del mtDNA total) se acumulan mutaciones con una frecuencia cinco veces superior a la del resto del DNA mitocondrial. En particular, la tasa de polimorfismos es máxima en dos pequeñas porciones del bucle D, denominadas por ello **regiones hipervariables**.



24.7

El análisis de polimorfismos en las regiones hipervariables mitocondriales encuentra gran aplicación en estudios familiares, antropológicos evolutivos y también de identificación. A este éxito contribuye la conveniencia de obtención y manejo de muestras mitocondriales, gracias a la existencia de un mayor número de copias (cientos de mitocondrias por célula, cada una con decenas de copias de mtDNA), lo que disminuye la probabilidad de que todas estén destruidas o dañadas, y se traduce, en definitiva, en una mayor estabilidad del DNA. Por otro lado, al ser las regiones hipervariables mucho más cortas en el mtDNA es improbable su degradación en todas las copias de DNA de la muestra.

24.5 ANÁLISIS DE GENES

Son muchos los aspectos de interés en el estudio del DNA génico, tanto desde el punto de vista básico (información sobre el propio gen) como, especialmente, para su aplicación al estudio de la enfermedad molecular con fines diagnósticos y terapéuticos. Aunque el conocimiento del factor genético no basta para resolver de manera definitiva la enfermedad, sí que es obligado conocer el aspecto molecular del gen o genes disfuncionales responsables (mutación o mutaciones) para poder identificarla. El número de enfermedades para las que ya se dispone de información génica es muy amplio; se pueden citar, por ejemplo, anemia drepanocítica, cáncer de mama y de colon hereditarios, corea de Huntington, distrofias musculares, enfermedad de Alzheimer, enfermedad de Tay-Sachs, fibrosis quística, neurofibromatosis, poliquistosis renal, síndrome de Marfan, síndrome del cromosoma X frágil, sordera hereditaria, etc.

El objetivo de la genómica estructural (pág. 239) es identificar la secuencia del genoma completo y, sobre todo, localizar los genes responsables de enfermedades. Aunque el número de genes localizados ha crecido exponencialmente desde la década de 1970 hasta el presente, aún quedan muchos por localizar y caracterizar.

Entre los aspectos estudiados cabe citar la localización de genes en los cromosomas, la identificación de la función de sus productos génicos en la célula, la búsqueda y detección de genes anómalos responsables de enfermedades conocidas o nuevas, la identificación de mutaciones responsables de las anomalías, el estudio del polimorfismo de secuencia, el análisis de las proteínas que resultan alteradas, las pruebas genéticas en el diagnóstico clínico de enfermedades monogénicas y multifactoriales, etc. El conocimiento de éstos y de otros aspectos no cesa de incrementarse con los avances en el conocimiento del genoma humano.

Bajo una expresión tan genérica como la de **análisis de genes** se estudian en este capítulo algunos de los planteamientos y herramientas básicas sobre las que se asienta la nueva medicina molecular o medicina genómica y en los que se apoya la industria farmacéutica en su intento por desarrollar fármacos que actúen sobre genes anómalos, con estrategias para sustituirlos o corregir su función. Nuestro enfoque, inevitablemente restringido, no se dirige al estudio de enfermedades concretas, más propio de tratados especializados (clínicos, metodológicos, etc.), sino a la descripción de algunos aspectos aplicados en el diagnóstico y la terapéutica, partiendo de la información aportada en los capítulos precedentes y siguiendo como objetivo los puntos incluidos en los capítulos siguientes.

24.5.1 Búsqueda de genes por análisis de ligamiento

Un primer paso en la búsqueda de genes es la determinación aproximada de su ubicación en los cromosomas. El análisis de ligamiento clásico, ya estudiado para la cartografía del genoma (págs. 241-243), permite averiguar la proximidad entre dos genes a partir de la frecuencia con que éstos aparecen recombinados en la descendencia. Se consigue una localización más precisa si el análisis del gen buscado se hace no con respecto a otro gen, sino a marcadores genéticos, en especial los polimórficos (pág. 243).

24.5.2 Empleo de métodos físicos para la identificación de genes

Estos proporcionan una localización más precisa de los genes, ya que se describen en distancias absolutas de pb (pág. 244). Se hace uso de métodos basados en algunas características de los genes, ausentes en las regiones no codificantes.

- **Islotes CpG.** Se conoce la presencia de islotes CpG hipometilados (págs. 283-284) en la región promotora de genes de vertebrados que se transcriben de manera activa; más del 50% de los genes humanos presentan dichos islotes. Es posible utilizar las genotecas (pág. 236) para buscar este tipo de secuencias y así identificar genes potenciales, aún no caracterizados.
- **Marcos de lectura abiertos.** Puesto que la región codificante de todo gen ha de empezar por un codón de inicio y finalizar en uno de terminación, la búsqueda en las genotecas de segmentos de DNA comprendidos entre estos tripletes, segmentos denominados **marcos de lectura abiertos** (ORF, de *open reading frame*; pág. 318), es otro método para encontrar genes potenciales.
- **Genotecas de cDNA.** Evidentemente, un gen que se expresa en un tipo celular determinado ha de estar representado por el mRNA transcrita a partir de él. El análisis de genotecas de cDNA (pág. 238) es, por tanto, otra fuente para la búsqueda de genes nuevos.

24.5.3 Identificación de genes por secuenciación

La secuenciación sería la forma más directa y sencilla de analizar un gen, de detectar un polimorfismo o una mutación. Sin embargo, en la práctica, este objetivo no es fácil dado el tamaño y complejidad del genoma humano y la especialización técnica que comporta la secuenciación a gran escala. Es preferible seleccionar o “cribar”, por alguno de los métodos anteriores, los fragmentos candidatos antes de secuenciarlos.

24.5.4 Clonación funcional

Esta aproximación se plantea en los casos en los que se ha caracterizado el producto génico antes de conocer la alteración del gen. Por ejemplo, la proteína responsable de una enfermedad (como la globina β^S en la anemia de células falciformes). El objetivo es, entonces, identificar la secuencia del gen alterado a partir de la secuencia de la proteína anómala, proceso conocido como **clonación funcional**. A partir de la secuencia de aminoácidos se deduce la secuencia de nucleótidos del gen y se diseña una sonda (pág. 184) que permita localizar por hibridación el alelo mutante asociado a la enfermedad.

24.5.5 Identificación de genes y bioinformática

La bioinformática o conjunto de tecnologías de la información aplicadas al estudio de los fenómenos biológicos ha alcanzando su cenit con ocasión del Proyecto Genoma Humano. Ha sido espectacular la tecnología desarrollada para dilucidar el ordenamiento de los tramos que se van secuenciando. Se habla ya de la bioinformática como una ciencia que recoge información de áreas de conocimiento antes tan separadas entre sí como la bioquímica, la biología molecular, la estadística, la farmacología, la física, la fisiología, la genética, la química física, etc., en colaboración con el desarrollo de las nuevas generaciones de ordenadores, internet, la robótica, etc.

A fin de identificar las secuencias que definen los genes, se han elaborado los llamados **programas predictores de genes**; se basan en distintas aproximaciones y combinan el empleo de sistemas de reglas, redes neuronales, inteligencia artificial, modelos estadísticos y métodos de análisis lingüístico.

24.6 DETECCIÓN DEL POLIMORFISMO DE SECUENCIA DE DNA

Para detectar las variaciones (polimorfismo) en la secuencia del genoma de distintos individuos se acude a estrategias diversas. En principio, la forma más directa de detectar un polimorfismo es la secuenciación del DNA, pero por su laboriosidad se suele acudir a otras técnicas, basadas en la hibridación de Southern (pág. 171) o en la PCR (capítulo 14).

24.6.1 Polimorfismo de un solo nucleótido

El caso conceptualmente más simple es aquel en el que el polimorfismo se manifiesta en forma de nucleótidos alternativos en una cierta posición del genoma; se habla de polimorfismos de un solo nucleótido o **SNP** (acrónimo inglés de *single nucleotide polymorphism*; coloquialmente, “*snip*”).

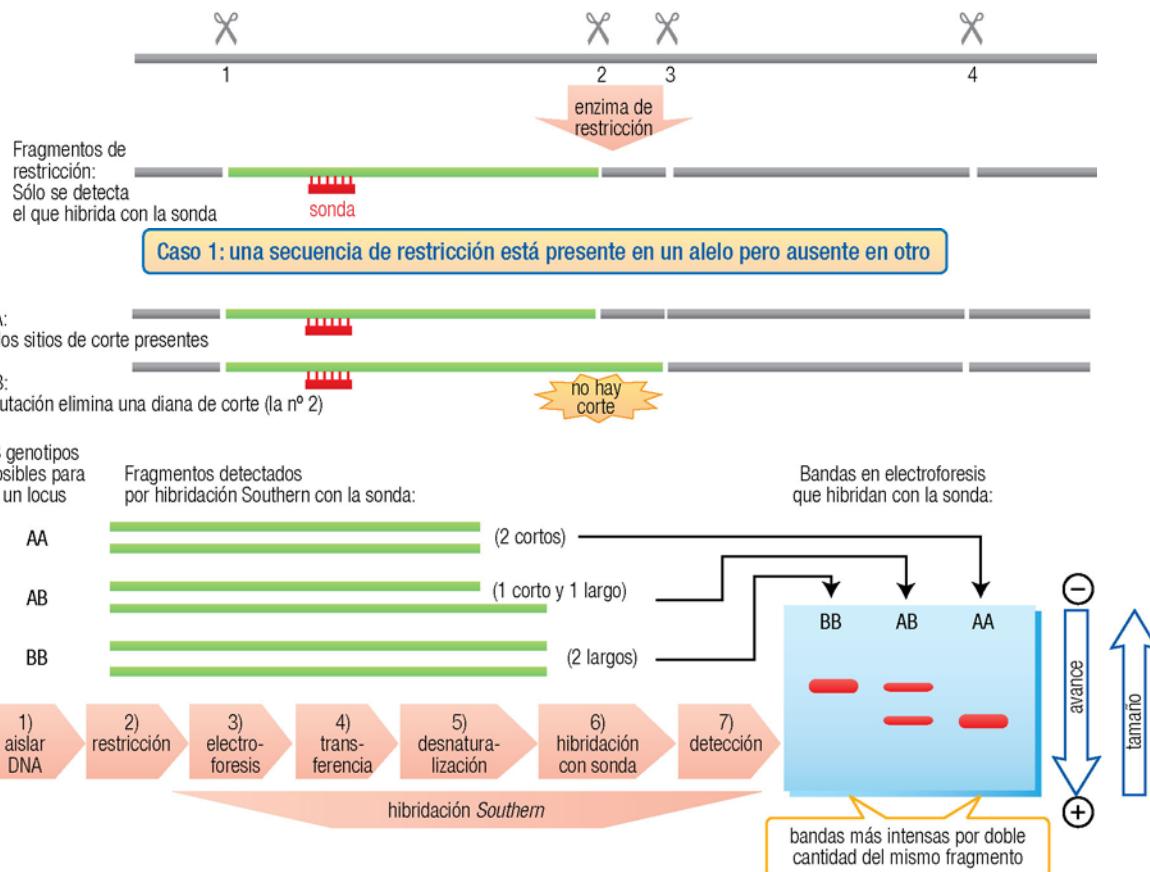
24.6.1.1 RFLP: detección de polimorfismo en la longitud de los fragmentos de restricción

Éste es uno de los primeros métodos desarrollados para la detección de polimorfismos, y el que más se ha empleado durante muchos años. Se basa en la detección de aquellas variaciones de la secuencia del DNA (codificante o no) que tienen como consecuencia un cambio en una diana de restricción. Casi siempre la variación afecta a un solo nucleótido. Por tanto, los fragmentos de restricción que se obtienen son diferentes dependiendo de qué alelo esté presente en la muestra; de ahí su nombre (*restriction fragment length polymorphism*).

Cada muestra de DNA generará múltiples fragmentos, de longitud diferente en función de los sitios de restricción presentes. Puesto que muchos de ellos no corresponden a la región que se pretende estudiar, generalmente se utiliza una sonda que hibrida en posición cercana al sitio polimórfico y se analizan los fragmentos por el método de Southern (pág. 171). Con el fin de poder utilizar el método sobre cantidades

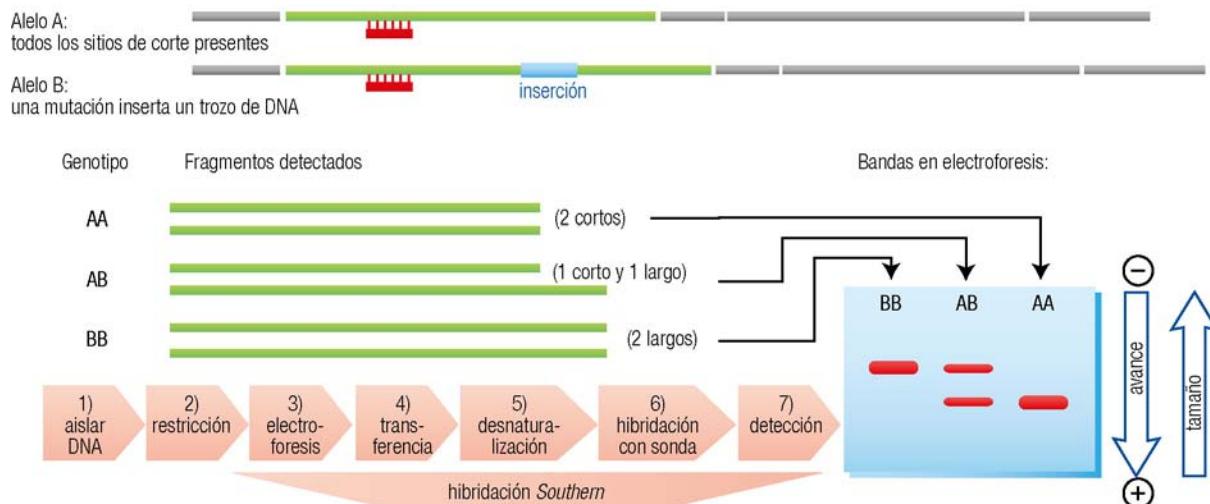
pequeñas de muestra, es común realizar una amplificación previa por PCR; ello exige disponer de un par de cebadores específicos para secuencias que flanquen la región polimórfica. En ese caso, puede ser innecesario el uso de la hibridación, pues la cantidad de DNA es suficiente para detectarlo en el gel con un método de tinción, y se podrá observar la diferencia en el patrón de fragmentación.

Cromosoma o región de DNA amplificada por PCR: corte con una enzima de restricción



24.8b

Caso 2: dentro del fragmento de restricción detectado existe inserción o delección en un alelo



24.8c



Web 24.4. Ejemplo: polimorfismo en la globina β , causante de la anemia de células falciformes.

Los perfiles de RFLP obtenidos sobre muestras de distintos individuos, dentro de una cierta similitud propia de la especie, reflejan la diversidad del genoma en la población. Al disponerse de centenares de enzimas de restricción y de la posibilidad de sintetizar sondas con cualquier secuencia, son enormes las posibilidades de combinación para el análisis genético; se han detectado así muchos sitios RFLP en el genoma humano. Por ello, adquieren el carácter de **marcadores genéticos** (pág. 240), con múltiples aplicaciones (pág. 424). Por convenio, como ya se ha indicado (pág. 406), esta variación sólo se califica estrictamente como polimorfismo si aparece en más del 1% de la población.

La técnica RFLP se ha visto desplazada en gran medida por otras, pero fue el primer método asequible para establecer perfiles de DNA, con aplicación en cartografía del genoma, ubicación de genes responsables de enfermedades, determinación del riesgo de enfermedades y ensayos de paternidad, entre otros.

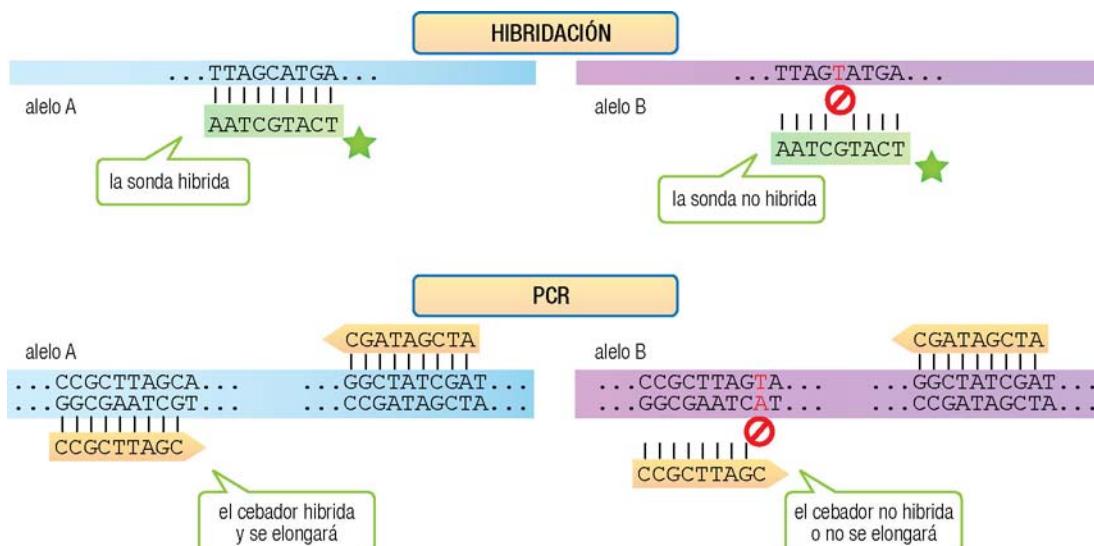
24.6.1.2 Otras formas de detectar polimorfismo SNP

Se han desarrollado infinidad de técnicas para el genotipado de polimorfismos SNP. Más que intentar una descripción exhaustiva, procede exponer los principios generales en los que se basan; los detalles experimentales son una cuestión especializada y en continua evolución.

El método último para detectar polimorfismos sería la secuenciación del DNA genómico pero, obviamente, esto es demasiado costoso en tiempo y dinero para usarse como método de rutina. Por ello, sólo se utiliza para la caracterización inicial del polimorfismo. Una vez conocido el nucleótido variable y la secuencia en la que está inmerso, se acude a otras técnicas para discernir cuál es la variante presente en una muestra.

La mayor parte de los métodos que no dependen de la alteración de dianas de restricción –si no todos– se basan en técnicas de hibridación o de PCR “específicas de alelo”. Se trata de diseñar oligonucleótidos complementarios a cada uno de los alelos (denominados **oligonucleótidos específicos de alelo** o ASO, *allele-specific oligonucleotides*). Éstos pueden emplearse como sondas de hibridación o bien como cebadores de PCR. La discriminación entre alelos, basada en una sola diferencia puntual de la secuencia, hace uso de diversos diseños experimentales.

- **Hibridación específica de alelo.** La detección directa de la hibridación requiere, evidentemente, el uso de condiciones de elevado **rigor** (pág. 168), de modo que el oligonucleótido no se empareje con la muestra si no hay complementariedad de bases perfecta en el punto polimórfico. Habitualmente, deben ensayarse en paralelo los oligonucleótidos con secuencia complementaria a cada uno de los alelos, pero puede ser factible su examen simultáneo si se emplean marcadores o etiquetas diferentes en cada uno (por ejemplo, fluorocromos distintos). Para la detección de la sonda marcada e hibridada se utilizan estrategias experimentales análogas a las descritas para la PCR en tiempo real (pág. 205).



- **Extensión del cebador.** Si se diseña un cebador cuyo extremo 3' coincide con el punto polimórfico, la DNA polimerasa en la PCR no lo elonga si su último nucleótido no empareja con el molde. En otras aproximaciones el cebador puede hibridar con una región común a ambos alelos y luego, empleando etiquetas marcadoras o incluso espectrometría de masas, se detecta qué nucleótidos se incorporan cuando se elonga el cebador sobre la secuencia polimórfica.
- **Enzimas sensibles al desemparejamiento parcial.** En algunas técnicas se tolera la hibridación imperfecta pero se aprovecha la selectividad de enzimas. Aparte de la propia DNA polimerasa, ya mencionada, se han utilizado ligasas para unir dos oligonucleótidos hibridados en posición contigua, sólo si están completamente emparejados, o también la endonucleasa FEN1 (web 11.9), que es muy sensible a las bases desemparejadas.
- **Propiedades del DNA bicatenario.** Tanto la hibridación de una sonda como la elongación de cebadores mediante PCR generan moléculas bicatenarias. Se pueden detectar éstas por la incorporación de intercalantes fluorescentes (análogo a lo descrito en págs. 205-206). Otras técnicas aprovechan las diferentes propiedades derivadas del emparejamiento perfecto o imperfecto de las cadenas. Por ejemplo, en la SSCP (*single strand conformation polymorphism*, polimorfismo de conformación de moléculas monocatenarias) los productos de PCR se desnaturizan con alta temperatura y formaldehído, y luego se permite que renaturalicen mientras avanzan por un gel de electroforesis. El emparejamiento intracatenario que se genera es diferente como consecuencia de la mutación y conduce a conformaciones alternativas que se separan en el gel. La técnica TGGE (*temperature gradient gel electrophoresis*, electroforesis en gel con gradiente de temperatura) aprovecha, asimismo, las diferencias de movilidad, en este caso resultantes de la desnaturalización gradual al avanzar por un gel cuya temperatura va aumentando. Las diferencias de secuencia entre alelos afectan a la temperatura de fusión del DNA (pág. 167), lo que se traduce en una movilidad diferente de las moléculas en el gel.

24.6.2 Polimorfismo de repetición

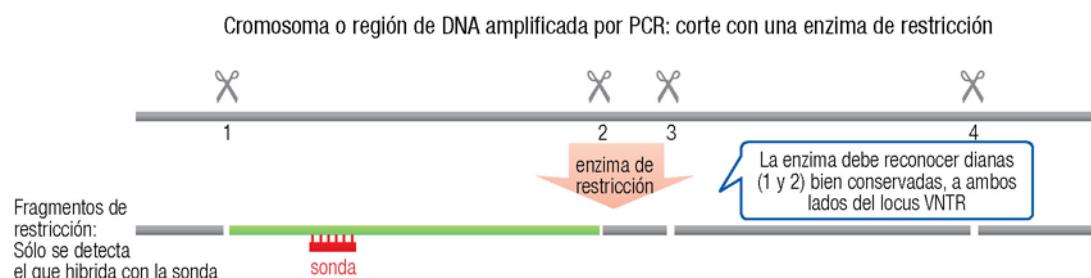
Este tipo de polimorfismo surge de la existencia en el genoma eucariótico de abundantes regiones de DNA repetitivo, en su mayor parte no codificante. Más concretamente, se trata de repeticiones en tandem (DNA satélite, minisatélite y microsatélite, págs. 116-118).

24.6.2.1 VNTR: detección de polimorfismo en el número de repeticiones en tandem

Este polimorfismo, entre individuos y entre los dos alelos de un mismo individuo, surge porque el número de unidades de repetición que forman cada bloque no es siempre el mismo; de ahí el nombre, polimorfismo de número variable de repeticiones en tandem, o simplemente VNTR (*variable number of tandem repeats*). La razón de que varíe con tanta facilidad el número de repeticiones se encuentra en la posibilidad de recombinación entre ellas con sobrecruzamientos desiguales (pág. 407).

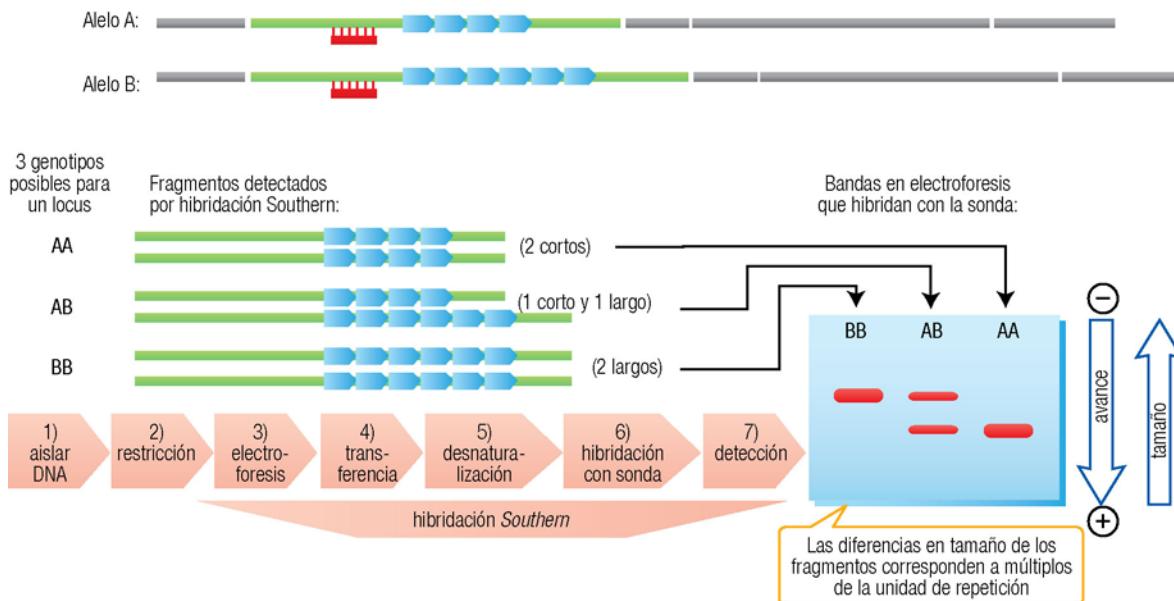
a) Detección por hibridación

La aproximación tradicional, al igual que en el caso de la técnica RFLP, se apoya en el uso de sondas de hibridación para detectar ciertos fragmentos de restricción y en el análisis del tamaño de éstos mediante electroforesis.



Caso 1: detección del fragmento de restricción que contiene el bloque de repeticiones:

Cada alelo contiene bloques con un número diferente de unidades de repetición. Por tanto, número variable de **secuencias repetidas** dentro del **fragmento de restricción** detectado por una **sonda** que hibrida fuera de las repeticiones



24.10b

Caso 2: detección del propio bloque de repeticiones:

En este caso, la **sonda** hibrida con la secuencia repetida. El resultado es similar al del caso 1

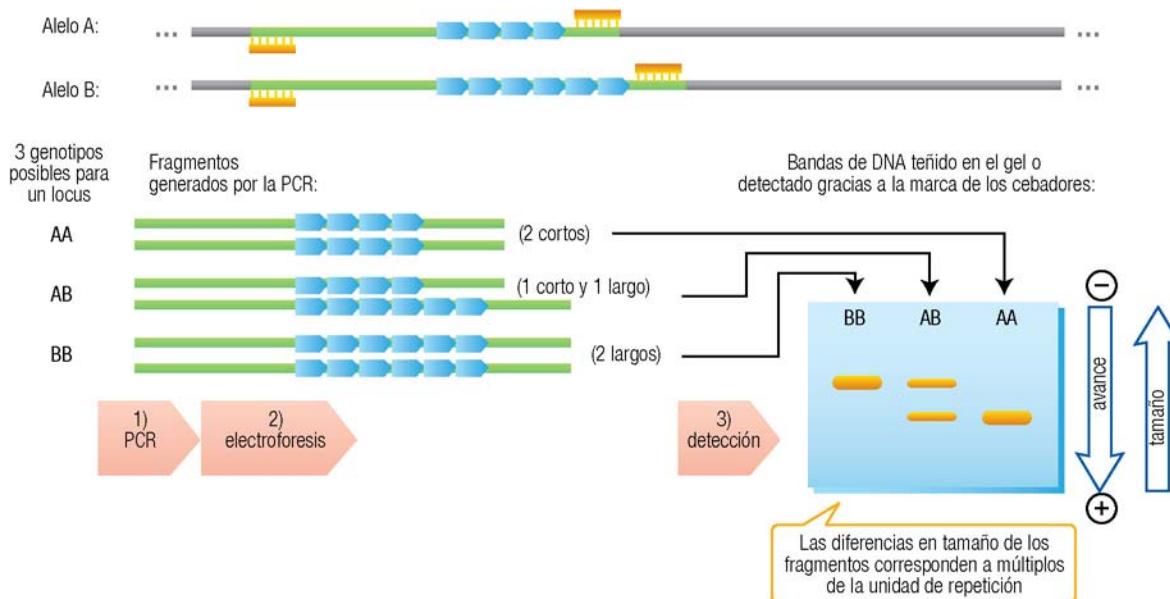


24.10c

Los polimorfismos VNTR que se emplean en la práctica corresponden a bloques moderadamente grandes con unidades de entre 5 y 64 pb, correspondientes a los **minisatélites hipervariables**. Otros tipos no suelen recibir ese nombre, como el polimorfismo de repeticiones sencillas en tandem (unidad de repetición de 1 a 4 pb, es decir, DNA microsatélite) y el polimorfismo de repeticiones en tandem asociado con grandes bloques de DNA satélite. Los **loci hipervariables** reciben este nombre por tener de forma característica muchos alelos. Los marcadores más informativos tienen decenas de alelos, por lo que es poco probable que dos individuos no emparentados compartan los mismos. Estos marcadores VNTR proporcionan mucha información para el análisis de ligamiento genético (mapas genéticos, pág. 239) y para la identificación de individuos (pruebas forenses y de paternidad, págs. 425-426).

b) Detección por PCR

Como alternativa al método de Southern, para detectar el polimorfismo VNTR puede aplicarse una amplificación por PCR (capítulo 14). En lugar de una sonda, se precisan entonces sendos cebadores que hibriden a ambos lados del bloque de repeticiones. Puesto que así sólo se amplifica la región que contiene el locus polimórfico, no es necesario purificar el DNA, ni digerirlo con enzimas de restricción, y en la electroforesis sólo se observarán (sin hibridación, basta la tinción normal de DNA, pág. 141) las bandas de los fragmentos que incluyen el bloque de repeticiones, de tamaños distintos según el alelo presente. Debido a estas ventajas, cada vez más se ha venido utilizando la PCR para la detección de los VNTR, con preferencia sobre la hibridación de Southern.



24.11

24.6.2.2 Polimorfismo STR

La elección de marcadores repetitivos más convenientes para el uso cotidiano ha ido modificándose a lo largo de los años, particularmente para una de sus aplicaciones más extendidas: las pruebas de identificación de individuos. Como resultado, los VNTR clásicos han ido dando paso a otras secuencias repetitivas polimórficas. El FBI (*Federal Bureau of Investigation*, Oficina Federal de Investigación de los Estados Unidos) ha liderado el desarrollo de las técnicas de tipificación de DNA para su uso en la identificación de autores de delitos violentos. En 1997 se seleccionaron 13 loci STR que forman el núcleo central de la base de datos nacional de los EE.UU., llamada CODIS (*Combined DNA Index System*). En Europa se utilizan sistemas similares que comparten varios de los marcadores STR como, por ejemplo, el grupo de loci ESS (*European Standard Set*) definido por el *European Network of Forensic Science Institutes*.

Los STR (*short tandem repeats*, repeticiones cortas en tandem) son secuencias repetitivas de 2 a 7 pb que se localizan en posiciones concretas de los cromosomas, formando bloques de no más de 500 pb. Se clasifican, pues, dentro de la categoría de microsatélites (pág. 118). Entre las ventajas que los han hecho populares destaca su elevada variabilidad en la población, con un gran número de alelos que se diferencian en el número de repeticiones (entre 7 y 38 unidades). Además, el pequeño tamaño de los STR permite analizarlos en muestras de DNA que hayan experimentado degradación parcial, reduce las posibilidades de que sufran mutación y facilita su identificación independientemente de otros loci cromosómicos. En la práctica habitual, la detección se realiza mediante PCR empleando cebadores diseñados específicamente para la región cromosómica que contiene cada uno de los loci. Mediante una combinación inteligente de la posición de los cebadores (que define el tamaño de los fragmentos amplificados) con el uso de cebadores etiquetados con varios fluorocromos, se hace posible la amplificación simultánea de todos los STR (técnica llamada de PCR **múltiplex**) y su análisis mediante electroforesis capilar. Por lo demás, el fundamento de la técnica es común con el de los VNTR descritos en la sección precedente.



Web 24.5. Ejemplo: detección de polimorfismo STR.

Los loci STR habituales (como los de CODIS) están repartidos por los autosomas y se heredan, por tanto, según las leyes clásicas mendelianas. Recientemente los resultados ofrecidos por estos marcadores se suplementan –o se sustituyen, dependiendo de las aplicaciones– con el empleo de marcadores ligados al sexo, con una herencia aún más sencilla de interpretar.

Se han desarrollado **marcadores STR del cromosoma Y**, que permiten investigar la línea de transmisión exclusivamente paterna. Poseen interés en la identificación de individuos cuando se dispone de muestras de referencia familiares a través de varias generaciones, así como en la identificación de restos celulares de varones mezclados con muestras de mujer (por ejemplo, vestigios de semen en agresiones sexuales).

También es posible investigar de manera selectiva la línea hereditaria materna mediante **marcadores de DNA mitocondrial**, secuencias polimórficas ubicadas en las regiones hipervariables del DNA mitocondrial (pág. 416).

24.7 APPLICACIONES DEL POLIMORFISMO GENÉTICO

El análisis del polimorfismo genético da pie a numerosas aplicaciones, tanto en investigación básica como en clínica, que se basan en los tipos de polimorfismos antes descritos.

- Los polimorfismos del DNA que **no tienen efecto fenotípico** se utilizan para la comprobación de la identidad (huella dactilar genética). Por ejemplo:
 - Diagnóstico de la paternidad biológica y seguimiento de árboles genealógicos.
 - Identificación de sospechosos en investigación policial y procedimientos penales, por comparación con el DNA procedente de diversos restos biológicos (sangre, semen, saliva, raíces de cabello, tejidos corporales diversos, piezas dentales, etc.).
 - Identificación post mórtem de individuos, en casos judiciales o catástrofes.
 - Otras aplicaciones de interés diverso. Por ejemplo, es posible aplicar estas metodologías a muestras clínicas con el fin de comprobar identidad de las biopsias, posibilidades de trasplante, inestabilidad genética característica de ciertos tumores, etc.
- El estudio de los polimorfismos que tienen una contribución a la **susceptibilidad** frente a ciertos procesos patológicos posee interés médico, especialmente en la prevención de enfermedades complejas.
- Los polimorfismos que desempeñan un **papel directo** en la aparición de un fenotipo patológico son, evidentemente, aplicables para el diagnóstico, así como para el estudio de los mecanismos moleculares de la enfermedad (patología molecular).

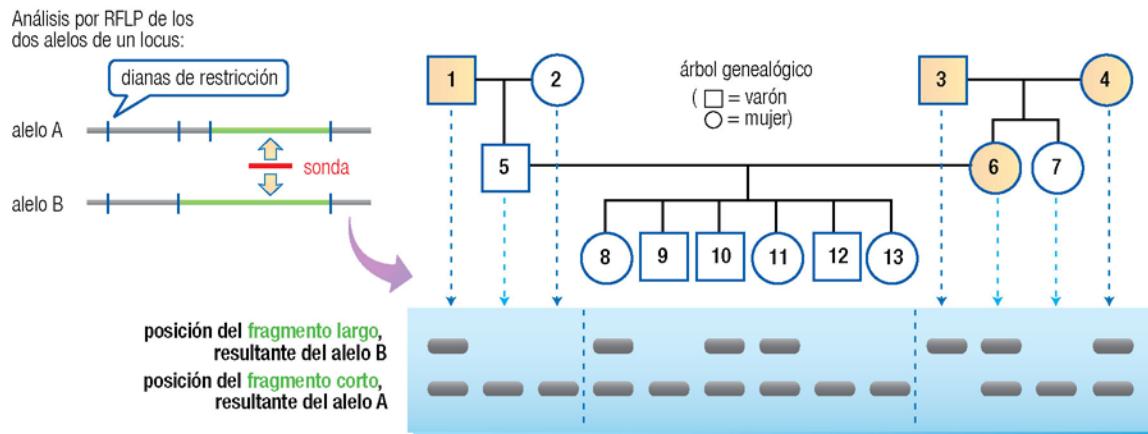
24.7.1 Mapa de polimorfismos SNP

Los polimorfismos SNP o de un solo nucleótido (págs. 418 y 420) constituyen alrededor del 90% de la diversidad del genoma humano. En promedio, existe un SNP por cada 200 pb en el genoma. Aunque la mayoría de ellos no originan directamente enfermedades, en ocasiones pueden localizarse muy cerca de mutaciones o polimorfismos involucrados en procesos patológicos, lo que los hace útiles como marcadores genéticos (pág. 243). Al irse caracterizando cada vez un número mayor de estos polimorfismos, se ha hecho posible la construcción de un **mapa de SNP**, lo que se añade a la utilidad propia de algunos como determinantes de enfermedades. Existen ya millones de polimorfismos SNP bien caracterizados, repartidos por el conjunto completo de los 24 cromosomas humanos (22 autosomas y 2 cromosomas sexuales).

24.7.2 Análisis del DNA en estudios familiares

El análisis de unos cuantos polimorfismos permite realizar un seguimiento de la herencia de los alelos respectivos y así obtener conclusiones sobre las relaciones familiares entre varias personas. De igual manera, se pueden realizar estudios evolutivos y antropológicos, tales como el seguimiento de las migraciones humanas desde tiempos prehistóricos y las relaciones entre los antecesores de la raza humana.

Se ilustra el fundamento de estos estudios con un ejemplo muy simplificado.

**Análisis:**

- 7 no es hija de 3 (le falta el alelo B paterno).
- Si B está asociado a una enfermedad autosómica recesiva (símbolos coloreados):
 - 3 la padeció (homocigótico), mientras que 1, 4 y 6 eran portadores (heterocigóticos).
 - 5 heredó de su padre (1) el alelo A, con lo que se extinguíó la herencia de la enfermedad por esa rama.
 - El gen portador de la enfermedad de 6 fue heredado de su madre (3), no de la madre (4).
 - De la última generación, 8, 10 y 11 son portadores, mientras que 9, 12 y 13 están sanos.

24.12

Una de las principales aplicaciones de estos estudios familiares ha sido relacionar la presencia de una enfermedad genética, causada por un defecto en un gen aún desconocido, con la de un polimorfismo determinado. Mediante un estudio amplio del árbol familiar, basado en la transmisión mendeliana del gen patológico y de los marcadores polimórficos, se puede llegar a encontrar un alelo polimórfico que esté ligado a la enfermedad (por la proximidad de sus loci respectivos en el genoma, págs. 241 y 244). Esto permite, en primer lugar, identificar de forma aproximada la posición del gen patológico y así facilitar su búsqueda y, por otro lado, el diagnóstico preventivo de la enfermedad mediante la detección del alelo polimórfico asociado. Este procedimiento se ha aplicado con algunas enfermedades humanas, como la fibrosis quística y la corea de Huntington.

En una aplicación más simple, el análisis familiar de marcadores polimórficos de DNA se utiliza ya rutinariamente para la realización de **pruebas de paternidad**. En este caso, la necesidad de asegurar los resultados con un grado elevado de certeza requiere que la probabilidad de una coincidencia de la distribución de alelos por puro azar se reduzca a valores suficientemente bajos. Esto se consigue eligiendo marcadores que sean altamente polimórficos en la población y aumentando el número de ellos que se analizan de forma simultánea (véase apartado siguiente).

24.7.3 Identificación de individuos: huella dactilar de DNA

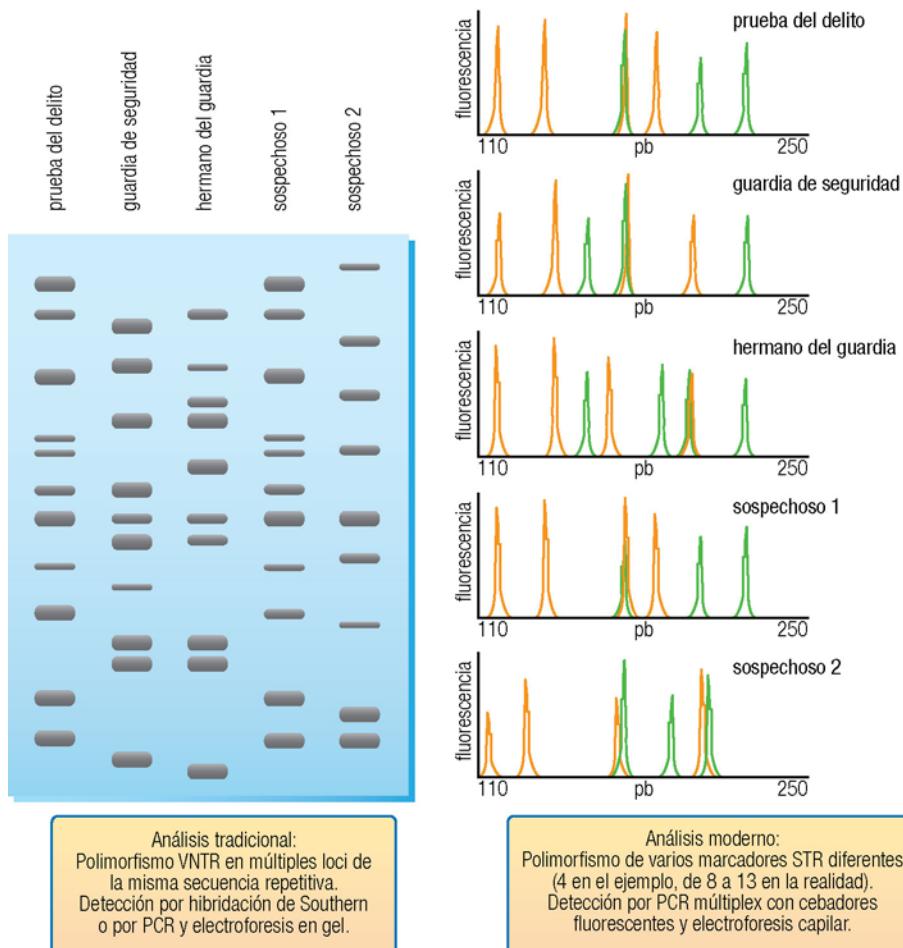
Como se ha visto hasta ahora, los polimorfismos se localizan con una sonda o un par de cebadores particulares, para examinar un locus único. Sin embargo, muchas secuencias estudiadas en los polimorfismos de repetición aparecen en múltiples zonas del genoma (recuérdese que los minisatélites y microsatélites son DNA moderadamente repetitivo y disperso, organizado en bloques dispersos de repeticiones en tandem, pág. 117). Las unidades de repetición presentes en puntos distintos del genoma pueden ser idénticas o bien lo suficientemente similares como para permitir su detección simultánea en todos esos loci. Si se diseña una sonda con una secuencia que hibride con todos los loci de una familia de DNA minisatélite o microsatélite, se puede observar de manera simultánea el patrón polimórfico de todos ellos que, dada la combinación de gran número de variaciones individuales, es prácticamente único para cada individuo. Por ello, el patrón resultante se denomina “huella dactilar de DNA” (*DNA fingerprint*). Sólo los gemelos idénticos presentan huellas indistinguibles.

Una situación similar se alcanza al estudiar marcadores polimórficos, como los STR, que están localizados en una única ubicación del genoma, si se analizan simultáneamente varios de ellos. Por ejemplo, analizando los 13 marcadores STR del CODIS (pág. 423) se observan entre 13 y 26 bandas de DNA (dependiendo de si el individuo es homocigótico o heterocigótico para cada locus). Se ha calculado que la probabilidad de que dos personas de raza blanca no emparentadas tengan una combinación idéntica de alelos en los 13 marcadores es de

uno en $5,75 \times 10^{14}$. Por muy baja que parezca esta cifra, significa que en una población de 100 millones de personas habrá 8 o 9 casos de dos personas que coinciden en los 13 loci.

La huella genética puede utilizarse en estudios familiares, pruebas de paternidad, identificación de recién nacidos, criminología, pruebas forenses, etc. Otra de sus aplicaciones es la detección rápida de cambios somáticos del genoma, por ejemplo, al comparar DNA de tejidos normal y neoplásico. En una alta proporción de tumores malignos, la huella genética sufre cambios, que consisten en la aparición o en la pérdida de bandas.

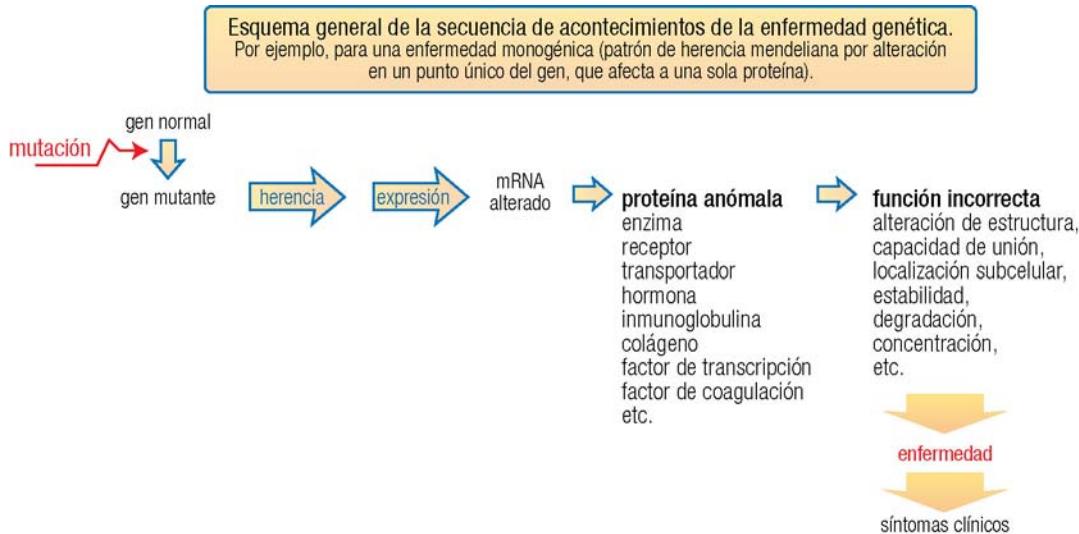
Ejemplo de aplicación de la huella genética (huella dactilar de DNA) a la investigación de un delito. Se ha producido el robo de una obra de arte en un museo. Al romper la vitrina el ladrón dejó una mancha de sangre a partir de la cual los investigadores han aislado el DNA. La policía interrogó al guardia de seguridad, a un hermano suyo y a dos delincuentes habituales, y obtuvo muestras de DNA de todos ellos. Aparte de identificar al probable autor del delito, puede observarse cómo la huella es más parecida entre familiares que entre individuos no relacionados.



24.13

24.7.4 Análisis de polimorfismos con fines diagnósticos

El análisis de polimorfismos es también un método óptimo para el diagnóstico, en especial por la posibilidad de detectar la enfermedad antes de que se desarrolle, o de evaluar la predisposición genética para ciertas enfermedades cuya causa no es única. Es de enorme interés, por ejemplo, el diagnóstico prenatal de enfermedades congénitas, es decir de “toda anomalía del desarrollo morfológico, estructural, funcional o molecular presente al nacer (aunque pueda manifestarse más tarde), externa o interna, familiar o esporádica, hereditaria o no, única o múltiple” (definición de la OMS).



24.14

Una vez realizada la exploración médica del paciente, es decir, delineando el fenotipo clínico, la contribución del laboratorio al diagnóstico puede hacerse a tres niveles: expresión y herencia de los genes (nivel génico o genoma), efecto sobre el producto génico anormal resultante (proteoma), y vía metabólica o desequilibrio químico afectado (nivel bioquímico). Cualquiera de estas situaciones constituye un objetivo y debe ser estudiada en patología molecular, una materia hoy día multidisciplinar.

El conocimiento de la causa de las enfermedades genéticas en familias se ha acelerado gracias al análisis de marcadores genéticos. El diagnóstico evoluciona hacia el conocimiento de un catálogo de variaciones genéticas mediante una aplicación tecnificada, en conjunción con la bioinformática. Se trata de buscar un perfil genético que permita identificar las enfermedades a las que podemos estar predispuestos, desde el punto de vista de las mutaciones. En definitiva, supone un gran paso hacia la **medicina predictiva**, que permite tomar medidas terapéuticas precoces para eliminar el riesgo de enfermedad.

24.7.4.1 Diagnóstico de errores congénitos del metabolismo

Los errores congénitos se pueden considerar como paradigma de los trastornos bioquímicos determinados genéticamente en la estructura y función de moléculas proteicas. La enfermedad es simplemente la ilustración más evidente de una variación genética que afecta a la salud.

A diferencia de los adultos, en quienes se suelen utilizar muestras de frotis bucales, sangre u otros tejidos, en el diagnóstico prenatal se emplean dos tipos de técnicas: no invasivas, que incluyen ultrasonidos y radiografía, e invasivas, como amniocentesis, biopsia de vellosidades coriónicas, cordocentesis o biopsias de hígado fetal (pág. 122). Si el contenido de DNA en la muestra no es suficiente, puede amplificarse la región de interés mediante PCR (Capítulo 14).

a) Valoraciones preliminares

Entre los parámetros de carácter básico o preliminar para la monitorización fetal y el diagnóstico prenatal se emplean la **gonadotropina coriónica** (hCG), una glicoproteína producida por el embrión en desarrollo, que se detecta como indicador de embarazo desde los primeros días de la concepción, el **lactógeno placentario** (hPL) sintetizado por la placenta, que aumenta en sangre materna hasta el término del embarazo, y el **estriol**, un esteroide sintetizado por el feto y la placenta. Estos ensayos han sido superados por la investigación ultrasónica y cardiotocográfica. Para establecer riesgos de malformación fetal se mide la **fetoproteína α** (AFP), sintetizada en saco vitelino e hígado fetal, que debido a su pequeño tamaño pasa del plasma a la orina fetal y, por tanto, al fluido amniótico y sangre materna. Un aumento por encima de lo normal es índice de anomalías congénitas, especialmente defectos del cierre del tubo neural (espina bífida, anencefalia, etc.). Además de estos parámetros, pueden también medirse, con distintos fines diagnósticos, los niveles de bilirrubina y la concentración de fosfolípidos en líquido amniótico, el pH, gases sanguíneos y concentración de lactato en sangre fetal, etc.

b) Ensayos genéticos

El ensayo genético detecta la patología buscando, como ya se ha descrito, el alelo asociado a la enfermedad, si se conoce su gen, o mediante otro polimorfismo cuyo locus muestre ligamiento con el locus patológico desconocido. Se aplican los métodos ya descritos, basados en el ligamiento de genes, polimorfismos, reacción de la PCR, etc. Como ejemplos cabe citar:

- Medida de un sitio de restricción alterado por la mutación, como en el polimorfismo RFLP con Mst II en la anemia de células falciformes (web 24.4).
- Empleo de oligonucleótidos sintéticos como sondas para la secuencia alterada, cuando el trastorno genético se debe a un cambio de un nucleótido específico. Un ejemplo es la deficiencia en antitripsina α_1 , que conduce a una mayor probabilidad de desarrollar enfisema pulmonar, debida a un cambio en un solo nucleótido (págs. 414-415). Un análisis de Southern con una sonda sintética permite determinar si el DNA contiene la secuencia normal o la mutada.
- Ensayos por PCR. Puesto que esta técnica permite amplificar la región de DNA potencialmente defectuosa, puede emplearse para el diagnóstico prenatal de enfermedades siempre que se conozca el sitio exacto de la mutación.

c) Incidencia de las enfermedades congénitas

Se estima que hay más de 10.000 enfermedades monogénicas, con una prevalencia global de 1% en el nacimiento. En centenares de estas enfermedades, tanto las comunes como las más infrecuentes, se ha reconocido el defecto bioquímico básico y, en muchos casos, se ha aislado y clonado el gen responsable. Esto quiere decir que existe información de algunos o varios de los tres niveles antes descritos, para multitud de trastornos con herencia mendeliana.

Es precisamente esta elevada incidencia de errores congénitos del metabolismo lo que ha motivado la aparición de **programas de cribado neonatal (newborn screening)** en todos los países desarrollados. Su objetivo es diagnosticar lo más precozmente posible la alteración metabólica, antes de que se manifieste en lesiones irreversibles. Esta detección supone iniciar un tratamiento precoz y más eficaz, tratando de favorecer el desarrollo de una vida normal. Incluso en casos, como la fibrosis quística, en los que la enfermedad no se puede evitar, no sólo se puede conocer anticipadamente sino que también es posible prevenir las infecciones o tratar la insuficiencia pancreática.

24.7.4.2 Diagnóstico de enfermedades poligénicas

Cuando no basta con examinar la secuencia de una única región del genoma –la mutación en un solo gen responsable de la enfermedad–, sino que ésta es el resultado de la confluencia de numerosas variaciones en distintos genes –que, además, no determinan inequívocamente la situación patológica sino que contribuyen parcialmente o predisponen (“factores de riesgo”)–, el diagnóstico genético resulta, obviamente, más complejo. En estas situaciones, uno de los principales avances deriva de la aparición de la tecnología de **matrices, micromatrices y chips de DNA** (pág. 172). La posibilidad de analizar una muestra del DNA genómico del paciente, o bien de los mRNA expresados en sus células (respectivamente, genoma y transcriptoma) en un único ensayo contra todas las variantes génicas (SNP o de otro tipo) cuya relación con la enfermedad ya se conoce supone un aporte de información impensable con los abordajes anteriores.

Como modelo de metodologías que permiten este tipo de planteamiento pueden citarse:

- La síntesis, mediante química combinatoria sobre un *chip* de DNA, de colecciones de moléculas de oligonucleótido con todas las secuencias posibles permite descubrir, sin información previa, secuencias asociadas a enfermedades. Dichas secuencias pueden entonces buscarse en las bases de datos del genoma para identificar su ubicación y, posiblemente, el gen al que pertenecen.
- La extracción de mRNA de personas sanas y enfermas permite, mediante su análisis comparado, encontrar los genes que se expresan de forma diferente en la situación patológica, y utilizarlos después para incluirlos en matrices de DNA que permitan diagnosticar a pacientes potenciales.

Se han desarrollado ya varias micromatrices o *chips* de DNA para su uso en diagnóstico, si no rutinario, al menos regular. Por ejemplo:

- Un *chip* para el diagnóstico de la hipercolesterolemia familiar, que permite detectar más de 250 mutaciones que afectan a los genes del receptor de lipoproteínas LDL, de la apolipoproteína B y de la proteína PCSK9, que interviene en la homeostasis del colesterol.
- Un *chip* para ayudar en la predicción de la respuesta individual a fármacos, detectando variantes en más de 30 genes y 90 polimorfismos SNP relacionados con enzimas que metabolizan los fármacos (enzimas de fase I, como la familia del citocromo P450, y enzimas de fase II, implicadas en reacciones de conjugación) o con transportadores, receptores, etc.
- “Oncochips” que incluyen fragmentos de entre 6.000 y 9.000 genes, seleccionados por su función directamente relacionada con procesos tumorales, por la alteración de su expresión en las células cancerosas o por su localización en regiones cromosómicas en las que se detectan alteraciones frecuentes. Estos chips son útiles no sólo para la detección temprana del cáncer, sino particularmente para discriminar entre tipos y subtipos de cáncer cuya evolución (por ejemplo, su potencial metastásico), pronóstico clínico o respuesta a terapias son diferentes.
- Aunque en este caso no hay enfermedad, también existe un *chip* para el genotipado de grupos sanguíneos de cara a definir la compatibilidad entre donante y receptor (no sólo de sangre, sino también para trasplante de órganos). Se caracterizan en él 128 polimorfismos relacionados con 10 sistemas de grupos sanguíneos, lo que incluye para los que no se dispone de antisueros, y antígenos plaquetarios.



Enfermedades moleculares

25.1 CONCEPTO Y CLASIFICACIÓN	431		
25.2 ENFERMEDADES EXÓGENAS	432		
25.3 ENFERMEDADES GENÉTICAS	432		
25.3.1 Extensión y frecuencia de las enfermedades genéticas	433	25.4.3.4 Enfermedades recesivas ligadas al cromosoma X	440
25.3.2 Clasificación de las enfermedades genéticas	433	25.4.3.5 Enfermedades ligadas al cromosoma Y	440
25.3.2.1 Tipo de célula que sufre la alteración	433	25.4.4 Enfermedades mitocondriales	441
25.3.2.2 Producto génico defectuoso que da lugar al trastorno	434	25.5 ENFERMEDADES POLIGÉNICAS O MULTIFACTORIALES	442
25.3.2.3 Extensión de la alteración	434	25.6 ENFERMEDADES CROMOSÓMICAS O CITOGENÉTICAS	442
25.3.2.4 Tipo de cromosoma afectado	434	25.6.1 Introducción	442
25.4 ENFERMEDADES MONOGÉNICAS	435	25.6.2 Clasificación y nomenclatura	443
25.4.1 Ejemplos de enfermedades monogénicas	435	25.6.3 Cromosomopatías numéricas	444
25.4.2 Transmisión en familias	435	25.6.3.1 Euploidía	444
25.4.3 Clasificación de los trastornos monogénicos	436	a) Triploidía	444
25.4.3.1 Enfermedades autosómicas dominantes	437	b) Tetraploidía	445
25.4.3.2 Enfermedades autosómicas recesivas	438	25.6.3.2 Aneuploidía	445
25.4.3.3 Enfermedades dominantes ligadas al cromosoma X	439	a) Mecanismo de generación de la aneuploidía	445
		b) Monosomía	447
		c) Trisomía	447
		25.6.3.3 Mixoploidía	447
		25.6.4 Cromosomopatías estructurales	447

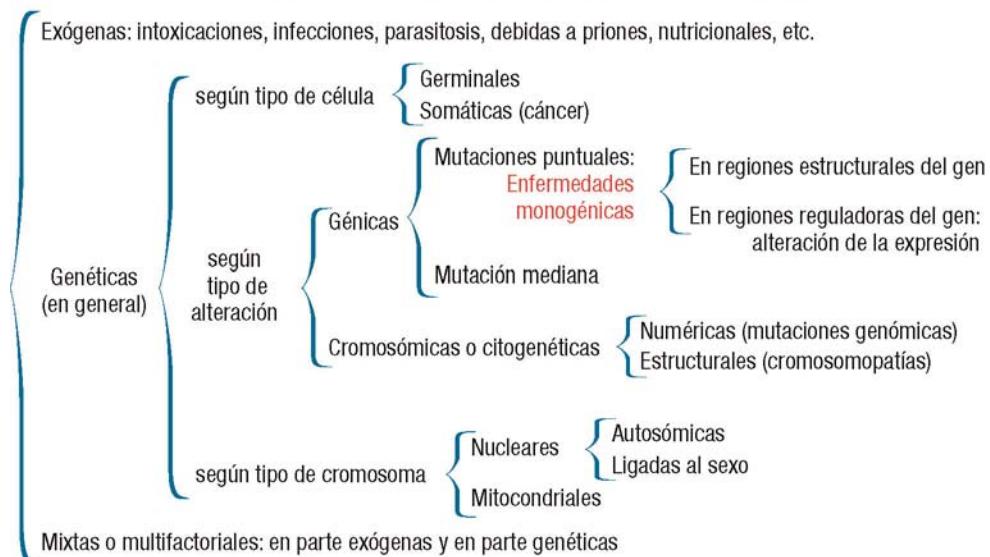
25.1 CONCEPTO Y CLASIFICACIÓN

Se admite hoy día que prácticamente todas las enfermedades tienen una **causa molecular**. Cualquier enfermedad, incluyendo las conocidas tradicionalmente, cursa en mayor o menor medida con alteraciones en la estructura, propiedades, metabolismo o función de las biomoléculas. Entre las causas que inducen la enfermedad se encuentran agentes físicos, ambientales o no; agentes biológicos como virus, bacterias, parásitos, etc., y trastornos genéticos, hormonales, inmunológicos, nutricionales, etc. Todos ellos se traducen en disfunciones celulares y tisulares, temporales o permanentes, cambios en el medio interno, variaciones en la función de órganos y sistemas, y se reflejan, en último término, en forma de enfermedad. Incluso las debidas a agentes externos (traumatismos por accidentes, catástrofes, etc.), que obviamente se consideran de causa no molecular, se convierten una vez iniciadas en enfermedades moleculares. De ahí que un estudio moderno de la enfermedad requiera la consideración de los aspectos moleculares que acompañan a las observaciones clínicas y se consideren válidos, complementarios y no excluyentes los dos tipos de clasificación de la enfermedad:

- Las **clasificaciones clínicas** se basan en el diagnóstico a partir de la observación médica (exploración) con el apoyo de los datos de laboratorio (magnitudes biológicas, isótopos, radiología, etc.). Desde el punto de vista genético, se apoyan en un criterio fenotípico, la manifestación bajo la forma de síndrome clínico. Este tipo de estudio es exclusivo de los profesionales clínicos.
- Las **clasificaciones moleculares**, o estudio de la enfermedad bajo un enfoque molecular, se basan exclusivamente en criterios genotípicos. Dependiendo de las causas que alteran la constitución o la expresión del genoma se distinguen tres tipos: exógenas, genéticas o mixtas. Al no conocerse todavía con precisión la causa molecular de muchas enfermedades, estas clasificaciones son aún imperfectas. Este tipo de estudio lo pueden realizar hoy día diversos profesionales especializados en el área de la biomedicina.

El objetivo de la **patología molecular** es precisamente el conocimiento de la enfermedad desde el punto de vista de su alteración molecular, para contribuir al diagnóstico y a la terapia. Empieza a observarse a comienzos del siglo XX, al establecerse el término **errores congénitos del metabolismo** para describir las alteraciones hereditarias, perennes durante toda la vida del paciente, que afectan a rutas específicas del metabolismo por deberse a la ausencia o inactivación de enzimas específicas (entre otras, albinismo, alcaptonuria, cistinuria y galactosemia). El inicio real y sistemático de la patología molecular, que tiene lugar en las décadas de 1940 y 1950, se desarrolla con las técnicas de estudio de los cromosomas humanos y su papel en el desarrollo sexual y en las anomalías cromosómicas relacionadas; en España, la patología molecular se introduce en la clínica en la década de 1960. El rastreo de determinadas enfermedades en recién nacidos y la ayuda de nuevas técnicas de manipulación y análisis del DNA, a partir de la década de 1970, transformó la genética médica. Al final de la década de 1980 se habían recopilado ya varios miles de alteraciones genéticas hereditarias distintas, casi siempre con aumento o disminución de la actividad de una enzima específica. Bajo el término **genómica** se incluye hoy día no sólo el estudio del genoma, sino también la aplicación de exámenes genéticos para identificar las alteraciones responsables de una enfermedad, detectar en una población las personas con riesgo y así contribuir al tratamiento de la enfermedad. De forma análoga, la **proteómica** está reafirmando el papel que desempeñaron las proteínas, en solitario, en el inicio de la patología molecular.

Clasificación molecular de las enfermedades humanas



25.2 ENFERMEDADES EXÓGENAS

Se denominan así las enfermedades causadas por agentes externos al individuo, de cualquier tipo: intoxicaciones por productos químicos, drogas o medicamentos, parásitos unicelulares o pluricelulares, toxinas biológicas, agentes infecciosos como bacterias y virus, etc. También se incluyen las debidas a alteraciones nutricionales (generalmente carenciales) y las causadas por priones (proteínas de plegamiento anómalo capaces de inducir el mismo en otras moléculas de proteína normal y por ello consideradas “proteínas infecciosas”). Ninguna de ellas se trata de forma específica en esta obra.

25.3 ENFERMEDADES GENÉTICAS

Constituyen el grupo de enfermedades moleculares por excelencia. Para su inclusión en este grupo, la enfermedad ha de cumplir las siguientes características:

- Debe estar determinada genéticamente, es decir, deberse a mutaciones que alteran la secuencia, la organización o la expresión del genoma codificante.

- Debe afectar a moléculas en cantidad, estructura, actividad o función. En especial, se alteran las proteínas, en cantidad o en actividad.
- La alteración bioquímica causante de la enfermedad debe ser el resultado de las alteraciones en las moléculas, que afectan a estructuras celulares o vías metabólicas en las que se encuentran implicadas.

25.3.1 Extensión y frecuencia de las enfermedades genéticas

En cuanto a la extensión de la enfermedad, debe diferenciarse entre individuos **portadores** de defectos genéticos, en los que no se hace evidente la enfermedad por la existencia de un alelo sano que suple la función del gen alterado, y **enfermos** o afectados por el trastorno genético. Este último caso corresponde a las enfermedades genéticas propiamente dichas, de gran relevancia globalmente a pesar de la baja frecuencia de cada una.

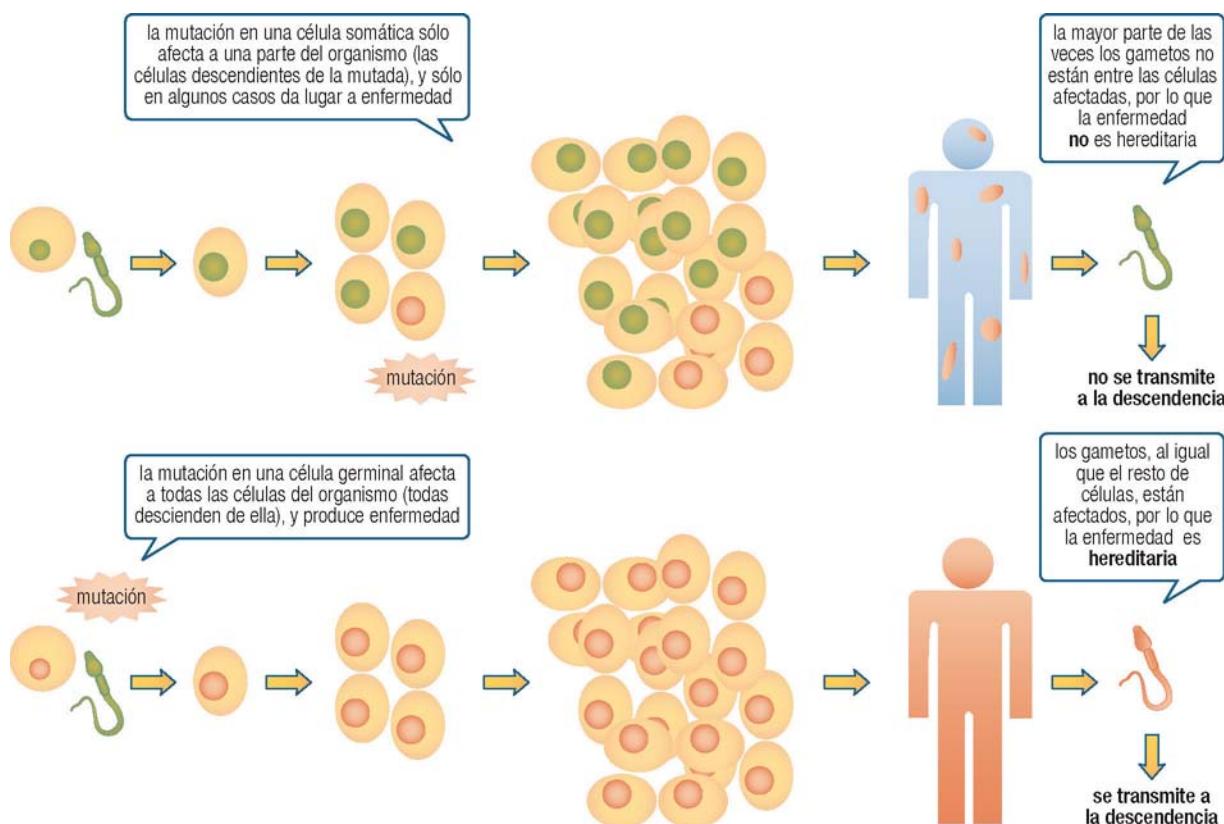
La frecuencia con la que aparece una enfermedad se expresa mediante dos magnitudes epidemiológicas: incidencia y prevalencia. La **incidencia** es la fracción del número de casos respecto al número de individuos nacidos vivos; puede resultar difícil de estimar si la enfermedad no se observa físicamente o no se diagnostica en el momento del nacimiento. La **prevalencia**, o frecuencia de casos en una población de individuos vivos, es aún más difícil de estimar, sobre todo cuando la población es amplia y la enfermedad, infrecuente.

25.3.2 Clasificación de las enfermedades genéticas

Se pueden seguir varios criterios paralelos:

25.3.2.1 Tipo de célula que sufre la alteración

Las enfermedades que se transmiten de generación en generación son de carácter hereditario debido a que se originan por una alteración en las **células germinales** (pág. 384). La alteración de las **células somáticas** no es hereditaria ni afecta a todo el organismo, y da lugar a pocas enfermedades. Entre éstas debe citarse el cáncer (Capítulo 26), que se estudiará como ejemplo de enfermedad somática, a la vez que de carácter mixto o multifactorial.



25.3.2.2 Producto génico defectuoso que da lugar al trastorno

Este criterio ha permitido estudiar numerosas enfermedades genéticas y establecer las clasificaciones propias de la patología molecular tradicional, con muchas variantes según el tipo de proteína alterada.

- La enfermedad se denomina **disenzimia** cuando se altera una enzima. Constituye el caso más frecuente de estudio de una vía metabólica alterada. Se modifican sus características cinéticas o su regulación (sensibilidad a activadores e inhibidores). Si disminuye la actividad, puede desaparecer el producto de la reacción catalizada y, como consecuencia, el producto final de la vía metabólica a la que pertenece la enzima, y también puede acumularse el sustrato correspondiente o un precursor.
- La anomalía puede afectar a una proteína implicada en procesos de **transporte** de moléculas por el organismo (citosol, medio extracelular o sangre) o a través de membranas (plasmática o intracelulares).
- Alteraciones de la **biosíntesis** o la **degradación** de otras proteínas, a través de mecanismos estructurales o procesos metabólicos, tales como transcripción, procesamiento postranscripcional, traducción, plegamiento, modificación postraduccional, distribución, agrupación de subunidades, formación de complejos multiproteicos, etc.
- Alteración de proteínas **estructurales**; por ejemplo, componentes de las fibras musculares o del tejido conectivo.
- Alteraciones de proteínas implicadas en procesos de **transmisión y transducción de señales** (hormonas, factores de crecimiento, neurotransmisores y neuromoduladores, receptores de membrana o intracelulares, etc.), en la **homeostasis** o en la **defensa** (inmunidad).

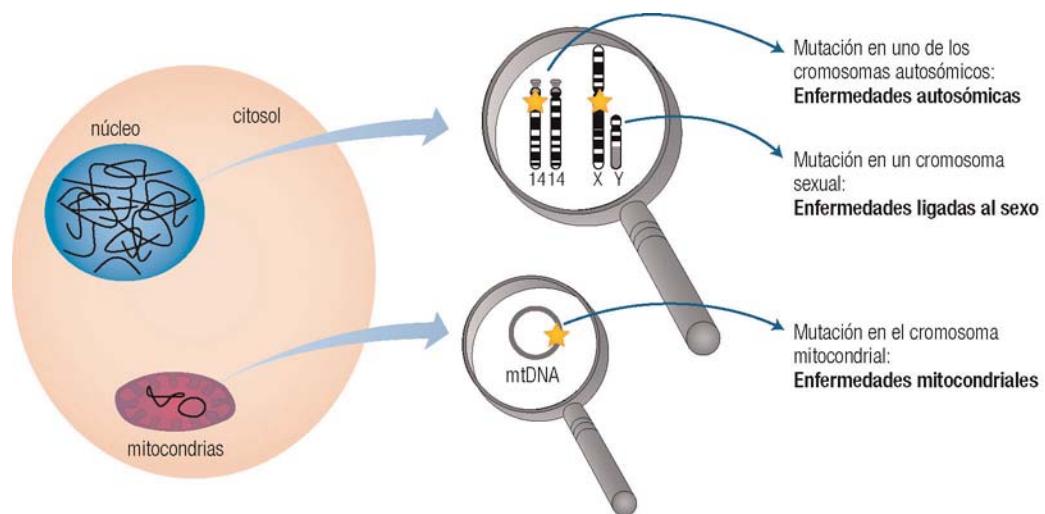
25.3.2.3 Extensión de la alteración

Aunque el tamaño de la alteración no conduce a una clasificación estricta este criterio, ya utilizado para describir las mutaciones (pág. 384), es una simplificación adecuada a efectos didácticos. Se distinguen tres categorías de alteraciones genéticas, y dentro de cada una se pueden considerar otros factores, como origen de la alteración, consecuencias funcionales, mecanismo patogénico, producto génico alterado y región génica afectada.

- La **alteración puntual**, debida a la mutación de un solo nucleótido, da lugar a las **enfermedades monogénicas**, que se describen a continuación.
- Una **alteración mediana** tiene lugar normalmente en secuencias repetidas, pero no se discutirán por ser poco conocidas.
- Las **alteraciones a gran escala, cromosómicas o citogenéticas** afectan a cromosomas completos o a grandes fragmentos cromosómicos. Su incidencia se ha estimado en un 0,65%. Se distingue entre anomalías numéricas, es decir, en el número de cromosomas existentes en la célula –debidas a segregación cromosómica errónea (mutación genómica)– y anomalías estructurales o cromosomopatías –debidas a reordenamientos cromosómicos–. Se describirán más adelante en este capítulo.

25.3.2.4 Tipo de cromosoma afectado

Las enfermedades moleculares pueden afectar al **genoma nuclear**, caso más frecuente por su mayor tamaño, o al **genoma mitocondrial**. Dentro de las nucleares, el defecto genético puede estar localizado en los **cromosomas sexuales** (1 par) o en los **cromosomas autosómicos** (22 pares). Esta distinción, combinada con el carácter dominante o recesivo con el que se expresa la enfermedad, es la clasificación más sencilla, la empleada más comúnmente en genética, y se utilizará para describir la transmisión por herencia de las enfermedades monogénicas (pág. 436).



25.3

25.4 ENFERMEDADES MONOGÉNICAS

Se pueden considerar como modelo de enfermedades genéticas, al ser las que se han estudiado mejor y con mayor rigor a escala molecular. Se deben a una mutación puntual que afecta a un único gen, originando un alelo anormal. La discusión se centrará en las alteraciones localizadas en el genoma nuclear, dejando el caso de las mitocondriales para un apartado específico. Si el alelo anormal aparece en los dos cromosomas homólogos, el individuo es homocigótico para dicho alelo; si aparece sólo en uno, con el alelo normal en el otro, el individuo es heterocigótico.

25.4.1 Ejemplos de enfermedades monogénicas

Las alteraciones monogénicas generan numerosas enfermedades. A título meramente informativo pueden citarse acondroplasia, albinismo, anemia de células falciformes (anemia drepanocítica o drepanocitemia), daltonismo (discromatopsia, ceguera para algunos colores), deficiencias enzimáticas de adenosina desaminasa, antitripsina α_1 (pág. 414), glucosa-6-fosfato deshidrogenasa, galactoquinasa, galactosa-1-fosfato uridililtransferasa (o UTP-hexosa-1-fosfato-uridililtransferasa, causante de la galactosemia, pág. 413) y fenilalanina-4-monooxigenasa (o fenilhidroxilasa, responsable de la fenilcetonuria), distrofia miotónica, distrofia muscular de Duchenne, corea de Huntington, enfermedades de Charcot-Marie-Tooth, de Tay-Sachs, de Von Gierke y de Wilson, mucoviscidosis (o fibrosis quística, pág. 392), hemofilia, hipercolesterolemia familiar, neurofibromatosis de tipo I (págs. 393 y 438), osteogénesis imperfecta, retinoblastoma (pág. 470), síndromes de Hunter y Hurler (mucopolisacaridosis II y I), síndrome del cromosoma X frágil, tumor de Wilms, etc. La información actualizada y más completa sobre caracteres mendelianos humanos (fisiológicos y patológicos) se encuentra en el catálogo de McKusick (*Mendelian Inheritance in Man*), publicado cada 2 años, y se centraliza en la base de datos en internet OMIM (*Online Mendelian Inheritance in Man*); en ambos se le asigna a cada gen un código numérico de 6 dígitos.

25.4.2 Transmisión en familias

Puesto que los trastornos monogénicos dependen del genotipo en un solo locus, sus patrones de herencia familiar son muy sencillos; se llaman **patrones mendelianos** porque se ajustan a lo descrito por la primera ley de Mendel (ley de la segregación). Los dos homólogos de cada par de cromosomas del gametocito primario ($2n$, $4C$) se separan o segregan en la meiosis I (pág. 106), de forma que cada cromosoma homólogo aparece en un gameto secundario (n , $2C$). En la meiosis II se separan al azar las dos cromátidas para formar los gametos (n , C). Globalmente ha tenido lugar una distribución independiente, aleatoria (pág. 107), de los cromosomas paternos y maternos y de sus cromátidas, con lo que cualquier individuo hereda de cada progenitor (y transmite

luego a cada descendiente) un cromosoma de cada par autosómico y un cromosoma sexual. Por tanto, sólo se transmite uno de los dos alelos del gen presente en cada locus.

Para establecer los patrones de herencia de estos trastornos se acude a la información facilitada por los antecedentes familiares del paciente, que se resumen en forma de **árbol genealógico**. En éste se emplean una nomenclatura y símbolos particulares:



Web 25.1. Simbología empleada en los árboles genealógicos.

25.4.3 Clasificación de los trastornos monogénicos

Como se ha comentado (pág. 434), es común clasificar las enfermedades monogénicas de acuerdo con el tipo de cromosoma afectado (sexual o autosómico) y el carácter dominante o recesivo (pág. 98) de su expresión. Si el locus patológico está situado en un cromosoma sexual, se habla de **enfermedad ligada al sexo**; en caso contrario, de **enfermedad autosómica**.

Los caracteres autosómicos y los ligados al cromosoma X se pueden manifestar de forma dominante o recesiva, mientras que los ligados al cromosoma Y no tienen esa doble posibilidad, pues sólo existe un alelo en cada individuo, que ha de manifestarse por sí mismo, sin interacción con un homólogo.

Recordatorio de dominancia de alelos:

- Un alelo es **dominante** sobre otros cuando su presencia se manifiesta siempre en el fenotipo, independientemente de cuál sea el otro alelo (uno paterno y el otro materno). El fenotipo es igual para los individuos homocigóticos del dominante (alelos idénticos en ambos homólogos) que para los heterocigóticos (alelos diferentes en cada uno). En éstos se anula, pues, la funcionalidad del otro alelo, el **recesivo**.
- Dos alelos son **codominantes** (o **semi-dominantes**) cuando el fenotipo heterocigótico refleja la presencia de ambos, siendo intermedio entre los dos fenotipos homocigóticos. Dicho de otro modo, se detecta la expresión de cada alelo en presencia del otro.

Aplicación a la clasificación de las enfermedades genéticas monogénicas:

- Se dice que **una enfermedad es dominante** o sigue un patrón de **herencia dominante** cuando el alelo patológico es dominante sobre el normal. La enfermedad (fenotipo dominante) se manifiesta igual en homocigotos que en heterocigotos; basta un alelo patológico para que se presente la enfermedad. Como ejemplo de enfermedad autosómica dominante, igualmente grave en el genotipo homocigótico y el heterocigótico, puede citarse la corea de Huntington.
- Las enfermedades de carácter semi-dominante o codominante se manifiestan de forma más leve en los heterocigóticos que en los homocigóticos.
- Las **enfermedades recesivas** o con **herencia recesiva** están causadas por un alelo recesivo frente al normal. Sólo se manifiesta la enfermedad cuando los dos alelos son anormales, es decir, en homocigotos.

Frecuencia de las enfermedades monogénicas

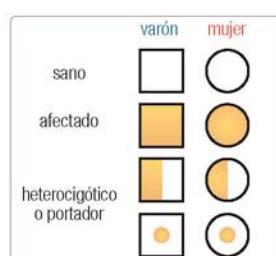
- La incidencia es de 1 caso por cada 100 nacimientos vivos.
- El 70% son dominantes, el 2,5% recesivas, 0,4% ligadas al cromosoma X y muy escasas las ligadas al Y.

25.4

A continuación se presentan los 5 tipos posibles de enfermedad monogénica, de acuerdo con su patrón de transmisión hereditaria.

Criterios de representación empleados para los árboles genealógicos de ejemplo

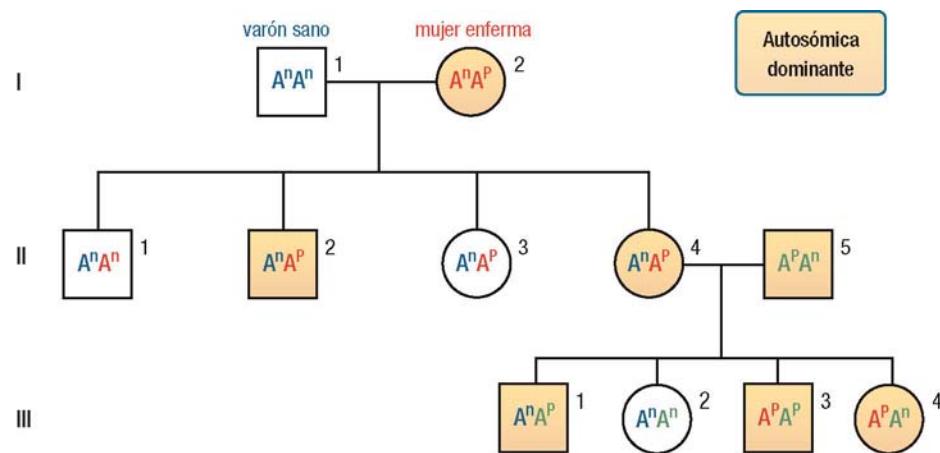
- Se consideran 3 generaciones (I, II y III) y sólo dos parejas (en las generaciones I y II; el 5º árbol es una excepción obligada). Ello supone, obviamente, una limitación en cuanto a posibilidades de emparejamiento y consiguiente aparición de individuos sanos o enfermos, homocigóticos o heterocigóticos, varones o mujeres. Entendemos que la exhaustividad de presentación de posibilidades de transmisión, al igual que el uso de datos familiares reales, va en detrimento de la claridad conceptual que se persigue.
- La generación I está formada por una pareja cuyos dos alelos (siempre referidos a un único locus génico) se muestran en color azul para la primera persona y en rojo para la segunda. La asignación de alelos a cada persona varía en los 5 árboles, para adecuarse a lo que se quiere ilustrar en cada caso.
- La generación II está formada por 4 descendientes, aquéllos resultantes de la combinación de los 4 alelos progenitores. De los 4 hijos, 2 son varones y 2 mujeres.
- La generación III está formada por los 4 descendientes resultantes de la combinación de alelos de la pareja formada por uno de los individuos de la generación II y una nueva persona.
- Para facilitar el seguimiento de la herencia, el genotipo de cada individuo se expresa con los siguientes colores y letras:
 - Alelos del primer miembro de la pareja I.
 - **Alelos del segundo miembro de la pareja I.**
 - Alelos del nuevo miembro de la pareja II (el ajeno a la familia inicial).
 - Letra N (o n) para el alelo normal y P (o p) para el patológico.
 - Letra A para alelos presentes en autosomas.
 - Letras X e Y para alelos en cromosomas sexuales.
 - Letra mayúscula (superíndice) para alelos dominantes.
 - Letra minúscula (superíndice) para alelos recesivos.



25.5

25.4.3.1 Enfermedades autosómicas dominantes

El gen responsable, localizado en uno de los autosomas, posee un alelo patológico (P) que es dominante sobre el normal (n).



Los dos sexos portan los alelos normal (n) y patológico (P) del gen. La enfermedad se transmite por igual a varones y a mujeres.	Por ser autosómica
La enfermedad se manifiesta en homocigotos (PP, III-3) y heterocigotos (Pn = nP). Todos los individuos no afectados son homocigóticos (nn). Cada individuo afectado tiene un progenitor afectado. Los hijos (II-1 a II-4) de pareja mixta entre afectado heterocigótico y no afectado tienen un 50% de probabilidad de heredar el rasgo. O lo que es igual, un individuo afectado heterocigótico (II-4) tiene un 50% de probabilidad de transmitir el rasgo patológico a cada descendiente.	Por ser dominante

25.6

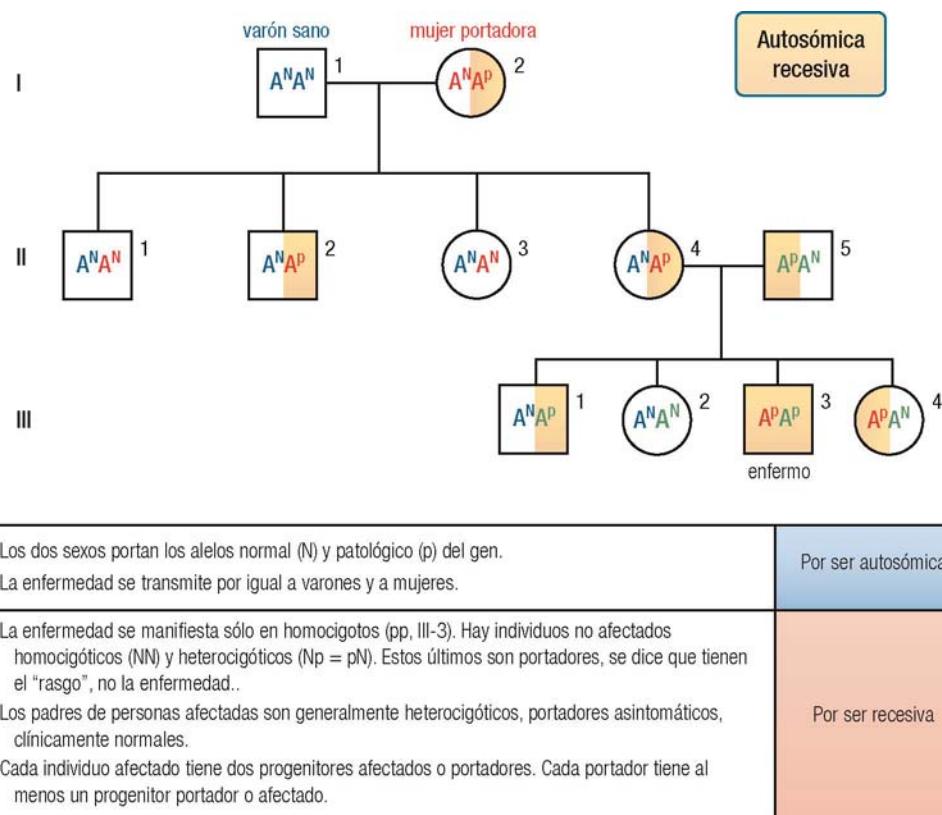
Estos trastornos constituyen cerca del 45% de las enfermedades monogénicas. Aunque muchos son poco comunes individualmente, son tan numerosos que su incidencia en conjunto es apreciable, y en ocasiones especialmente alta en áreas geográficas específicas.

Ejemplos de enfermedades autosómicas dominantes

Enfermedad	Proteína afectada	Locus del gen	Descripción
Hipercolesterolemia familiar de tipo IIA	Receptor de lipoproteínas LDL	19p13.2	
Corea de Huntington	“Huntingtina”	4p16.3	
Neurofibromatosis de tipo I	Neurofibromina	17q11.2	pág. 393
Síndrome de Marfan	Fibrilina-1	15q21.1	

25.4.3.2 Enfermedades autosómicas recesivas

En este caso, el gen responsable, localizado en uno de los autosomas, posee un alelo patológico (p) que es recesivo con respecto al normal (N). En conjunto, este tipo de trastornos representa alrededor del 25% de las enfermedades monogénicas.



25.7

La anemia de células falciformes es un ejemplo típico de trastorno autosómico recesivo que fue, además, la primera prueba experimental de una enfermedad causada por mutación (en la década de 1950). El síndrome clínico es consecuencia de la alteración de un nucleótido en el gen de la globina β (pág. 391), conduciendo a una proteína β^S . Los homocigotos para el gen de β^S manifiestan el trastorno (**anemia**), mientras que los heterocigotos $\beta^A\beta^S$ se dice que muestran el **rasgo falciforme**: producen tanto hemoglobina normal (A) como anómala (S), una parte de sus eritrocitos son falciformes y padecen una anemia leve. Por tanto, la dominancia es incompleta. El diagnóstico es sencillo, por ejemplo mediante RFLP (pág. 418).

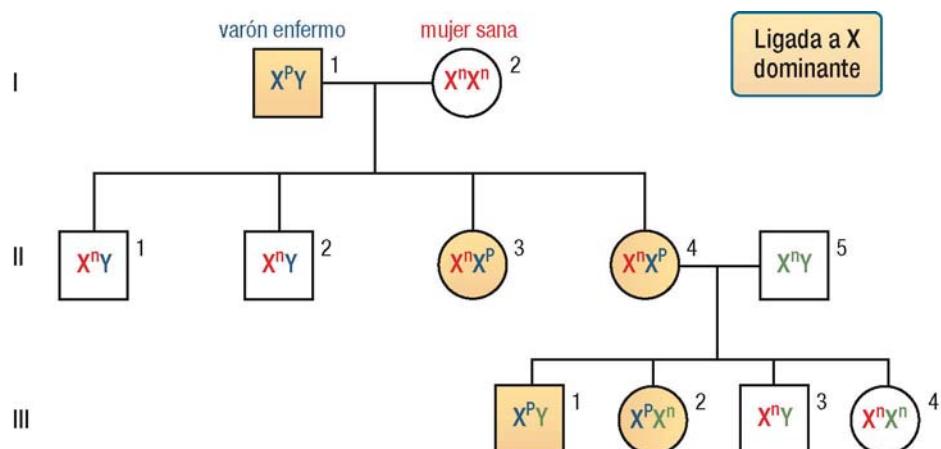
Ejemplos de enfermedades autosómicas recesivas

Enfermedad	Proteína afectada	Locus del gen	Descripción
Deficiencia de ADA	Adenosina-desaminasa	20q13.11	
Enfisema hereditario	Antitripsina α_1	14q32.1	pág. 414
Mucoviscidosis (fibrosis quística)	Regulador de la conductancia transmembranaria (CFTR)	7q31.2	pág. 392
Enfermedad de Tay-Sachs	Subunidad α de la isoenzima A de la β -N-acetilhexosaminidasa	15q23-q24	pág. 391-392
Fenilcetonuria	Fenilalanina-4-monoxigenasa	12q24.1	
Talasemia α	Globina α	16pter-p13.3	
Talasemia β	Globina β	11p15.5	

25.4.3.3 Enfermedades dominantes ligadas al cromosoma X

Las mujeres pueden ser homocigóticas o heterocigóticas para un gen que esté situado en un locus del cromosoma X, pues tienen dos copias de este cromosoma sexual. Sin embargo los varones, al poseer un solo X, no pueden ser homocigóticos ni heterocigóticos para ese gen o rasgo; se dice que el varón es **hemicigótico**.

Si el alelo patológico (P) es dominante sobre el normal (n), las mujeres padecerán la enfermedad tanto si poseen uno como ambos alelos P, mientras que los varones la padecerán si su único alelo es P.



Los dos sexos portan los alelos normal (n) y patológico (P) del gen, pero las mujeres portan dos copias y los varones sólo una.

La enfermedad se transmite a varones y mujeres, pero afecta con mayor frecuencia a éstas.

Las mujeres afectadas (II-4) transmiten la enfermedad por igual a hijos e hijas (III-1 a III-4), con una probabilidad del 50% si son heterocigóticas (Pn) y del 100% si fueran homocigóticas (PP).

Por el contrario, los varones afectados (I-1) sólo la transmiten a sus hijas (II-3 y II-4).

Por ser ligada al X

La enfermedad se manifiesta en mujeres homocigóticas (PP) y heterocigóticas (Pn), y en varones hemicigóticos (P-). Los individuos no afectados son mujeres homocigóticas (nn) y varones hemicigóticos (N-).

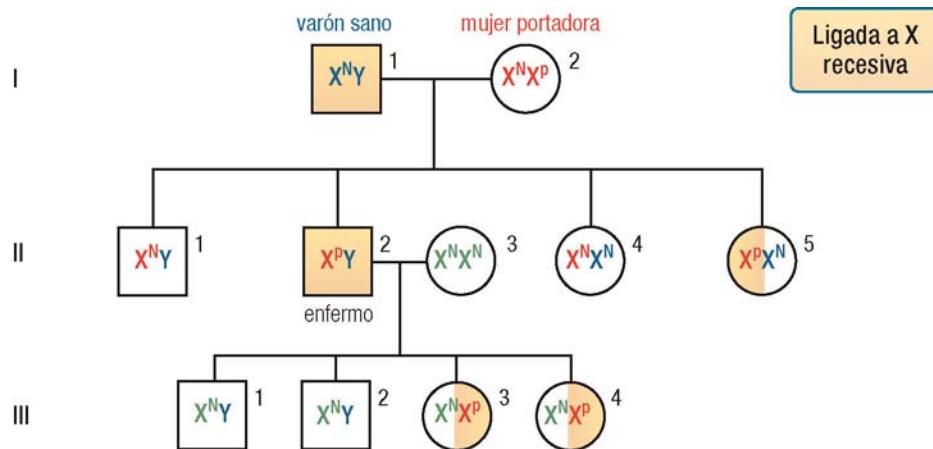
Por ser dominante

Cada individuo afectado tiene un progenitor afectado.

Para compensar la presencia de dos copias del gen en las mujeres (“doble complemento”), sus células sólo expresan los alelos de un cromosoma X (el otro está inactivado como heterocromatina, págs. 88 y 285). En consecuencia, las mujeres heterocigóticas pueden estar o no afectadas, dependiendo de que se exprese el alelo patológico o el normal; en general, esta manifestación es incluso diferente en unas células y otras, pues la selección de cuál de los cromosomas X resulta inactivado tiene lugar durante una etapa embrionaria temprana pero multicelular y es aleatoria en cada una de sus células; una vez establecida la inactivación, todas las células descendientes posteriores mantienen la selección.

25.4.3.4 Enfermedades recesivas ligadas al cromosoma X

Al ser en este caso dominante el alelo normal (N) sobre el patológico (p), las mujeres padecen la enfermedad sólo cuando son homocigóticas y, por tanto, con una frecuencia muy baja. Los varones la padecerán, de forma similar al caso anterior, cuando posean el alelo patológico.



<p>Los dos性s portan los alelos normal (N) y patológico (p) del gen, pero las mujeres portan dos copias y los varones sólo una. La enfermedad se transmite a varones y mujeres, pero afecta con mayor frecuencia a aquéllos. Las mujeres portadoras (I-2) transmiten la enfermedad por igual a hijos e hijas (II-1,2,4 y 5), con una probabilidad del 50%. Las mujeres enfermas (pp) también, pero con un 100%. Por el contrario, los varones enfermos (II-2) sólo la transmiten a sus hijas (III-3 y III-4).</p>	<p>Por ser ligada al X</p>
<p>La enfermedad se manifiesta en mujeres homocigóticas (pp) y varones hemicigóticos (p-). Los individuos no afectados son mujeres homocigóticas (NN) o heterocigóticas (Np) y varones hemicigóticos (N-). Los varones afectados tienen progenitores sanos; la madre es generalmente portadora asintomática.</p>	<p>Por ser recesiva</p>

25.9

Como ejemplo característico, la **hemofilia A** se debe a un defecto en el factor VIII de la coagulación. El gen de esta proteína está situado en el cromosoma X (banda Xq28), por lo que la transmisión sigue el patrón ligado al sexo. Se han identificado varias alteraciones distintas del gen que causan la enfermedad. Sus alelos se expresan de forma recesiva con respecto al alelo normal del factor VIII, por lo que las mujeres sufren hemofilia con mucha menor frecuencia que los varones (se suele decir que la hemofilia A sólo la padecen los varones y no las mujeres, pero esto no es exacto, como se discute a continuación). La frecuencia de alelos patológicos en la población es de alrededor de uno por cada 10.000 alelos normales, por lo que la incidencia en varones (hemicigóticos para el gen) es de 1/10.000. Las mujeres sólo padecen la enfermedad cuando son homocigóticas para alelos patológicos, por lo que la incidencia sería de $1/10^8$; con una población mundial de 7.000 millones, esto equivale a 35 casos vivos de mujeres afectadas en todo el mundo, frente a 350.000 varones; de ahí que se asuma que las mujeres no padecen la enfermedad. En cualquier caso, si es relevante la frecuencia de mujeres portadoras, heterocigóticas para el gen del factor VIII, que pueden transmitir la enfermedad a sus hijos varones (1/10.000). Sólo la descendencia de una mujer portadora ($X^N X^p$) con un varón hemofílico ($X^p Y$) puede dar lugar a mujeres enfermas ($X^p X^p$).

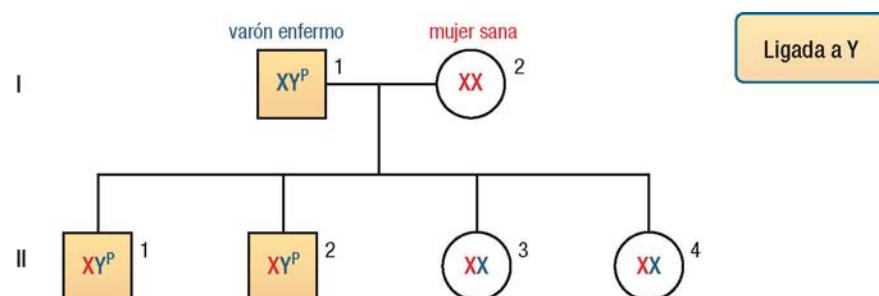
Otros ejemplos de enfermedades recesivas ligadas al cromosoma X son la distrofia muscular de Duchenne ($Xp21.2$) y el daltonismo ($Xq28$).

25.4.3.5 Enfermedades ligadas al cromosoma Y

Se expone este caso por razones didácticas, a pesar de que apenas existen enfermedades ligadas a Y. (Algunas variantes de retinitis pigmentaria se han asociado a la herencia del cromosoma Y. Ciertas patologías con infertilidad están ligadas a Y, pero más que monogénicas son anomalías cromosómicas de tamaño mediano.)

La primera razón de su escasez es el pequeño tamaño del cromosoma Y (en realidad, aún menor, pues parte del Y tiene región homóloga en el X y esos loci se comportan como los autosómicos); por ello, se conocen escasos ejemplos de caracteres asociados a Y. La segunda razón es que para que exista una patología debe haber un gen cuya función sea importante y, teniendo en cuenta que las mujeres carecen de tales loci ligados al Y, no podrían desarrollar esa función.

Aparte de estas consideraciones, en el caso hipotético de una enfermedad o trastorno ligado al cromosoma Y, las mujeres no padecerán la enfermedad ni la transmitirán. Los varones la padecerán siempre que tengan el alelo patológico (P) y la transmitirán a todos sus hijos varones. No tiene sentido el concepto de dominancia, pues nunca coinciden dos alelos en la misma persona.



Sólo los varones son portadores del gen, para los alelos normal (N) y patológico (P).

La enfermedad se transmite sólo a varones, y a todos ellos.

La enfermedad se manifiesta sólo en varones hemicigóticos (P-). Son sanos los varones hemizigóticos (N-).

Cada varón afectado tiene su padre afectado.

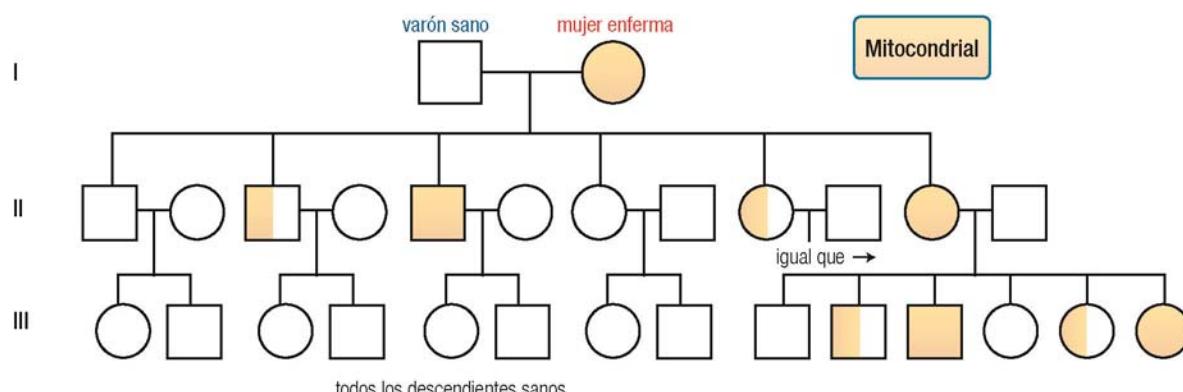
Por ser ligada al Y

25.10

25.4.4 Enfermedades mitocondriales

Estrictamente, algunas enfermedades mitocondriales (es decir, donde se ve afectada la función de estos orgánulos) están causadas por defectos en la parte del genoma nuclear que codifica proteínas mitocondriales. Sin embargo, nos centraremos en aquellas cuya causa se encuentra en el genoma mitocondrial. La herencia mitocondrial es una de las excepciones a los principios de transmisión mendeliana de las enfermedades monogénicas. Esto se debe a dos características peculiares: la herencia por línea materna y la heterogeneidad genética o heteroplasmia.

Herencia materna. Todos los individuos heredan el DNA mitocondrial (mtDNA) de su madre, ya que el óvulo es muy rico en mitocondrias (entre 100.000 y 300.000 por óvulo), mientras que el espermatozoide tiene muchas menos (unos cientos) y, además, se degradan selectivamente tras la fertilización. Por tanto, una mutación patológica en el genoma mitocondrial de una mujer se puede transmitir a toda su descendencia, mientras que el varón en raras ocasiones transmite una mutación análoga.



25.11

Heteroplasmia. Cada célula tiene varias mitocondrias, y cada una de éstas, múltiples copias del mtDNA. Es decir, en una célula hay muchas moléculas independientes de mtDNA, que no siempre tienen exactamente la misma secuencia, en especial debido a su elevada tasa de mutación (pág. 416). Esta heterogeneidad del mtDNA en cada célula, y también entre las distintas células de un individuo, se denomina heteroplasmia. Además, cuando la célula se divide por mitosis, sus mitocondrias se reparten de modo aleatorio entre las dos células hijas. Como consecuencia de todo esto, la proporción de genomas mitocondriales mutados respecto a los originales es variable dentro de cada mitocondria, cada célula, cada tejido y cada individuo. Esto hace que la manifestación de la enfermedad asociada a la mutación sea muy variable y complica los estudios clínico y genealógico. Las enfermedades mitocondriales poseen síntomas muy variados, de gravedad diversa, y afectan a distintos tejidos, incluso para una misma causa genética.

Entre las enfermedades asociadas a mutaciones del DNA mitocondrial se pueden citar las siguientes:

- Neuropatía óptica hereditaria de Leber (*LHON*, *NOHL*). Fue la primera enfermedad en la que se demostró la herencia por vía materna de un defecto en el genoma mitocondrial.
- Oftalmoplejía externa progresiva crónica (*CPEO*).
- Síndrome de Kearns-Sayre (*KSS*), una oftalmoplejía externa con complicaciones varias.
- Síndrome de epilepsia mioclónica asociada a fibras rojas rasgadas (*MERRF*).
- Encefalomielopatía mitocondrial con acidosis láctica e incidentes similares a apoplejía (*MELAS*).

Además, es probable la implicación de mutaciones mitocondriales somáticas o adquiridas como uno de los factores que contribuyen al proceso de envejecimiento y a enfermedades neurodegenerativas, incluyendo la enfermedad de Alzheimer.

25.5 ENFERMEDADES POLIGÉNICAS O MULTIFACTORIALES

Muchas enfermedades genéticas no están causadas únicamente por el defecto de un gen y constituyen un grupo de trastornos conocidos bajo nombres diferentes, de difícil delimitación o caracterizados de forma incompleta: enfermedades poligénicas, multifactoriales, mixtas o complejas. A diferencia de las enfermedades monogénicas, la manifestación clínica es un reflejo del efecto combinado o interacción acumulativa de alteraciones producidas por factores genéticos (múltiples mutaciones, a veces simultáneas en un número impreciso de genes), epigenéticos (modificaciones de la cromatina que varían la expresión génica, pág. 282) y ambientales o exógenos (de todo tipo: nutricional, agentes tóxicos, estrés, etc.). En cualquier caso, la complejidad del origen explica el conjunto de efectos diversos producidos y los distintos nombres que reciben. Estas enfermedades multifactoriales no deben confundirse con los trastornos cromosómicos (enfermedades genéticas a gran escala, descritas más adelante en este capítulo), aunque en ambas pueda tener lugar una alteración compleja del genoma.

Al no ser trastornos de un gen único, la herencia de estas alteraciones no sigue la genética clásica mendeliana. A pesar de ello, sí se observa una mayor transmisión de estas enfermedades dentro de una familia que entre individuos no emparentados. Para que el síndrome se transmita a otro miembro de la misma familia es necesario que se reúna una combinación de cambios genéticos.

Entre estas enfermedades pueden citarse varias malformaciones congénitas, como el labio leporino, las cardiopatías congénitas y los defectos del tubo neural (anencefalia, espina bífida y otras), y ciertas enfermedades de la edad adulta, como algunos tipos de cáncer y las enfermedades de las arterias coronarias, en las que se ha observado una predisposición genética. De todas ellas sólo se considerará aquí el cáncer, la más estudiada desde el punto de vista molecular, además de por su trascendencia clínica (Capítulo 26).

25.6 ENFERMEDADES CROMOSÓMICAS O CITOGENÉTICAS

25.6. I Introducción

En los apartados siguientes se describirán las enfermedades cromosómicas, uno de los tipos de enfermedad molecular (pág. 434). Como su nombre indica, se originan por la alteración del número o la estructura de los cromosomas; de ahí que se denominen también alteraciones cromosómicas a gran escala, aberraciones, trastornos o anomalías cromosómicas. Al poder observarse los cambios con el microscopio óptico, se las considera también enfermedades citogenéticas.

Son alteraciones esporádicas y de baja frecuencia, que raramente se perpetúan por ser casi incompatibles con la supervivencia y la reproducción. Sin embargo, cuando lo hacen muestran manifestaciones clínicas, físicas o psíquicas muy graves, a veces dramáticas, por lo que pueden considerarse las principales enfermedades genéticas. De hecho, son responsables de una gran proporción de abortos espontáneos, malformaciones congénitas, retrasos mentales y tumores.

Como parte de las consecuencias de estas enfermedades, aparecen también alteraciones en el metabolismo y en la síntesis o actividad de proteínas. Estas variaciones bioquímicas pueden considerarse secundarias a la enfermedad, a diferencia de las que aparecen en las enfermedades monogénicas, pero son también objeto de estudio de la patología molecular.

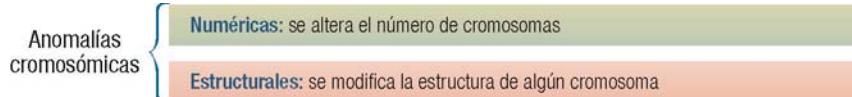
Las alteraciones son mucho más extensas que las producidas en los trastornos monogénicos y poligénicos, y se caracterizan por grandes cambios en el número de nucleótidos (miles o millones), en la estructura de los cromosomas o en la reubicación de genes por inserción o delección, duplicación, translocación, inversión, etc.

Pueden afectar tanto a los autosomas como a los cromosomas sexuales; en ambos casos, pueden ser constitutivas o adquiridas.

- Son **cromosomopatías constitutivas** las que afectan a todas las células del organismo, por aparecer durante la división meiótica de las células germinales (gametogénesis) o bien durante la primera división mitótica del cigoto. Como consecuencia, son las de mayor gravedad.
- Las **cromosomopatías adquiridas** o somáticas sólo ocurren durante las mitosis poscigóticas. Se inician, por tanto, en células somáticas, sin afectar a todas las células del organismo; de ahí la menor gravedad (a veces mínima) de sus consecuencias.

25.6.2 Clasificación y nomenclatura

25.12



Entre los abortos por anomalías cromosómicas durante el primer trimestre de embarazo, el 96% se debe a alteraciones numéricas y el 4% a estructurales. De forma similar, en madres mayores de 35 años, el 85% de los abortos debidos a cromosomopatías corresponde a las anomalías numéricas y el 15% a las estructurales.

En la descripción de las anomalías cromosómicas, la **nomenclatura** utilizada se basa en la de la dotación cromosómica normal, que se expresa como **46,XY** en células diploides del varón y **46,XX** en la mujer. Se añaden abreviaturas para expresar aspectos relativos a la morfología del cromosoma o para indicar las causas de la anormalidad. Los detalles de aplicación de la nomenclatura se describen en cada anomalía

46,XY	Célula masculina normal, diploide: 46 cromosomas, incluyendo un X y un Y	46,XX	Célula femenina normal, diploide: 46 cromosomas, incluyendo dos X
47,XXY	Un cromosoma X adicional	69,XXX	Célula femenina triploide
p	Brazo corto (pág. 89)	pat	Origen paterno
q	Brazo largo	mat	Origen materno
cen	Centrómero	dic	Cromosoma dicéntrico
ter	Extremo (terminal, telómero)	der	Cromosoma derivado
del	Deleción	i	Isocromosoma
ins	Inserción	r	Cromosoma anular (<i>ring</i>)
dup	Duplicación	mar	Cromosoma marcador
inv	Inversión	fra	Sitio frágil
t	Translocación	+	Ganancia
rcp	Translocación recíproca	-	Pérdida
rob	Translocación de Robertson	:	Rotura
/	Mosaicismo	::	Rotura y unión

25.6.3 Cromosomopatías numéricas

También llamadas **alteraciones genómicas**, implican un cambio en el número de cromosomas, bien pérdida o ganancia, pero sin alteración de la organización de los cromosomas. Son las anomalías más frecuentes y se originan en gran parte durante la división meiótica (pág. 105), con lo que se transmiten tras la fecundación a todas las células del hijo o hija. En general son letales, e impiden que se supere el desarrollo embrionario, con abortos espontáneos tras la concepción.

Las anomalías cromosómicas numéricas se clasifican así:

Nombre de la anomalía cromosómica		N.º de cromosomas	Ejemplos (humanos)	
Euploidía $(= x n)$	Nuliploidía	0	Eritrocitos y plaquetas	
	Haploidía (normal)	n	23,X	23,Y
	Diploidía (normal)	2n	46,XY	46,XX
	Poliploidía	3n	69,XXY	
	Tetraploidía	4n	92,XXYY	
...		xn		
Aneuploidía $(\neq x n)$	Nulisomía	2n-2	44,XY,-21,-21	
	Monosomía	2n-1	45,X	45,XX,-18
	Disomía (normal, no es aneuploidía)	2n		
	Trisomía	2n+1	47,XXY	47,XX,+21
	Polisomía	2n+x	49,XXXXX	
Mixoploidía	Mezcla de dos poblaciones celulares con distinta dotación cromosómica		46,XX/47,XX,+8	

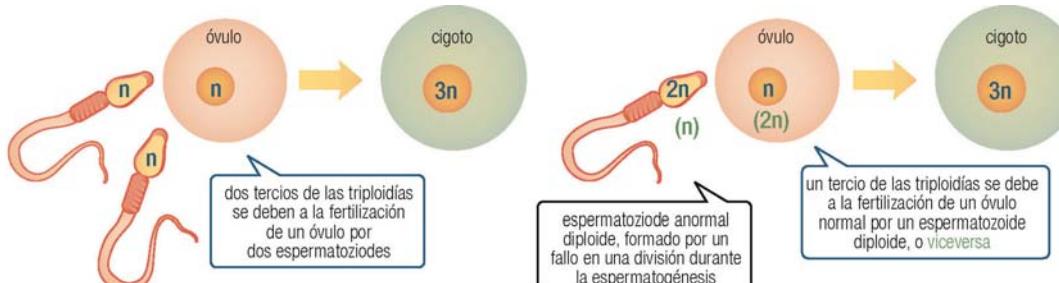
Se habla de **heteroploidía** en todas las situaciones en las que el número de cromosomas es diferente al normal.

25.6.3.1 Euploidía

Recibe este nombre (del griego *eu*, bien, verdadero: ploidía correcta o verdadera) la situación en la que existe un número entero de dotaciones cromosómicas haploides completas, es decir, el número de cromosomas es un múltiplo de *n*. Para los casos con más cromosomas de lo normal ($2n$) se habla de **poliploidía**. En el extremo opuesto se encuentran aquellas células diferenciadas como plaquetas y eritrocitos, que carecen de núcleo y se dice que son **nuliploides**.

a) Triploidía

Es la anomalía en la que las células contienen 3 juegos cromosómicos ($3n$; en humanos, $3 \times 23 = 69$), que incluyen 3 cromosomas sexuales. Es la poliploidía más común, y llega a constituir el 2% de todas las concepciones. Sin embargo, éstas se suelen perder por aborto espontáneo, por lo que la incidencia es muy baja.



3 posibilidades:

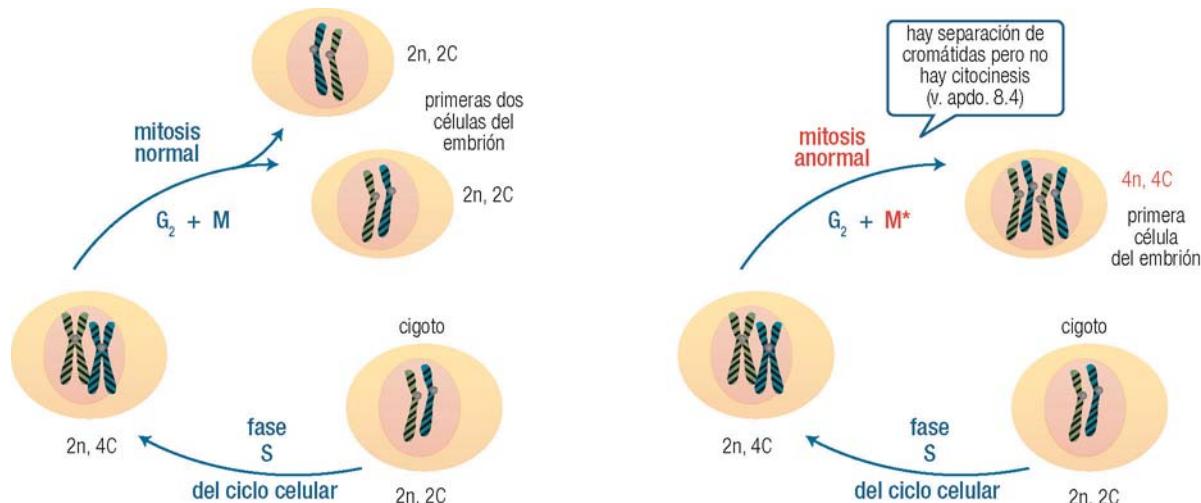
$$\begin{array}{l} 23,X + 23,X + 23,X \longrightarrow 69,XXX \\ 23,X + 23,Y + 23,X \longrightarrow 69,XXY \\ 23,Y + 23,Y + 23,X \longrightarrow 69,XYY \end{array}$$

2 posibilidades:

$$\begin{array}{l} 46,XX + 23,X \longrightarrow 69,XXX \\ 46,XY + 23,X \longrightarrow 69,XXY \end{array}$$

b) Tetraploidía

En este caso ($4n$, $4 \times 23 = 92$) la célula contiene dos dotaciones cromosómicas extras; son siempre 92,XXXX o 92,XXYY. Se origina usualmente por un fallo en la primera mitosis del cigoto: tras una replicación normal del DNA no se produce la división celular subsiguiente, la que debería formar las dos primeras células embrionarias.



25.6.3.2 Aneuploidía

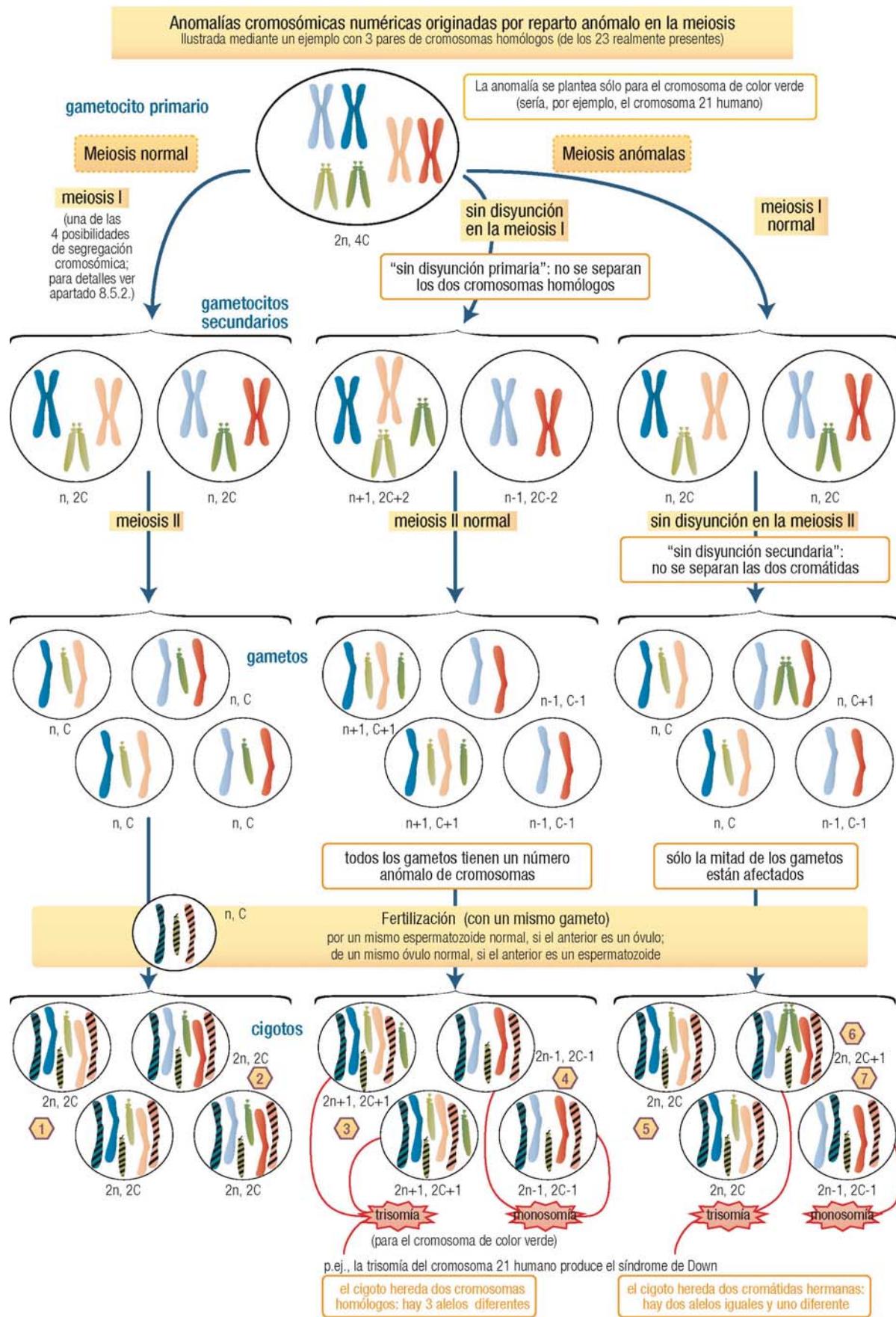
Este nombre se aplica a las anomalías en las que se pierde el carácter euploide, es decir, el número de cromosomas no es múltiplo del número haploide n . Ello se debe a la variación del número de copias de un solo cromosoma, no de juegos completos de cromosomas. En general, se presenta una copia de más o de menos de un solo cromosoma, específico en cada alteración, produciéndose, respectivamente, una **trisomía** (3 cromosomas homólogos) o una **monosomía** (una sola copia, no hay homólogo). En consecuencia, la dotación cromosómica total es de $2n + 1$ o $2n - 1$ (respectivamente, $46 + 1 = 47$ cromosomas en la trisomía humana y $46 - 1 = 45$ cromosomas, monosomía humana). El cromosoma extra se indica con un signo "+" y el ausente con un signo "-".

Son las anomalías numéricas más frecuentes y con relevancia clínica; aparecen en el 3-4% de los embarazos, y se asocian con un desarrollo físico o mental anómalos. En las células cancerosas también aparecen a menudo aneuploidías extremas, con múltiples anormalidades cromosómicas. Cuando falta un par de homólogos, la anomalía es letal, y se denomina **nulisomía**; por ejemplo 44,XY,-21,-21 (varón) o 44,XX,-21,-21 (mujer).

a) Mecanismo de generación de la aneuploidía

La aneuploidía se genera por una falta de disyunción durante la meiosis, es decir, un fallo en la separación normal de un par de cromosomas o de cromátidas. Las consecuencias del proceso son diferentes dependiendo de la división meiótica (I o II) en la que se produzca el fallo.

- La ausencia de disyunción cromosómica durante la meiosis I hace que los dos homólogos del cromosoma afectado pasen juntos a uno de los gametocitos secundarios, y el otro gametocito quede sin ese cromosoma. Los $2(n-1)$ cromosomas restantes han sufrido la disyunción normal y, por tanto, han separado sus homólogos y cada gametocito secundario contiene $n-1$. En la subsiguiente meiosis II, con disyunción normal, se separan todas las parejas de cromátidas. Los gametos maduros tienen una cromátida (ahora ya cromosoma hijo) de más, o bien una de menos, siempre con los restantes $n-1$ cromosomas correctos.
- La falta de disyunción de las cromátidas del cromosoma afectado durante la meiosis II (tras una meiosis I normal) hace que ambas pasen a un mismo gameto, que queda así con dos copias del cromosoma afectado, mientras el otro no contiene ninguna; se forman también otros dos gametos normales. El resto de cromosomas ($n-1$) separan sus cromátidas correctamente.



b) Monosomía

Como se ha indicado, es la situación en la que hay pérdida de un solo homólogo de un par de cromosomas (en total $2n-1$, 45 en humanos). El ejemplo más frecuente es el **síndrome de Turner**, una monosomía en cromosomas sexuales, con cariotipo 45,X y, por tanto, fenotipo femenino. Se origina por la falta de disyunción de los cromosomas sexuales (casos número 4 y 7 en la figura; el cromosoma de color verde rayado sería en este caso el X, procedente del gameto normal). Su incidencia es de 1 por cada 5.000 niñas nacidas vivas (la mayoría de las concepciones se pierden antes del nacimiento). Otros casos con el mismo fenotipo son **mosaicos** 45,X/46,XX o 45,X/46,XY (definidos a continuación, pág. 447).

c) Trisomía

En este caso existe un homólogo adicional de uno de los cromosomas, es decir, 3 copias de dicho cromosoma (en total $2n+1$, 47 en humanos). Puede afectar a cualquiera de los 23 pares. Existen varios ejemplos moderadamente frecuentes.

El **síndrome de Down** (originalmente denominado mongolismo) presenta en el 95% de los casos una copia extra del cromosoma 21 (**trisomía 21**), por lo que existen tres homólogos de este cromosoma. La nomenclatura es 47,XY,21 para los varones y 47,XX,21 para las mujeres. Se origina habitualmente por la falta de disyunción del cromosoma 21 (casos números 3 y 6 en la figura; los cromosomas de color verde claro y oscuro serían los homólogos 21). Es el tipo de trisomía más común. Presenta, entre otros síntomas, retraso mental moderado y unos rasgos físicos peculiares. No existe anomalía alguna a nivel génico. Una pequeña parte de los casos de síndrome de Down se originan por translocación: si una parte del cromosoma 21 se fusiona con otro, en alguno de los gametos existirá un cromosoma 21 completo y una copia adicional de esa parte del 21.



Web 25.2. Mecanismos de generación de trisomía 21.

El **síndrome de Edwards** es la **trisomía 18**, representada 47,XX,+18 (varón) y 47,XY,+18 (mujer), con una incidencia de uno por cada 6.000 nacidos vivos.

El **síndrome de Patau**, o **trisomía 13**, tiene una incidencia de uno de cada 10.000 nacidos vivos. Pueden sobrevivir a término, pero presentan malformaciones congénitas múltiples y graves. La nomenclatura es 47,XY,+13 (varón) y 47,XX,+13 (mujer).

El **síndrome de Klinefelter**, con una incidencia de uno por cada 1.000 nacidos vivos, es una trisomía en cromosomas sexuales que produce varones 47,XXY. Con una incidencia parecida se encuentran también trisomías 47,XXX (mujer) y 47,XYY (varón).

Se ha observado una clara correlación entre la frecuencia de trisomías y la edad de la madre. Por ejemplo, el síndrome de Down presenta una incidencia en promedio de uno por cada 800 nacidos vivos, pero la frecuencia es de 1/400 en madres de 35 años y 1/100 a los 40 años. Las causas de esta correlación no están completamente claras, pero debe recordarse (pág. 105) que cuando la mujer alcanza la pubertad todos sus oocitos primarios están ya formados y se mantienen detenidos en profase de la meiosis I hasta que llegan los estímulos hormonales de la ovulación; a mayor edad materna, más tiempo han pasado los oocitos en ese estado latente y mayores son, al parecer, las probabilidades de que la disyunción de sus cromosomas se vea alterada.

25.6.3.3 Mixoploidía

Recibe este nombre la situación en la que el organismo contiene células con distinto número de cromosomas. Aquéllas constituyen dos (o más) líneas celulares, cada una derivada de una célula troncal (o célula madre). Se puede distinguir entre **mosaico**, cuando las dos líneas derivan de un solo cigoto, en general por mutación en una etapa inicial del embrión, y **quimera**, cuando el origen son dos células troncales independientes que se han unido en un solo embrión. Como ejemplo, 46,XX/47,XX,+8 indica una mujer con dos poblaciones celulares, una con cariotipo normal y otra con trisomía del cromosoma 8.

25.6.4 Cromosomopatías estructurales

Este segundo tipo de anomalías cromosómicas, menos frecuentes que las numéricas, corresponde a cambios en la estructura de uno o varios cromosomas, sin que se altere su número (46 en el ser humano). El cambio se debe

a la rotura del cromosoma, comúnmente seguida de una reconstitución de la molécula en una disposición diferente, anómala. Son particularmente frecuentes en la anemia de Falconi, la ataxia-telangiectasia y el síndrome de Bloom.

La descripción de los diferentes tipos de anomalías estructurales tiene en cuenta el **número de puntos de rotura del cromosoma** y el carácter **equilibrado** (no hay pérdida de material cromosómico) o bien **desequilibrado** (se pierde parte del cromosoma).



Web 25.3. Descripción e ilustración de las cromosomopatías estructurales.



Bases moleculares del cáncer

26.1 INTRODUCCIÓN	449
26.2 EL CÁNCER COMO ENFERMEDAD GENÉTICA	450
26.3 ETAPAS CARACTERÍSTICAS EN EL DESARROLLO DEL CÁNCER	451
26.3.1 Tumor primario benigno	451
26.3.2 Cáncer <i>in situ</i>	451
26.3.3 Tumor maligno o cáncer propiamente dicho	451
26.3.3.1 Escape celular o invasividad	451
26.3.3.2 Vascularización o angiogénesis	452
26.3.3.3 Formación del tumor secundario o metástasis	452
26.4 MECANISMOS DE TRANSFORMACIÓN DE CÉLULA NORMAL EN TUMORAL	452
26.4.1 Equilibrio entre proliferación y muerte celular	452
26.4.2 Genes responsables del cáncer	454
26.4.2.1 Mutación de protooncogenes: oncogenes	454
a) Concepto	454
b) Tipos de protooncogenes	454
c) Características de la expresión de los oncogenes	456
26.4.2.2 Mutación de genes oncosupresores	456
a) Concepto	456
b) Tipos de genes oncosupresores	456
c) Características de la expresión de los genes oncosupresores	458
26.5 CONTROL DE LA PROLIFERACIÓN POR SEÑALES EXTRACELULARES	458
26.5.1 Tipos de señal	459
26.5.2 Transducción de la señal extracelular al interior	460
26.6 REGULACIÓN DEL CICLO CELULAR	462
26.6.1 Puntos de control en el ciclo de división celular	462
26.6.2 Proteínas que controlan el ciclo celular	463
26.6.2.1 Ciclinas	463
a) Ciclinas de la fase G ₁ o ciclinas de inicio	463
b) Ciclinas mitóticas o de la fase G ₂	464
26.6.2.2 Proteína quinasas dependientes de ciclinas (Cdk)	464
26.6.2.3 Actividad de los complejos Cdk-ciclina	466
26.6.3 Regulación del ciclo celular por los complejos Cdk-ciclina y alteraciones en el cáncer	467
26.6.3.1 Estímulos de progresión por el ciclo celular y oncogenes que los alteran	468
26.6.3.2 Estímulos de detención del ciclo celular	469
26.6.3.3 El gen supresor del retinoblastoma (Rb)	470
26.6.3.4 El gen oncosupresor p53	471
26.7 APOPTOSIS O MUERTE CELULAR PROGRAMADA	472
26.7.1 Etapas y control del proceso de apoptosis	472
26.7.1.1 Inducción	472
26.7.1.2 Ejecución	472
a) Proteasas: caspasas y su papel en el control de la apoptosis	473
b) Nucleasas	474
26.7.1.3 Fase final de la apoptosis	474
26.7.2 Relación entre ciclo celular y apoptosis: p53 como punto de enlace	474
26.8 HERENCIA DEL CÁNCER	475
26.9 MARCADORES TUMORALES	475
26.9.1 Alteraciones locales y generalizadas	476
26.9.2 Alteraciones tumorales y tisulares	476
26.9.2.1 Marcadores tumorales	476
a) Marcadores tumorales de secreción o marcadores "clásicos"	476
b) Citoquinas o mediadores solubles de origen celular	478
26.9.2.2 Marcadores tisulares	478
26.9.3 Marcadores genotípicos o moleculares	479

26.1 INTRODUCCIÓN

El cáncer, neoplasia o tumor puede considerarse una enfermedad genética (véase la clasificación molecular de las enfermedades en pág. 432) que se desarrolla en organismos superiores, en la mayoría de los tejidos y en todo tipo de células somáticas (como excepción, también en las células germinales en el caso de cánceres hereditarios, pág. 475). Actualmente se considera que también existe en el cáncer una contribución importante de alteraciones epigenéticas (pág. 287).

Bajo el nombre genérico de cáncer se engloba un conjunto de enfermedades que tienen en común un crecimiento celular desordenado (tumor) y una colonización tisular (metástasis), todo ello determinado por la acumulación de mutaciones en el genoma y por la alteración de las marcas epigenéticas. Cada cáncer es una situación distinta con peculiaridades dependientes del tipo de célula donde se origina, sus causas (etiología) y su

mecanismo, el grado de malignidad y otros factores. Por consiguiente, se diferencia claramente de otras enfermedades genéticas, en especial las monogénicas y las anomalías cromosómicas ([Capítulo 25](#)), y se distingue de las enfermedades multifactoriales (pág. 442) por la necesidad de que en éstas exista una confluencia de factores genéticos y ambientales para iniciar la enfermedad. Al igual que todas las enfermedades moleculares, el cáncer se manifiesta en alteraciones bioquímicas diversas en las células y tejidos y, en particular, por la aparición de marcadores tumorales y tisulares (pág. 475).

Es, posiblemente, la enfermedad estudiada más intensamente y desde más puntos de vista diferentes, y se ha pasado en los últimos años de un conocimiento casi nulo a un gran avance en numerosos aspectos moleculares, desde los que explican la proliferación excesiva y la pérdida de células por apoptosis (muerte celular) hasta los mecanismos por los que el cáncer genera las metástasis. Nuestro objetivo es presentar una visión global e integrada de todos esos aspectos, que ya han beneficiado a la biología molecular básica y son necesarios para comprender las aplicaciones, presentes y futuras, en especial para configurar una creciente capacidad clínica en cuanto a pronóstico, diagnóstico y tratamiento. Las primeras terapias desarrolladas, mediante fármacos para el tratamiento del cáncer por quimioterapia, se complementan ya –y lo harán más en el futuro– con las que van dirigidas a inhibir de manera específica la angiogénesis, la metástasis y la apoptosis, así como con la inhibición de enzimas responsables de las modificaciones epigenéticas de la cromatina. Además, sigue vigente la expectativa de desarrollar acciones más específicas sobre los genes alterados utilizando la terapia génica.

26.2 EL CÁNCER COMO ENFERMEDAD GENÉTICA

La participación de los genes en la carcinogénesis puede seguir dos tipos de mecanismo:

- Muchos cánceres se deben a **cambios genéticos** inducidos por mutaciones. La célula inicial afectada se divide con mayor rapidez de lo normal y, si no existe un control de ese crecimiento, llega a formar un **tumor primario** en el tejido circundante (**neoplasia**) o un **tumor secundario** (**metástasis**) en tejidos lejanos.
- Por otro lado, también es frecuente la contribución de los **cambios epigenéticos**, que afectan a la expresión de algunos genes y pueden no sólo permanecer durante toda la vida de la célula, sino también transmitirse cuando ésta se divide e, incluso –en el caso de las células germinales– pasar a la descendencia (dado que, por definición, las marcas epigenéticas son heredables, pág. 282).

Las mutaciones del genoma responsables de la enfermedad las producen agentes cancerígenos, de muy variadas naturaleza y características (pág. 396), así como errores espontáneos o inducidos en la replicación y reparación del DNA; por ello, el análisis de la carcinogenia se hace a la luz del efecto mutágeno. Se conocen bien como causa de cáncer las mutaciones que afectan esencialmente a genes de dos tipos: los **protooncogenes** y los **genes oncosupresores** (o **genes supresores de tumores**). Esto justifica una de las características del cáncer, el ser una **enfermedad poligénica**, lo que no debe confundirse con su carácter monoclonal (pág. 451). Como resultado de la mutación, los genes expresan proteínas alteradas, sobreexpresan las normales o no las expresan, según los casos; cualquiera de estos efectos modifica funciones de la célula tan básicas como el control de su división (pág. 462), lo que se manifiesta en una proliferación excesiva y continua o en una reducción de la muerte celular programada o apoptosis (pág. 472). De forma alternativa –o sinérgica–, las alteraciones en la regulación epigenética de esos mismos genes hará que, aunque no hayan sufrido mutaciones, se vea alterada su expresión, conduciendo a resultados similares.



26.3 ETAPAS CARACTERÍSTICAS EN EL DESARROLLO DEL CÁNCER

El cáncer, una vez originado, se manifiesta de forma variable según las circunstancias, el tipo de cáncer, la edad del individuo, etc. A grandes rasgos, su desarrollo suele agruparse en tres fases, con posibles subetapas de difícil delimitación: la formación del tumor primario benigno, su progresión *in situ* y la invasión y aparición de metástasis o tumor maligno.

26.3.1 Tumor primario benigno

Esta primera etapa del desarrollo del tumor, denominada **tumorigénesis, carcinogénesis tumoral o transformación neoplásica**, es la más estudiada y la mejor conocida. Se inicia con la llamada “célula progenitora del cáncer”, que surge en un determinado tejido por una mutación en algún gen (pág. 450). Su proliferación da lugar a un clon (pág. 193) de células mutadas, un **clon neoplásico**. Este comienzo monoclonal es la primera característica del cáncer, que lo diferencia de las enfermedades multifactoriales (pág. 442); a pesar de ese origen común, las células del clon inicial se hacen genéticamente distintas por la acumulación de nuevas mutaciones. Así, la mutación iniciadora confiere a la célula la **susceptibilidad tumoral**, una predisposición selectiva para la proliferación desmedida. La acumulación en las células hijas de sucesivas mutaciones en genes implicados en el control del ciclo celular o de la apoptosis, que favorezcan la proliferación o la inmortalidad, conduce de forma progresiva a las propiedades genéticas anormales del cáncer.



Web 26.1. Eventos en la generación del cáncer.

26.3.2 Cáncer *in situ*

Corresponde a la situación en la que las células del tumor primario no han escapado del tejido donde se originaron. Puesto que el tumor carece de capacidad invasiva, el pronóstico clínico es en general benigno y con frecuencia puede corregirse por simple extirpación quirúrgica. Sólo resulta grave si por su localización afecta a vasos sanguíneos, músculos o nervios críticos, en cuyo caso, aun sin ser formalmente un cáncer, podría considerarse como una neoplasia maligna.

26.3.3 Tumor maligno o cáncer propiamente dicho

La etapa que define y caracteriza el cáncer, responsable de la conversión de la célula neoplásica en cancerosa, es su progresión o propagación tumoral. Puede dividirse en **varias subetapas**, de difícil subordinación.

26.3.3.1 Escape celular o invasividad

Las células constituyentes de tejidos normales se mantienen doblemente adheridas entre sí gracias a uniones intercelulares y a uniones de las células con el **espacio o matriz extracelular**. Ésta es una estructura formada por fibrillas de proteínas fibrosas (colágeno y elastina), secretadas por las propias células, que actúan a modo de soporte para empaquetar o separar las células o para proporcionar elasticidad o fuerza de sostén. En estas interacciones intervienen proteínas transmembranarias, llamadas **integrinas**, de varios tipos; entre ellas, la mejor descrita es la fibronectina. Por su dominio intracelular, las integrinas se unen a los filamentos de actina del citoesqueleto de cada célula, mientras que a través de su dominio extracelular tienen afinidad por proteínas receptoras de las células vecinas (adhesión célula-célula) y por proteínas fibrosas de la matriz extracelular (adhesión célula-matriz).

Ambos tipos de adhesión se ven alterados en el cáncer y, como consecuencia, la célula puede desprenderse del tejido, convirtiéndose en una célula invasora de otros tejidos. La liberación de las células es consecuencia también de la acción de enzimas proteolíticas o **proteasas**, específicas, que destruyen la matriz extracelular, y permiten que sea atravesada por la célula despegada de sus vecinas. Entre ellas destacan las denominadas **metaloproteasas de matriz extracelular (MMP)**, que degradan casi todos los componentes de la matriz extracelular, por lo que intervienen en procesos fisiológicos de remodelación tisular, como desarrollo embrionario,

crecimiento óseo y cicatrización de heridas. La actividad proteasa aumenta sobre todo en las neoplasias malignas, posiblemente como consecuencia de un aumento de la biosíntesis de estas enzimas. En consecuencia, la inhibición de las proteasas adquiere interés con fines terapéuticos, para el tratamiento del cáncer. Algunos de estos inhibidores actúan sobre la activación de las proteasas a partir de sus precursores inactivos (proenzimas o zimógenos, pág. 370), mientras que otros bloquean directamente la acción de la proteasa activa.

26.3.3.2 Vascularización o angiogénesis

Este proceso consiste en la capacidad de un tejido para formar una red vascular propia mediante el desarrollo de nuevos vasos sanguíneos. Las células tumorales producen agentes angiogénicos que actúan sobre células endoteliales vecinas para dar lugar a la formación de una nueva red de venas, arterias y capilares. Ello permite que el tumor crezca mucho y muy rápidamente, al asegurar la llegada de oxígeno y nutrientes; la vascularización puede, asimismo, favorecer la migración de algunas células cancerosas al torrente circulatorio para difundirse a otros tejidos, donde generarán una metástasis. La angiogénesis es una de las dianas de los tratamientos contra el cáncer, haciendo uso de la existencia de proteínas que inhiben el proceso, como endostatina y angiotensina.

26.3.3.3 Formación del tumor secundario o metástasis

Mediante la invasión o diseminación celular, las células del tumor primario pueden fijarse en tejidos normales, para multiplicarse y diferenciarse en ellos, dando lugar al llamado **tumor secundario o metástasis**. Esta etapa es de enorme gravedad clínica y constituye la nota distintiva letal, irreversible, del tumor maligno, porque supone la destrucción del tejido sano que rodea a las células tumorales y la consiguiente dificultad de eliminación por intervención quirúrgica.

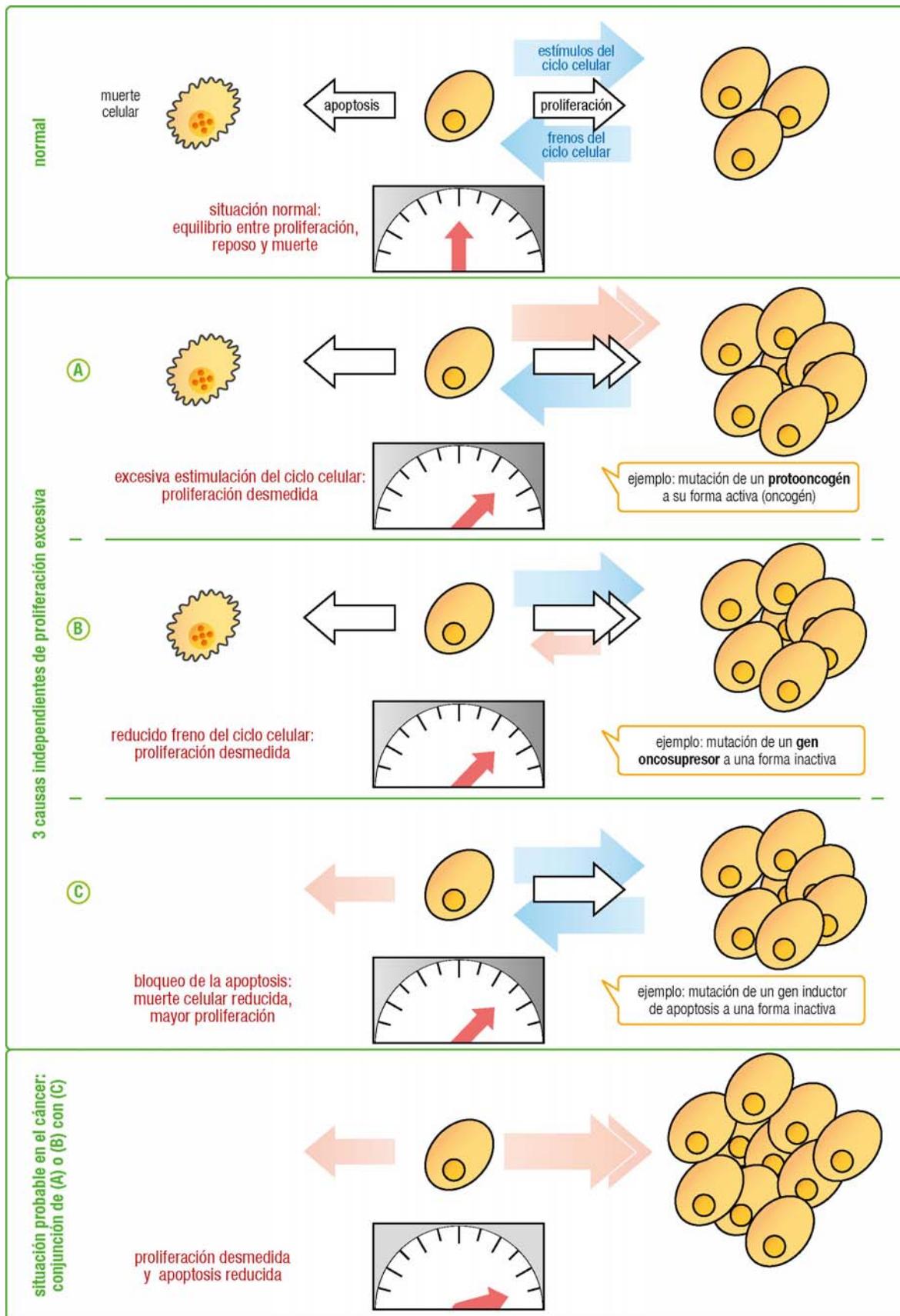
26.4 MECANISMOS DE TRANSFORMACIÓN DE CÉLULA NORMAL EN TUMORAL

26.4.1 Equilibrio entre proliferación y muerte celular

La vida media de la célula, o período de tiempo transcurrido desde su formación hasta su destrucción, es muy variable (desde horas hasta años). Las células mantienen controlada esta capacidad intrínseca de crecimiento tanto en la embriogénesis y el desarrollo como a lo largo de la vida del individuo adulto. Algunas, como las precursoras hematopoyéticas o las epiteliales, con vida media muy corta, experimentan una proliferación constante para dar lugar a otras muchas células; otras, como las neuronas, dejan de dividirse una vez diferenciadas. Salvo éstas, todas las células se renuevan continuamente en un proceso de “recambio celular”. En condiciones fisiológicas se requiere una estricta regulación genética para mantener esta homeostasis o equilibrio celular de forma adecuada a las necesidades de cada tejido, mediante la contraposición de dos procesos:

- La formación por mitosis de nuevas células (parte del ciclo de división celular, Capítulo 8) determina la tasa de **proliferación**. Se asocia a un proceso simultáneo de diferenciación celular, particular para cada tejido o estado del desarrollo (aspecto que no se considera en esta obra). El control de la proliferación comprende dos grandes aspectos, relacionados pero generalmente no considerados de la forma global necesaria para facilitar su comprensión: las señales extracelulares (pág. 458) y la regulación del ciclo celular (pág. 461).
- La **apoptosis**, muerte celular programada o suicidio celular es un mecanismo fisiológico de eliminación de células al final de su vida activa, como parte del recambio celular necesario para contrarrestar la formación continua por división celular. Asimismo, interviene en la formación de tejidos (por ejemplo, células del cristalino y de la piel), en la morfogénesis (por ejemplo, formación de los dedos) y en la eliminación de células con daño genético. Los detalles de la apoptosis se analizan posteriormente (pág. 472).

La alteración del equilibrio en condiciones patológicas (neoplasias) da lugar a un crecimiento incontrolado, que puede deberse a una excesiva proliferación o a una reducida muerte celular. Parece que habitualmente el cáncer se origina por una conjunción de ambas causas.



26.4.2 Genes responsables del cáncer

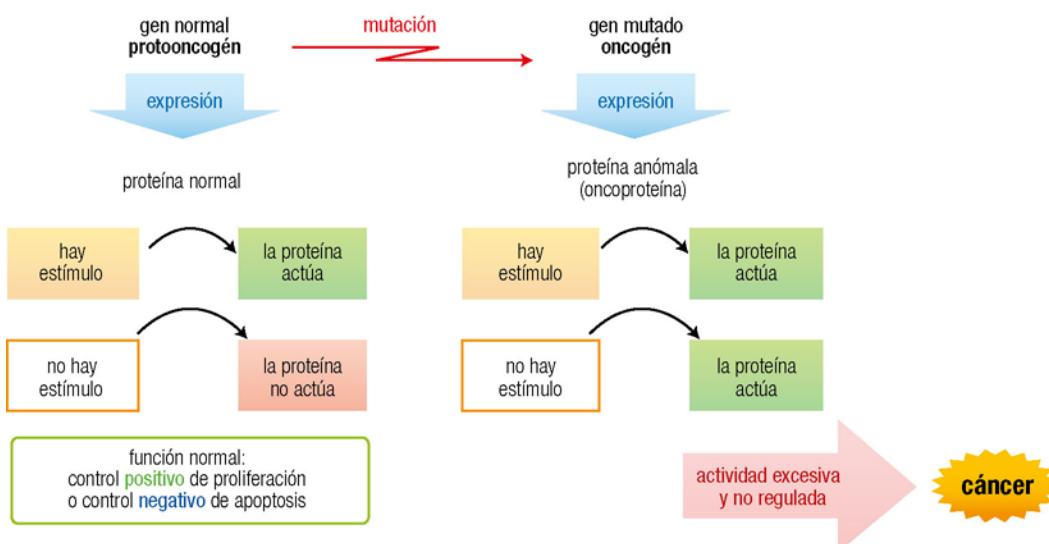
Las mutaciones, presentes en las células cancerosas, que provocan la aparición de las características tumorales afectan a genes cuyos productos proteicos regulan la proliferación celular, la apoptosis o la senescencia celular (envejecimiento). La proliferación y la apoptosis se estudiarán progresivamente en este capítulo, en el que se describen tanto los mecanismos normales de regulación como sus alteraciones conducentes al cáncer.

En primer lugar, debe definirse la existencia de dos tipos de genes cuya alteración se relaciona con el proceso canceroso: oncogenes y genes oncosupresores (y sus productos respectivos, oncoproteínas y proteínas oncosupresoras).

26.4.2.1 Mutación de protooncogenes: oncogenes

a) Concepto

Recibe el nombre de **oncogén** la forma mutada de un gen normal o **protooncogén** que causa cáncer como consecuencia de una “ganancia de función” de la proteína expresada por el gen. El protooncogén codifica una proteína normal, generalmente relacionada con la proliferación celular o la apoptosis, que actúa sólo cuando recibe señales reguladoras específicas. Por el contrario, la forma mutada u **oncogén** expresa una proteína anormal (**oncoproteína**) que se mantiene activa con independencia de las señales reguladoras, por lo que el resultado es una ganancia de función de la proteína. Se puede decir, en consecuencia, que los oncogenes son el resultado de una mutación activadora de la función normal de los protooncogenes. Ésta conduce, según los casos, a una proliferación descontrolada o a una apoptosis reprimida y, en definitiva, a la aparición del cáncer.



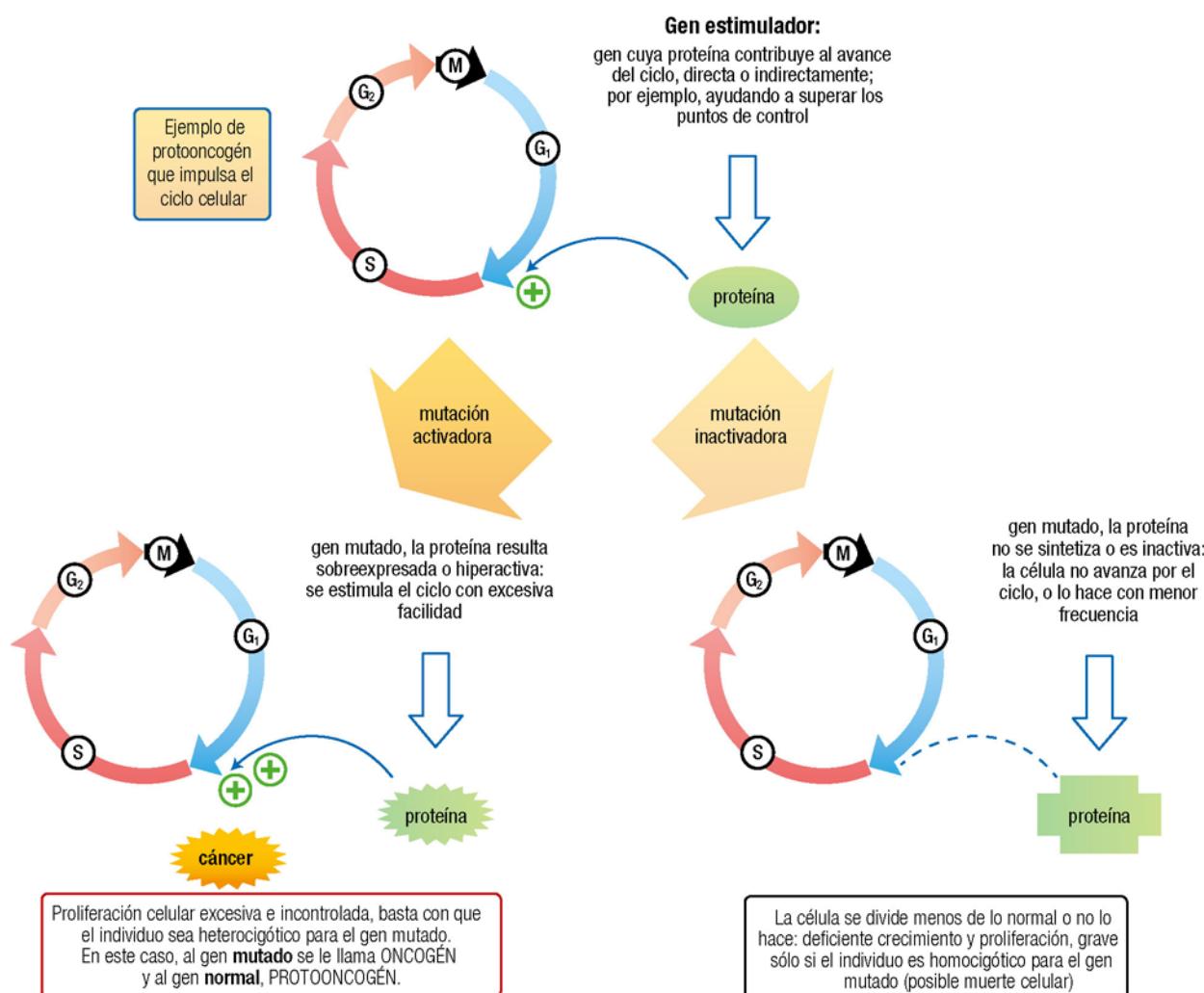
26.3

b) Tipos de protooncogenes

Se han identificado protooncogenes que codifican proteínas muy diversas:

- **Factores estimuladores del crecimiento celular.** La mutación oncogénica origina una producción excesiva del factor proteico o una mayor actividad de éste.
 - Ejemplo: el protooncogén *sis* codifica la cadena B del factor de crecimiento derivado de plaquetas (PDGF).
- **Receptores de factores de crecimiento o de hormonas,** presentes en la membrana o dentro de la célula. Para el caso de un factor de crecimiento u hormona estimuladores de la proliferación, la forma oncoproteíca del receptor puede ser aquella con una conformación que la mantiene activa, aunque no se le une el ligando. Si se trata del receptor de un factor u hormona inhibidores del crecimiento, la oncoproteína es la que no responde a la unión del ligando.

- Ejemplo: el protooncogén *erbB* codifica el receptor de membrana para el factor de crecimiento epidérmico EGF.
- **Proteínas citoplasmáticas que intervienen en los sistemas de transducción de señales.** La mutación hace que el oncogén exprese una proteína que se mantiene activa sin necesidad de que le llegue la señal.
 - Ejemplo: el protooncogén *ras* codifica una proteína G monomérica (pág. 460).
- **Factores de transcripción** (pág. 288) que controlen la expresión de genes que codifican a su vez proteínas implicadas en la señalización, el control del ciclo celular o la apoptosis.
 - Ejemplos: protooncogenes *fos*, *jun* y *myc*.
 - El protooncogén *erbA* codifica el receptor intracelular para hormonas tiroideas, que es un factor de transcripción.
- Finalmente, **proteínas responsables de la activación directa del ciclo celular** (tales como las ciclinas o las quininas y fosfatasas activadoras de las Cdk, págs. 463-464) o de la **inhibición de la apoptosis** (pág. 474). Los oncogenes respectivos expresan proteínas en mayor cantidad o con función aumentada.
 - Ejemplos: el protooncogén *bcl-1* codifica la ciclina D1.
 - El protooncogén *cdk1* codifica la quinasa dependiente de ciclina Cdk1 (pág. 465).
 - El protooncogén *mdm-2* codifica un antagonista de la proteína p53 (pág. 471).
 - El protooncogén *bcl-2* codifica una proteína mitocondrial que bloquea la apoptosis al inhibir las caspasas (págs. 473-474).



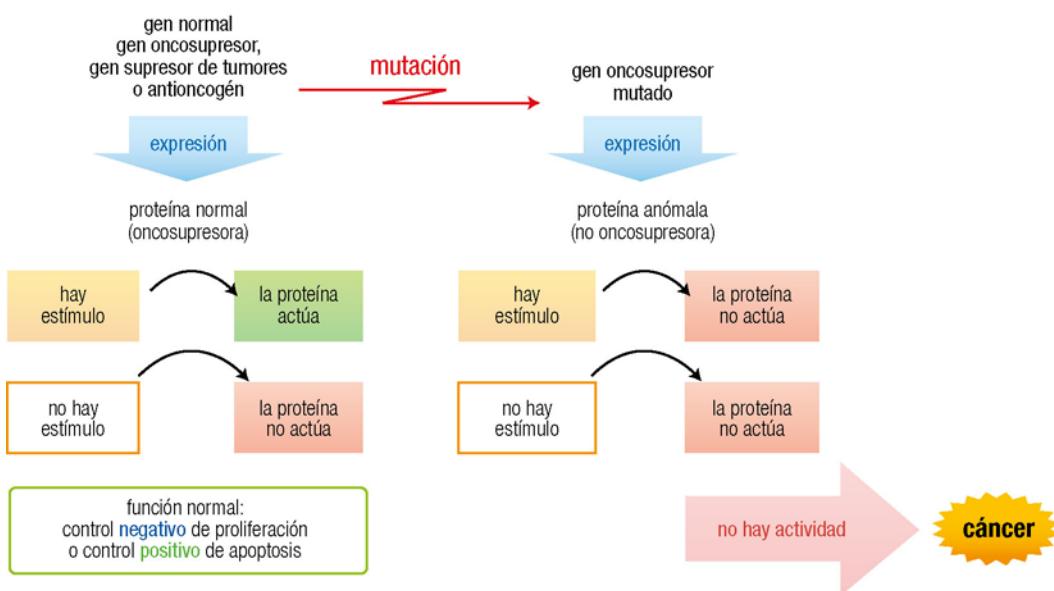
c) Características de la expresión de los oncogenes

Puesto que existe ganancia de función en la oncoproteína, basta con que esté mutado uno de los dos alelos del protooncogén para que se exprese el efecto de la oncoproteína anormal. En consecuencia, los oncogenes se expresan con carácter dominante (págs. 100 y 437), y los individuos heterocigóticos para un oncogén (genotipo protooncogén-oncogén, o nP) sufren sus efectos de forma similar a los homocigóticos (genotipo oncogén-oncogén, o PP). La herencia del oncogén concordará con este carácter, pero sólo si la mutación del protooncogén afecta a células germinales (pág. 475).

26.4.2.2 Mutación de genes oncosupresores

a) Concepto

Al igual que los protooncogenes, se trata de genes que causan cáncer como consecuencia de una mutación experimentada por la forma normal, denominada **gen oncosupresor**, **gen supresor de tumores** o **antioncogén** (porque su acción es contraria a la de los oncogenes, aunque la ejercen por un mecanismo distinto). La forma normal del gen (**gen oncosupresor**) codifica una proteína normal (**proteína oncosupresora** o **antioncogénica**) que actúa deteniendo la proliferación o bien induciendo la apoptosis. Cuando una mutación del gen conduce a la síntesis de una proteína no funcional o impide esa síntesis (“**pérdida de función**”), la proliferación deja de estar controlada o la apoptosis nunca tiene lugar, por lo que aparece una proliferación excesiva, a menudo con acumulación de daños genéticos en las células. Dicho de otro modo, la **versión mutada** de un gen oncosupresor pierde su función normal, reguladora de la proliferación, y da lugar al cáncer.



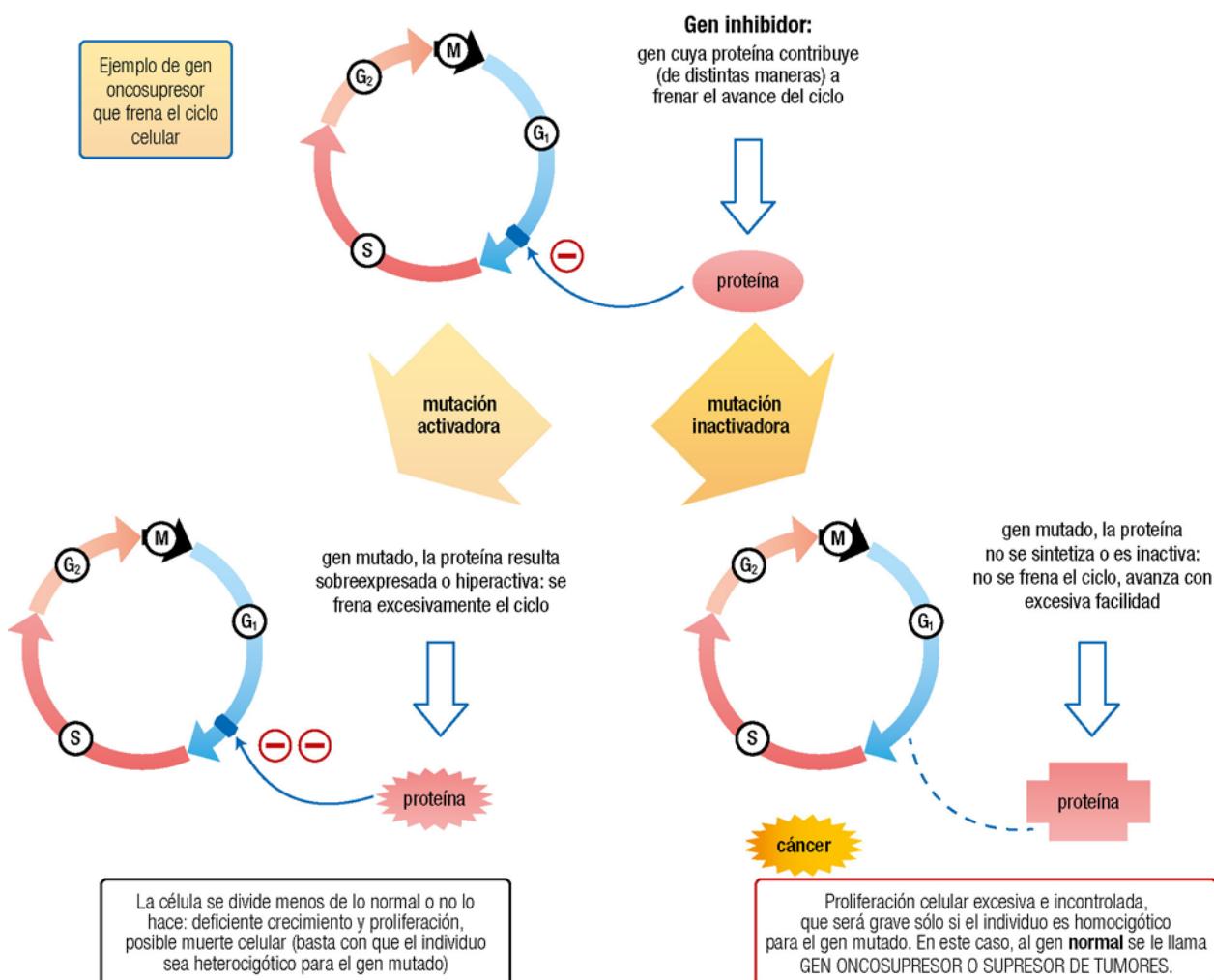
26.5

b) Tipos de genes oncosupresores

Al igual que los protooncogenes, los genes oncosupresores codifican proteínas que actúan en distintos puntos de las rutas de señalización, del control del ciclo celular y de la apoptosis:

- **Factores inhibidores del crecimiento celular.** La mutación del gen oncosupresor correspondiente anula la funcionalidad de la proteína sintetizada.
 - Ejemplo: el gen oncosupresor *dcc* codifica una proteína de adhesión celular, ausente en el carcinoma de colon.

- Receptores de esos factores inhibidores o de hormonas que frenan el crecimiento celular. La forma mutada del gen oncosupresor codifica un receptor insensible a su ligando o que no transmite la señal al interior de la célula.
 - Ejemplo: el gen oncosupresor que codifica el receptor de membrana para el factor β de crecimiento transformante, TGF- β .
- Proteínas citoplasmáticas que intervienen en los sistemas de transducción de señales acoplados a los receptores anteriores. La situación varía según cuál sea la función concreta de la proteína.
 - Ejemplo: el gen oncosupresor *nf1* codifica la neurofibromina, proteína que estimula la actividad GTPasa de Ras, en consecuencia inactivando ésta (pág. 460). En la enfermedad de Recklinghausen, el gen *nf1* está mutado (pág. 393), por lo que la neurofibromina es defectuosa, la proteína Ras no libera el GTP y queda en forma activa más tiempo del normal, aumentando la proliferación.
- Factores de transcripción que dirijan la expresión de genes cuyos productos proteicos frenan el ciclo celular o producen apoptosis. La forma mutada del gen oncosupresor expresa una proteína no funcional.
 - Ejemplos: los genes oncosupresores *Rb* y *p53*, que se estudian en detalle más adelante (págs. 470-471).
- Proteínas que frenan el ciclo celular o producen apoptosis. El gen mutado no expresa la proteína o da lugar a una forma no funcional.

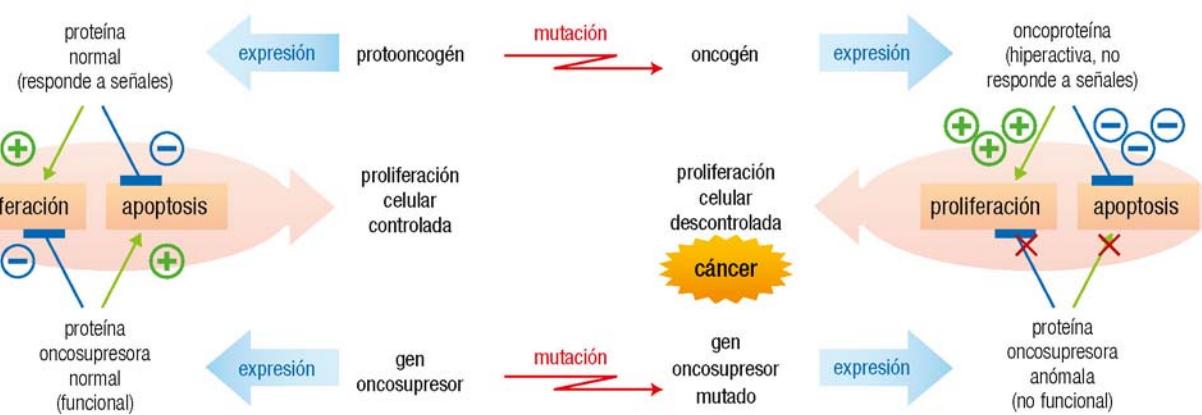


c) Características de la expresión de los genes oncosupresores

A diferencia de los oncogenes, en los que un solo alelo mutante afecta al fenotipo de la célula, la pérdida de la función de un gen oncosupresor sólo se produce cuando han mutado los dos alelos (individuos homocigóticos, genotipo oncosupresor mutado-oncosupresor mutado, o pp). Es decir, la forma mutada de un gen oncosupresor se expresa con carácter **recesivo** (págs. 100 y 438). Esto es así puesto que la mutación produce pérdida de función en la proteína, pero las moléculas de proteína sintetizadas a partir del alelo normal pueden suplir esa carencia, al menos parcialmente. Al igual que en el caso de los oncogenes, la herencia de los genes oncosupresores mutados es la correspondiente a su carácter recesivo, pero sólo tiene lugar cuando la mutación ha afectado a células germinales (pág. 475).

La probabilidad de mutación somática en los dos alelos de una misma célula es baja, por lo que la homocigosis generada *de novo* es muy infrecuente; sin embargo, los individuos heterocigóticos en un gen oncosupresor (que han heredado un alelo mutado) tienen predisposición a la aparición de cáncer, ya que sólo necesitan una mutación somática en el otro alelo para perder totalmente la función de la proteína oncosupresora.

Resumen de la actuación de los genes responsables del cáncer



Corresponde estudiar a continuación con detalle las tres rutas que regulan la proliferación celular y cuya alteración conduce a cáncer:

1. La transducción de señales extracelulares al interior celular hasta conectar con el ciclo celular.
2. La regulación del ciclo celular (proceso estudiado previamente en el Capítulo 8).
3. La apoptosis o muerte celular programada.

26.5 CONTROL DE LA PROLIFERACIÓN POR SEÑALES EXTRACELULARES

La señalización iniciada por moléculas externas a la célula constituye, en general, el mecanismo de control de la proliferación celular, desencadenante de la regulación del ciclo celular por otras moléculas intracelulares. Los mecanismos mediante los cuales las señales extracelulares dan lugar a una respuesta en el interior de la célula se conocen genéricamente como **transducción** de la señal, constituyen un área de gran significado en biología celular e intervienen no sólo en el control de la proliferación celular, sino también en la respuesta fisiológica a una gran cantidad de sustancias, como parte esencial de la función de los distintos tipos celulares. Se estudian aquí, de forma sucinta, las señales procedentes del exterior que van a determinar las respuestas intracelulares que regulan en último extremo el ciclo celular y, con ello, la tasa de proliferación. Aunque se conocen también como **señales mitógenas**, debería reservarse este término para las sustancias que actúan estrictamente sobre la mitosis.

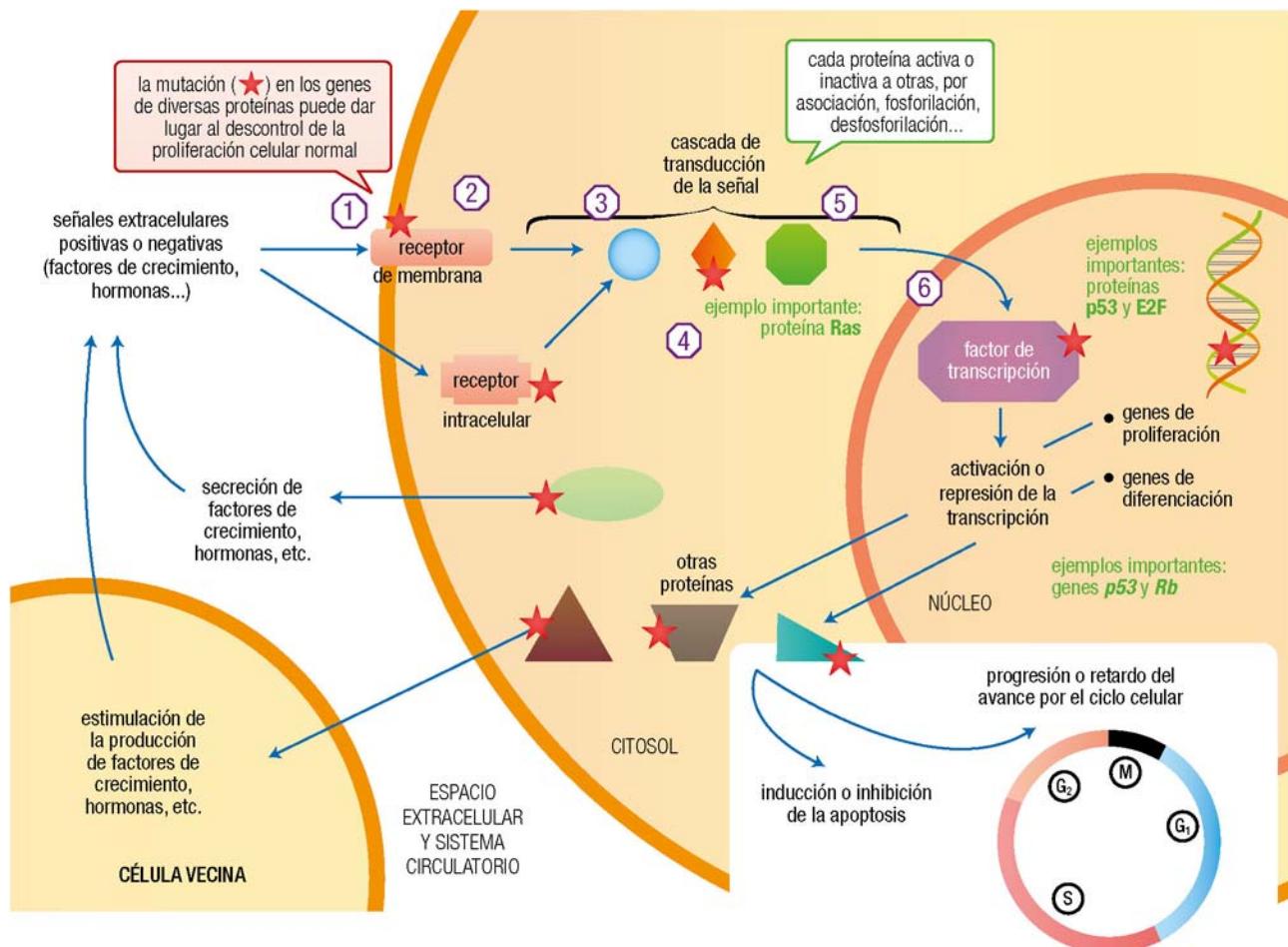
26.5.1 Tipos de señal

Las señales que favorecen la actividad del ciclo celular se conocen como **positivas o estimuladoras**. Existen también señales extracelulares **negativas o inhibidoras**, que actúan frenando la actividad del ciclo. En ambos casos, son muy variados los tipos y características de las moléculas que al incidir sobre su célula diana actúan como señal, a través de su interacción con receptores específicos.

Tipos de señalización	
Por moléculas secretadas	
Señalización autocrina	La diana es la propia célula secretora
Señalización paracrína	La célula secretora y la célula diana son adyacentes
Señalización endocrina (u hormonal)	La célula secretora y la célula diana están alejadas; la molécula señal (hormona) se transporta por la sangre
Señalización intracrína	La molécula señal interacciona con un receptor intracelular en la propia célula
Señalización sináptica	La molécula señal (neurotransmisor) es secretada por una neurona e interacciona con el receptor en otra neurona o en una célula muscular que están en estrecha proximidad (sinapsis)
Por moléculas unidas a la membrana (señalización yuxtapuesta o yuxtarcrína)	Interacción directa de moléculas de membrana con receptores de membrana de la célula diana
Por uniones comunicantes	Conexiones directas entre el citoplasma de las dos células
Moléculas señal	
Factores de crecimiento	Factor de crecimiento epidérmico (EGF) y factor de crecimiento derivado de plaquetas (PDGF)
Hormonas	
Péptidos y proteínas	Hormonas pancreáticas insulina y glucagón
Esteroides	Hormonas adrenales y sexuales
Derivados de aminoácidos	Hormonas tiroideas
Neurotransmisores	Glutamato, γ -aminobutirato (GABA), dopamina, serotonina, acetilcolina, etc. (sólo son relevantes en las neuronas)
Otros	Nucleótidos, retinoides, derivados de ácidos grasos, óxido nítrico, monóxido de carbono, etc.
Tipos de receptor	
Intracelulares	Proteínas situadas en citosol y núcleo, que unen moléculas señal hidrófobas
De superficie celular	Proteínas integrales de membrana, que unen moléculas señal hidrófilas
Con actividad enzimática intrínseca	Ejemplo: tirosina quinasas, como el receptor de PDGF y el de insulina
Asociados a canales iónicos	Ejemplo: receptor de acetilcolina
Asociados a enzimas	Ejemplo: receptor de la hormona de crecimiento
Asociados a proteínas G triméricas	Ejemplos: receptores de adrenalina y de glucagón. Las proteínas G están acopladas, a su vez, a canales iónicos o a enzimas

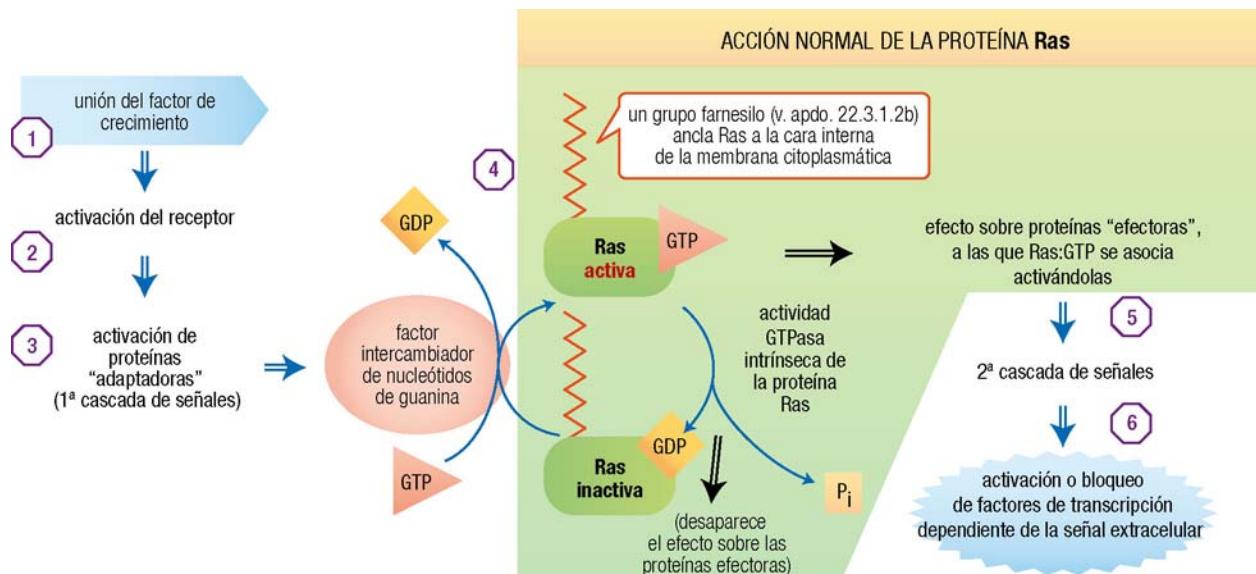
26.5.2 Transducción de la señal extracelular al interior

Aunque, como se ha indicado, existen muchas posibilidades de transmisión de la señal, por simplicidad la descripción se centrará en el mecanismo de activación en el que interviene la proteína Ras, causante de numerosos procesos cancerosos. Su activación es un paso esencial en la transducción de señales producidas por muchos factores de crecimiento.



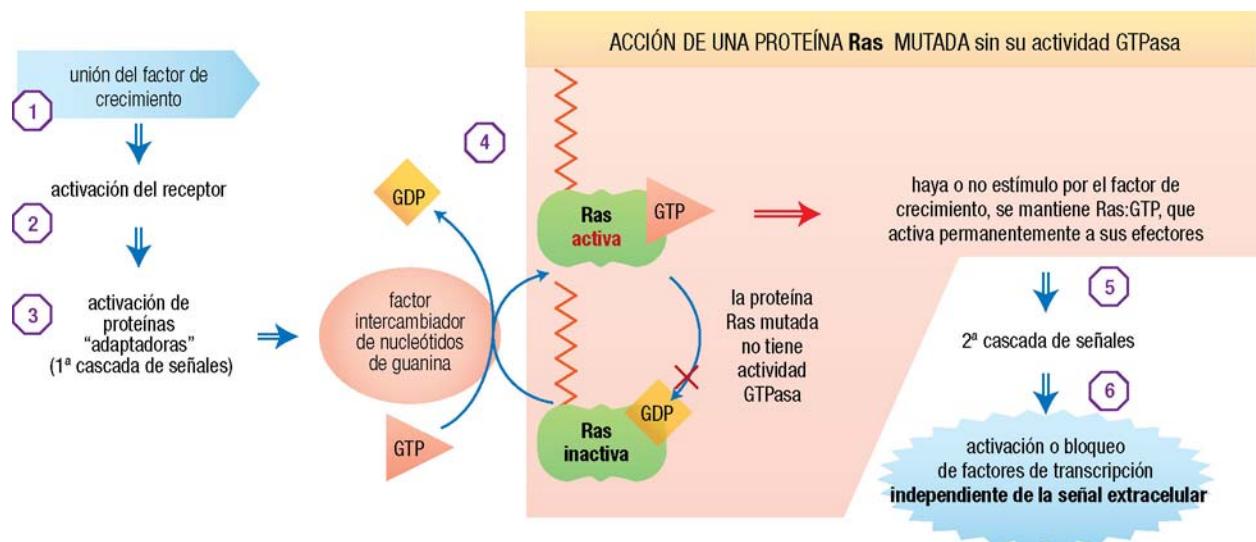
1. La molécula señal actúa como ligando al interactuar con gran afinidad y de modo reversible con el dominio extracelular de una **proteína receptora** específica, cuyo dominio transmembranario hidrófobo está insertado en la bicapa lipídica. El número de receptores de un tipo concreto en cada célula es muy variable, desde cientos a millones. En el caso de la ruta que activa a Ras, la señal procede de diversos receptores: para el factor de crecimiento derivado de plaquetas (PDGF), el factor de crecimiento epidérmico (EGF), la insulina, etc.
2. Al unirse el ligando al receptor, éste se estimula. En el caso del receptor de PDGF, forma un dímero y se activa una actividad intrínseca como quinasa de proteínas, concretamente del tipo tirosina quinasa (pág. 360 y web 22.4), con lo que fosforila, con gasto de ATP, el dominio intracelular propio o bien otras proteínas.
3. El paso siguiente, que llamaremos “primera cascada de transmisión de la señal”, puede realizarse por varias vías en las que, en general, se producen modificaciones en la conformación de una proteína que a su vez permiten la interacción con otras proteínas “adaptadoras” o su fosforilación.
4. La propagación de la señal continúa mediante la activación de la proteína **Ras**, una proteína G monomérica (también llamada p21Ras, por su masa de 21 kDa). Las **proteínas G** actúan en numerosos procesos como “interruptores moleculares”, al poder pasar de su forma inactiva, unida a GDP, al estado activo, unida a GTP (véanse, como ejemplo, algunos factores de traducción: eIF-2, eEF-1A, eEF-2, eRF (págs. 336, 338 y 341). En este caso, el complejo activo Ras:GTP transmite la señal mediante su interacción con otras quinásas de proteínas citoplasmáticas (proteínas “efectoras”), activándolas, con lo que comienza una

“segunda cascada de transmisión de la señal”. La consiguiente hidrólisis del GTP por Ras devuelve a ésta al estado inactivo, interrumriendo así la transmisión de señal hasta que haya un nuevo estímulo. Debe añadirse que, aunque Ras fue la primera descubierta y es la más importante, existen otras proteínas relacionadas que forman la familia Ras e intervienen en varios procesos celulares (como Rab, Rho y Arf).



26.9

Una alteración frecuente en una proporción elevada de tumores (en especial de colon y páncreas) es la mutación del gen *ras*, dando lugar a una proteína Ras que es incapaz de hidrolizar el GTP, por lo que permanece continuamente en su forma activa Ras:GTP, desencadenando una señal de proliferación continua y no regulada, responsable de dichos cánceres.



26.10

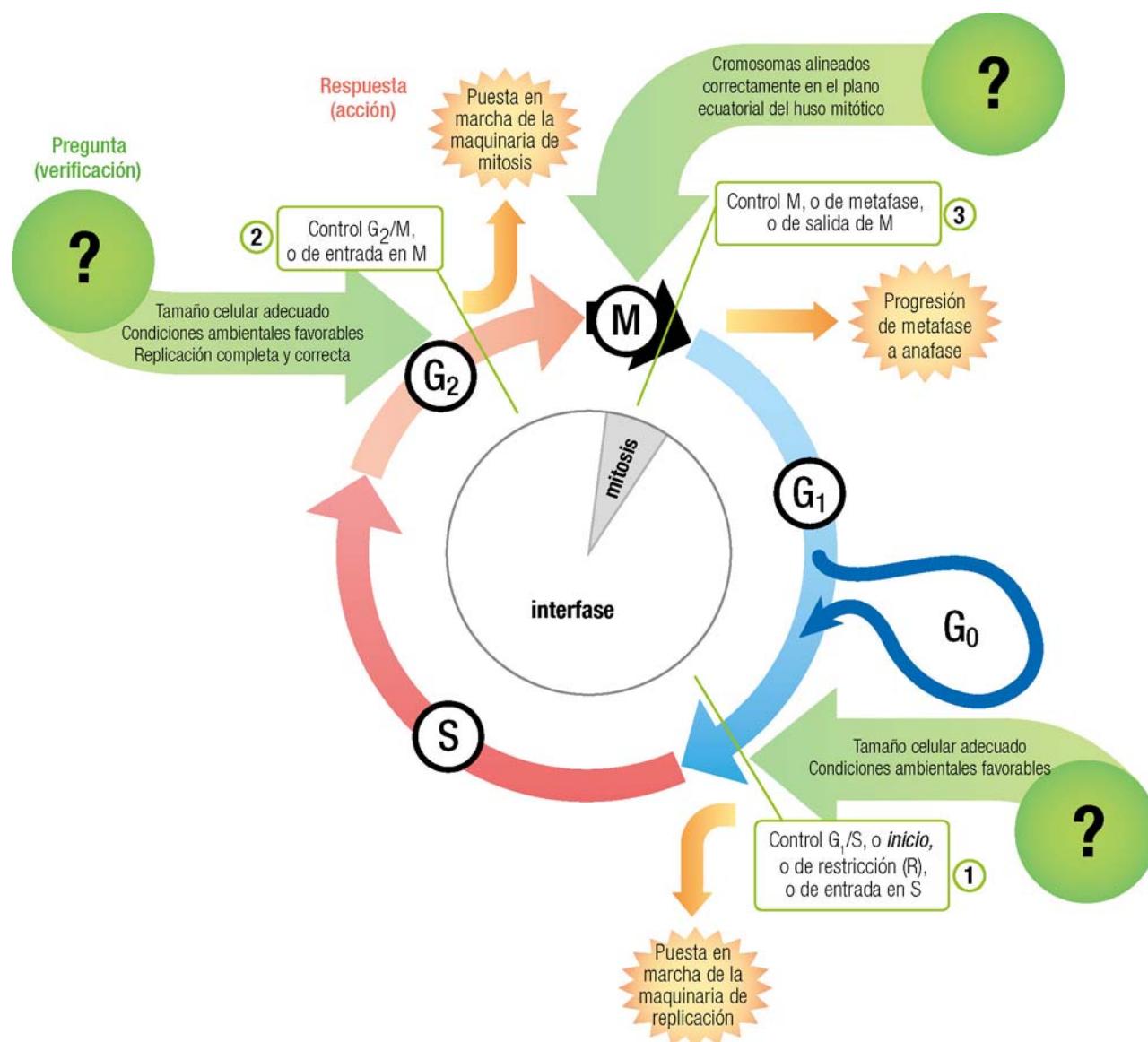
- Cada proteína efectora activada por Ras produce a continuación cambios de conformación o fosforilaciones en sucesivas proteínas componentes de la “segunda cascada de señales”, entre las que cabe citar Raf, MEK y MAP quinasas.
- Finalmente, algunas de estas quinasas de proteínas de la segunda cascada fosforilan un factor de transcripción. Como consecuencia, éste cambia su conformación, lo que le permite unirse al DNA y activar o reprimir la transcripción de genes específicos que codifican proteínas que regulan el ciclo celular y la apoptosis.

26.6 REGULACIÓN DEL CICLO CELULAR

Antes de describir los detalles de la regulación, debe indicarse que la repetición ordenada y controlada del ciclo celular define su duración característica, es decir, determina el grado de proliferación normal de las células en cada tejido. Cuando la proliferación celular es continua, excesiva y descontrolada se origina el cáncer. Como se ha indicado, esto puede ser consecuencia de la acumulación sucesiva (a veces durante años) de alteraciones o mutaciones en algunos de los genes anteriormente descritos (por ejemplo, *ras*). Analizaremos previamente los aspectos moleculares del proceso normal.

26.6.1 Puntos de control en el ciclo de división celular

La regulación del ciclo se ejerce en tres puntos bien definidos, que dependen en su conjunto de las condiciones ambientales más favorables (nutrientes y temperatura, entre otras) y de la presencia de los factores de crecimiento celular necesarios. Los dos puntos de control principales se sitúan en la interfase, en concreto en las transiciones de G₁ a S y de G₂ a M. El tercer punto se localiza en plena mitosis. Gracias a este triple control, las células sólo inician una nueva fase de división cuando se han cumplido las condiciones necesarias en cada punto y se ha finalizado la fase anterior. Utilizando un símil, cada punto de control actúa a modo de “freno” de la etapa previa. Cualquier circunstancia anormal en uno de ellos hace cesar la división y la consiguiente proliferación celular.



- El primero se denomina **punto de control en G₁/S**. Ocurre cerca del final de G₁ ($2n, 2C$), antes de entrar en fase S, y una vez superado se dispara la replicación del DNA. Este punto se llama *inicio* en levaduras de gemación (como *Saccharomyces cerevisiae*) y *punto R o de restricción* en mamíferos, donde es el control principal del ciclo. Al ser la fase G₁ la más prolongada del ciclo (entre 6 y 12 h), es en ella donde más intervienen las condiciones extracelulares (por ejemplo, del medio de cultivo), el único punto donde el ciclo responde a señales externas y progresiona sin dependencia de estímulos mitóticos. El paso por este punto lo regulan las proteínas intracelulares llamadas **ciclinas de G₁** (pág. 463).
- El segundo, o **punto de control en G₂/M**, ocurre al final de G₂ ($2n, 4C$), antes del inicio de la mitosis M. Si la célula no supera este punto, permanece con ese complemento doble de dotación cromosómica. Su desregulación puede ser importante en la transformación neoplásica. El paso por este punto lo regulan las proteínas intracelulares llamadas **ciclinas de G₂ o mitóticas** (pág. 464).
- El tercer control, **punto M**, se ejerce durante la mitosis ($2n, 4C \rightarrow 2n, 2C$), entre la metafase (cromosomas condensados dispuestos en el plano ecuatorial de la célula) y la anafase (separación de las cromátidas hermanas unidas al huso mitótico hacia cada polo celular). Asegura que la célula no se divide si hay errores en la formación del huso acromático o en la alineación de los cromosomas en la placa ecuatorial. Es el punto de control cuyo mecanismo se conoce menos y no se detallará en este texto, pero posiblemente depende también de las ciclinas mitóticas.

GLOSARIO particular para la regulación del ciclo celular:

Los **genes** se nombran con cursiva y minúsculas: *cdc2*
o todo en mayúsculas (dependiendo de la especie): **CDC2**

(de *cell division cycle*,
ciclo de división celular)

Las **proteínas** se nombran sin cursiva, con la primera letra en mayúscula: **Cdc2**

Cdk = *cyclin-dependent kinase*, quinasa dependiente de ciclina. Se trata de una familia de proteínas con actividad proteína-quinasa sobre residuos de serina o treonina de determinadas proteínas sustrato. Sólo son activas como quinasas cuando están asociadas con una ciclina.

Ciclina = miembro de una familia de proteínas que se unen a las proteínas Cdk activándolas. Su nombre procede de las variaciones cíclicas que experimenta su concentración en la célula, en paralelo con el ciclo de división.
Se clasifican en **ciclinas de G₁** o de inicio y **ciclinas de G₂ o mitóticas**.

26.12

26.6.2 Proteínas que controlan el ciclo celular

En los puntos de control G₁/M y G₂/S intervienen dos tipos característicos de proteínas, las **ciclinas** y las **proteína quinasas Cdk**, que se asocian formando complejos Cdk-ciclina para controlar la actividad del ciclo celular. Se estudiarán sucesivamente los dos tipos de proteínas y la activación del complejo formado por ambas, para poder abordar después la regulación del ciclo celular por los complejos Cdk-ciclina.

26.6.2.1 Ciclinas

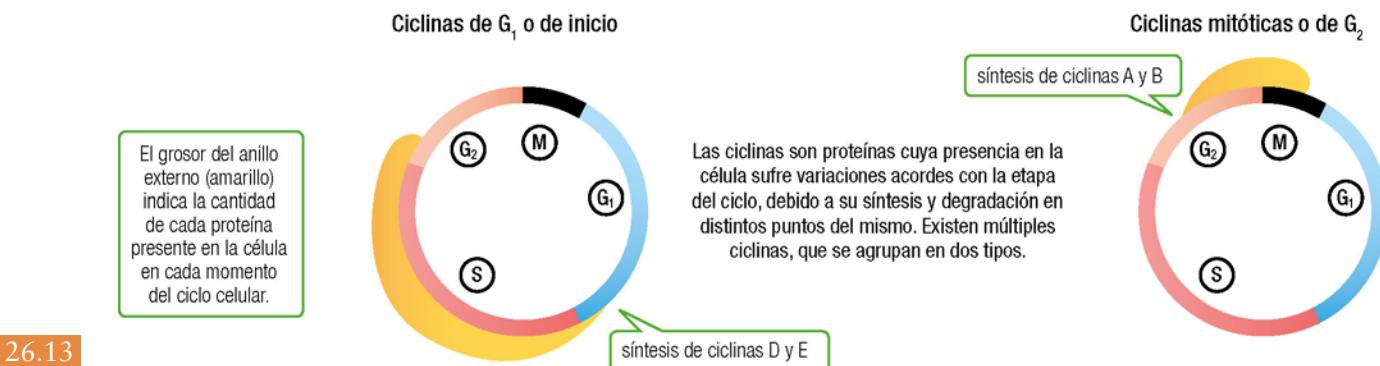
El nombre de este grupo de proteínas se debe a que su concentración en la célula fluctúa de acuerdo con la etapa del ciclo celular, de modo que están presentes en una etapa concreta y desaparecen durante el resto del ciclo.

Existen numerosas ciclinas; las principales en humanos son las ciclinas A, B, D y E. Algunas de ellas están relacionadas entre sí; por ejemplo, las ciclinas B1 y B2, o las D1, D2, D3 y D4. La función exacta de otras ciclinas (C, F, G, H e I) se conoce peor; la ciclina F parece estar implicada en el control de la transición entre las fases S y G₂.

a) Ciclinas de la fase G₁ o ciclinas de inicio

Son las que intervienen regulando el punto de control G₁/S. En humanos pertenecen a este grupo las ciclinas del tipo D (D1, D2, D3 y D4) y la ciclina E. Ambas se sintetizan durante la fase G₁ (por ejemplo, al entrar la célula

en el ciclo desde G₀) y su concentración celular se regula mediante transcripción. Debido a su rápida degradación, tienen una vida media corta (unos 25 min para las D). El balance entre síntesis y degradación hace que aparezcan en la célula al final de G₁ y desaparezcan durante la fase S o al acabar ésta.

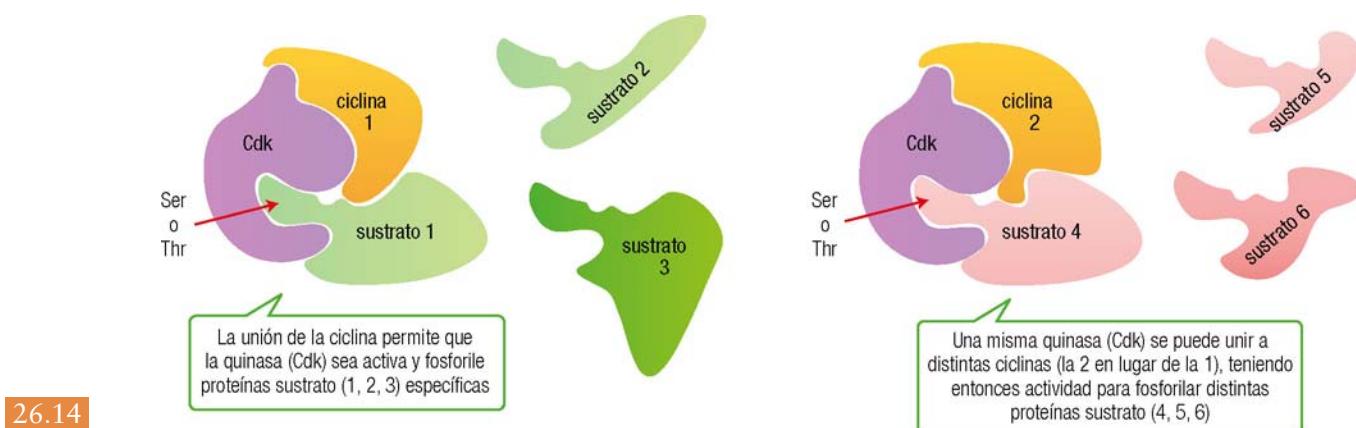


b) Ciclinas mitóticas o de la fase G₂

Este grupo, formado por las ciclinas A, B1 y B2, regula el punto de control G₂/M. En este caso se sintetizan durante las fases previas a la mitosis (S o G₂) y se degradan rápida y específicamente durante la mitosis. La ciclina A participa en el control de la replicación del DNA y en la entrada en mitosis. La ciclina B, que sólo participa en la mitosis, se considera la ciclina por antonomasia, pues es la homóloga en mamíferos de la primera ciclina descubierta cuando comenzó el estudio del control del ciclo celular en los oocitos de la rana *Xenopus laevis*, la que forma parte del *factor promotor de la maduración o de la fase M* (MFP en inglés).

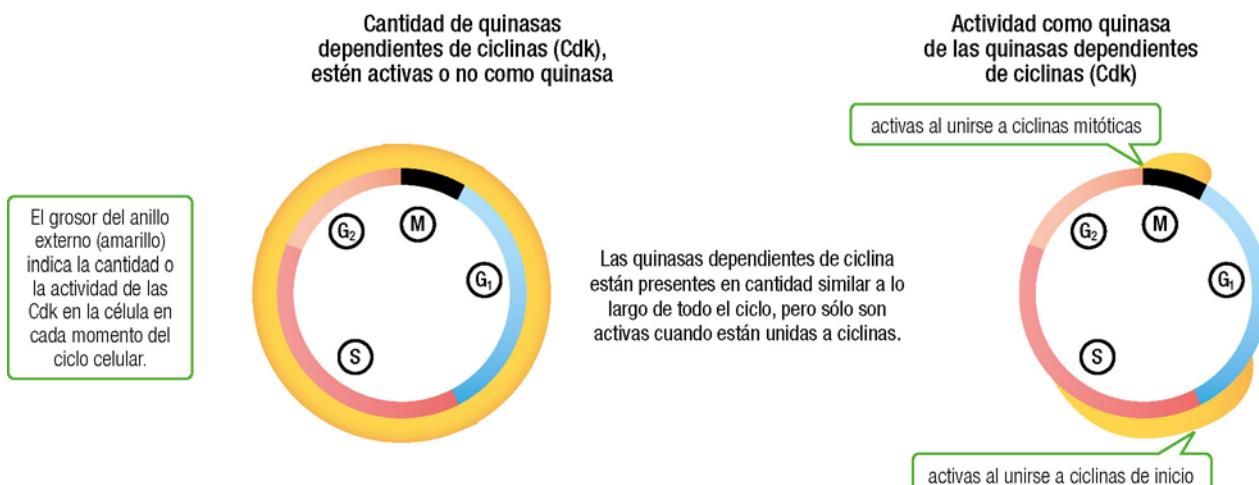
26.6.2.2 Proteína quinasas dependientes de ciclinas (Cdk)

Constituyen una familia de proteínas con una gran homología de secuencia (superior al 50%), que fosforilan residuos de serina y treonina en proteínas sustrato específicas. Se trata, por tanto, de serina/treonina quinasas (pág. 360 y web 22.4). Sin embargo, sólo pueden ejercer esa actividad enzimática cuando están asociadas con una ciclina; por ello, reciben el nombre de quinasas dependientes de ciclinas, Cdk o CDK (de *cyclin-dependent kinases*). De acuerdo con ello, se considera que en el complejo activo (Cdk-ciclina) la Cdk es la subunidad catalítica y la ciclina la subunidad reguladora.



Una vez fosforiladas por las Cdk, las proteínas sustrato pueden perder los grupos fosfato por acción de fosfoproteína fosfatases igualmente específicas. Esta reversibilidad de la fosforilación es crucial para el control del ciclo celular.

A diferencia de las ciclinas, la concentración de las Cdk no varía de forma marcada durante el ciclo celular, aunque puede estar regulada por ciertas señales estimuladoras o inhibidoras del crecimiento.



En algunos organismos (como las levaduras) existe una sola Cdk, que actúa tanto en el punto de control G₁/S como en el G₂/M (uniéndose a distintas ciclinas), mientras que en otros (mamíferos) esas funciones se distribuyen entre varias Cdk. En el ser humano se han encontrado al menos 8 miembros de esta familia, denominados Cdk1 a Cdk8, cada uno de los cuales forma complejos quinasa activos con ciclinas diferentes.

Resumen de las distintas ciclinas y quinasas dependientes de ciclina (Cdk) conocidas en humanos. Se indica las parejas que se asocian y los puntos de control del ciclo en los que actúan.

Observaciones:

- La asignación exacta de puntos de control difiere según las fuentes; las indicadas entre paréntesis están menos confirmadas.
- Cdk1 se llama comúnmente Cdc2, por ser homóloga de la Cdc2 de levaduras.
- El complejo Cdk2/ciclina A parece actuar a lo largo de toda la fase S, mientras que los demás actúan principalmente en las transiciones entre G₁ y S o entre G₂ y M (puntos de control).
- Cdk1/B parece actuar también durante la fase M.
- Cdk7 y ciclina H son subunidades componentes de TFIIH, factor de transcripción (apdo. 18.4.3.1) que además interviene en la reparación del DNA (apdo. 23.6.3.1).

	Ciclinas de G ₂ A B	Ciclinas de G ₁ D E	Otras ciclinas menos conocidas C F G H I
Proteína quinasa	G ₂ /M G ₂ /M S asociaciones de ciclinas de G ₂ que controlan el ciclo en el punto indicado	(G ₁ /S) G ₁ /S G ₁ /S (G ₁ /S) G ₁ /S	asociaciones de ciclinas de G ₁ que controlan el ciclo en el punto indicado asociaciones de otras ciclinas, cuyo punto de control no se conoce ✓ S/G ₂ control por una Cdk desconocida
Cdk1			
Cdk2			
Cdk3			
Cdk4			
Cdk5			
Cdk6			
Cdk7			
Cdk8	?		

26.16

26.6.2.3 Actividad de los complejos Cdk-ciclina

La activación de las Cdk como quinasas de proteínas no consiste simplemente en su unión con una ciclina, sino que su regulación depende de otros mecanismos:

a) La posibilidad de que se forme el complejo Cdk-ciclina depende de la disponibilidad de ambas proteínas; puesto que las Cdk están presentes de forma continua en la célula, son las ciclinas las que regulan.

a1) **Tasa de síntesis de las ciclinas:** como se ha indicado (pág. 464), la síntesis está controlada en función de la etapa del ciclo, aumentando en ciertos puntos del mismo, diferentes para las ciclinas A y B que para las D y E.

a2) **Tasa de degradación de las ciclinas:** determina, junto con la síntesis, la concentración disponible para su unión a las Cdk.

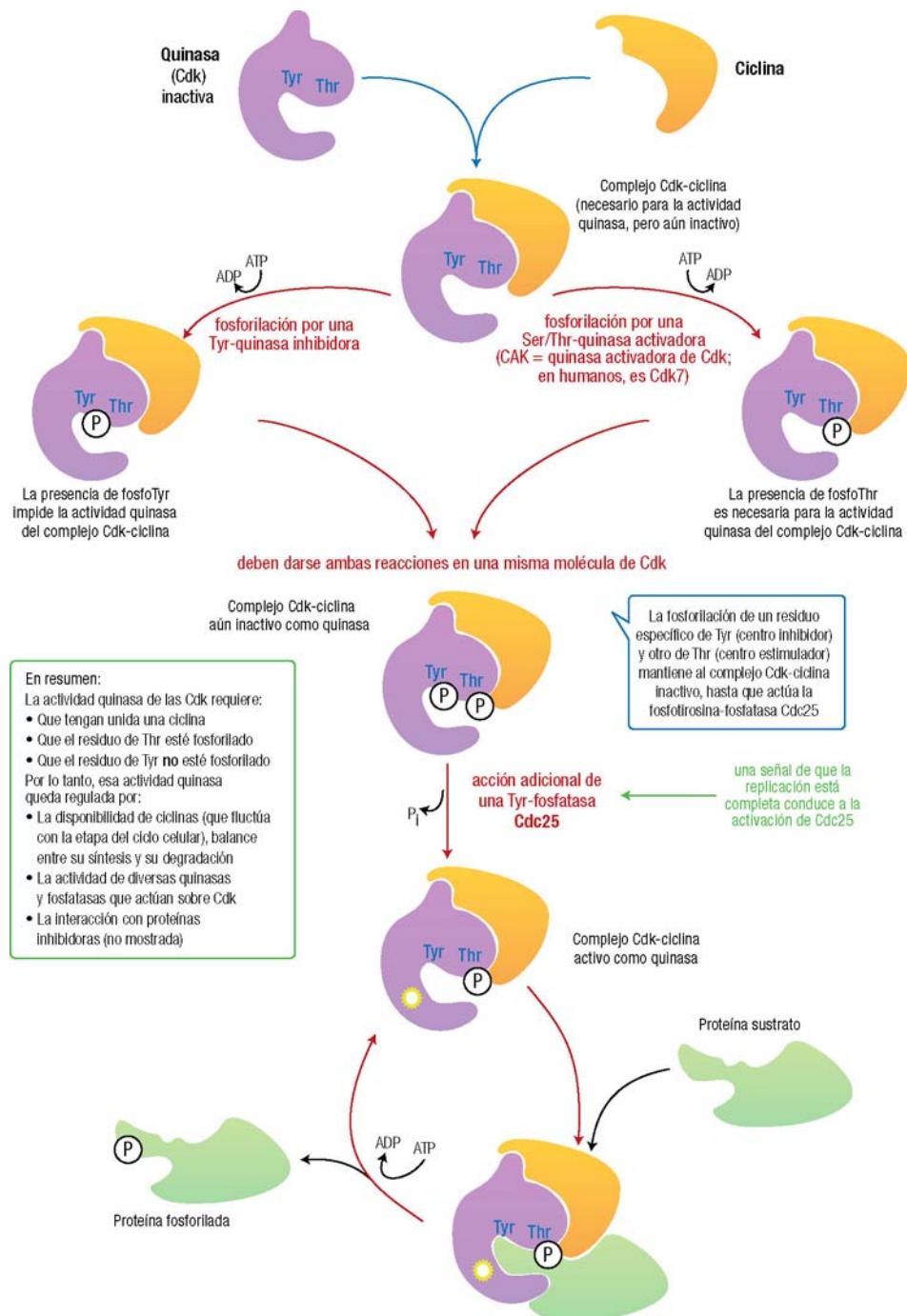
b) Una vez formado el complejo Cdk-ciclina, su actividad como quinasa está sometida a regulación por otras proteínas.

b1) **Modificación de la Cdk por fosforilación y desfosforilación:** la actividad del complejo Cdk-ciclina precisa la fosforilación de ciertos residuos aminoácidos de la Cdk y no de otros, lo cual está regulado por distintas quinasas y fosfatasas que, a su vez, están reguladas por la propia Cdk activa (véase figura en pág. 467).

b2) **Interacción de las quinasas con proteínas inhibidoras:** el último mecanismo regulador de la actividad de los complejos Cdk-ciclina se basa en la existencia de muchas proteínas que se unen a las Cdk, inhibiendo directamente su actividad; se denominan *inhibidores de los complejos Cdk-ciclina*, o CKI. Su efecto es, por tanto, bloquear el ciclo de división, por lo que funcionan como oncosupresoras (pág. 456).

Los CKI se agrupan en dos familias, en función de su estructura (existe similitud de sus secuencias), modo de acción y especificidad de sustrato:

Familia	Nombre de la proteína inhibidora	Otros nombres	Acción inhibidora
KIP	p21	CIP1	Preferentemente sobre los complejos de Cdk con ciclina de G ₁ y S. En menor medida, sobre complejos de Cdk1 con ciclina de G ₂
	p27	KIP1	
	p57	KIP2	
INK4	p16	INK4a	Específicamente sobre los complejos de ciclinas de G ₁ con Cdk4 o Cdk6, pues compiten con las ciclinas D
	p15	INK4b	
	p18	INK4c	
	p19	INK4d	



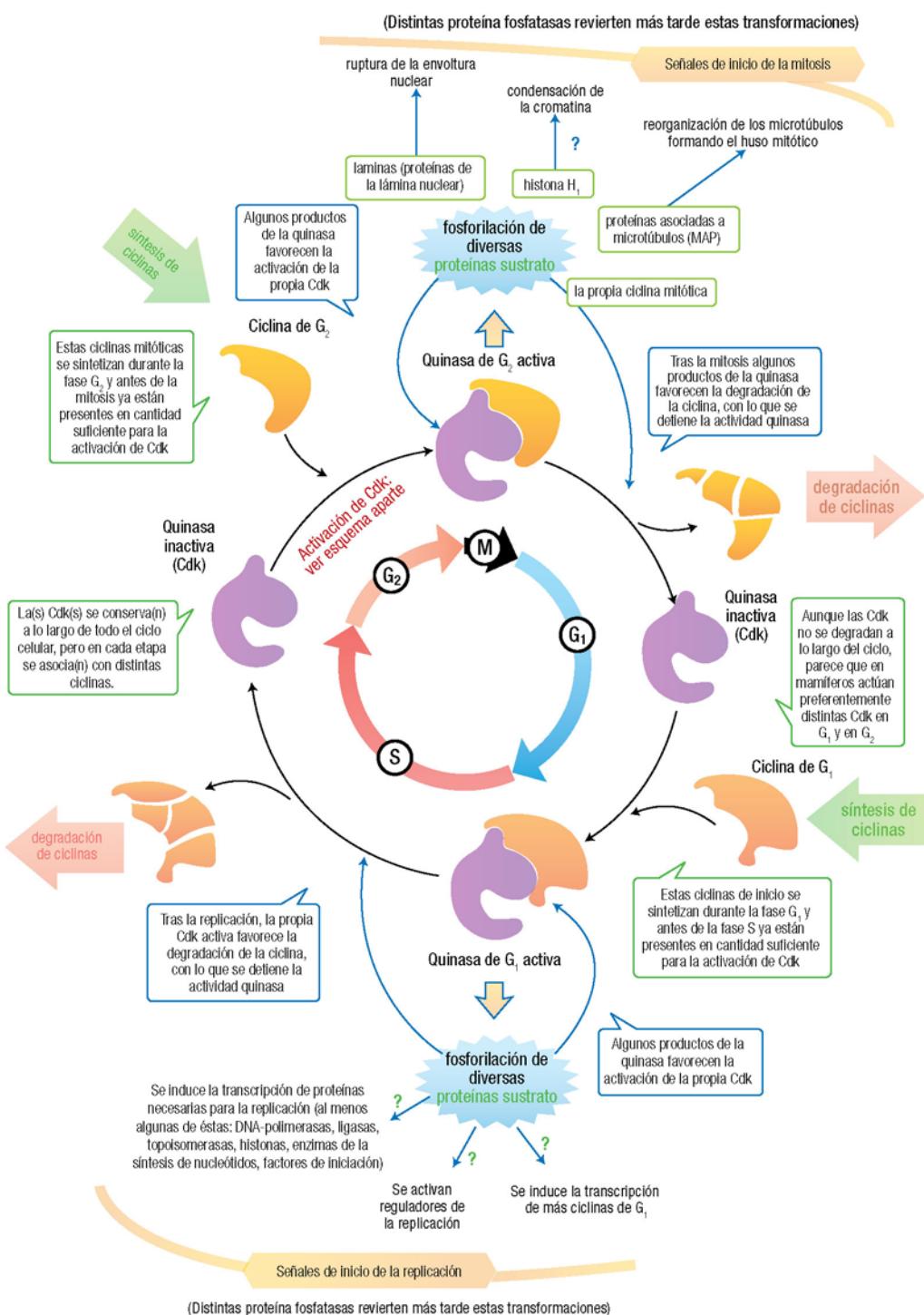
26.17

26.6.3 Regulación del ciclo celular por los complejos Cdk-ciclina y alteraciones en el cáncer

El descubrimiento de la relación entre diversos procesos neoplásicos y las alteraciones en algunos genes, especialmente los que poseen un papel clave en el control del ciclo celular, ha permitido comprender la trascendencia y los mecanismos de funcionamiento de este control. Distintas alteraciones en las proteínas que regulan el ciclo de división celular pueden tener como consecuencia la división celular excesiva o incontrolada, dando lugar a una proliferación celular desmedida con acumulación de anomalías genéticas y fisiológicas en las células, que pasan a constituir el tumor.

26.6.3.1 Estímulos de progresión por el ciclo celular y oncogenes que los alteran

Como ya se ha descrito, el avance por el ciclo celular se impulsa a grandes rasgos en dos puntos: la síntesis y degradación de ciclinas de G₁, seguida de la activación de sus complejos Cdk-ciclina, y la etapa equivalente en G₂. La activación de estos complejos conduce a respuestas variadas, mediadas por las proteínas diana que resultan fosforiladas por la actividad quinasa. Se muestra a continuación una visión global de estos dos puntos de activación del ciclo.

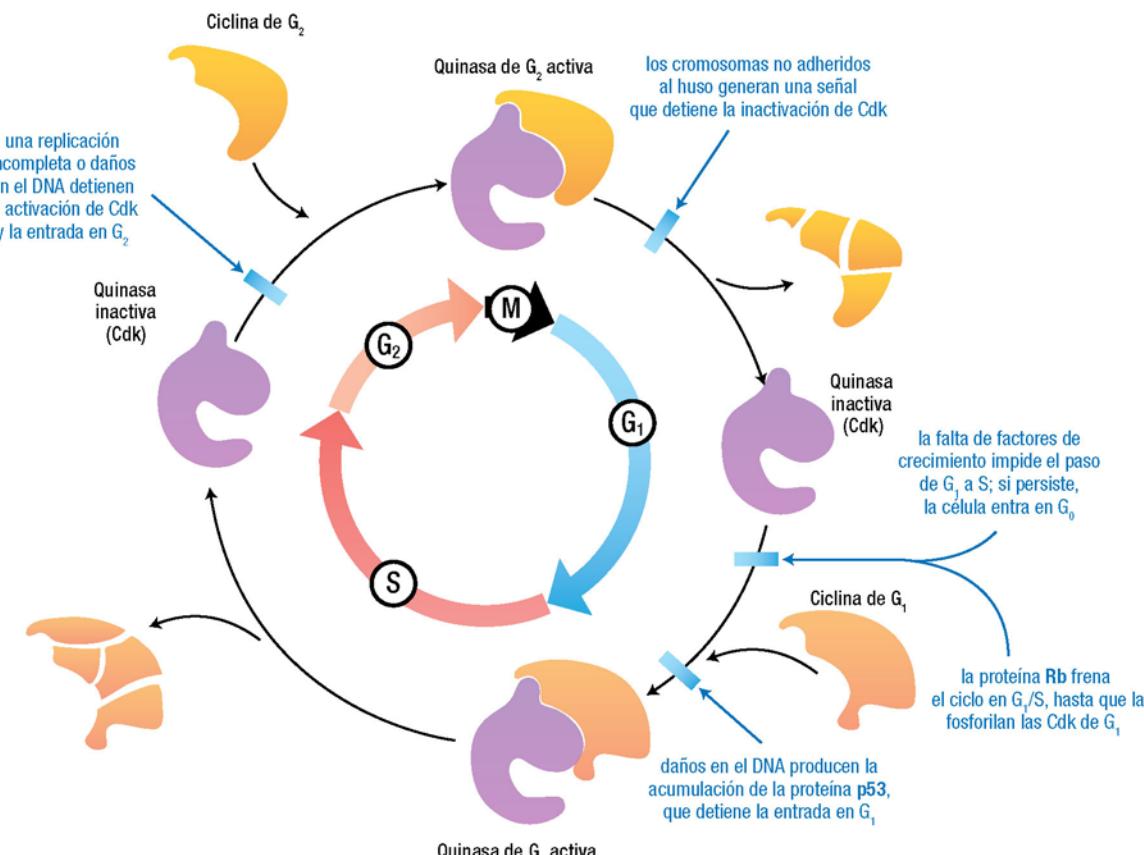


Esta estimulación normal de la progresión por el ciclo puede alterarse y conducir a un cáncer. La mutación de protooncogenes que codifican estas proteínas estimuladoras del ciclo los pueden transformar en oncogenes. Esto ocurre cuando las oncoproteínas respectivas actúan sin obedecer a la regulación, provocando una estimulación continuada del ciclo que conduce a la proliferación excesiva causante del cáncer.

Ejemplo de proteína reguladora del ciclo celular, codificada por un protooncogén	Efecto oncogénico resultante de la mutación
Ciclina	Síntesis excesiva Síntesis no regulada de acuerdo con el ciclo
Cdk	Síntesis excesiva Actividad quinasa excesiva o independiente de la unión de la ciclina
Proteína quinasa o proteína fosfatasa activadoras del complejo Cdk-ciclina	Síntesis o actividad excesivas
Proteína quinasa o proteína fosfatasa inhibidoras del complejo Cdk-ciclina	Proteína no funcional

26.6.3.2 Estímulos de detención del ciclo celular

La regulación normal del ciclo depende del balance entre los estímulos positivos que lo impulsan y los negativos o frenos que detienen el ciclo en los tres puntos de control, G₁/S, G₂/M y M (pág. 462), para asegurar la correcta preparación de la célula para la fase siguiente (por ejemplo, que se haya completado la replicación antes de que se condense el genoma y se divida la célula). Los mecanismos de actuación de estos estímulos negativos se conocen con menos detalle molecular que aquellos por los que tiene lugar el control positivo; sin embargo, conceptualmente sus implicaciones en la regulación normal y alterada en el cáncer son similares en ambos casos.

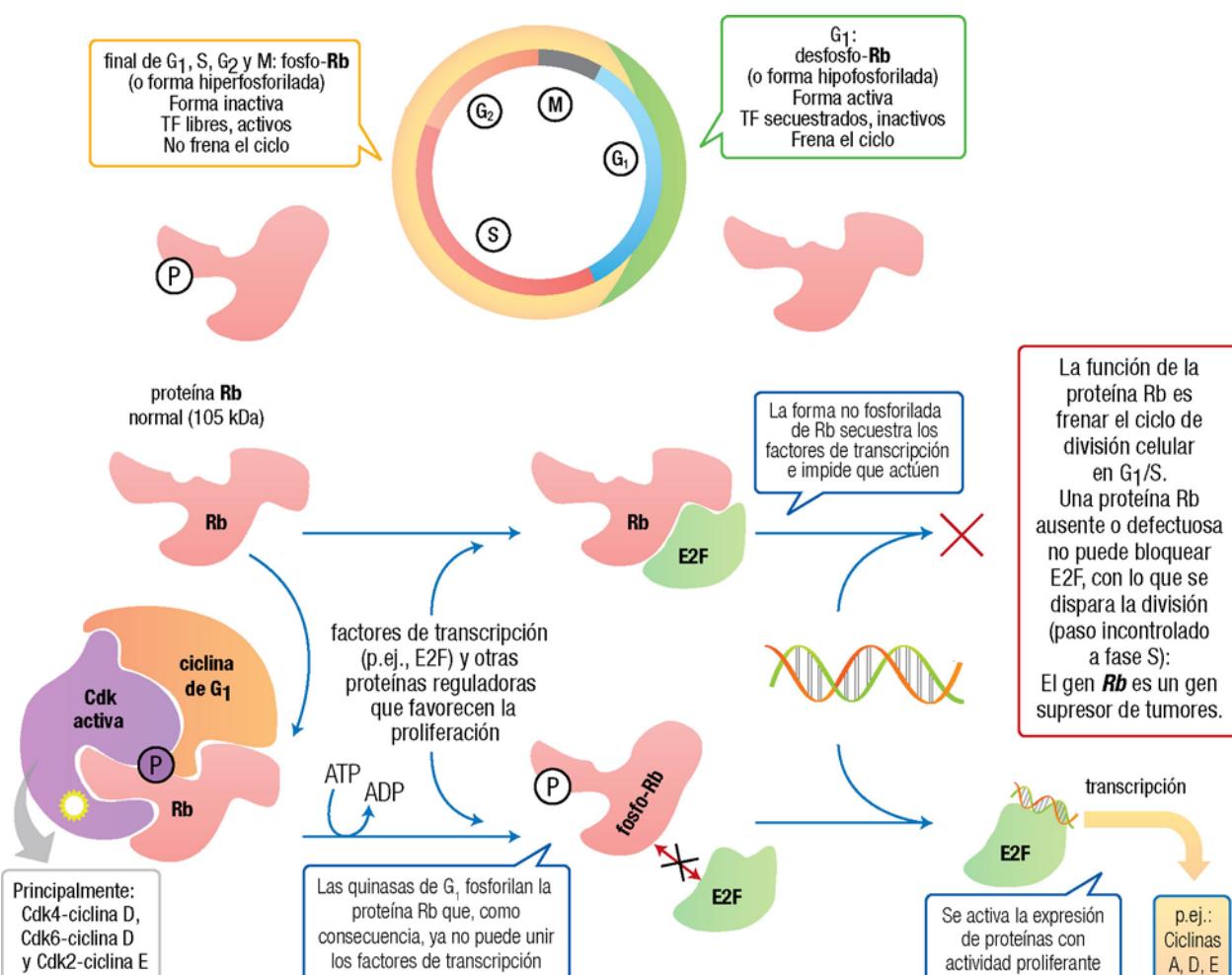


La aparición de cáncer por alteración del control negativo corresponde a genes oncosupresores, o supresores de tumores, cuya función normal es frenar el ciclo y, al verse mutados, dejan de hacerlo (pág. 456). Como consecuencia, la célula replica su DNA y se divide por mitosis sin atender a los puntos de control y, en consecuencia, prolifera excesivamente y acumula defectos genéticos (por ejemplo, por una mitosis prematura antes de que se haya completado la replicación). Entre los múltiples genes supresores de tumores que se han descrito, los mejor estudiados son *Rb* y *p53*, cuyos productos proteicos actúan sobre la regulación del ciclo. A continuación se detallan su función normal y su implicación en el cáncer.

26.6.3.3 El gen supresor del retinoblastoma (*Rb*)

El descubrimiento de una predisposición hereditaria para sufrir *retinoblastoma*, un tipo infrecuente de cáncer que afecta a los ojos, permitió identificar un gen (denominado *Rb*) relevante en el control del ciclo celular. Su producto génico, la proteína del retinoblastoma (*Rb*, pRb o p105-Rb, por su masa de 105 kDa), se encuentra presente y actúa en el control de todo tipo de células normales; de ahí su importancia general, en contraste con el carácter restringido de actuación que pudiera deducirse de su nombre.

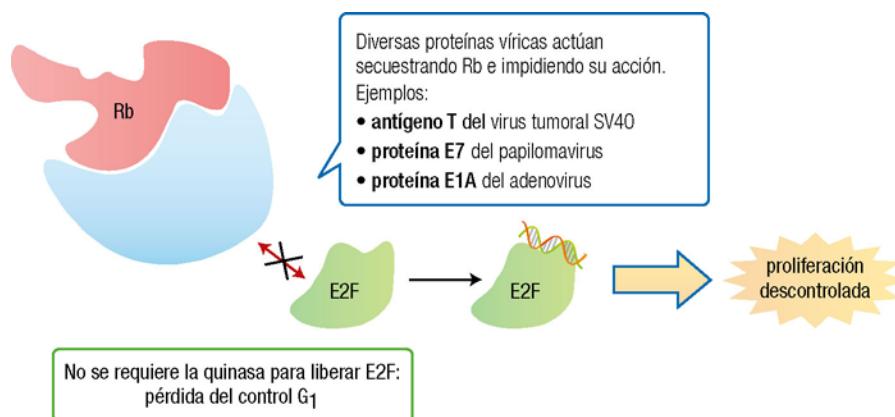
La proteína Rb se caracteriza por experimentar cambios de fosforilación a lo largo del ciclo celular, que modulan su actividad como freno en el punto de control G₁/S. Esta actividad se debe a la asociación de Rb no fosforilada con multitud de proteínas necesarias para la proliferación celular, en especial factores de transcripción inducibles (pág. 293) como Myc y E2F.



En una célula normal, la proteína Rb está siempre presente, pero su estado de fosforilación (y, por consiguiente, su actividad) varía a lo largo de las etapas del ciclo. Los factores estimuladores del crecimiento actúan como señal que favorece la activación de las proteínas Cdk, con lo que éstas pueden fosforilar Rb y anulan así su efecto inhibidor sobre la progresión del ciclo en G₁/S.

La pérdida o mutación de las dos copias (alelos) del gen *Rb* anula la función de la proteína Rb, suprime el control por Cdk y provoca un paso prematuro a fase S. Aunque la inactivación del gen *Rb* es responsable de cánceres en diversos tejidos, afecta en especial a las células de la retina inmadura, conduciendo en ella a una proliferación celular excesiva que causa el retinoblastoma. Los niños que heredan una copia defectuosa del gen *Rb* poseen un único alelo normal, por lo que es más probable que una mutación anule por completo la funcionalidad de Rb y tengan mayor susceptibilidad al desarrollo de tumores en la retina.

Finalmente, la acción de la proteína Rb normal se puede ver también impedita como consecuencia de ciertos virus, que tienen por ello acción promotora del cáncer independiente de la mutación del genoma humano.



26.21

26.6.3.4 El gen oncosupresor p53

Aproximadamente el 50% de los tumores humanos presentan mutaciones en un gen relacionado con el control del ciclo celular, denominado *p53* por codificar una proteína cuya masa molecular es de 53 kDa. Esto no quiere decir que la proteína p53 anormal sea la única causa del 50% de los tumores, pues se acumulan mutaciones en otros genes cuyo efecto combinado es la enfermedad. Sin embargo, las lesiones del gen *p53* tienen un efecto claro sobre la progresión neoplásica.

El carácter oncosupresor del gen *p53* se debe a las funciones de la proteína normal que codifica, que consisten en **inducir la apoptosis** y actuar como **freno en el punto de control G₁/S**. La concentración de la proteína p53 en células normales es baja, por lo que sólo adquiere relevancia en situaciones especiales. Así, se ha comprobado un importante aumento de p53 cuando existe daño en el DNA (por ejemplo, tras la exposición a radiaciones), cuando se altera la concentración de nucleótidos en la célula, en situaciones de hipoxia, bajo estímulos oncogénicos, etc. En esas condiciones, aumenta la concentración de la proteína p53 y se prolonga su vida media (pasando de 5-10 minutos hasta 150 minutos). La presencia de daños en el DNA es detectada por varias proteínas, entre ellas una quinasa de proteínas que fosforila a p53. La forma fosforilada de p53 es un factor de transcripción (pág. 288) que induce la expresión de algunos genes y reprime la de otros. Entre las proteínas cuya expresión aumenta destaca, en primer lugar, **p21**, que inhibe la actividad quinasa del complejo Cdk2-ciclina E, necesaria para superar G₁/S. En segundo lugar, la proteína **GADD** (*growth arrest on DNA damage*, detención del crecimiento por daños en el DNA) que, al igual que p21, sequestra el PCNA, factor de procesividad de la DNAPol δ (v. Web 11.5), e impide, de este modo, la replicación. En consecuencia, y de forma similar a Rb, la función normal de p53 es frenar la progresión del ciclo, evitando que la célula replique ese DNA dañado. Más aún, es posible que la propia proteína p53 estimule la reparación del DNA o participe en ella.

Los defectos homocigóticos en *p53* conducen, por tanto, a una falta de control en el punto G₁/S y una reducción de la entrada en apoptosis, lo que da como resultado una proliferación celular descontrolada que ocasiona tumores. En individuos heterocigóticos producen el síndrome de Li-Fraumeni, caracterizado por una predisposición al desarrollo de tumores. Por último, de manera análoga a lo que ocurre sobre Rb, diversas

proteínas víricas generan tumores a través del bloqueo de p53 (*antígeno T* de SV40, *proteína E6* de papilomavirus y *proteína E1B* de adenovirus, por ejemplo).

26.7 APOPTOSIS O MUERTE CELULAR PROGRAMADA

La apoptosis es un proceso fisiológico, controlado genéticamente, que se da en todo tipo de células de organismos multicelulares y por tanto en todos los órganos y sistemas, mediante el cual las células dañadas o no necesarias (que han cumplido su ciclo vital, pág. 452) activan mecanismos que conducen a su propia destrucción de forma ordenada y regulada; de ahí que se le haya llamado “suicidio celular”. Éste es un proceso muy diferente de la necrosis, causada por el efecto nocivo directo de agentes externos, de naturaleza mecánica (rotura), biológica (virus, bacterias), química (drogas, toxinas, venenos), etc., que conducen a pérdida de la capacidad de regulación osmótica, hinchamiento, rotura de la membrana, liberación del contenido celular y, finalmente, fagocitosis de los residuos celulares.

Se observa una reducción de la apoptosis en situaciones fisiológicas caracterizadas por una gran proliferación celular (desarrollo embrionario y posnatal temprano, metamorfosis, mecanismo de acción de hormonas y factores paracrinos, etc.), así como en situaciones patológicas tan diversas como el cáncer (la más representativa de una excesiva proliferación), enfermedades autoinmunes e infecciones víricas. De ahí la importancia del estudio del proceso apoptótico, su control y los fallos asociados a procesos patológicos.

El estudio de la apoptosis ha de hacerse en conexión con el de la proliferación celular. Como ya se ha indicado (pág. 452), la célula cancerosa se considera una célula con una proliferación incontrolada debida probablemente a la suma de una activación de la proliferación y un bloqueo de la apoptosis. De hecho, la propia quimioterapia del cáncer está imbuida de esta idea: el impacto más bien modesto de los tratamientos anti-proliferativos, posiblemente causado por una baja capacidad de proliferación de muchos tumores, debe complementarse con tratamientos que induzcan la apoptosis.

26.7.1 Etapas y control del proceso de apoptosis

Desde un punto de vista celular y molecular, los cambios en la apoptosis parecen transcurrir en tres etapas relativamente definidas, denominadas inducción, ejecución y fase final.

26.7.1.1 Inducción

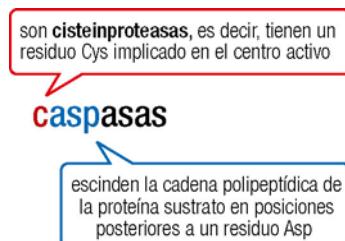
Se da este nombre a la reacción ante un estímulo o señal. Al igual que en la proliferación celular, es importante indicar que la inducción de la apoptosis requiere estímulos adecuados (señales positivas), tanto extracelulares como procedentes de la propia célula. Existen también señales negativas que actúan en sentido contrario, frenando la apoptosis.

26.7.1.2 Ejecución

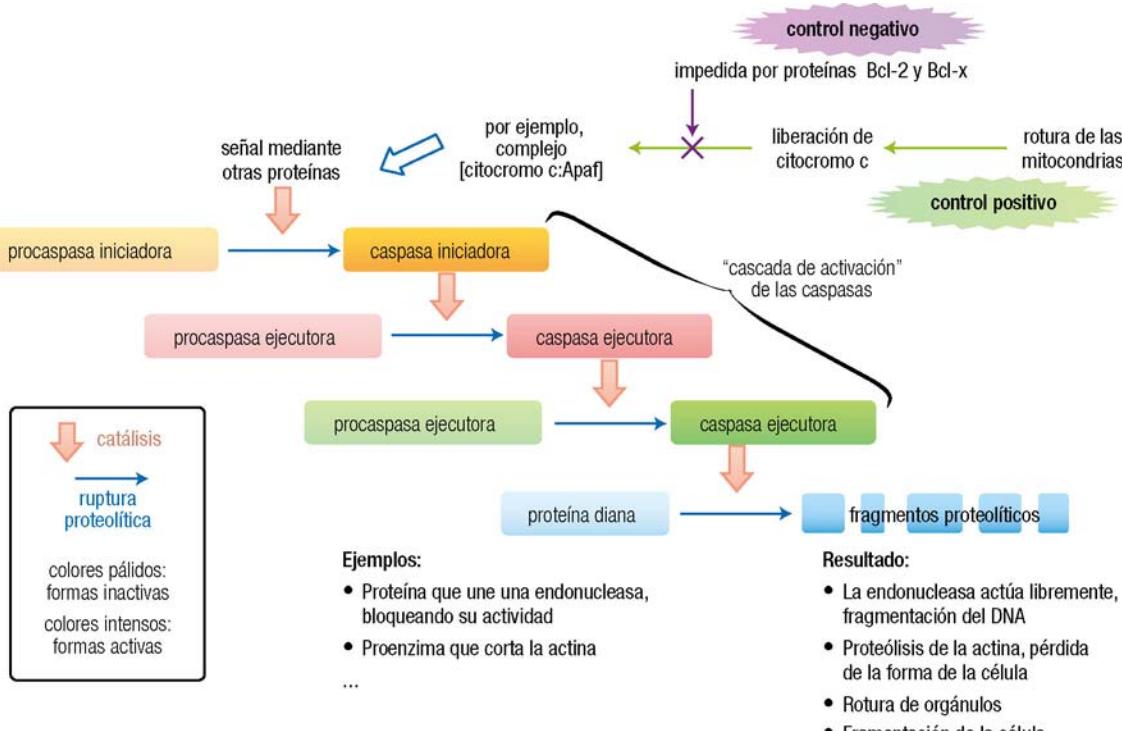
Tras la inducción comienza el verdadero proceso de apoptosis, que consta de alteraciones diversas que conducirán a la muerte de la célula: cambios morfológicos y de la membrana con disminución de la adhesión a otras células y a la matriz extracelular, contracción del citoplasma, condensación de la cromatina, fragmentación del DNA, disminución del volumen nuclear seguida de su desintegración, rotura de mitocondrias con vertido de su contenido al citoplasma, aparición de protuberancias en la membrana celular (fenómeno conocido como *zeiosis*), etc. Estas alteraciones se deben a la activación de varias proteasas y nucleasas, que se convierten así en las principales moléculas responsables de la apoptosis.

a) Proteasas: caspasas y su papel en el control de la apoptosis

26.22



Las caspasas son una familia de proteasas citosólicas (dentro de la clase EC 3.4.22, cisteína endopeptidasas). En humanos se han identificado 10, clasificadas en tres grupos de acuerdo con su especificidad de sustrato. Producen la proteólisis de proteínas diana o sustrato, entre las cuales cabe citar una proteína nuclear con actividad de reparación del DNA (PARP o poli(ADP-ribosa)polimerasa) y un factor de transcripción (SREBP). Las caspasas no sólo actúan degradando proteínas sustrato, sino que también activan a otras caspasas por proteólisis limitada (pág. 369). Dependiendo de su posición en la cascada de activación, se distingue así entre caspasas "iniciadoras" y "ejecutoras".



La apoptosis se controla mediante proteínas implicadas en varios niveles de señalización celular, codificadas por protooncogenes y genes oncosupresores. Al igual que el ciclo celular, el control de la apoptosis puede alterarse, generando situaciones patológicas entre las que se incluye el cáncer. Se conocen ya algunos genes que codifican proteínas reguladoras de la apoptosis y que resultan alterados por mutaciones.

El **control positivo** consiste en el inicio de la cascada de caspasas; una de las principales señales inductoras es la liberación al citoplasma de los componentes de las mitocondrias rotas, en especial el **citocromo c** (integrante de la cadena respiratoria). Éste se une a la proteína **Apaf** (factor activador de las proteasas apoptóticas) formando un complejo que activa la procaspasa iniciadora. Algunas proteínas, como **Bax**, estimulan estas etapas de inicio; en ciertos cánceres **Bax** está inactiva, anulando su efecto de control positivo. Este control positivo de la apoptosis es análogo al ejercido en la proliferación celular por la inducción en cascada de la actividad proteína quinasa de los complejos Cdk-ciclina (pág. 466).

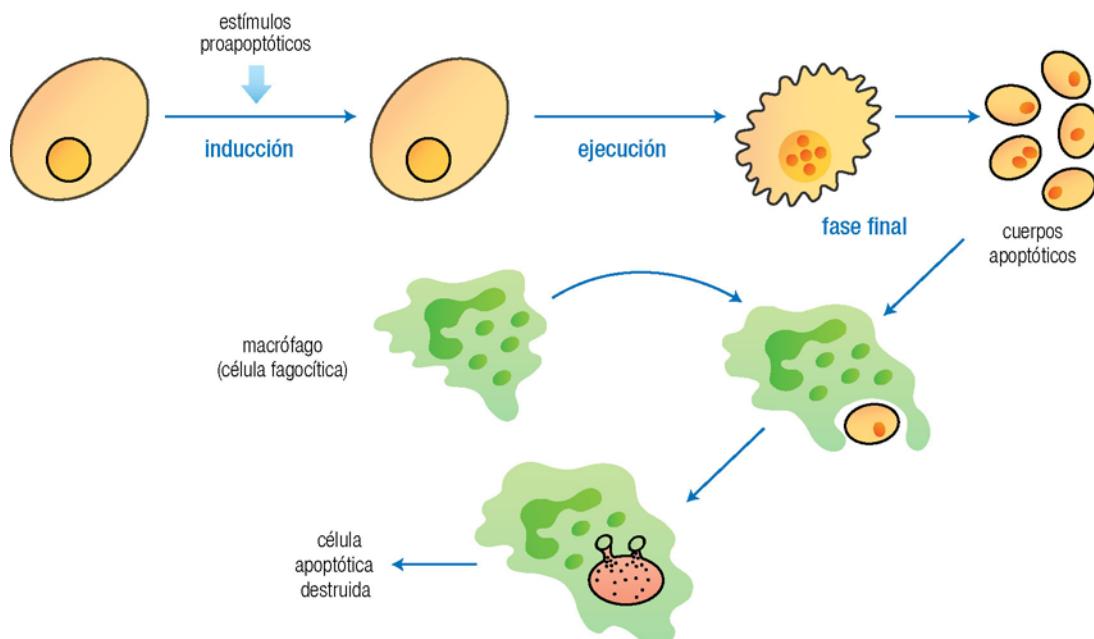
El control negativo de la apoptosis, por freno de la activación de las caspasas, viene determinado por proteínas como Bcl-2 y Bcl-x, que aseguran que el proceso apoptótico permanezca bloqueado bajo condiciones normales. En algunos cánceres se sobreexpresa Bcl-2, reduciendo aún más la apoptosis; en otros, se sobreexpresan proteínas que bloquean las caspasas, como las denominadas *survivinas*. En general, este control es similar al ejercido sobre el ciclo celular por la activación de proteínas que inhiben la actividad proteína quinasa de los complejos Cdk-ciclina.

b) Nucleasas

Se encargan de la rotura del DNA cromosómico, en general en sus posiciones más accesibles, el DNA espaciador o internucleosómico, de forma que se originan fragmentos cuyo tamaño es múltiplo de 200 pb (longitud aproximada del DNA de cada nucleosoma, pág. 88). Por ello, uno de los síntomas de apoptosis es la observación, al estudiar el DNA por electroforesis, de un patrón de bandas “en escalera” (v. web 7.3).

26.7.1.3 Fase final de la apoptosis

Las protuberancias externas de la célula, formadas durante la etapa de ejecución, terminan por separarse formando los llamados **cuerpos apoptóticos**; éstos son característicos del proceso y constituyen una de las diferencias respecto a la necrosis celular. Los cuerpos apoptóticos son reconocidos y fagocitados por otras células, tanto vecinales o del propio tejido como especializadas (macrófagos), de forma que toda la célula apoptótica se destruye sin que se llegue a liberar material intracelular al medio y, en consecuencia, sin la aparición de una reacción inflamatoria (a diferencia de la necrosis, en la que la célula revienta, libera su contenido y provoca esa reacción). Obviamente, todo este proceso va acompañado de otros cambios bioquímicos.



26.24

26.7.2 Relación entre ciclo celular y apoptosis: p53 como punto de enlace

Como se ha indicado previamente, el gen oncosupresor *p53* codifica la proteína nuclear *p53*, reguladora de la transcripción de ciertos genes relacionados con la progresión del ciclo celular (pág. 471). Por otra parte, *p53* también provoca la apoptosis, cuando las lesiones del DNA son imposibles de reparar, bien sea por su magnitud o por defectos en los sistemas de reparación. Se elimina así la posibilidad de que se transmita el DNA gravemente

alterado (con gran número de mutaciones, con reorganizaciones cromosómicas o con aneuploidía, por ejemplo). Los mecanismos por los cuales p53 produce esa activación de la apoptosis son múltiples y se conocen aún de forma incompleta, pero incluyen tanto la activación de la transcripción de genes como otras actividades no transcripcionales de la proteína p53. La alteración de estas funciones de p53, debido a la mutación de su gen, impide la entrada de la célula en apoptosis, lo que, unido a la desaparición de la función de freno del ciclo celular, tiene como consecuencia que las células continúen proliferando aunque posean daños en el DNA o defectos genéticos, y ello aumenta la probabilidad de que éstos se acumulen y favorece la aparición de un cáncer.

26.8 HERENCIA DEL CÁNCER

Las alteraciones inducidas en el cáncer se transmiten en la mitosis de unas células a otras (células somáticas), pero no se transmiten por herencia de padres a hijos. Para ello habría de existir una predisposición o susceptibilidad a la mutación en la línea germinal, una **susceptibilidad hereditaria**. Existen, sin embargo, casos de cáncer hereditario, debidos a la mutación de protooncogenes o genes oncosupresores en el óvulo o el espermatozoide de los padres. La causa y mecanismo moleculares del desarrollo del cáncer hereditario son los mismos que para el cáncer somático, sólo difieren en el tipo de célula en la que se inicia (germinal y somática, respectivamente). Por otra parte, una vez heredado el cáncer, el individuo afectado se convierte en portador y lo transmitirá a sus descendientes de acuerdo con las reglas de herencia mendeliana y dependiendo del carácter dominante o recesivo del gen alterado causante del cáncer (por lo común, dominante si es un oncogén y recesivo si se trata de un gen oncosupresor mutado, págs. 456 y 458).

Uno de los cánceres hereditarios mejor conocidos y con especial trascendencia es el de mama, que tiene este carácter en un 5 a 10% de los casos. Su causa principal es la mutación de los genes oncosupresores BRCA1 del cromosoma 17 o BRCA2 del 13 (de *breast cancer*, cáncer de mama). El primero de estos genes se expresa en mama y ovario, codificando un factor de transcripción que posee un motivo dedo de zinc (pág. 62) en su región N-terminal. Al tratarse de un gen oncosupresor, la transmisión hereditaria se debe, en general, a un alelo mutado, que sólo se manifestará en la hija desarrollando un tumor si también sufre mutación posteriormente el otro alelo, heredado como forma normal (pág. 458). La presencia de BRCA1 anormal se asocia en mujeres a un 50-80% de probabilidad de padecer cáncer de mama y a una predisposición al cáncer de ovario.

Existen otros defectos genéticos hereditarios, en los que se transmite la predisposición a desarrollar un tipo específico de cáncer; entre ellos, xerodermia pigmentosa (pág. 404), enfermedad de Li-Fraumeni, cáncer cerebral, sarcoma, leucemia, ataxia-telangiectasia, linfoma no hodgkiniano y leucemia linfocítica aguda. En general, la aparición del cáncer se demora bastante, pues es necesario que se acumulen mutaciones somáticas adicionales.

26.9 MARCADORES TUMORALES

Las alteraciones de las células tumorales pueden analizarse en diversas escalas. Lo tradicional y más utilizado en clínica son aquellas que, aun sin estar directamente implicadas o ser responsables del proceso neoplásico, son consecuencia de la existencia del tumor o incluso de su evolución o respuesta a una determinada terapia. Estas alteraciones constituyen la base de los denominados **marcadores tumorales y tisulares**, establecidos con fines clínicos y en especial relacionados con el grado de malignidad del tumor. Se pueden observar tanto *in vitro* (por ejemplo en células en cultivo infectadas con virus oncogénicos en el proceso llamado *transformación maligna*) como *in vivo*, en numerosos casos clínicos.

El uso de los marcadores tumorales sólo es válido si para su determinación se cumplen ciertos requisitos analíticos. En cuanto a sus posibilidades de utilización clínica, es preciso exponer algunas consideraciones básicas. Por un lado, raramente se emplean los datos de marcadores para establecer un diagnóstico, sino que éste ha de basarse en datos clínicos, radiológicos, biopsias, etc.; en general, los marcadores sólo sirven para confirmar la evidencia del cáncer o como alerta temprana para iniciar un diagnóstico más detallado. En cuanto al pronóstico de una enfermedad, se exige una correlación entre la concentración del marcador y la masa del tumor. Son adecuados para el seguimiento del cáncer, en cuyo caso debe confirmarse la eficacia del tratamiento siguiendo la concentración del marcador a lo largo del tiempo, para averiguar, por ejemplo, si el aumento de un marcador es indicio de recurrencia de la enfermedad y se requiere una terapia más enérgica o un cambio de estrategia.

Las alteraciones de las células cancerosas son tan variadas que pueden estudiarse desde distintos puntos de vista: locales y generalizadas, tumorales y tisulares y, finalmente, genotípicas o moleculares.

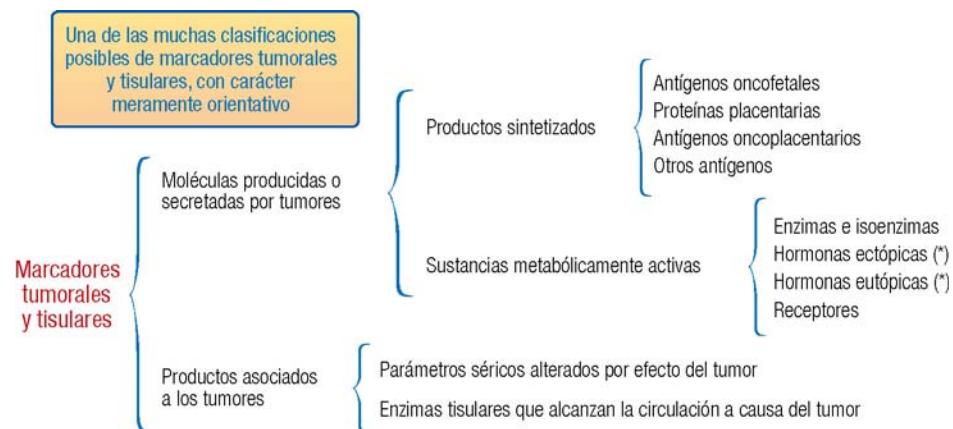
26.9.1 Alteraciones locales y generalizadas

El proceso canceroso ejerce una influencia directa sobre la propia célula y su vecindad (efecto local) o sobre la totalidad del organismo (efecto generalizado o sistémico).



26.25

26.9.2 Alteraciones tumorales y tisulares



Ectópica = producida por un tejido que normalmente no sintetiza esa hormona

Eutópica = producida (por lo común, en mayor cantidad o de forma anormal) en el propio tejido donde se sintetiza normalmente esa hormona

26.26

26.9.2.1 Marcadores tumorales

Se da este nombre a cualquier sustancia en líquidos biológicos (en especial, plasma o suero) relacionada con la existencia o ausencia de un tumor. Su naturaleza molecular es muy heterogénea; las más comunes tienen función hormonal o enzimática, o son antigenicas. Su aparición puede deberse a alteraciones de la morfología celular, del citoesqueleto (microtúbulos, microfilamentos y filamentos intermedios), de la superficie de la membrana (antígenos asociados, pérdida de inhibición del crecimiento por contacto, receptores), de diversos parámetros bioquímicos o de la dotación cromosómica (cariotipo anormal, Capítulo 25). Son los marcadores que más se analizan desde un punto de vista general. Se pueden distinguir dos grupos de marcadores tumorales: los de secreción y las citocinas (o citoquinas).

a) Marcadores tumorales de secreción o marcadores "clásicos"

Pueden clasificarse en función de los factores (tisulares, metabólicos, fisiológicos, metodológicos, etc.) que determinan su concentración en líquidos biológicos, principalmente en la sangre. La lista de marcadores de secreción es amplísima y, a efectos prácticos, también pueden clasificarse en función del material biológico en el que se pueden estudiar: marcadores séricos, urinarios, de líquido cefalorraquídeo y de metástasis óseas, por ejemplo.

A continuación se describen algunos de los marcadores de secreción mejor caracterizados.

Fetoproteína α (AFP) o antígeno oncofetal**Cáncer de hígado**

Glicoproteína de cadena sencilla (591 aminoácidos), con estructura similar a la globulina α_1 , codificada por un gen del cromosoma 4. Se detecta a alta concentración en suero y líquido amniótico del feto normal de 7 a 9 semanas, donde se sintetiza en saco vitelino e hígado, con una función similar a la de la albúmina en adultos: mantener la presión oncótica y transportar sustancias. También pasa al suero materno, a través de la placenta. Es específica para el diagnóstico precoz del cáncer hepático y de tumores germinales que afectan a componentes del saco vitelino.

Subunidad β de la gonadotropina coriónica (β -hCG)**Cánceres de la línea germinal**

Glicoproteína expresada en la placenta durante el embarazo, formada por una cadena α (92 aminoácidos), codificada en el cromosoma 6, y otra β (145 aminoácidos), expresada por 7 genes del cromosoma 19. Su misión es estimular el cuerpo lúteo, una vez implantado el óvulo fecundado. La subunidad β libre se detecta en suero y, en ocasiones, en líquido cefalorraquídeo. Sirve para diagnosticar embarazo normal, gestaciones patológicas y ectópicas, cánceres testiculares de origen germinal y coriocarcinomas; aumenta también en el síndrome de Down. Se mide por métodos muy variados, principalmente inmunoenzimáticos, tanto cualitativos como cuantitativos.

Microglobulina β_2 (β_2 mc)**Linfomas y leucemias**

Polipéptido pequeño (99 aminoácidos) sintetizado en casi todos los tejidos, forma parte de un antígeno de superficie de todas las células nucleadas. Puede pasar en forma soluble a la circulación y se filtra en los túbulos renales. Se detecta en suero y en orina. Se eleva en síndromes proliferativos, como mielomas, linfomas y leucemias, y es en general de utilidad en todo proceso con disfunción renal tubular.

Antígeno carcinoembrionario (CEA)**Cáncer de colon**

Familia de glicoproteínas expresadas por genes del cromosoma 19q. Ejercen funciones en adhesión celular, reconocimiento de bacterias, iniciación temprana del tumor, capacidad de metástasis y regulación del efecto inmunológico. Se presenta en varios carcinomas y sus metástasis, así como en tejidos embrionarios (de ahí el nombre). Se determina en suero, como marcador heterogéneo, poco específico (asociado a muchos carcinomas); se usa más en clínica para tumores derivados del endodermo (colorrectales, estómago, pulmón y mama).

Antígeno carbohidrato 125 (CA 125)**Cáncer de mama**

Glicoproteína de elevado tamaño, reconocida por un anticuerpo monoclonal, presente en varios procesos no tumorales y tumorales (secretada en células epiteliales del ovario), y asociada inespecíficamente a tumores de mama, pulmón, páncreas y colorrectal. Es específica para el carcinoma no mucinoso ovárico, y un factor de pronóstico para el cáncer de pulmón no microcítico. Se utiliza como marcador de elección en tumores epiteliales de ovario (excepto los mucinosos).

Calcitonina**Carcinomas medulares de tiroides**

Hormona peptídica de tamaño pequeño (32 aminoácidos) segregada por las células C del tiroides (o parafoliculares). El aumento de su secreción produce un descenso de la concentración de calcio libre en plasma, debido a la inhibición de su resorción ósea y su reabsorción tubular. Es útil para el diagnóstico precoz y el seguimiento del cáncer medular de tiroides.

Fosfatasa ácida prostática (PAP o PSAP)**Cáncer de próstata**

Enzima específica de órgano, concretamente la isoenzima 2, implicada en el metabolismo de fosfolípidos y presente en el epitelio acinoso prostático. Se detecta en suero. Se utiliza para detectar y controlar a pacientes con cáncer de próstata. Su determinación se ha visto desplazada en gran medida por la del antígeno prostático específico.

Antígeno prostático específico (PSA)**Cáncer de próstata**

Glicoproteína de 34 kDa producida casi exclusivamente por la glándula prostática. Es una serinproteasa codificada en el cromosoma 19 (19q13). Se detecta en suero y es útil para detectar y controlar a pacientes con cáncer de próstata. Se eleva también en prostatitis, hiperplasia.

Antígeno asociado al carcinoma de células escamosas (SCC)**Cáncer de útero**

Glicoproteína aislada de carcinoma del cuello uterino, presente en suero y específica como marcador epidermoide.

Tiroglobulina (TGB)**Cáncer de tiroides**

Proteína rica en residuos de tirosina, muy abundante en el tiroides como precursora de síntesis y almacén de hormonas tiroideas. La sintetiza el tirocito o célula tiroidea, en varios grados de yodación. Se determina en sangre, estando aumentada en el bocio tóxico, la tiroiditis subaguda y la enfermedad de Basedow. Es específica para cánceres diferenciados como el cáncer papilar y del folículo tiroideo.

Queratinas CYFRA 21.1**Carcinoma escamoso de pulmón**

Queratina del tipo 19, presente en el citoesqueleto celular. Su valoración en suero es sensible a hepatopatías crónicas, insuficiencias renales y carcinomas escamosos de pulmón. Es también un indicador inespecífico de proliferación celular.

Otras enzimas**Cánceres diversos**

Isoenzimas de la creatinquinasa (CK): BB o CK1, en cerebro; MB o CK2, en músculo cardíaco, y MM o CK3, en músculo esquelético. La isoenzima BB aparece en neoplasias malignas de próstata, pulmón y sangre.

Isoenzimas IV y V de la lactato deshidrogenasa (LDH). Cáncer de próstata metastásico.

Ribonucleasa: cáncer de próstata.

Hexosa fosfato isomerasa: carcinoma hepático, de tubo digestivo o de pulmón.

5'-Nucleotidasa: cáncer de ovario y metástasis hepáticas.

Aspartato aminotransferasa y γ -glutamiltranspeptidasa: metástasis hepáticas.

26.27b

b) Citocinas o mediadores solubles de origen celular

De acuerdo con su acción específica, también se subdividen en: linfocinas, monocinas, interferones, factores de necrosis tumoral, factores estimulantes de colonias y factores de crecimiento. La concentración sérica de varias de ellas aumenta en determinadas situaciones clínicas, incluidos algunos cánceres.

26.9.2.2 Marcadores tisulares

Aunque son a veces indistinguibles de los marcadores tumorales, están más relacionados con la propia biología neoplásica y vienen dados por parámetros de definición imprecisa, no referidos a alteraciones bioquímicas. Su detección se basa, en general, en el empleo de anticuerpos monoclonales. Aunque tienden a usarse de forma creciente, no se ha llegado a un reconocimiento generalizado de su utilidad clínica.

Ejemplos:

Oncogenes	(pág. 454)
Genes oncosupresores	(pág. 456)
Receptores de hormonas esteroideas	Los más conocidos son los receptores citosólicos de estrógenos y de progesterona, que se ven alterados en los tumores malignos de mama.
Proteínas asociadas a los receptores de esteroideos	Algunas proteínas de choque térmico (muchas de ellas carabinas moleculares que facilitan el plegamiento adecuado de otras proteínas, pág. 372). Proteína pS2, inducida por estrógenos.
Componentes del citoesqueleto	Filamentos intermedios. Queratinas de tipo 8 (antígeno polipeptídico tisular, TPA). Queratinas de tipo 18 (antígeno polipeptídico específico tisular, TPS). Están presentes en el citoesqueleto de células epiteliales, normales y malignas, pero se determinan en líquidos biológicos (suero). En general, no son específicas de ningún tipo histológico, pero son útiles como indicador de proliferación y para el diagnóstico y seguimiento de muchos tumores.
Antígenos de membrana	Oligosacáridos Le ^x , sialil-Le ^x , sialil-Le ^a , sialil-Tn y sialosil-Tn. Glicoproteína P.
Factores de crecimiento	TGF α , FGF β , IGF-1 y TGF- β 1
Receptores de los factores de crecimiento	
Moléculas de adhesión	Integrinas, cadherinas y selectinas
Proteasas y sus inhibidores	

26.9.3 Marcadores genotípicos o moleculares

En la búsqueda del marcador tumoral “perfecto” se ha planteado una **clasificación molecular** de los tumores basada en la detección directa de los **cambios genotípicos** responsables del cáncer. En definitiva, ello supone un análisis molecular de las mutaciones en todo tipo de secuencias del genoma, tanto génicas (oncogenes, genes oncosupresores) como no génicas (secuencias repetidas) e incluso en las regiones teloméricas (pág. 96). Esto ofrece un elevado potencial para el diagnóstico (confirmación y seguimiento del desarrollo del tumor), la prevención (aparición precoz, cáncer asintomático, evaluación del riesgo y la susceptibilidad hereditaria) y la terapia (tratamiento farmacológico y su seguimiento). En particular, este análisis molecular de marcadores tumorales genéticos se ha hecho posible como consecuencia del desarrollo de los biochips (pág. 174). Concretamente, se han desarrollado ya varias micromatrizes (denominadas *oncochips*, págs. 428-429) en las que se fija una genoteca de cDNA cuyas secuencias se ha comprobado que guardan relación con diversos tipos de cáncer. Por ejemplo, como consecuencia del análisis de expresión diferencial (pág. 176) entre pacientes sanos y pacientes con cáncer, o entre muestras de dos subtipos diferentes de cáncer con un pronóstico clínico diferenciado.

Índice alfabético



Notas relativas a la ordenación alfabética de los términos:

Se han ignorado letras griegas, números y otros indicadores de posición del sustituyente (tales como N- y O-).

Por ejemplo, '2-acetamidofluoreno' y 'N-acetilglucosamina' aparecen en la 'a'.

Los ácidos se citan en la posición de su nombre, no en la 'a'. Por ejemplo, 'fusídico (ácido)'.

Las enfermedades específicas se citan por su nombre, no en la 'e'. Por ejemplo, 'Alzheimer (enfermedad de)'.

Los nombres de enzimas se citan completos, no como subentradas de la primera palabra. Por ejemplo, 'DNA polimerasa' no forma parte de la entrada 'DNA'. En general se emplea el nombre común y sin guiones intermedios, siguiendo el uso internacional en inglés y la recomendación de 1994 de la Sociedad Española de Bioquímica Clínica y Patología Molecular.

A

A1AT, 414. V. también *antitripsina α₁*.

aaRS, 327. V. también *aminoacil-tRNA sintetasa*.

abrazadera deslizante, 152

absorción óptica

variaciόn en la desnaturalización del DNA, 166

absortividad

de ácidos nucleicos, 167

de bases nitrogenadas, 167

de nucleósidos y nucleótidos, 167

ABTS, 189

acenocumarol, 359

acéntrico, 448. V. también *cromosoma*.

2-acetamidofluoreno, 398, 400

acetilación

de aminoácidos, 359

de histonas, 285

y expresión génica, 285

N-acetilgalactosamina, 364

N-acetilglucosamina, 364

N-acetylneuramílico (ácido), 364

β-N-acetylhexosaminidasa. V. *hexosaminidasa A*.

acetiltransferasas

de histonas, 285

N^α-acetiltransferasas, 359

aciclovir, 15

ácidos, efecto sobre la desnaturalización, 165

acilación

de aminoácidos en proteínas, 361

acortamiento de los extremos del DNA, 158, 159

acridina, 278, 398

naranja de, 132

actina

escisión por caspasas, 473

familia multigénica, 115

actinomicina D, 277

activación

de aminoácidos

energética, 328, 333

esquema global, 326

mecanismo, 328

reacción, 327

de grupos funcionales, 182, 183

activadores, 290

actividad transcripcional, 94, 281

ADA, 438. V. también *adenosina desaminasa*.

adaptador, 209

adelantada (hebra), 156, 157

adenilato, 21

adenina fosforribosiltransferasa, 246. V. también *APRT*.

S-adenosilhomocisteína, 214, 215, 283

S-adenosilmetionina, 19, 213, 214, 283, 301

adenosina, 18

como centro de ramificación, 305

adenosina desaminasa, 306

deficiencia de, 438

adenovirus, 471, 472

adhesión

celular, 129

entre célula y matriz extracelular, 449, 451

alteración en la apoptosis, 472

intercelular, 449, 451

alteración en la apoptosis, 472

A-DNA, 48

geometría de los pares de bases, 49

modelos moleculares, 49

ADP-ribosilación, 369

de histonas, 286

adriamicina, 278

afinidad

métodos, 129

para extracción de mRNA, 302

por la cola de poli(A), 302

aflatoxina, 398

AFP, 476

agar, 170

- agarosa, 141
 aglutinación, 410
 aisladores, 290
 alcohol isoamílico: en la extracción del DNA, 136
 aldolasa
 familia multigénica, 115
 alelos, 97, 99
 coheredados, 241
 ligados, 242
 segregados, 241
 y polimorfismo, 406
 alfoide, 117
 alquilación, 397
 reparación, 402
 alquilguanina, 400. V. también *O-metilguanina*.
 alquiltransferasa
 mecanismo de catálisis, 400
 alquiltransferasas, 400
 Alu, 118, 357
 Alzheimer (enfermedad de)
 relación con el plegamiento, 372
 α -amanitina, 270, 271
 estructura, 277
 aminas aromáticas, 398
 aminoácidos
 abreviaturas de 1 y 3 letras, 316
 acetilación, 359
 acilación, 361
 activación, 326
 básicos, 84
 carboxilación, 359
 fosforilación, 359
 hidroxilación, 360
 metilación, 361
 modificación en las proteínas, 358
 prenilación, 362
 unión por enlace peptídico, 339
 aminoácido-tRNA ligasa, 327. V. también
 aminoacil-tRNA sintetasa.
 aminoacil-2'-tRNA, 328
 aminoacil-3'-tRNA, 328
 aminoacilación del tRNA, 327
 aminoacil-AMP, 327, 328
 verificación, 330
 aminoacil-tRNA, 315
 ubicación en el sitio A, 337
 verificación, 330
 aminoacil-tRNA sintetasa, 327
 características, 329
 centros activos, 330
 corrección de errores, 330
 de clase I, 328, 329
 de clase II, 328, 329
 especificidad, 329
 reconocimiento de aminoácidos, 329, 330
 reconocimiento de tRNA, 329
 aminoglicósidos, 343, 344
 aminopterina
 en medio HAT, 246
 amnios, 122
 AMP cíclico, 25
 ampicilina
 gen de resistencia, 227, 228, 234
 amplificación genética, 191. V. también *clonación*.
 amplificadores, 290
 amp^r, 227, 228, 234
 anafase, 101
 anclaje
 de proteínas a membranas a través de
 lípidos, 361
 andamiaje nuclear, 88. V. también *esqueleto nuclear*.
 anemia de células falciformes
 detección por RFLP, 420
 herencia, 438
 mutación causante, 390
 polimorfismo, 420
 aneuploidía, 444
 mecanismo, 445, 447
 anfítriona, 230. V. también *célula*.
 angiogénesis, 452
 anomalías cromosómicas, 442
 antibióticos, 271. V. también *rifampicina*.
 de tipo I, 276
 de tipo II, 277
 de tipo III, 278
 puromicina, 19
 que inhiben la traducción, 343
 resistencia, 223, 271. V. también α -amanitina.
 anticodón, 66, 315
 emparejamiento con el codón, 319
 número necesario, 320
 relación con el número de codones, 321
 anticuerpos
 anti-digoxigenina, 189
 aplicación en ensayo Western, 178
 monoclonales, 129
 que existen naturalmente, 410
 antifólicos, 246
 antígeno, 470
 A, 411
 asociado al carcinoma de células escamosas, 475
 B, 411
 carbohidrato 125, 475
 carcinoembrionario, 475
 celular de células en proliferación, 152.
 V. también *PCNA*.
 H, 411
 oncofetal, 475
 prostático específico, 475
 T del virus SV40, 471, 472
 antioncogén, 393, 454, 456. V. también *oncosupresores*.
 antiparalelismo, 42
 antisentido, 74, 267
 α_1 -antitripsina, 372
 herencia, 438
 isoforma Z, 415
 mutaciones causantes, 414, 415
 α_1 -AT, 414. V. también *antitripsina* α_1 .
 antivirales, 15, 19
 Apaf, 471
 apilamiento, 44, 166
 de bases, 16

- apirimidínico, 252
 sitio, 395
 apoB100, 306
 apoB48, 306
 apobec-1, 306
 apolipoproteína B, 306
 apoptosis, 452
 concepto, 472
 control
 negativo, 454, 473
 positivo, 456, 473
 cuerpos apoptóticos, 474
 ejecución, 472
 etapas, 472, 474
 fagocitosis, 474
 fase final, 474
 inducción, 472
 iniciación por citocromo c, 471, 473
 papel de p53, 471
 relación con proliferación, 472
 señales
 negativas, 472
 positivas, 472
 APRT, 246
 apurínico, 252
 sitio, 395
 AraA o arabinosiladenina, 19
 AraC o arabinosilcitosina, 19
 árbol genealógico, 424
 nomenclatura y símbolos, 436
 Arf (proteína), 461, 466
 ARS. V. *secuencia de replicación autónoma*.
 ARS/CEN/TEL, 230
 arsénico, 398
 asa de desdoblamiento o de desplazamiento, 161.
 V. también *bucle D*.
 asbestos, 398
 ascorbato, 360
 ASO, 420. V. también *oligonucleótido específico de alelo*.
 aspartato aminotransferasa, 476
 ataxia telangiectasia, 448
 ATP como cosustrato de la ligasa, 218
 AttoPhos, 189
 autoiniciadora, 269
 automatización
 de la PCR, 205
 de la secuenciación enzimática de DNA, 255
 de la síntesis de oligonucleótidos, 183
 autoradiografía, 170, 189
 autosomas, 95
 Avery, MacLeod y McCarty, 4
 avidina, 129
 ayuste
 alternativo, 298, 310
 autoayuste, 72
 centros de, 303
 de proteínas, 371
 mecanismo, 305
 reconocimiento, 304
 secuencias, 303
 ayustosoma, 304
 azúcares
 activados con nucleótido, 367
 componentes de ácidos nucleicos, 14
 configuración anomérica, 17
 azul de tripano, 131
- B**
- BAC, 227, 230
 bacterias
 cápsula, 4
 colonias, 170, 235
 lisis *in situ*, 170
 bacteriófagos, 228
 λ, 228, 229
 M13, 228
 P1, 230
 T2, 4
 baculovirus, 228
 balanceo, 321, 329
 de guanina, 320
 de inosina, 320
 de uracilo, 319
 bandas cromosómicas G, Q, R y T, 93, 94.
 V. también *cromosomas*.
 bandeo o bandeados, 93
 de alta resolución, 95
 de cromosomas, 94
 de profase, 95
 de prometafase, 95
 barrera, métodos de, 125
 basal, 289. V. también *promotores*.
 bases, efecto sobre la desnaturalización, 165
 bases nitrogenadas
 absorción en el ultravioleta, 16, 167
 carácter ácido-base, 16
 complementariedad, 38
 componentes de ácidos nucleicos, 14
 derivados metilados, 15, 65, 215, 282
 desaminación, 394
 eliminación selectiva, 252
 estructura, 14
 infrecuentes, 15
 modelos moleculares, 16
 numeración de los átomos, 14
 propiedades
 ácido-base, 16
 fisicoquímicas, 16
 tautomería, 16
 vías de síntesis, 246
 Bax, 471
 bcl-1 (gen), 453
 bcl-2 (gen), 453
 Bcl-2 (proteína), 471
 Bcl-x (proteína), 471, 474
 B-DNA, 38, 47
 Bdp1, 295
 benzo[a]pireno, 398
 benzoiladenina, 183
 benzoilcitosina, 183
 BER. V. *escisión de bases*.
 berilio, 398

- biblioteca de DNA. V. *genoteca*.
 biciatenario, 40
 bidireccional
 replicación, 148
 biochips, 479, 172
 bioinformática, 418
 biolística, 231, 232
 biología molecular, 3
 biopsia, 120
 biotina, 189, 369
 bisfosfatos, 13, 25
 blastocele, 121
 blastocisto, 121, 194, 200
 blastodermo, 121
 blastómeros, 121
 Bloom (síndrome de), 448
 bloques de repetición. V. *DNA repetitivo*.
 bombardeo, 231
 Bombay (fenotipo), 413
 brazos
 en la estructura del tRNA, 66
 BRCA (genes), 473, 475
 BrdUrd o bromodesoxiuridina, 134
 BRE, 289
 Brf1, 295
 Brf2, 295
 bromocloropropano, 137
 bromodomino, 285
 bromouracilo, 398
 bromuro de etidioetidio, 141. V. también *etidio*.
 bucle, 62
 bucle D, 161, 276, 416
 burbuja, 62
 de transcripción, 272, 273
 desplazamiento, 274
 bZIP (proteínas), 59
- C**
- C₂H₂, 58. V. *dedo de zinc*.
 C₄, 58. V. *dedo de zinc*.
 CA 125, 475
 CAAT (secuencia o caja), 289, 292
 cabeza de martillo, 73
 cadmio, 398
 cafeína, 15
 caja
 A (promotor), 295
 B (promotor), 295
 C (promotor), 295
 CAK. V. *quinasa activadora de Cdk*.
 calcio
 cloruro, 231
 fosfato, 231
 calcitonina, 475
 calmodulina, 361
 calpaínas, 379
 cambio de α a δ o ϵ
 cAMP. V. *AMP cíclico*.
 cáncer
 angiogénesis, 450
 como enfermedad genética, 450
- concepto, 449
 de colon, 475
 de hígado, 475
 de la línea germinal, 474, 475
 de mama, 475
 hereditario, 473, 475
 de ovario, 473, 475
 de próstata, 475
 de tiroides, 475, 476
 de útero, 475
 diagnóstico, 429
 etapas, 451
 germinal y somático, 473, 475
 hepático, causado por deficiencia de
 antitripsina, 414
 hereditario, 473
 herencia, 475
in situ, 451
 invasivo, 451
 origen
 monoclonal, 451
 por mutación, 451
 vascularización, 452
 y metilación del DNA, 287
 y telomerasa, 161
 caótropos, 136
 CAP. V. *proteína receptora de AMP cíclico*.
 caperuza
 5', 300
 metilación, 301
 papel
 en el inicio de la traducción, 335
 protector, 301
 tipos 0, 1 y 2, 301
 de guanina. V. caperuza 5'.
 de guanosina. V. caperuza 5'.
 cápsida, 4
 cápsula bacteriana, 4
 carabinas, 372, 373
 actividad ATPasa, 374
 función, 376
 Hsp pequeñas, 376
 Hsp10, 375
 Hsp26, 376
 Hsp27, 376
 Hsp40, 374
 Hsp60, 375
 Hsp70, 374
 Hsp90, 375
 multiméricas, 375
 tipos, 374
 caracteres, 98
 β -carboxiaspartato, 359
 γ -carboxiglutamato, 359
 carboxilación
 de aminoácidos, 359
 carboxilasa, 359
 cariograma, 92. V. también *cariotipo*.
 de flujo, 247
 cariotipo
 análisis, 92

- cariotipo (*cont.*)
 molecular, 177
 preparación, 92
- cartografía, 240. V. también *mapa*.
 cascada
 de caspasas, 473
 de señales, 460
- caseína, fosforilación, 359
- caspasa
 activación en cascada, 473
 ejecutora, 473
 inhibición, 474
 iniciadora, 473
- casquete. V. *caperuza*.
- catepsinas, 377
- CBP, 335, 347
- Cdc25, 465, 467
- Cdk, 462
 activación, 467
 cíclica, 464, 468
 por quinasa, 465
 actividad quinasa, 465–467
 complejo con ciclina, 463–465, 467
 especificidad de sustrato, 464
 fosforilación, 465, 467
 de Rb, 471
 inhibición por quinasa, 465
 inhibidores, 466
 quinasa activadora de, 467
 quinasa inhibidora de, 467
 regulación, 466, 467
 del ciclo celular, 467
 por fosforilación, 464, 466
 síntesis, 465
 sustratos, 468
 tipos, 465
 unión con ciclinas, 464, 465, 467
- cdk1 (gen), 455
- Cdk1
 como oncoproteína, 455
- Cdk2, 469
 inhibición por p21, 471
- cDNA, 174, 207, 263
 como inserto, 222
 genotecas de, 238
 preparación, 222
 y EST, 248
- CEA, 475
- cebado al azar, 187
- cebadores, 148, 156, 269. V. también *primasa*.
 eliminación, 158
 empleo en PCR, 202
- ceguera a los colores, 440
- celadoras. V. *carabinas*.
- células
 competentes, 231
 en cultivo, como muestras, 123
 híbridos
 monocromosómicos, 246
 subcromosómicos, 246
 métodos de separación, 124
- recambio, 452
 vida media. V. *apoptosis*.
- células anfítrionas, 219
E. coli como prototipo de, 230
 tipos, 230
- células germinales, 99, 102, 196, 198, 200
 ciclo celular, 104
 primordiales, 104, 105
- células hospedadoras. V. *célula anfítriona*.
- células madre, 197. V. también *células troncales*.
- células somáticas, 99
 ciclo celular, 104
 híbridas muestras para preparar, 244
- células troncales
 de adulto, 198, 200
 EG, 198, 200
 embrionarias, 122, 196, 198, 200
 ES, 198, 200
 multipotentes, 198, 200
 obtención, 198
 PG, 198
 pluripotentes, 121, 196, 198, 200
 totipotentes, 191, 198, 199
- centimorgan, 243
- centrifugación
 de equilibrio de sedimentación, 126
 de velocidad de sedimentación, 126
 diferencial, 125, 135
 fundamento, 125
 isopícnica, 81, 117, 125, 126, 140, 147
 para elutriación, 126
 ultracentrifugación, 139
 zonal, 125, 126, 140
- centro
 de ayuste 3' y 5', 303
 de ramificación, 303
- centrómeros, 89, 90, 91, 95, 101
 en vectores de clonación, 230
 y DNA satélite, 116
- cesio, 81. V. también *centrifugación isopícnica*.
- CFTR. V. *regulador de la conductancia transmembrana*.
- cGMP. V. *GMP cíclico*.
- Chambord (castillo de), 44
- chaperonas. V. *carabinas*.
- Chargaff (reglas de), 36, 41
- chips de DNA. V. *biochips*.
- choque térmico, 374
- ciclinas, 464
 A, 464
 actividad quinasa, 466
 B, 464
 complejo con Cdk, 463, 464, 465, 467
 D, 461, 463
 D1, como oncoproteína, 455
 de G₁, 461, 463
 de G₂, 461, 463, 464
 de inicio. V. *ciclina de G₁*.
 degradación, 466, 468
 E, 461, 463
 F, 463
 mitótica. V. *ciclina de G₂*.

- ciclinas (*cont.*)
 síntesis, 464, 468
 tipos, 463
- ciclo
 de elongación en la traducción, 340
 del factor eIF-2, 336
 interferencia, 352
- ciclo celular, 99
 estímulos
 de detención, 469, 471
 de progresión, 468
 fases, 100, 104
 freno por p53, 471
 oncogenes relacionados, 469
 puntos de control, 462
 regulación, 462
 por Cdk, 467
 y condensación, 104
- ciclofosfamida, 398
- cicloheximida, 344
- cigoto, 104, 105
- cinetocoro, 95
- CIP1, 464
- CIP1 (proteína), 466
- ciprofloxacina, 278
- circularización, 209, 226
- cirrosis, 414
- cis, 355. V. también *Golgi*.
 elementos, 288
 factores, 288
 secuencias, 288
- cisplatino, 398, 400
- cisteína, en centro activo de la alquiltransferasa, 400
- cisteína endopeptidasas, 473
- cistrón, 324
- citidilato, 21
- citidina, 18
- citidina desaminasa, 306
- citocinas, 478, 476
- citocinesis, 101
- citocromo c, 361
 como señal activadora de apoptosis, 471
- citofluorímetro, 127
- citogenética, 90, 91, 95
- citómetro de flujo, 127, 247
- citoquinesis. V. *citocinesis*.
- citosina
 metilada, 282
- CKI. V. *Cdk, inhibidores*.
- clasificación de proteínas, 354.
 V. también *tráfico*.
- clon
 concepto, 191
 neoplásico, 449
- clonación
 introducción general, 191
 objetivos, 219
- clonación acelular. V. *PCR*.
- clonación celular, 193, 211
 esquema general, 219
 etapas, 219
- para obtención de sondas, 180
 preparación del inserto, 221
- clonación funcional, 418
- clonación terapéutica, 197
- clones, 235
 biblioteca de, 236
- cloranfenicol, 344
- clorobioicina, 278
- cloroformo: en la extracción de DNA, 136
- cloroplastos, 5, 355
- cloruro cálcico, 231
- cloruro de cesio, 81, 116. V. también
centrifugación isopícnica.
- cM. V. *centimorgan*.
- CMP-siálico, 367
- coactivadores, 294
- cobalamina
5'-desoxiadenosil-, 19
- cocarabinas, 374, 375
- codificación
 relación de, 314
 unidad de, 314
- codificante, 267
- código genético, 313
 degeneración, 318
 determinación experimental, 315
 diferencias, 321
 mitocondrial humano, 321
 representaciones
 en tabla, 316
 rectangulares, 316
- segundo código, 329
 universalidad, 321
- CODIS, 421
- codominante, 98, 436
- codón, 313, 315
 AUG, 317, 333
 de inicio, 317
 de paro. V. *codón de terminación*.
 de terminación, 317, 340
- emparejamiento con el anticodón, 319
- número
 necesario, 314
 por aminoácido, 318
- sin sentido. V. *codón de terminación*.
- sinónimos, 318
- solapantes, 314
 superpuestos, 314
 tamaño, 314
- coeficiente de sedimentación, 63, 76
- coenzima A, 25
- coenzimas
 nucleosídicas, 19
 nucleotídicas, 25
- cofactores generales, 294
- cohesivos
 telómeros, 158
 asociación de extremos, 217
- cohesivos (extremos), 215
- cola de poli(A), 139, 222, 223, 302
 efecto protector, 347

- colágeno, 451
hidroxilación, 360
colágenos, 363
colchicina, 93
colonias, 233
azules, método de, 235
colonias. V. bacterias y clones.
compactación, 79. V. también *condensación*.
del DNA, 79
compartimentos subcelulares, 354
compatibles, 216
competente, 231
complejidad del genoma, 111
complejo
abierto, 273
de empalme. V. *ayustosoma*.
de iniciación 80S, 335, 336
de inicio de la transcripción, 291
modelos, 291
de preiniciación 40S, 335
de reparación, 402
de translocación, 358
hierro-azufre, 350
proteinasa multicatalítico, 379
complementariedad, 38
imperfecta, 169
perfecta, 169
concepto, 454
condensación, 83, 94, 101
del DNA, 79
en la apoptosis, 472
grado de, 85
niveles, 85, 104
relación con la transcripción, 268
y acetilación de histonas, 285
y ciclo celular, 104
y regulación de la transcripción, 281
conformación
anti y *sin* de los nucleósidos en el DNA, 50
C2'-endo y C3'-endo, 49, 62
de la desoxirribosa en A- y B-DNA, 49
no superenrollada. V. *conformación relajada*.
relajada, 80
superenrollada, 81
congénitas, diagnóstico, 427
conservadora. V. *mutación por sustitución*.
constitutivo (gen), 292
constricciones secundarias, 90
cóntigos, 250
controles G₁/S, G₂/M y M, 461-463
cordocentesis, 122
corea de Huntington. V. *Huntington*.
corion, 122
corrección
de errores en la traducción, 330
de pruebas, 150
cortadoras infrecuentes, 237, 248, 250
corte
y empalme. V. *ayuste*.
y pegado. V. *ayuste*.
cortes tisulares, hibridación, 176
cos, 229
cósmidos, 229
cotranscripción, 272
cotransformación. V. *transformación*.
CPEO, 442
CPRG, 189
CPSF, 302
creatiquinasa, 476
cremalla de leucina, 59
Creutzfeld-Jacob (enfermedad de)
relación con el plegamiento, 372
cribado neonatal, 428
Cro (proteína del fago lambda), 58
cromátidas, 85, 90, 101
separación anómala, 448
y ciclo celular, 104
cromatina, 6, 84
actividad transcripcional, 89, 281
condensación, 101
descondensación, 101
extendida, cartografía por FISH, 247
remodelado, 281
cromo, 398
cromodominio, 285
cromosoma, 6, 84
acéntrico, 95, 448
acrocéntrico, 90, 91, 448
anomalías. V. *cromosomopatías*.
anular, 448
artificial, 227, 230
asas, 85, 88
bandas, 94
brazos, 90, 101
clasificación, 91
derivado, 448
dicéntrico, 448
hibridación, 177
homólogos, 96
interfásico, 85, 101
cartografía por FISH, 247
isocromosomas, 448
metacéntrico, 90, 91
metafásico, 85, 90, 101
cartografía por FISH, 247
morfología, 90, 91
número, 7, 8, 95
pares, 96
pintado de, 177
recombinantes, 241
reparto en la división celular, 107
rotura, 448
segregación, 106, 107
sexuales, 95, 96
submetacéntrico, 90, 91
tamaño, 90, 91
X, 95
inactivación, 89
por acetilación, 285
Y, 95
cromosomopatías, 442
adquiridas, 443

- cromosomopatías (*cont.*)
 concepto, 442
 constitutivas, 443
 estructurales, 447
 nomenclatura, 443
 numéricas, 444
 somáticas, 443
- CRP. V. proteína receptora de AMP cíclico.**
- cruciforme, 55
- CSB**, 274
- CTD**, 272
 asociación con la caperuza 5', 302
 en inicio y elongación de la transcripción, 273
 fosforilación, 273, 291
- CTF**, 292
- cuerpo**
 de empalme. V. *ayustosoma*.
 de inclusión, 373
 polar, 105
- cultivo de células anfítrionas para clonar, 233
- cumarol, 359
- cumermicina, 278
- curvas
 Cot, 118
 de desnaturalización, 167
 de fusión. V. *curvas de desnaturalización*.
- D**
- dactinomicina, 398
- daltonismo, 440
- DAPI**, 94, 176, 177
- daunomicina, 278
- daunorrubicina, 278
- dcc (gen), 456
- ddNTP. V. *didesoxinucleótidos*.
- DEAE-dextrano, 231
- dedo de zinc, 58
- dedos de zinc, 288
- defectos nutricionales, para la selección de clones recombinantes, 234
- degeneración
 del código genético, 318
 aplicación al diseño de sondas, 184
 e impacto sobre las mutaciones, 318, 319
- delección, 387, 388
 cromosómica intersticial, 448
 cromosómica terminal, 448
- densidad
 gradiente, 125
- DEPC. V. dietilpirocarbonato**.
- desacetilasas de histonas, 285
- desaminación, 394
 de bases nitrogenadas, 306
 de la citosina, reparación, 402
 reparación de, 402
- desarrollo embrionario, 310
- descondensación, 84, 85, 101
- desenrollamiento, 154, 268
 en la transcripción, 273, 274
- desequilibrado, 448
- desfosforilación, 359
 del extremo 5' del mRNA, 300
- deshidrogenasas, 133
 en ensayos de viabilidad, 133
- desmetilasas, 283
 en regulación de la expresión, 286
- desnaturalización
 agentes causantes, 164
 curvas de, 167
 del promotor, 273
 en la PCR, 202, 203
 local por la unión de proteínas, 57
 reversible. V. *renaturalización*.
 y contenido en G+C, 167
 y propiedades del DNA, 165
- desoxirribonucleasas. V. *DNAsas*.
- desoxirribosa, 14
- desplazamiento de la mella. V. *traslado de la mella*.
- despurinación, 395
- destino
 de proteínas, 354
- detección
 de clones recombinantes, 234
 directa, 185
 indirecta, 185
- dextrorso, 43
- diagnóstico prenatal, 121, 426
- diana, 168
 de restricción, 212
- dicéntrico. V. *cromosoma*.
- Dicer**, 309
- dictiosomas, 355
- didesoxinucleótidos, 253
- dietilestilbestrol, 398
- dietilpirocarbonato, 135
- diferenciación, 310
- diferencias entre genético y físico, 240
- difosfatos, 12
- difracción de rayos X
 para la estructura del DNA, 35
 para la estructura del tRNA, 67
- DIG. V. digoxigenina**.
- digestión
 doble, 248
 enzimática, 124
 parcial del DNA genómico, 237
- digoxigenina, 189
- dihidrouridina, 65
- dímeros de timina, 400
 reparación, 400
- dimetilnitrosamina, 397
- dimetilsulfato, 397
 para secuenciación, 252
- dimetoxitritilo, 183
- dinucleósido polifosfatos, 25
- dioxetano, 189
- dioxigenasa, 360
- diploidía, 7, 95, 102, 104, 444
- disenzimia, 434
- disgregación
 de tejidos, 135
 enzimática, 124
 química, 124

- disomía, 444
displasia, 451
distal. V. *promotores.*
distrofia muscular de Duchenne, 440
disyunción, 106
 defecto de, 445, 447
diversidad genética. V. *polimorfismo.*
división
 meiótica. V. *meiosis.*
 mitótica. V. *mitosis.*
 reductora. V. *meiosis.*
DMT, 183
DNA, 3
 agrupado, 110, 111
 altamente repetitivo, 115
 análisis de tamaño por electroforesis, 143
 basura, 69, 111
 carácter polianiónico, 84
 cartografía por FISH, 247
 chatarra, 111
 circular, 148
 codificador, 9, 110, 261, 264, 267
 comparación entre las formas A,
 B y Z, 47
 complementario. V. *cDNA.*
 composición. V. G + C.
 daño y proteína p53, 471
 de copia única, 9
 densidad, 140
 detección, 141
 disperso, 110, 111
 escalera, 141, 143, 474
 espaciador, 87, 281, 474
 esqueleto, 84
 fetal, 122
 forma
 A. V. *A-DNA.*
 B. V. *B-DNA.*
 mutante, 383
 silvestre, 383
 Z. V. *Z-DNA.*
 fragmentación
 en la apoptosis, 472, 474
 análisis electroforético, 472
 hipermetilado, 287
 hipometilado, 286
 iniciador, 157
 interacción con proteínas, 56
 intergénico, proporción en el genoma, 110
 internucleosómico, 474
 ligero, 147
 marcaje, 147
 masa en la célula, 7
 metabolismo, 145
 metilación, 282
 microsatélite, 118, 421
 minisatélite, 117, 421
 mitocondrial, 110, 276
 hebra
 ligera (L), 161
 pesada (H), 161
 marcadores, 424
 polimorfismo, 424
 utilidad, 416
 moderadamente repetitivo, 115
 molécula(s)
 progenitora, 145
 hijas, 145
 no codificador, 9, 110, 264, 267
 pesado, 147
 proporción de bases, 36
 reasociación. V. *renaturalización.*
 recombinante, 219
 reparación. V. *reparación del DNA.*
 repetitivo, 9, 110, 111
 agrupado, 111
 codificador, 112
 codificador disperso, 115
 disperso, 111
 no codificador, 115, 421
 satélite, 89, 116, 140, 421
 tinción, 141
 tipos en el genoma, 110
 variantes estructurales, 47
 velocidad de síntesis, 152
DNA glicosilasa. V. *N-glicosilasas de DNA.*
DNA ligasa. V. *ligasa.*
DNA metiltransferasa. V. *metilasa.*
DNA polimerasas, 85, 149, 157
 actividades enzimáticas, 151
 centros activos, 150
 del fago T7, 252
 empleo para el marcaje de sondas, 187
 eucarióticas, 151
 fragmento de Klenow, 150
 masa molecular, 152
 Pfu, 205, 208
 procarióticas, 150, 151
 propiedades, 151
 subunidades, 150, 152
 Taq, 202, 206, 252
 termoestable, 202
 tipos, 151
 Tth, 208
 ubicación subcelular, 151
 velocidad, 151
I, 150, 151
I, aplicación para la preparación de sondas, 187
I, en la síntesis de cDNA, 222, 223
I, en marcaje de sondas, 187
I, fragmento de Klenow, 187, 252
II, 150, 151
III, 150, 151
III, modelo molecular, 152
III, núcleo, 152
α, 151, 154, 156, 157
α, en la síntesis de telómeros, 160
β, 151
β, actividad fosfodiesterasa, 402
γ, 151
δ, 151, 157
ε, 151, 157
ζ, 151
η, 151

- DNA polimerasas (*cont.*)
 θ, 151
 υ, 151
 κ, 151
 λ, 151
 μ, 151
- DNApol. V. *DNA polimerasas*.
 DNAsas, 86, 212
 I pancreática, 187
 II, 212
- DNMT. V. *DNA metiltransferasa*.
 doble hélice
 apilamiento, 44
 carácter
 hidrófilo, 44
 hidrófobo, 44
 curvatura, 51
 de tipo A
 en el DNA. V. *A-DNA*.
 en el RNA, 62
 enlaces de hidrógeno, 44
 formación de triple hélice, 55
 fuerzas estabilizadoras, 44
 parámetros, 43
 posición de los fosfatos, 44
 relación con la replicación, 46
- dogma central de la genética molecular, 4, 323
 comparado en citosol y mitocondria, 261
- dolicol, 367
 fosfato, 368
 pirofosfato, 368
- dominante, 98, 436
- dot-blot*, 171
- Down (síndrome de), 447
- doxorrubicina, 278
- DPE, 289
- drepanocitosis, 420, 438
- Drosha, 309
- dRPasa. V. *desoxirribonucleasa*.
- DSE, 295
- DSIF, 274
- Duarte. V. *galactosemia*.
- Duchenne (distrofia muscular de), 440
- duplicación, 407
- E**
- E1A (proteína), 471
- E1B (proteína), 472
- E2F (proteína), 468, 470
- E6 (proteína), 472
- E7 (proteína), 471
- edad materna
 y riesgo de trisomía, 447
- edición del RNA, 305
- EDTA, 135
- Edwards (síndrome), 447
- eEF. V. *factores de elongación de la traducción*.
- EGF. V. *factor de crecimiento epidérmico*.
- eIF. V. *factores de iniciación de la traducción*.
- elastasa
 inhibición por antitripsina, 414
- elastina, 414, 451
- electroforesis, 171, 178
 calibración, 143
 en campo pulsante, 143, 248
 en gel con gradiente de temperatura, 413
 fundamento, 141
 preparación de geles, 141
 resolución, 141
 tipos de gel y tamaño de DNA, 141
- electroporación, 231, 232
- electrostática (interacción), 188
- elementos
 de respuesta. V. *promotores distales*.
 de respuesta a hierro, 350
 de respuesta a hormonas, 293
 específicos de regulación. V. *promotores distales*.
 móviles. V. *transposones*.
 nucleares dispersos
 cortos. V. *SINE*.
 largos. V. *LINE*.
 transponibles. V. *transposones*.
- ELL, 274
- elongación
 de la replicación, 156
 etapas, 157
 mecanismo, 157
- de la traducción, 337
 translocación, 339
 transpeptidación, 339
 ubicación, 337
- de la transcripción
 complejo de, 274
 factores, 274
 en la PCR, 203
- elonguina, 274
- elutriación centrífuga, 127
- embriogénesis
 metilación del DNA, 286
- empaqueamiento, 83
- emparejamiento
 codón-anticodón, 319
 de bases, 38. V. también *complementariedad*.
 entre 3 bases, 66
 intracatenario, 62, 66
 temporal, 269
 incorrecto, 403
- encefalomiopatía mitocondrial con acidosis láctica e
 incidentes similares a apoplejía, 442
- endocitosis, 355
- endomitosis, 96
- endonucleasas
 de restricción. V. *enzimas de restricción*.
 de tipos 'a' y 'b', 212
 en la reparación, 402
 FEN1. V. *FEN1*.
 generación del extremo 3' del mRNA, 302
- endosomas, 355
- enfermedad(es), 427, 448. V. también bajo el nombre
 específico
 asociadas a la reparación del DNA, 404
- autosómicas

- enfermedad(es) (*cont.*)
concepto, 434, 436
dominantes, 437
recesivas, 438
citogenéticas, 442. V. también cromosomopatías.
clasificación molecular, 432
complejas, 442
congénitas
diagnóstico, 426, 427
cromosómicas, 434, 442. V. también
cromosomopatías.
de Huntington. V. *Huntington*.
dominantes ligadas a X, 439
genéticas, 409, 417, 432
búsqueda del gen responsable, 425
cromosoma afectado, 434
productos génicos, 434
hereditarias, 384
búsqueda de genes por cartografía genética, 244
ligadas a Y, 441
ligadas al sexo, 439-441
concepto, 434, 436
mitocondriales, 441
concepto, 434
mixtas, 442
moleculares, 431
monogénicas, 434, 435
ejemplos, 435
herencia, 435
multifactoriales, 442
poligénicas, 442
recesivas ligadas a X, 440
enfisema, 414, 438
relación con el plegamiento, 372
enlace
de hidrógeno
de tipo Hoogsteen, 56, 66, 319, 320
de tipo Watson y Crick, 319, 320
disulfuro, 369
fosfoanhídrido, 12
fosfodiéster, 13, 149, 269
atípico 5'-2', 305
fosfoéster, 11
N-glicosídico, 17, 364
N-glicosídico, hidrólisis, 395, 402
O-glicosídico, 364
S-glicosídico, 364
isopeptídico, 378
peptídico, formación, 339
tioéster con ubicuitina, 377
topológico, 82
trifosfato 5'-5', 300
enrollamiento
en la transcripción, 274
enteroquinasa, 370
enucleación, 194
envejecimiento
y telómeros, 160
enzima
activadora de ubicuitina, 377
como marcador de sondas, 185
hidrolítica en lisosomas, 377
portadora de ubicuitina, 377
enzima de restricción, 211
aplicación a la cartografía, 248
características generales, 212
clasificación, 213
como mecanismo de defensa en las bacterias, 213
cortadoras infrecuentes. V. *cortadoras*.
de tipo I, 213
de tipo II, 213, 215
bacteria de origen, 216
ejemplos concretos, 216
nomenclatura, 216
secuencias diana, 216
de tipo III, 213
en la preparación de genotecas, 237
especificidad, 212
estructura, 213
frecuencia de reconocimiento, 248
nomenclatura, 212
para linealización de plásmido, 181
para síntesis de sondas, 187
sitios de reconocimiento, 212
epigenético, 282, 448
epilepsia mioclónica asociada a fibras rojas rasgadas, 442
equilibradas
hebras del DNA, 43
equilibrado, 448
equivalente genómico, 238
erbA (gen), 453
erbB (gen), 453
eRF. V. *factores de terminación de la traducción*.
eritromicina, 344
errores congénitos del metabolismo, 432
escalera doble helicoidal, 44
escalonados (extremos), 215
escinucleasa
estructura, 401
escisión
de bases nitrogenadas, 400
mecanismo, 402
de nucleótidos, 400
mecanismo, 400
escorbuto, 360
esparsomicina, 344
especies reactivas de oxígeno, 416
espectrofotometría, 189
espermatíidas, 104
espermatocitos, 104
espermatogénesis, 102, 104
espermátogonios, 102, 104
espermatozoides, 85, 104
espermidina, 31
espermina, 31
esqueleto nuclear, 85, 88
EST, mapa físico, 248
estabilidad del RNA, 347
estearato, 361
esterasas, en ensayos de viabilidad, 133
estimuladores, 290
estreptavidina, 129, 189

estreptolidigina, 276
 estreptomicina, 343
 estructura
 terciaria relación con la secuencia, 373
 tridimensional de proteínas, predicción, 372
 estructural. V. también *gen, secuencias.*
 secuencias génicas, 110
 estudios familiares, 424
 etanolamina, 367
 etidio, 141, 278
 etilmetanosulfonato, 397
 etiquetas, 357. V. también *péptido señal.*
 de secuencias expresadas. V. EST.
 eucariontes. V. *eucariotas.*
 eucariotas, 3
 diferencias con procariotas, 5
 eucromatina, 85, 88, 286
 y actividad transcripcional, 269
 euploidía, 444
 exclusión, ensayos de, 131
 exocitosis, 355
 exones, 299
 concepto, 110, 264, 299
 disposición en el gen, 264
 proporción en el genoma, 110
 exonucleasas, 212
 3' y 5', 150
 5' en marcaje de sondas, 187
 expresión
 diferencial, 174
 genes de, 293
 génica, 4
 basal, 293
 diferencial, 279
 niveles de regulación, 280
 regulación, 279
 postranscripcional, 309
 regulada, 293
 exteína, 371
 extremos
 3', modificación, 302
 5', modificación, 300
 cohesivos, 215, 217. V. también *cohesivos.*
 compatibles, 216
 romos, 215, 217. V. también *romos.*

F

FACS, 128, 247
 factor(es), 157
 activador de las proteasas apoptóticas, 473
 asociados a TBP, 291, 294, 295
 de coagulación
 carboxilación, 359
 VIII de la coagulación, 440
 de crecimiento
 derivado de plaquetas, 452, 454
 epidérmico, 453, 455
 β de crecimiento transformante, 457
 de terminación mitocondrial, 276
 estimulador de la poliadenilación
 por escisión, 302

inhibidores del crecimiento, 454, 456
 promotor de la maduración o de la fase M, 464
 factores de elongación de la traducción
 eEF-1A, 337
 regeneración, 337
 eEF-1B, 337
 eEF-2, 339
 factores de iniciación de la traducción, 334
 eIF-2, 335, 336
 actividad GTPasa, 336
 regeneración, 336
 eIF-2B, 336
 eIF-3, 335
 eIF-4A, 335
 eIF-4B, 335
 eIF-4C, 335
 eIF-4E, 335, 347
 eIF-5, 335, 336
 eIF-6, 335
 factores de terminación de la traducción
 eRF, 340
 factores de traducción, 334
 factores de transcripción, 56, 85, 272
 asociación en el complejo de inicio, 293
 coactivadores, 294
 cofactores generales, 294
 como oncoproteínas, 453
 como oncosupresores, 455
 de cadena arriba, 292
 de la elongación, 274
 definición, 288
 E2F, 468, 470
 función, 288
 generales, 291
 inducibles, 293
 interacción con la RNA polimerasa, 293
 mitocondrial, 276
 motivos estructurales, 288
 Myc, 468, 470
 p53, 471
 proximales, 292
 SREBP, 473
 TFI, 295
 definición, 288
 TFII, 290
 definición, 288
 TFIIA, 291
 TFIIB, 291
 TFIID, 291
 TFIIE, 273, 291
 TFIIF, 273, 291
 TFIIFH, 273, 291, 404
 como helicasa, 291
 como proteína quinasa, 291
 en la reparación, 401, 402
 subunidades, 402
 TFIIL, 295
 definición, 288
 TFIILA, 58, 295
 TFIILB, 295
 TFIILC, 295

- factores de transcripción (*cont.*)
 TFIIK, 291. V. también *TFIIC*.
 TFIIS, 274
 FAD, 25
 FADH₂
 coenzima de la fotoliasa, 399
 fagos, 228. V. también
bacteriófagos.
 λ, 228
 M13, 228
 falciforme, 390, 420, 438.
 V. también *anemia*.
 Falconi (anemia de), 448
 FAM, 256
 familias
 génicas, 112
 multigénicas, 112, 407
 fármacos, respuesta individual, 429
 farnesilo, 362
 farnesiltransferasa
 y terapia antitumoral, 362
 far-western, 179
 fase
 analítica, 119
 de reposo, 100
 G₀, G₁, G₂ y M, 100
 postanalítica, 119
 preanalítica, 119
 S, 100
 Fast Red TR, 189
 FCR V. fuerza centrífuga relativa.
 fecundación, 104, 108
 FEN1, 158, 413
 en la reparación, 402
 en la síntesis de telómeros, 160
 fenilalanina-4-monoxigenasa, 438
 fenilcetonuria, 438
 fenol, en la extracción de DNA, 136
 fenotipo, 98
 fenoxazona, 277
 ferritina, 350
 feto
 DNA, 122
 sangre, 122
 fetoproteína a, 477
 fetoscopia, 122
 fibra
 básica de cromatina, 87. V. también fibra
 de 10 nm.
 de 10 nm, 85, 87
 y control pretranscripcional, 281
 y transcripción, 269
 de 30 nm, 85, 88
 fibrilina, 437
 fibrosis quística
 herencia, 438
 mutación causante, 392
 relación con el plegamiento, 372
 Ficoll, 125
 FISH, 95, 176, 177. V. también *hibridación in situ con fluorescencia*.
- cromosómica, 246
 de alta resolución, 247
 fitohemaglutinina, 93
 fluctuación. V. *balanceo*.
 fluoresceína, 186, 256
 diacetato de carboxifluoresceína, 133
 fluorimetría, 189
 fundamento, 127
 fluorocromo, 177, 127, 189, 206, 247, 252
 para el marcaje de sondas, 185, 186
 para secuenciación del DNA, 256
 fluoróforo, 127
 fMet. V. *N-formilmetionina*.
 FMN, 25
 formaldehído, 165
 formamida, 165, 169
 formamidopirimidina-DNA glicosilasa, 403
 formazano, 133
 formiato, para secuenciación, 252
 formiato-tetrahidrofolato ligasa, 334
 fórmico (ácido), para secuenciación, 252
 formilación de metionina, 334
N-formilmetionina, 334
 formiltetrahidrofolato, 334
 fos (gen), 453
 fosfatasa ácida prostática, 475
 fosfatasa alcalina, 187, 189, 226
 isoenzimas por duplicación génica, 407
 fosfatasa inorgánica, 328
 fosfato cálcico, 231
 fosfato inorgánico (P_i), 11
 fosfato α en NTP. V. *nucleofílico, marcaje*.
 fosfato γ en NTP. V. *marcaje*.
 fosfodiesterasa
 I, de veneno de serpiente, 212
 II, de bazo bovino, 212
 fosfoproteína fosfatasa, 462
 fosforamidita, 183
 fosforilación, 462
 como mecanismo regulador, 359
 de aminoácidos en las proteínas, 359
 de eIF-2, 352
 de histonas, 286
 del CTD, 273
 del dominio CTD, 272
 reversible, 359, 464
 fosforilasa fosfatasa, 359
 fosforilasa quinasa, 359
 fosfoserina/fosfotreonina fosfatasa, 359, 464
 fosfotirosina fosfatasa, 359, 464
 fotodímeros, reparación, 399
 fotoliasa, 399
 fotoproductos, 398
 fracción de recombinación. V. *frecuencia*.
 fracciones subcelulares, 135
 fragmentos de restricción, 215
 frecuencia de recombinación, 242
 interpretación, 243
 FRET, 206
 fucosa, 364
 fucosiltransferasa, 411

fuerza centrífuga relativa, 125

funiculocentesis, 122

fusídico (ácido), 344

fusión

celular, 244

del DNA, 167

G

G+C (contenido), 38, 139, 167

galactosa, 364

galactosa-1-fosfato uridiltransferasa, 413

galactosaminiltransferasa, 411

galactosemia

bioquímica, 413

descripción, 413

genética, 413

genotipos y fenotipos, 413

localización cromosómica, 413

mutaciones causantes, 414

β -galactosidasa, 189, 228, 234

galactosiltransferasa, 411

GALT. V. *galactosemia*.

gametogénesis, 102

gametos, 97, 99, 102

ganciclovir, 15

GC (caja), 292

GC (secuencia o caja), 289

GDP-fucosa, 367

GDP-manosa, 367

gen

AB0, 410

análisis, 417

búsqueda. V. *genes, análisis*.

comparado con cistrones, 325

con varios productos génicos, 265, 298

constitutivos, 292

de clase I, definición, 288

de clase II, definición, 288

de clase III, definición, 288

de expresión diferencial, 293

de mantenimiento, 292

de resistencia a antibiótico, 223

definición, 263

actual, 265

casos límite, 265

matizaciones, 264, 298

eucariótico

componentes, 264

promotores, 264

secuencias

estructurales, 264

reguladoras, 264

evolución del concepto, 263, 265

expresión. V. *expresión génica*.

H, 410.

fragmentos de, 114

y transcripción, 267

inducibles, 290, 293

marcadores, 223, 234

monomorfos, 409

nomenclatura, 266

oncosupresor, 454. V. también *oncosupresor*.

NF1, 393

organización, 299

polimorfos, 409

regiones

no traducidas 5' y 3', 264, 298, 299

no traducidas internas. V. *intrones*.

traducida, 264

transcrita, 264

transcritas eliminadas antes de la traducción, 264

regulables, 293

secuencias

estructurales, 110

reguladoras, 110

terminadoras, 264

solapantes, 265

supresor de tumores, 393, 454, 456. V. también *oncosupresor*.

truncados, 114

variantes, 114, 409

genealogía, 424

genético, 240

genochips. V. *biochips*.

genoma, 3, 5

codificante, 267

diversidad. V. *polimorfismo*.

extranuclear, 6

fracción transcrita, 267

nuclear, 6

tamaño, 6, 8

tipos de DNA, 110

genoma humano

datos, 6

número

de cromosomas, 6

de genes, 9

tamaño, 6

genoma mitocondrial

características, 10

magnitud, 10

genómica, 239, 432

estructural, 417

genosensores. V. *biochips*.

genoteca, 236

de cDNA, 238

en la búsqueda de genes, 417

genómicas, 237

número de clones, 238

genotipo, 98

gentamicina, 344

geometría de los pares de bases

comparada en A- y B-DNA, 49

geranilgeranilo, 362

geranilo, 362

Giemsa (tinción), 94

girasa, 156

inhibición por antibióticos, 278

glicanos. V. *oligosacáridos*.

N-glicanos

estructura central, 365

reacciones de biosíntesis, 368

- N-glicanos (*cont.*)
 ricos en manosa, complejos e híbridos, 366
 tipo complejo, 366
 tipo híbrido, 366
- O-glicanos, 364
 glicolípido
 en la bioíntesis de N-glicanos, 368
 glicoproteínas. V. *glicosilación*.
 glicosaminoglicanos, 363
 glicosilación
 en N-, 365
 en O-, 364
 localización subcelular, 367
 propiedades que induce en la proteína, 363
 tipos, 364
 ubicación subcelular, 367
- glicosilasa, 283
 N-glicosilasas de DNA, 403
 glicosilfosfatidilinositol, 367
 glicosiltransferasas, 367
 A y B, 411
 de grupo sanguíneo AB0, 410
- globina
 expresión en el desarrollo animal, 115
 familia multigénica, 114
 genes de, 114
 herencia, 438
 origen evolutivo, 114
 polimorfismo, 420
 S, 420
 tipos, 115
 velocidad de traducción, 352
 y talasemia, 438
 β^S , 390
- glóbulo fundido, 373
 glucógeno fosforilasa, 359
 glucosa, 364
 γ -glutamiltranspeptidasa, 476
 glutaraldehído, 188
 glutatión, 369
 GMP cíclico, 25
 Golgi (complejo o aparato de), 355, 368
 gonadotropina coriónica, 475
 GPI, anclaje, 367
 Gppp, 300
 gradiente de densidad, 125, 140, 147
 gránulos de secreción, 371
 Griffith, 4
 gRNA. V. *RNA guía*.
 GroEL, 375
 GroES, 375
 Grunberg-Manago y Ochoa, 269
 grupo(s)
 sanguíneo AB0
 anticuerpos, 410
 antígenos, 410
 bioquímica, 411
 ensayo serológico, 410
 fenotipos, 410, 412
- genética, 410
 genotipos, 412
 localización cromosómica, 411
 puntos polimórficos, 411
 resumen, 412
 prostéticos, 369
- GTPasa
 proteína Ras, 461
 guanidina, 136, 137
 guanilato, 21
 guanosilación
 del extremo 5' del mRNA, 300
 guanosina, 18
 guía (hebra). V. *adelantada*.
 Guthrie (tarjeta), 122
- H**
- haplogrupo, 416
 haploide, 7, 95, 102, 104
 haploidía, 444
 haplotipo, 242
 HAT (medio de cultivo)
 composición, 246
 fundamento para la selección de células, 246
 HAT. V. *histona acetiltransferasas*.
 hCG, 475
 HCI, 352
 HDAC. V. *histona desacetilasas*.
 H-DNA, 55
 hebra
 adelantada, 156, 157
 antisentido, 267
 codificador, 267
 con sentido, 267
 H. V. *DNA mitocondrial*.
 informativa, 267
 L. V. *DNA mitocondrial*.
 molde, 267. V. también *molde*.
 negativa, 267
 no codificador, 267
 no informativa, 267
 no molde, 267
 no transcrita, 267
 positiva, 267
 transcrita, 267
 helicasa, 153, 154
 en la transcripción, 273
 TFIIC, 291
 hélice de reconocimiento, 57
 hélice-bucle-hélice, 58
 hélice-giro-hélice, 57
 hélice-vuelta-hélice, 57
 helicidad del DNA, 43
 hemicigótico, 439
 hemo, 369
 como regulador de la traducción, 352
 hemofilia A, 440
 hemoglobina
 de células falciformes, 420. V. también *hemoglobina S*.
 glucosilada, 363
 S, 390, 420

- hemoglobinopatías, 390
 herencia, 4, 436. V. también *árbol genealógico*.
 de las mutaciones, 383
 del cáncer, 473, 475
 materna, 441
 patrón
 mendeliano, en enfermedades monogénicas, 435
 no mendeliano, 441
 Hershey y Chase, 4
 heterocarionte, 244
 heterocigosis
 y polimorfismo, 406
 heterocigótico, 98
 heterocromatina, 85, 88, 90, 116, 286
 constitutiva, 89, 95
 facultativa, 89
 heterodúplex, 168
 heterogamético, 97
 heteroplasmia, 441
 heteroploidía, 444
 hexosa fosfato isomerasa, 476
 hexosaminidasa A
 herencia, 438
 mutación, 391
 HGPRT
 deficiencia, 246
 hibridación, 163
 análisis molecular, 168
 de DNA, 171
 de RNA, 172
 de Southern. V. *Southern*.
 dot-blot, 171
 en fase líquida, 169
 en la PCR, 202, 203
 en soporte sólido, 170., 172. V. también
 sobre *biochips*.
 específica de alelo, 424
 factores que afectan, 168
 fundamento, 168
 in situ, 175
 con fluorescencia, 95, 177. V. también *FISH*.
 métodos, 169
 Northern. V. *Northern*.
 para la detección de clones recombinantes, 234
 rigor, 168
 slot-blot, 171
 sobre biochips, 172
 sobre cortes tisulares, 176
 sobre preparaciones cromosómicas, 177
 híbridos, 222, 223, 275
 DNA-DNA, 163, 168
 DNA-RNA, 163, 267
 mRNA-cDNA, 222, 223
 RNA-cDNA, 207
 RNA-DNA, 62, 274
 desestabilización en la terminación
 de transcripción, 275
 RNA-RNA, 163
 hidrato de pirimidina-DNA glicosilasa, 403
 hidrazina para secuenciación, 252
 hidrocarburos aromáticos policíclicos, 398
 hidrofobia
 de las bases nitrogenadas, 16
 efecto hidrófobo, 16, 373
 influencia en el plegamiento, 373
 hidrólisis
 alcalina del RNA, mecanismo, 32
 química de ácidos nucleicos, 31, 395
 hidroxiapatito, 169
 hidróxido (ion), 396
 8-hidroxiguanoína, 396, 398
 hidroxilación
 de aminoácidos, 360
 hidroxilasas, 360
 hidroxilo, 398
 radical, 396
 hidroxiprolina, 360
 hierro
 captación, 351
 reserva, 350
 transporte, 351
 hipercolesterolemia familiar, 174, 437
 diagnóstico, 428
 hipercrómico, 166
 hiperplasia, 451
 hipervariables. V. *regiones, minisatélites*.
 hipocrómico, 166
 hipótesis
 un gen - un polipéptido, 263
 un gen - un producto génico, 265
 un gen - un RNA, 265
 un gen - una enzima, 263
 un gen - una proteína, 263
 hipoxantina, 19
 como resultado de la desaminación, 395
 en medio HAT, 246
 hipoxantina/guanina fosforribosiltransferasa. V. *HGPRT*.
 histona, 31, 87
 acetilación, 284, 359
 ADP-ribosilación, 286
 fosforilación, 286
 genes de, 112
 H1, 84, 88
 H2a, 84, 88
 H2b, 84, 88
 H3, 84, 88
 H4, 84, 88
 acetilación, 359
 metilación, 361
 H5, 84, 88
 metilación, 285
 octámero, 88
 reducción de la carga positiva, 284
 sumoilación, 286
 tipos, 84
 ubiquitinación, 286
 histona acetiltransferasas, 285
 histona desacetilasas, 285
 histona desmetilasas, 285
 histona metiltransferasas, 285
 HLH. V. *hélice-bucle-hélice*.
 HMG (proteínas), 85, 87

- HMT. V. *histona metiltransferasas*.
HNPP, 189
hnRNA, 63, 270, 298
hnRNP, 310
Hoechst, 94
Hogness (caja de). V. TATA.
holoenzima. V. *modelo*.
homeodomino, 58, 288
homocarionte, 244
homocigótico, 98
homodúplex, 168
homogamético, 97
homogeneización, 135
homología de secuencia e hibridación, 168
Hoogsteen. V. *enlaces de hidrógeno de tipo Hoogsteen*.
HOP, 376
hormonas
 esteroides, receptores de, 293
 tiroideas, receptores de, 293
horquilla, 54, 62, 157
 de replicación, 148, 154, 156. V. también *retardada*.
 de replicación C, 157
 de RNA, 350
hospedadora. V. *célula anfitriona*.
HRE. V. *elementos de respuesta a hormonas*.
Hsp, 374. V. también *proteínas de choque térmico*.
Hsp pequeñas, 376
Hsp10, 375
Hsp26, 376
Hsp27, 376
Hsp40, 374
Hsp60, 375
Hsp70, 374
Hsp90, 375
hSPT4, 274
hSPT5, 274, 301
HTH. V. *hélice-giro-hélice*.
hTR. V. *telomerasa, componentes*.
huella
 dactilar genética, 416, 425
 de DNA
 y DNA minisatélite, 118
 genética, 416
humo, 400
 productos mutágenos, 398
Huntington (enfermedad de), 437
huso mitótico, 101
- ICR, 295
identidad, pruebas de, 425
idiograma, 92
IF. V. *factores de iniciación de la traducción*.
impronta genómica, 286
inactivación insercional, 234
incidencia, 433
indicadores, 185, 188
 tipos, 189
individualidad genética, 406
inducible
 gen, 293
información genética, 3, 261
informativa, 267
ingeniería genética, 3, 211
inhibidor(es)
 controlado por hemo, 352
 promotores, elementos o secuencias, 290
iniciación
 aleatorias, 187
 de la traducción, energética, 336
 de la transcripción, modelos, 272
inicio (punto de control del ciclo celular), 461
inicio
 de la replicación, 153
 de la traducción, 333
 especificidad, 333
 etapas, 334
 por el extremo N, 324
 de la transcripción, 272
 como punto de control, 287
 complejo de, 272, 291
 determinación del punto de inicio, 291
 modelos, 273
 de la holoenzima, 292
 escalonado, 291
 transición a elongación, 273, 291
 proteínas de, 153
INK4, 466
inmunoanálisis, 129
inmunofilina, 376
inmunoglobulina M
 gen y ayuste alternativo, 310
inosina, 19, 65
 en el anticodón, 320
inositol, 367
Inr, 272, 289
insecto, 228
inserción, 387, 389
 método de preparación de virus recombinantes, 229
inserto, 219
 de cDNA, 222
 tamaño, 221
insulina, 370
integrinas, 451
inteína, 371
intensificadores, 290
intercalantes, 94, 142, 188, 247, 277, 398
interfase, 99, 101
 y transcripción, 269
interferón, 344
intergénico, 110
intrágénico, 110
intrón, 299
 concepto, 72, 110, 264, 298
 disposición en el gen, 264
 eliminación por autocatálisis, 72
 en forma de lazo, 304
 en pre-tRNA, 308
 magnitud, 303
 número, 303
 polimorfismo, 409
 proporción en el genoma, 110

intrón (*cont.*)
 tipos I y II, 72, 305
 tipos III y IV, 305

invasividad, 451

inversión
 cromosómica, 448
 paracéntrica, 448
 pericéntrica, 448

IRE, 350

IRE-BP, 350

IRP, 350

islotes CpG, 167, 215
 en la búsqueda de genes, 417
 metilación, 282

isobutirilguanina, 183

isocoros, 94

isocromosomas, 448

isoenzimas, 407

isoesquizómeros, 216

isoformas, 407

isoleucil-tRNA sintetasa
 corrección de errores, 331

isopeptidasa, 379

isopeptídico, 378

isopícnica. V. *centrifugación*.

isótopos radiactivos, 189, 252
 como marcadores de sondas, 185
 para estudiar la traducción, 324

izozimas, 407

IVS. V. *intrón*.

J

JOE, 256

jun (gen), 453

K

kanamicina, 344

Kaposi, 404. V. también *xerodermia*.

Kearns-Sayre (síndrome), 442

Khorana, 182

KIP, 464, 466

kirromicina, 344

Klenow (fragmento de), 150. V. también *DNA polimerasa I*.
 para la síntesis de sondas, 187

Klinefelter (síndrome), 447

Kozak (secuencia de), 333

Kpn, 118

KSS, 442

L

L. V. *número de enlace*.

L1. V. *LINE*.

L-19 IVS, 72

lactato deshidrogenasa, 476

lacZ, 227, 228, 234

LA-PCR, 208

lazo. V. *intrón*.

LCR, 114

Leber (neuropatía), 442

leucemia, 475

leucina. V. *cremallera*.

LHON, 442

Li-Fraumeni (síndrome), 471

ligamiento, 242

en la búsqueda de genes, 417

ligasas, 158, 217, 218

en la síntesis de telómeros, 160

mecanismo de reacción, 218

para formar el rDNA, 225, 226

LINE, 118

LINE-1, 408

linfoma, 475

lipofección, 231

liposomas, 231

líquido

amniótico, 121

articular, 124

ascítico, 123

cefalorraquídeo, 123

pericárdico, 123

peritoneal, 123

pleural, 123

seroso, 123

sinovial, 124

lisis celular, 134

lisosomas, 355

en la degradación de proteínas, 377

primarios, 377

Lk. V. *número de enlace*.

lobucavir, 15

locus/loci, 97

distancia entre loci, 242

Los Ángeles. V. *galactosemia*.

L-PCR, 208

luciferasa, 189

luciferinas, 189

lumen. V. *retículo endoplásmico, Golgi*.

luminol, 189

luz visible y reparación del DNA, 399

Lymphoprep, 125

M

maduración

de proteínas, 358

de mRNA, 300

de rRNA, 307

de rRNA y tRNA, 72

postranscripcional, 261, 262, 264, 297

diferencias entre procariotas y eucariotas, 298

en la mitocondria, 309

regulación, 309

magnético

microesferas, 129

manosa, 364

MAP quinasas, 461

mapas

de cíntigos, 250

de ligamiento. V. *mapa genético*.

de restricción, 248

de SNP, 424. V. también *polimorfismo SNP*.

físico

de alta resolución, 247

de baja resolución, 244

- mapas (*cont.*)
 de EST, 248
 de restricción, 248
 de STS, 248
 empleo de Southern, 248
 interpretación, 240, 248
 resolución, 248
 genético, 240
 aplicación para localizar un gen, 244
 obtención, 243
 genómico, 240
 de cónigos, 250
 marca de secuencia expresada. V. *EST*.
 marcadores, 185., 240., 247. V. también *mapa*.
 de DNA mitocondrial, 424
 EST. V. *EST*.
 genético, 240
 génicos, 248
 no génicos, 240, 248
 no radiactivos, 189
 polimórficos, 243
 STR, 421
 STS. V. *STS*.
 tipos, 189
 tumorales
 clasificación, 473, 476
 de secreción, 476
 VNTR, 422
- marcaje
 con azufre radiactivo, 4
 con cuatro fluorocromos, 255
 con DNA polimerasa, 187
 con fósforo radiactivo, 4
 con indicador, 188
 con polinucleótido quinasa, 187
 de NTP en fosfato α , 187
 de NTP en fosfato γ , 187
 de sondas no nucleotídico, 188
 detección, 189
 directo, 186
 en cultivo, 185
in vitro, 185
 indirecto, 188
 para la proteólisis, 378
 por cebado al azar, 187
 por traslado de la mella, 187
 terminal, 187, 252
 por relleno, 187
 marco de lectura, 314, 391
 desplazamiento, 315
 marco de lectura abierto, 318
 en la búsqueda de genes, 417
 Marfan (síndrome de), 437
 material genético
 extranuclear, 5
 nuclear, 5
 matriz
 extracelular, 124, 449, 451
 nuclear, 85, 88. V. también *esqueleto nuclear*.
 Max (proteína de ratón), 58
 Maxam y Gilbert, 252
- May-Grünwald-Giemsa (tinción de), 88
 β_2 mc, 475
 MCS. V. *sitio de clonación múltiple*.
 mdm-2 (gen), 453
 MeCP1, 284
 MeCP2, 284
 mediador, 294
 medicina predictiva, 428
 médula ósea
 muestras, 122
 punción, 122
 meiosis, 99, 101, 104, 107
 comparación con mitosis, 106
 no-disyunción, 445, 447
 MEK, 461
 MELAS, 442
 mella, 187, 217, 218. V. también *marcaje*.
 Mendel
 primera ley, 106
 segunda ley, 107
 MERRF, 442
 Meselson y Stahl, 147
 metabolismo oxidativo, 396, 416
 metafase, 101
 metaloproteasas de matriz extracelular, 449
 metástasis, 449, 451, 452
 metil lisina, 361
 metilación
 de aminoácidos, 361
 de histonas, 285
 de la caperuza 5', 301
 de mantenimiento, 286
 de novo, 286
 del DNA, 282
 en el desarrollo embrionario, 286
 regulación de la expresión génica, 283
 y cáncer, 287
 metilación-restricción, 213
 metiladenina-DNA glicosilasa, 403
 N-metil-4-aminoazobenceno, 398
 N-metil-4-antraceno, 398
 metilasas, 283, 301
 asociadas a las enzimas de restricción, 213
 O⁶-metilguanina
 reparación de, 400
 metilnitrosoguanidina, 397
 metilnitrosourea, 397
 metiltransferasas, 361
 metionil-tRNA
 iniciador, 334
 metionil-tRNA-formiltransferasa, 334
 metionina
 incorporación interna y N-terminal, 334
 métodos de introducción del rDNA, 231, 232
 metotrexato, 246
 metrizamida, 125
 MFP, 464
 microarreglo de DNA, 172. V. también *biochips*.
 microesferas magnéticas, 129, 138
 microglobulina β 2, 475
 microinyección, 194, 231, 232

- micromatriz de DNA, 172. V. también *biochips*.
 para el diagnóstico, 428
- microRNA, 63
- microsatélites, 118. V. también *DNA microsatélite*.
 aplicación a la huella genética, 425
- microscopía, 130, 189
 de fluorescencia, 94, 177
- migración de células tumorales, 451
- minisatélites, 117. V. también *DNA minisatélite*.
 aplicación a la huella genética, 425
 hipervariables, 421
- mioinositol, 367
- miristato, 361
- miristilación de proteínas, 362
- miRNA, 63. V. también *microRNA*.
 biosíntesis, 309
- mitocondrias, 5, 10, 355
 división, 161
 expresión génica, 261
 generación de mutágenos, 396
 rotura en la apoptosis, 472
- mitógenos, 455
- mitosis, 99-101, 104
 comparación con meiosis, 106
 regulación, 463, 464
 señales de inicio, 468
 sin citocinesis, 445
- mixoploidía, 444, 447
- MMP. V. *metaloproteasas de matriz extracelular*.
- modelo
 de la holoenzima para el inicio de la transcripción, 292, 273, 292
 escalonado para el inicio de la transcripción, 273, 291
- modificación
 de aminoácidos, 369
 en proteínas, 358
- postraduccional, 353, 358. V. también *procesamiento postraduccional*.
- módulos de control. V. *promotores distales*.
- molde, 148, 261, 269
- moléculas señal, 459
- mongolismo, 447
- monocistrónico, 324
- monofocal, 148
- monomorfos, 409
- monosomía, 444
 mecanismo, 447
- mórula, 121
- mosaico, 385, 447
- mostazas nitrogenadas, 397
- motivos estructurales
 en proteínas ligantes de DNA, 57, 288
- mRNA, 63
 bloqueo, 350
 como muestra
 para clonación, 222
 para PCR, 208
- degradación, 347
- estabilidad, 347
- extracción por afinidad, 139
- maduración, 300
- polimerasa que los transcribe, 270
 precursores de, 270
- protección frente a la degradación, 351
 transporte, regulación, 346
- mRNA(guanina-N7-) metiltransferasa, 301
- mRNA(nucleósido-2'-O) metiltransferasa, 301
- MTE, 289
- mTERF, 276
- MTS (proteínas), 464, 466
- mtTFA, 276
- muerte celular
 equilibrio con proliferación, 452
 programada. V. *apoptosis*.
- muesca. V. *mella*.
- muestras
 bucales, 122
 de DNA mitocondrial, 416
 disociación, 124
 tipos, 120
- multifocal, 148
- mupiroicina, 343
- mutaciones, 145
 a gran escala, 386
 a pequeña escala, 386
 asintomática. V. *silenciosa*.
 clasificación, 384
 con cambio de marco, 391
 con desfase o desplazamiento de marco.
 V. *con cambio de marco*.
 con sentido. V. *silenciosa*.
 con terminación prematura. V. *mutación sin sentido*.
 con terminación retrasada, 394
- conservadora, 409
 de aminoácido, 390
 de sentido alterado, 390
 de sentido equivocado o falso. V. *de sentido alterado*.
 de terminación prematura, 392
 efecto beneficioso o perjudicial, 390
 en DNA mitocondrial, 416
 endógenas, 394
 espontáneas, 394
 exógenas, 396
 frecuencia, 386
 génica, 384
 germinal, 384
 heredable, 384
 incidencia, 386
 inducidas. V. *mutaciones exógenas*.
 inducidas por agentes físicos, 398
 magnitud, 386
 materia prima para cambios evolutivos, 384
 neutral. V. *silenciosa*.
 no heredable, 385
 no silenciosa, 390
 no sinónimas. V. *de sentido alterado*.
 origen del cáncer, 386
 por delección, 387, 388
 por desaminación de bases, 394
 por inserción, 387, 389
 por sustitución, 387
 prematura de la traducción, 392

mutaciones (*cont.*)
 puntual, 387
 silenciosa, 389, 409
 sin cambio de marco, 392
 sin sentido, 392
 sinónima. V. *silenciosa*.
 somática, 385
 sustitución
 conservadora, 391
 no conservadora, 391
 terminadora, 392
 transición, 387
 transversión, 387
 y degeneración del código, 319
 y metabolismo, 396
 y polimorfismo, 406
 y sustancias ambientales, 396
 mutagénesis, 396
 mutágenos
 eliminación, 399
 endógenos, 396
 exógenos, 396
 myc (gen), 453
 Myc (proteína), 468, 470

N

NAD⁺, 25
 como cosustrato de la ligasa, 218
 NADP⁺, 25
 naftol, 189
 nailon, 170
 NAO, 410
 naranja de acridina. V. *acridina*.
 NBT, 189
 necrosis, 472
 NELF, 274
 neomicina, 344
 neoplasia. V. *cáncer*.
 NER. V. *escisión de nucleótidos*.
 neurofibromatosis de tipo 1, 457
 herencia, 437
 mutación causante, 393
 neurofibromina, 437, 457
 mutada, 393
 neuronas, 99
 neuropatía óptica hereditaria de Leber, 442
 nf1 (gen), como oncosupresor, 455
 NF1, 288, 292
 nitratos, 398
 nitritos, 398
 nitrocelulosa, 170
 nitrógeno
 isótopos 14N y 15N, 147
 nitrosaminas, 397
 nitroso (ácido), 398
 no-disyunción, 445, 447
 norfloxacina, 278
 Northern, 172
 comparado con Southern y Western, 178
 novobiocina, 278
 NPP, 189

N-terminal. V. *proteínas, extremo*.
 NTP
 marcado en fosfato α , 187
 marcado en fosfato γ , 187, 252
 nucleasas
 aplicación a la cartografía, 248
 aspectos generales, 212
 de restricción. V. *enzima de restricción*.
 en la apoptosis, 472
 S1, 169
 nucleicos (ácidos), 3
 absorción en el ultravioleta, 32, 167
 apirimidínicos, 31
 apurínicos, 31
 carácter
 hidrófilo, 31
 polianiónico, 31
 componentes, 11
 esqueleto, 28
 estructura primaria, 28
 extracción, 136
 con microesferas magnéticas, 139
 función, 27
 hidrólisis química, 31
 localización, 27
 niveles estructurales, 28
 propiedades físicoquímicas, 30
 purificación, 136
 reactividad, 31
 representaciones lineales, 29
 solubilidad, 136
 tipos, 29
 viscosidad, 31
 nucleína, 4
 núcleo, 6
 celular, 4
 nucleofílico/nucleófilo, 149, 218, 269, 305, 339
 nucleoide, 6
 nucléolo, 270
 nucleoproteidos, 27, 75
 nucleósidos
 absorción en el ultravioleta, 167
 componentes, 17
 de los ácidos nucleicos, 18
 -difosfato, 20
 hidrólisis química, 24
 infrecuentes, 19
 en el rRNA, 65, 68
 en el tRNA, 65, 307
 -monofosfato, 20
 propiedades físicoquímicas, 23
 que forman los ácidos nucleicos, 18
 tipos, 18
 -trifosfato, 20
 nucleosomas, 84-86
 desensamblado, 281
 fibra de, 87
 liberación, 281
 posición y control de la expresión, 281
 5'-nucleotidasa, 476

- nucleótidos
 absorción en el ultravioleta, 167
 cílicos, 25
 como grupo activador de azúcares, 367
 complejos, 25
 complejos con Mg²⁺, 24
 componentes, 19
 didesoxi-. V. *didesoxinucleótidos*.
 hidrólisis química, 24
 nomenclatura, 21
 posición del enlace fosfato, 19
 propiedades
 fisicoquímicas, 23
 iónicas, 23
 sencillos, 20
 tipos, 19
 unidos a azúcares, 367
 nucléolo, 194, 195
 nuliploidía, 444
 nulisomía, 444
 numeración de los nucleótidos en genes y RNA, 267
 número de enlace, 82
 modificación por topoisomerasas, 156
 número *n*, 95
 variación en la fecundación, 108
 y ciclo celular, 104
 Nycoprep, 126
- O**
- Oct-1, 295
 octámero de histonas, 88
 octapéptido, 277
 oftalmoplejía externa progresiva crónica, 442
 Okazaki (fragmentos de), 151, 156, 157
 maduración, 158
 progreso de la síntesis, 157
 unión de fragmentos, 158
 oleato, 361
 oligo-dT, 139, 222, 223
 oligonucleótido, 187
 empleo como sondas, 182
 empleo en PCR, 202
 específico de alelo, 420
 síntesis química, 182
 oligosacáridos, 363
 estructura química, 365
 precursor del grupo AB0, 411
 unidos por N, 365
 unidos por O, 364
 olivomicina, 95
 oncochip, 429, 477
 oncogenes
 comparación con gen oncosupresor, 456, 458
 concepto, 454, 455
 expresión dominante, 453, 456
 expresión recesiva, 454
 que estimulan el ciclo celular, 469
 oncoproteína, 452, 457
 oncosupresor (gen), 454
 comparación con oncogén, 456, 458
 concepto, 456, 457
 expresión recesiva, 458
 tipos, 456
 oncosupresor (proteína), 454, 457
 concepto, 456, 457
 oocito, 105, 194, 195
 secundario, 195
 oogénesis, 102, 105
 oogonios, 102, 105
 operón, 272, 326
 ORF. V. *marco de lectura abierto*.
 organización tridimensional, 157
 orgánulos, 6, 354
 origen
 de cambios evolutivos, 384
 de replicación. V. *replicación, origen*.
 mitocondrial, 161
 plasmídico, 227, 228
 de transcripción
 como referencia de numeración, 267
 orina, muestras, 124
 oro
 coloidal, 189
 microesferas para clonación, 231, 232
 óvulos, 105
 2-oxoglutarato, 360
- P**
- p105-Rb (proteína). V. *Rb (proteína)*.
 p21 (proteína), 469
 p21Ras. V. *Ras (proteína)*.
 p23, 376
 p53 (gen y proteína), 471
 actuación sobre el ciclo celular, 467, 469
 como factor de transcripción, 471
 inactivación por proteínas víricas, 472
 incidencia en el cáncer, 469, 471
 papel dual en ciclo celular
 y en apoptosis, 474
 PAC, 227
 pactamicina, 344
 palíndromos, 51
 como diana de restricción, 215
 monocatenarios, 52
 palmitato, 361
 PAP, 475
 papilomavirus, 471, 472
 parentesco, grados de, 436
 pares de bases. V. *complementariedad*.
 geometría, 42
 modelo molecular, 42
 pares A-T, A-U y G-C, 40
 PARP, 473
 partícula
 central, 87
 de reconocimiento de la señal, 357
 Patau (síndrome), 447
 paternidad, pruebas de, 424
 patología molecular
 concepto, 432
 pauta de lectura. V. *marco*.
 PBP, 295

- PCNA, 152, 157
 secuestro por GADD, 471
 PCNA, 403. V. también *DNA polimerasas δ y ε.*
 PCR, 193
 anidada, 208
 aplicaciones, 201
 asimétrica, 209
 ciclo, 203
 con adaptadores, 209
 con comienzo en caliente, 208
 con transcriptasa inversa, 206
 cuantitativa, 205
 de RNA, 206
 duración, 204
 en tiempo real, 205
 específica de alelo, 424
 especificidad, 204
 etapas, 203
 fidelidad, 205
 fundamento, 202
 grado de amplificación, 204
 inversa, 208
 LA-PCR, 208
 larga, 208
 límite de detección, 204
 L-PCR, 208
 múltiplex, 423
 número
 de ciclos, 204
 de copias conseguidas, 204
 objetivos, 201
 para la detección de polimorfismos, 422
 para obtención de sondas, 181
 para obtener DNA monocatenario, 209
 reactivos, 202
 rendimiento, 204
 RT-PCR, 206
 PDGF. V. *factor de crecimiento derivado de plaquetas.*
 PEG. V. *polietilenílico*
 penciclovir, 15
 peptidasa, 379
 de la señal. V. *proteasa del péptido señal.*
 peptidiltransferasa, 74
 centro activo, 79
 peptidil-tRNA
 hidrólisis, 342
 péptido C, 370
 péptido señal, 356, 370
 eliminación, 358
 partícula de reconocimiento de, 357
 peptidoglicanos, 363
 periodicidad en la estructura del DNA, 36
 permanganato potásico para secuenciación, 252
 peroxidasa, 188, 189
 peroxisomas, 355
 PEST, 379
 PFGE. V. *electroforesis en campo pulsante.*
 PGH. V. *Proyecto Genoma Humano.*
 pH
 efecto sobre el emparejamiento de bases, 41
 efecto sobre la desnaturalización, 165
 pintado de cromosomas, 177
 piperidina para secuenciación, 252
 pirimidina, 14
 pirofosfatasa, 149, 269
 pirofosfato, 149, 269
 pistolas de genes, 231, 232
 plantilla. V. *molde.*
 plásmido, 6
 artificial, 227
 características, 227
 circularización, 208, 226
 como vector de clonación, 219
 F, 230
 linealización, 181
 origen de replicación, 227, 228
 pSP64, 181
 plegamiento de proteínas, 261, 372, 373.
 V. también *carabinas.*
 deficiencia en la fibrosis quística, 392
 enfermedades, 372
 mecanismo, 372
 poli(A). V. *cola de poli(A).*
 poli(A) polimerasa, 302
 poli(ADP-ribosa)polimerasa, 473
 poli(dT), 303
 poli(T), 303
 poli(U), 303
 poliacrilamida, 141, 178
 poliaminas, 31
 policationes, 188, 231
 policistrónico, 324
 policonector, 224, 227, 228, 234
 polietilenglicol, 231, 244
 polietilenimina, 188
 polimorfismo, 97, 99, 243, 420. V. también
 marcadores polimórficos.
 aplicaciones, 424
 diagnóstica, 426, 427
 con efecto fenotípico, 409
 concepto, 406
 de conformación de moléculas monocatenarias, 413
 de mini- y microsatélites, 425
 de oligosacáridos, 410
 de un solo nucleótido. V. *polimorfismo SNP.*
 detección, 418
 en el número de repeticiones en tandem V. *VNTR.*
 en genoma mitocondrial, 416
 en intrones, 409
 en la longitud de los fragmentos de
 restricción. V. *RFLP.*
 fisiológico y patológico, 406
 genético, 405
 no proteico, 390
 génico, 405, 409
 ejemplos, 409
 génico no codificante, 415
 grado de, 406
 mecanismos de aparición, 407
 no génico, 415
 proteico, 390, 409, 410
 RFLP. V. *RFLP.*

- polimorfismo (*cont.*)
 sin efecto fenotípico, 409
 SNP, 405, 418, 424
 detección, 418
 STR. V. STR.
 uso forense, 421
 VNTR. V. VNTR.
 y heterocigosis, 406
 y mutación, 406
 polinucleótido fosforilasa, 269
 polinucleótido quinasa, 187, 252
 polioma, 81
 polipéptido naciente, 79
 poliploidía, 96, 444
 polirribonucleótido, 269
 polirribosomas, 75
 polisomas. V. *polirribosomas*.
 polisomía, 444
 poli-U, 139
 poliubicitina, 378
 Polymorphoprep, 126
 poro nuclear, 297, 346, 356
 postranscripcional
 control de la expresión génica, 309
 potenciadores, 290
 pRb (proteína). V. *Rb (proteína)*.
 precipitación salina diferencial, 137
 prefoldina, 375
 pre-miRNA, 309
 prenilación de aminoácidos en proteínas, 362
 preparación de muestras, 119
 esquema unificado, 120
 preproinsulina, 370
 péptido señal, 357
 pre-pro-proteína, 357, 370
 pre-proteína, 357
 prequimotripsinógeno, 370
 pre-RNA, 261, 264, 297
 con varios productos génicos, 298
 maduración. V. *maduración*
 postranscripcional.
 mitocondriales, 276, 309
 mixto para tRNA y rRNA, 308
 precursor grande del RNA ribosómico, 68
 productos génicos, 270
 ribosomal, 76
 de *Tetrahymena*, 72
 grande, 270, 308
 pequeño, 270, 308
 transferente, 308
 pretraduccional
 control de la expresión génica, 309
 pretranscripcional
 control de la expresión génica, 281
 pretripsinógeno, 370
 prevalencia, 433
 Pribnow (caja de), 289. V. también TATA.
 primasa, 149, 151, 156
 pri-miRNA, 309
 primosoma, 156
 priones, 372. V. también *PrP*.
- probando, 436
 procariotes. V. *procariotas*.
 procariotas
 diferencias con eucariotas, 5
 genoma, 5
 procesamiento
 postranscripcional. V. *maduración postranscripcional*.
 proteolítico de proteínas. V. *proteólisis*.
 procesividad, 151, 152, 274
 procolágeno, 360
 producto génico, 264
 nomenclatura, 266
 profase, 101
 proflavina, 398
 prohormona convertasa PC2 y PC3, 370
 proinsulina, 371
 proliferación, 233
 control
 negativo, 456
 por señales extracelulares, 454, 458
 positivo, 454
 equilibrio con muerte celular, 452
 monoclonal, 451
 prolil hidroxilasa, 360
 prolina
 hidroxilación, 360
 prometafase, 101
 promotores, 288
 basales, 272, 289
 complejo plurimolecular, 291
 reconocimiento por TBP, 291
 de clases I, II y III, definición, 288
 de tipo III, 295
 desnaturalización, 273
 distales, 290
 en vectores de expresión, 236
 Inr, 272
 liberación del, 273
 metilación, 282
 mitocondriales, 276
 potenciadores, 290
 proximales, 289
 silenciadores, 290
 TATA, 272
 pronúcleo, 108
 propagación celular, 233
 propidio, 132, 133
 propositus, 436
 pro-proteínas, 357, 370
 protaminas, 31, 85
 proteasas, 124, 379, 451
 cisteinproteasas, 473
 del péptido señal, 357, 358
 digestivas, 370
 en la apoptosis, 472
 inhibición por antitripsina, 414
 inhibidores, 452
 lisosomales, 377
 proteasoma, 379
 protección de grupos funcionales, 182, 183

proteínas, 4
 activación proteolítica. V. *proteólisis*.
 anclaje a membranas mediante lípidos, 367
 como producto génico. V. *producto génico*.
 de adhesión, 451, 476
 de anclaje, 357
 de choque térmico, 374
 de fusión, 236
 de membrana, 357
 de pausa, 275
 de replicación A. V. *RPA*.
 de secreción, 357, 370
 degradación, 376
 citosólica, 377
 lisosómica, 377
 destino, 354
 detectora, 185, 188
 distribución, 354
 extremo
 amino o N-terminal, 324
 carboxilo o C-terminal, 324
 funcional, 353
 G, 461
 glicosilación, 363
 interacción con DNA, 56
 ligante, 185
 de DNA monocatenario, 154, 155
 de IRE, 350
 de la caperuza, 335. V. *CBP*.
 de TATA, 291, 295. V. *TBP*.
 función múltiple, 296
 del telómero, 96
 localización, 354
 metabolismo, 354
 naciente, 324, 340, 373
 no histónicas, 85
 oncosupresora, 454
 plegamiento, 261, 372. V. también *plegamiento*.
 priónica, 372
 procesamiento proteolítico. V. *proteólisis*.
 receptora
 de AMP cíclico, 57
 de SRP, 357
 del ribosoma, 357
 recombinantes, 236
 reguladoras del hierro, 350
 secreción, tráfico, 357
 secretadas, 357, 370
 síntesis de, 323. V. también *traducción*.
 SR, 311
 supresora de tumores, 454, 456. V. también
 oncosupresores.
 terminadoras de la transcripción, 275
 tránsito, 354
 transportadora de cloruro, 392, 438.
 V. *regulador de la conductancia
 transmembrana*.
 ubicutinadas, 377
 vida media, 377
 proteína-disulfuro isomerasa, 369, 372
 proteína fosfatasa, 359

proteína quinasa, 359, 464
 dependiente de ciclinas, 462, 464. V. también *Cdk*.
 TFIIH, 291
 proteoglicanos, 363
 proteólisis, 369, 473. V. también *proteasas*.
 de la DNAPol I, 150
 de proteínas marcadas con ubicitina, 379
 ubicitina como marcador, 378
 proteoma, 239
 proteómica, 239
 protooncogén
 concepto, 454, 455
 tipos, 454
 proximal. V. *promotores*.
 Proyecto Genoma Humano, 257
 características, 240
 PrP, 372
 PSA, 475
 PSAP, 475
 PSE, 295
 pseudogenes, 114
 y transcripción, 267
 pseudomónico (ácido), 343
 pseudouridina, 19, 65
 psoraleno, 400
 pSP64, 181
 P-TEFb, 274
 PTF, 295
 pUC19, 227
 puentes disulfuro. V. *enlaces*.
 punto de restricción, 461
 punto R, 461
 purina, 14
 puromicina, 19, 344
 mecanismo de acción, 345
 peptidil-, 345

Q

qPCR, 205
 quiescencia, 100
 quimera, 447
 quimioluminiscencia, 189
 quimotripsina, 370
 inhibición por antitripsina, 414
 quimotripsinógeno, 370
 quinacrina, 94
 quinasa
 activadora de Cdk, 465, 467
 inhibidora de Cdk, 465, 467
 quinolonas, 278

R

Rab (proteína), 458
 Rab, 461
 radiaciones
 rayos gamma, 398
 rayos X, 398
 radicales libres, 396, 399
 de oxígeno
 reparación, 402
 Raf, 461

- ramificación, centro de, en ayuste, 303
 ras (gen), 453
 Ras (proteína), 455
 - en neurofibromatosis tipo 1, 393
 - farnesilación de, 362
 - función, 461
 - mutada, 461
 rasgo falciforme, 438
 rasgos, 98
 razón de asimetría, 37
 Rb (gen), 470
 - inactivación y cáncer, 471
 Rb (proteína)
 - actuación sobre el ciclo celular, 469
 - como sustrato de Cdk, 468, 471
 - fosforilación, 470
 - inactivación por proteínas víricas, 471
 rDNA, 219. V. también *DNA recombinante*.
 - formación, 225
 - introducción en la célula anfitriona, 231, 232
 - subproductos, 225
 reacción en cadena de la polimerasa, 201. V. también *PCR*.
 reasociación del DNA, 118. V. también *curvas Cot.*;
renaturalización.
 recambio proteico, 376
 receptor, 453
 - de factores
 - de crecimiento, 454
 - inhibidores del crecimiento, 457
 - de hormonas tiroideas, 453, 455
 - de transferrina, 351
 recesivo, 98, 436
 Recklinghausen (enfermedad de), 393
 recombinación, 145, 242
 - fracción de. V. *frecuencia*.
 - homóloga, 108, 241
 - meiótica, 241
 - no homóloga o desigual, 407
 - relación con la cartografía, 241
 - y segregación de alelos, 241
 recuento de células viables, 131
 recuperación de bases
 - vía metabólica, 246
 reemplazamiento
 - método de preparación de virus
 - recombinantes, 229
 región
 - estructural. V. *gen, secuencias*.
 - reguladora. V. *gen, secuencias*.
 - organizadora del nucleolo, 95
 regiones
 - hipervariables en el DNA mitocondrial, 416
 - internas de control, 295
 regulación
 - por fosforilación reversible, 462, 464
 - traduccional, 347
 regulador de la conductancia transmembrana, 438
 reguladora (secuencias génicas). V. *gen, secuencias*.
 relación de codificación, 314
 relleno, 187
 renaturalización, 118, 163, 164, 167
 rendimiento
 - de la PCR, 204
 reparación, 145
 - actividad 5'-exonucleasa, 150
 - de bases metiladas, 400
 - de emparejamientos incorrectos, 403
 - del DNA, 402. V. también *escisión*.
 - por escisión, esquema global, 400
 - sistemas enzimáticos, 398
 - tipos, 399
 - del uracilo en el DNA, 402
 - por recombinación, 403
 repeticiones, 111. V. también *DNA repetitivo*.
 - cortas en tandem, 421. V. también *STR*.
 - directas, 54
 - especulares, 54
 - invertidas. V. *palíndromos*.
 - terminal. V. *telómeros*.
 replicación, 145
 - autónoma, 223
 - bifocal, 161
 - bloqueo por GADD o p21, 471
 - cofactores, 148
 - comparación entre procariota, eucariotas y mitocondrias, 161
 - complejo de. V. *replisoma*.
 - conservadora, 146, 147
 - continua, 161
 - del RNA, 4, 149
 - demostración del mecanismo, 147
 - dirección, 149
 - dispersante, 146
 - elongación, 156
 - en el ciclo celular, 100
 - en la PCR, 202
 - enzimología, 148
 - errores, 394, 403
 - final. V. *terminación*.
 - hipótesis, 146, 147
 - horquilla, 148, 154
 - inicio, 153
 - mecanismo, 146
 - de la reacción, 149
 - mitocondrial, 161
 - monofocal, 148
 - multifocal, 148
 - no simultánea, 161
 - origen, 95, 148, 153, 230
 - órigenes múltiples en un cromosoma, 158
 - reacción
 - comparada con transcripción, 269
 - global, 149
 - relación con el ciclo celular, 104
 - semiconservadora, 146, 147
 - señales de inicio, 468
 - sustratos, 148
 - terminación. V. *terminación*.
 - unidad de. V. *replicón*.
 - unidireccional, 161
 - velocidad, 152
 - y ciclo celular, 146

- replicador, 153
 replicón, 148, 153, 158
 replisoma, 154, 156, 157
 reprogramación del genoma, 286
 resazurina, 133
 resistencia a antibióticos, 223, 227, 228, 234, 343
 resorufina, 133
 restricción. V. *enzimas de restricción*.
 restricción-modificación, 213
 restrictasas. V. *enzimas de restricción*.
 retardada (hebra), 156, 157
 retículo endoplásmico
 lumen, 357
 rugoso, 354
 reticulocitos
 para estudiar la traducción, 324
 retinoblastoma. V. *Rb*.
 retrasada (hebra). V. *retardada*.
 retrotransposición, 408
 retrotransposones, 408
 retrovirus, 4, 70, 228, 263
 revelado del DNA tras electroforesis, 142
 reversión directa de la mutación, 399
 RFC, 157, 403. V. también *DNA polimerasas δ y ε*.
 RFLP, concepto, 418
 Rho (proteína), 458, 461
 riboedición, 305
 riboforina, 357
 ribointerferencia, 69
 ribointerruptor, 69
 ribonucleasa, 212. V. también *RNAs*.
 como marcador tumoral, 476
 como proteína modelo de plegamiento, 372
 ribonucleoproteínas nucleares
 pequeñas. V. *snRNP*.
 riborregulador, 69. V. también *ribointerruptor*.
 ribosa, 14
 ribosilitmina, 18
 ribosoma
 centros activos, 79
 composición, 76, 79
 deslizamiento
 en el inicio de la traducción, 335, 336
 en la elongación, 339
 disociación, 342
 durante la traducción, 335
 en mitocondrias y cloroplastos, 75
 interacción con mRNA, 79
 reasociación de subunidades, 336
 separación de subunidades, 335
 sitios A, P y E, 337
 subunidades, 76
 mayor, 335
 menor, 333, 335
 tamaño, 75, 76
 ribosondas. V. *sondas de RNA*
 ribotimidina, 18, 65
 ribovirus, 70
 ribozimas, 71
 autoayuste, 305
 autocatalíticas, 305
 RNA autocatalítico, 305
 transpeptidasa, 339
 Rich-Dickerson, 47
 ricina, 344
 rifamicinas, 276
 rifampicina, 271, 276
 rigor, 168
 RISC, 309
 RNA
 7SL, 118, 357
 abundancia, 63
 catalítico. V. *ribozimas*.
 citoplásmico pequeño. V. *scRNA*.
 como producto génico. V. *producto génico*.
 composición de bases, 61
 cromosómico, 4
 diversidad de estructuras, 62
 en orgánulos, 63
 estabilidad, 302, 347
 estructura, 61
 secundaria, 62, 66
 funcional, 264, 297. V. también *RNA maduro*.
 genómico, 5, 79
 guía, 306
 interferencia por. V. *ribointerferencia*.
 interferente pequeño, 63. V. también *siRNA*.
 interferentes, 309
 maduración. V. *maduración postranscripcional*.
 maduro, 261, 264, 270
 mensajero. V. *mRNA*.
 mitocondrial, 276
 maduración, 309
 naciente, 269, 274
 nuclear
 heterogéneo. V. *hnRNA*.
 pequeño. V. *snRNA*.
 pequeños, 69
 poliadénilado. V. *cola de poli(A)*.
 precursor, 68, 264, 297.
 V. también *pre-RNA*.
 secuenciación, 257
 tamaño, 63
 tipos, 61, 63
 transporte, 261
 al citoplasma, 302
 ubicación en la célula, 63
 vida media, 302
 RNA polimerasa, 85, 261
 activación por el complejo de inicio, 291
 cloroplásticas, 270
 en la transposición, 408
 especialización, 270, 288
 holoenzima, 292
 inhibición
 por amanitina, 270
 por rifampicina, 271
 localización subcelular, 270
 mitocondrial, 270, 276
 pausas, 275
 procariótica, 269

- RNA polimerasa (*cont.*)
 enzima mínima o núcleo, 271
 subunidad σ , 271
- SP6, 181
 subunidades, 271
 tipos, 269
 transcritos primarios sintetizados, 270
- I, 270, 295
 II, 270, 291
 dominio C-terminal, 272. V. CTD.
 fosforilación, 272
- III, 270, 295
- RNA replicasas, 4, 149
- RNA ribosómico. V. rRNA.
- RNA transferente. V. tRNA.
- RNApol. V. RNA polimerasas.
- RNAsas, 212
 A, como proteína modelo del plegamiento, 372
 como contaminación, 172
 D, 307
 de *Bacillus cereus*, 257
 degradación del mRNA, 347
 H1, 158
 en la síntesis de telómeros, 160
 P, 71, 307
 Phy M, 257
 T1, 257
 U2, 257
 I, 307
 de páncreas bovino, 212
 III, 307, 309
- ro (proteína)
 análogos eucarióticos, 275
- Robertson. V. translocación.
- rodamina, 256
- rodopsina, 361
- rojo neutro, 131
- romos (extremos), 215
- ROS. V. especies reactivas de oxígeno.
- ROX, 256
- RPA, 154, 155
- rRNA
 5, 8S, 63, 68, 76, 308
 5S, 63, 68, 76, 308
 12S, 308
 16S, 63, 68, 76, 308
 18S, 63, 68, 76, 95, 308
 23S, 63, 68, 75, 76, 308
 28S, 63, 68, 75, 76, 95, 308
 28S, actividad transpeptidasa, 339
 genes de, 112
 maduración, 307
 polimerasa que los transcribe, 270
 precursores de, 270
 tipos, generación por maduración, 308
- RT-PCR, 205, 206. V. PCR con transcriptasa inversa; PCR en tiempo real.
- S**
- saco amniótico, 121
 salientes 5' y 3' en los extremos de restricción, 215
 saliva, muestras, 122
- Sanger, 252
 sangre
 fetal, 122
 manchas, 122
 periférica, muestras, 122
 sarcosina, 278
- satélite. V. DNA satélite.
 cromosómico. V. cromosoma.
- SCC, 475
- scRNA, 63
 miRNA, 69
- secuena, 252
- secuencia
 cis. V. cis.
 cos, 229
 de etiquetado. V. péptido señal.
 de replicación autónoma, 230
 estructurales. V. gen, secuencia.
 génicas. V. gen, secuencia.
 Inr, 272
 intercaladas. V. intrón.
 líder. V. péptido señal.
 no traducidas 5' y 3', 299
 promotoras. V. promotor.
 reguladoras. V. gen, secuencia.
 señal, 355, 356. V. péptido señal.
 subteloméricas, 96
 TATA, 272
 teloméricas. V. telómeros.
 trans. V. trans.
- secuenciación, 240
 de RNA, 257
 didesoxi. V. secuenciación, método enzimático.
 en la búsqueda de genes, 418
 enzimática, variantes, 256
 método
 enzimático, 252
 químico, 252
 por terminación de cadena. V. secuenciación, método enzimático.
- secuencial, replicación, 148
- sedimentación, 124
- segregación
 de alelos, 241
 de cromosomas y cromátidas, 106, 107, 241
- segundo código genético, 329
- selección, 246
 de células hibridas, 246
 de clones recombinantes, 234
 fenotípica, 234
 negativa, 129
 positiva, 129
- sellado genético, 286
- semen, 122
- semiconservadora. V. replicación.
- semidiscontinua, 157
- señalización, tipos, 459
- separación, métodos, 124
- serina/treonina quinasa, 359, 464
- Shine-Dalgarno (secuencia de), 333
- showdomicina, 344
- siálico (ácido), 364

- sigma. V. *RNA polimerasa procariótica*.
 silenciadores, 290
 simultánea, replicación, 148
 SINE, 118
 sinistrorso, 43
 Z-DNA, 49
 síntesis *de novo* de nucleótidos, 246
 siRNA, 63, 69
 biosíntesis, 309
 sis (gen), 452
 sistema endomembranoso, 354
 sitio
 A, 337, 339
 aminoacílico. V. *sitio A*.
 de clonación múltiple, 224
 de restricción, 212
 ejemplo en un plásmido, 227, 228
 de salida. V. *sitio E*.
 de secuencia marcada. V. *STS*.
 E, 339
 EA, 337
 marcados únicos. V. *STS*.
 P, 337, 339
 peptidílico. V. *sitio P*.
 SL1, 295
slot-blot, 171
 SNAPc, 295
 snip. V. *polimorfismo SNP*.
 SNP. V. *polimorfismo SNP*.
 snRNA, 63, 69, 270, 295
 tipos, tamaño y función, 304
 snRNP
 tipos, tamaño y función, 304
 sobre cruzamiento. V. *recombinación*.
 solapamiento de codones, 314
 solenoide, 85. V. también fibra de 30 nm.
 sonda, 168, 179, 206
 clasificación, 179
 de DNA, 180
 de RNA, 181
 diseño, 184
 fluorescentes, 177, 246
 marcaje, 185
 obtención
 por clonación, 180
 por PCR, 181
 para detectar anemia de células falciformes, 418
 preparación, 179
 síntesis, 187
 sintéticas, 182
 soporte sólido
 unión a, 182
 Southern, 171
 aplicación a la cartografía, 248
 comparado con Northern y Western, 178
 interpretación de resultados, 172
 para detectar
 anemia de células falciformes, 418
 RFLP, 418
 VNTR, 421
 south-western, 179
 SP1, 292
 SP6, 181
 SREBP, 473
 SRP, 357
 SRP-R, 357
 SSB. V. *proteínas ligantes de DNA monocatenario*.
 SSCP, 413
 Staf, 295
 STR, 421, 425
 en el cromosoma Y, 424
 stRNA, 69
 STS, mapa físico, 248
 subclonación, 237
 subtilisina, proteólisis de la DNAPol I, 150
 suicidio celular. V. *apoptosis*.
 sulfatación, 369
 SUMO, unión a histonas, 286
 sumoilación, 369
 superenrollamiento, 79, 83, 88
 aspectos teóricos, 82
 demostración experimental, 81
 negativo, 81, 156
 parámetros cuantitativos, 82
 plectónémico, 82
 positivo, 81
 toroidal, 82
 superhélice, 80
 superóxido, 396, 399
 superóxido dismutasa, 399
 surco
 mayor, 43, 45
 unión de proteínas al, 57
 menor, 43, 45
 unión de proteínas al, 291
 susceptibilidad tumoral, 449
 sustancia H. V. *antígeno H*.
 sustitución, 387. V. también *mutación*.
 método de preparación de virus recombinantes, 229
 SV40, 228
 svedberg. V. *coeficiente de sedimentación*.
 SII, 274
 SIII, 274
- T**
- T_f, 167. V. también T_m.
 T_m
 en la PCR, 203
 y contenido en G+C, 167
 TAF, 291, 294. V. *factores asociados a TBP*
 de tipo I, 295
 de tipo II, 291
 de tipo III, 295
 talasemia
 alfa, 438
 beta, 438
 tallo, 62
 TAMRA, 256
 tandem, 111, 421
 TATA, 272, 289, 295
 reconocimiento por TBP, 291
 tautomería
 amina-imina, 16

- tautomería (*cont.*)
 ceto-enólica, 16
 de bases nitrogenadas, 16
 lactama-lactima, 16
 relación con la mutación, 394
Tay-Sachs (enfermedad de), 438
 mutación causante, 391
TBP, 291, 295. V. también *proteína ligante de TATA; proteína ligante del telómero*.
 función múltiple, 296
telofase, 101
telomerasa, 149, 159
 ausencia, 160
 componentes, 160
 función, 160
 y cáncer, 161
 y proliferación celular, 161
telómero, 89, 95, 96
 acortamiento, 158–160
 en vectores de clonación, 230
 funciones, 96
 replicación, 158
 secuencia, 160
 y DNA minisatélite, 118
 y DNA satélite, 116
 y envejecimiento, 160
temperatura
 de fusión del DNA. V. T_m .
 en la PCR, 203
 y desnaturalización, 164
TER. V. *telomerasa, componentes*.
terapia antitumoral
 y telomerasa, 161
terminación
 de cadena. V. *secuenciación enzimática*.
 de la replicación, 158
 de la traducción, 340
 de la transcripción, 275
 mitocondrial, 276
 retrasada de la traducción, 394
terminador, 275
 de la transcripción, 272
termociclador, 202
termoestable, 202, 252
terpenos, 362
tetraciclinas, 344
 gen de resistencia, 234
tetrahidrofolato, 246, 334
Tetrahymena, 72
tetraploidía, 97, 444, 445
tetrazolio, 189
 sales de, 133
TF. V. *factores de transcripción*.
TFIIS, 274
TGB, 476
TGF-β, 457
TGGE, 413
timidilato, 21
timidina, 18
 en medio HAT, 246
 tritiada, 134
timidina quinasa, 246
timina
 dímeros, 398
timina glicol, 396
 reparación de, 402
tinción, 93-95
 de la cromatina, 88
 del DNA, 142
 NOR, 95
tiouridina, 65
tiroglobulina, 476
tirosina quinasa, 359, 464
TK. V. *timidina quinasa*.
TMB, 189
topogénesis, 354
topoisomerasas, 80, 85, 88, 154
 de tipo I, 155
 de tipo II, 156
 en la transcripción, 274
topología, 82
toxina diftérica, 344
TR. V. *telomerasa, componentes*.
traducción, 4, 261, 264
 características, 323
 conexión espacial y temporal con la transcripción, 262
 elongación. V. *elongación*.
 energética, 342
 fases o etapas, 326
 frecuencia, 348
 inhibidores, 343
 inicio. V. *inicio*.
 macromoléculas implicadas, 324
 regulación, 345
 regulada
 por bloqueo del mRNA, 350
 por estabilización del mRNA, 351
 por fosforilación, 352
 resumen, 326, 342
 sentido, 324
 tasa de error, 332
 terminación. V. *terminación*.
 velocidad, 345
tráfico, 354. V. también *tránsito*.
 de proteínas, 354, 356
 secreción, 357
trans, 355. V. también *Golgi*.
 elementos, 288
 factores, 288
 secuencias, 288
transcripción, 4, 261, 264
 asimetría, 266
 características comparadas con replicación, 268
 cofactores, 269
 comparación entre procariotas y eucariotas, 268
 conexión espacial y temporal con la traducción, 262
 del genoma mitocondrial, 276
 dirección, 269
 e interfase, 269
 elongación. V. *elongación*.
 energética, 269
 enzimología, 269
 etapas, 272
 hebra

- transcripción (*cont.*)
codificante y no codificante, 267
con sentido y antisentido, 267
informativa y no informativa, 267
molde y no molde, 267
negativa, 267
positiva, 267
positiva y negativa, 267
transcrita y no transcrita, 267
inhibidores, 276
inicio. V. *inicio*.
mecanismo de la reacción, 269
mitocondrial, 276
origen. V. *origen de transcripción*.
promotor. V. *promotor*.
reacción
comparada con replicación, 269
global, 269
nicial, 273
regulación, 279
por señales, 290
positiva y negativa, 290
postranscripcional, 309
terminología, 288
y accesibilidad del DNA, 281
relación con la condensación del DNA, 268
sentido contrario en cada hebra, 266
sustratos, 269
terminación. V. *terminación*.
terminología, 267
unidad de, 272
y euromatina, 269
y fibra de 10 nm, 269
transcripción inversa, 4
transcriptasa. V. *RNA polimerasas*.
transcriptasa inversa, 149, 263
empleo para la secuenciación, 252
empleo para PCR, 207
en la transposición, 408
para preparar cDNA, 222
para secuenciar RNA, 257
transcrita, 267
transcrito primario. V. *pre-RNA*.
transducción de señales, 455
esquema global, 460
transesterificación, 305, 328
transfección, 229, 231
transferasa terminal, 222, 223
transferencia, 171, 178
nuclear, 193
transferrina
receptor de, 351
transformación, 230, 231
bacteriana, 230
transfosforilación, 305
transfusión
compatibilidad, 410
transición, 387
tránsito de proteínas, 354, 355
translocación
de proteínas del citosol al retículo endoplásmico, 357
de Robertson, 448
del ribosoma, 339
papel de las carabinas, 372
recíproca, 448
translocasa. V. *eEF-2, riboforina*.
transpeptidación, 339
transpeptidasa. V. *peptidiltransferasa*.
transporte
a través de poros nucleares, 261
de mRNA, 346
regulado, 356
transmembrana, 356
vesicular, 356
transposasas, 408
transposición, 407
transposones, 407
transversión, 387
traslado de la mella, 187
trébol, estructura del tRNA, 66
TRiC/CCT, 375
trifosfatos, 13
triple hélice, 55
triplete. V. *codón*.
triploidía, 97, 444
tripsina, 370
inhibición por antitripsina, 414
proteolisis de la DNAPol I, 150
tripsinógeno, 370
triptófano
deficiencia de biosíntesis, 234
trisfosfatos, 25
trisomía, 444
13, 447
18, 447
21, 385, 447
en cromosomas sexuales, 447
mecanismo, 447
trilito, 183
tRNA, 63
especificidad por aminoácido, 64
estructura terciaria, 67
isoaceptores, 329
maduración, 307
modelo molecular, 67
nomenclatura, 64
número de moléculas diferentes, 320
papel adaptador en la traducción, 315
polimerasa que los transcribe, 270
precursores de, 270
separación del polipéptido al terminar la traducción, 342
sinónimos, 64, 329
tRNA_f^{Met}, 334
tRNA_i^{Met}, 334
tRNA_t^{Met}, 334
complejo ternario con eIF-2 y GTP, 335
tRNA nucleotidiltransferasa, 307
trofoblasto, 121, 200
trombina, carboxilación, 359
troponina T gen y ayuste alternativo, 310
TRP1, 234
TRT. V. *telomerasa, componentes*.
tumor, 449. V. también *cáncer*.
maligno, 451

tumor (*cont.*)
 primario benigno, 449, 451
 secundario, 450, 451
 tungsteno, microesferas para clonación, 231, 232

U

U1, U2, U3, U4, U5, U6. V. *snRNA y snRNP*.

UBF, 295

ubicuitina, 377

 unión a histonas, 286

ubicuitina proteína ligasa, 377

ubicuitinación, 369, 377

 inducida

 por aminoácido N-terminal, 378

 por secuencias PEST, 379

UCE, 295

UDP-galactosa, 367

UDP-glucosa, 367

UDP-N-acetilgalactosamina, 367

UDP-N-acetilglucosamina, 367

ultracentrifugación, 81, 140

 isópénica, 81

unidad

 de codificación

 tamaño, 314

de repetición. V. *DNA repetitivo*.

de replicación. V. *replicón*.

de transcripción, 272

 mitocondrial, 276

unidireccional, replicación, 161

unión

 célula-matriz extracelular, 451

 intercelular, 451

universalidad. V. *código genético*.

uracilo

 reconocimiento en el DNA, 402

urea, 136, 165, 169

úrico (ácido), 15

uridilato, 21

uridina, 18

3'UTR y 5'UTR. V. *secuencias no traducidas*.

V

valil-tRNA sintetasa

 corrección de errores, 332

valor C, 7, 97

 variación

 en la fecundación, 108

 en la meiosis, 106

 y ciclo celular, 104

variabilidad genética. V. *polimorfismo*.

vascularización, 452

vector

 ARS/CEN/TEL, 230

 BAC, 230

 clasificación, 226

 cósmidos, 229

 de clonación, 219, 223

 de expresión, 224, 236

 de inserción, 224, 229

 de propagación, 224

 de sustitución, 229

plasmídicos, 227
 preparación, 224
 propagación, 227
 tamaño del inserto admitido, 227
 tipos, 226
 víricos, 228, 231

YAC, 230

vellosidades coriónicas, 122

velocidad de sedimentación, 125, 126

vesículas

 de secreción, 354
 de transición, 354
 de transporte, 354

viabilidad celular, 131

vida media

 de proteínas, 377

 del RNA, 347

viroides, 73

virus, 228. V. también *bacteriófagos*.

 adenovirus, 471, 472

 bacteriófago T2, 4

 cápsida, 4

 como vectores, 228, 231

 del mosaico del tabaco, 4

 del papiloma, 471, 472

 empleo como vectores. V. *vectores víricos*

 recombinantes, 228

 SV40, 471, 472

 tumorales, 471, 472

viscosidad

 variación en la desnaturización del DNA, 165

vitamina

 B₁₂. V. *cobalamina*.

 5'-desoxiadenosil-, 19

 C, 360

 K, 359

VNTR, 244, 421

W

WAF1, 464

WAF1 (proteína), 466

Watson y Crick, 4

 modelo para el DNA, 38

Western

 comparado con Southern y Northern, 178

X

xantina, 15

 como resultado de la desaminación, 395

xerodermia de Kaposi, 404

X-gal, 189, 234

Y

YAC, 227, 230

yoduro de propidio, 132. V. también *propidio*.

Z

Z-DNA, 49

 modelos moleculares, 50

zeiosis, 472

zimógenos, 370

zinc. V. *dedo de zinc*.