

Biologia molecolare della cellula 2

Giacomo Fantoni

Telegram: @GiacomoFantoni

Github: <https://github.com/giacThePhantom/BioMolCellula2>

24 ottobre 2020

Indice

1	Struttura e funzione dei cromosomi	4
1.1	Organizzazione dei cromosomi	4
1.1.1	Ploidia	4
1.1.2	Ulteriore DNA presente nelle cellule	5
1.2	Impacchettamento del DNA cromosomale	5
1.2.1	Il nucleotide	5
1.2.2	DNA eucariotico	6
1.2.3	Il ciclo cellulare e la dinamicità dei cromosomi	6
1.3	Impacchettamento del DNA cromosomale degli eucarioti	6
1.3.1	Istoni	6
1.3.2	Livelli di compattazione	7
1.4	Modifiche covalenti degli istoni	8
1.4.1	Epigenetica	8
1.4.2	Acetilazione	8
1.4.3	Metilazione	9
1.4.4	Fosforilazione	9
1.4.5	Ubiquitinazione e sumoilazione	9
1.4.6	Codice istonico	9
1.5	Complessi rimodellatori dei nucleosomi	10
1.5.1	Ruoli dei rimodellatori della cromatina	10
1.5.2	Sottofamiglie	10
1.6	Variazione nella struttura cromatinica	11
1.6.1	Eucromatina	11
1.6.2	Eterocromatina	11
1.6.3	Effetti della cromatina	11
1.6.4	Nucleolo	11
1.7	Metilazione del DNA	12
1.7.1	DNA metilasi	12
1.7.2	Effetti della metilazione	12
1.7.3	La disattivazione del cromosoma <i>X</i> è un esempio di silenziamento epigenetico della metilazione di cromatina nei mammiferi	13
1.7.4	Imprinting genetico	13
1.8	La separazione dei domini cromatinici da elementi barriera	13
1.8.1	Variegazione da effetto di posizione	13
1.8.2	Elementi di barriera	14
1.9	Elementi richiesti per la funzione dei cromosomi	14

1.9.1	Origine di replicazione	14
1.9.2	Centromeri	14
1.9.3	Telomeri	15
2	Replicazione	17
2.1	Replicazione del DNA semi-conservativa	17
2.1.1	Modelli di replicazione	17
2.2	Il modello dei repliconi	18
2.2.1	Scoperta del modello	18
2.2.2	Origini di replicazione	19
2.3	Identificazione delle origini di replicazione	19
2.3.1	Esperimento	19
2.3.2	Origini di replicazione negli eucarioti	19
2.4	Panoramica della replicazione del DNA	20
2.4.1	Le fasi della replicazione del DNA	20
2.5	Iniziazione	21
2.5.1	Svolgimento dell' <i>Ori</i> nei procarioti - E. coli	21
2.5.2	Svolgimento dell' <i>Ori</i> negli eucarioti	22
2.5.3	DNA elicasi	23
2.5.4	Sintesi di primer a RNA o RNA-DNA dalla DNA primasi	24
2.6	Allungamento	24
2.6.1	Pinza scorrevole	24
2.6.2	Sintesi del DNA	25
2.6.3	DNA polimerasi	25
2.6.4	Fedeltà della polimerizzazione del DNA	25
2.6.5	Sintesi del DNA discontinua	27
2.6.6	Attività del replisoma alla forcella di replicazione	27
2.7	Terminazione	28
2.7.1	Terminazione nei batteri	28
2.7.2	Terminazione negli eucarioti	28
2.8	Replicazione dei telomeri	29
2.8.1	Telomerasi	29
2.8.2	Mantenimento della lunghezza	29
2.8.3	Problemi nel mantenimento della lunghezza del telomero	29
2.8.4	Il limite di Haflick	29
2.9	Correzione degli errori post-replicativa	30
2.10	Mantenimento delle modifiche istoniche	30
2.11	DNA polimerasi specializzate	30
2.11.1	DNA polimerasi batteriche	30
2.11.2	DNA polimerasi specializzate	31
3	Trascrizione	32
3.1	Panoramica della trascrizione	32
3.1.1	Il processo di trascrizione	32
3.1.2	Nomenclatura dei geni	33
3.1.3	Regolazione della trascrizione	33
3.2	L'enzima centrale della RNA polimerasi	33
3.2.1	Struttura	34

3.3	Riconoscimento dei promotori	34
3.3.1	Batteri	35
3.3.2	Eucarioti	35
3.3.3	Formazione del complesso di pre-inizio	37
3.4	Iniziazione della trascrizione e transizione a un complesso di allungamento	37
3.5	Allungamento della trascrizione	37
3.6	Terminazione della trascrizione	37
3.7	Principi della regolazione della trascrizione	37
3.8	Domini leganti il DNA in proteine che regolano la trascrizione	37
3.9	Meccanismi per regolare l'iniziazione della trascrizione nei batteri	37
3.10	L'operone <i>lac</i> in <i>E. coli</i>	37
3.11	L'operone triptofano <i>trp</i> in <i>E. coli</i>	37
3.12	Regolazione della trascrizione da parte di riboswitches trascritti	37
3.13	Regolazione dell'espressione genica del batteriofago λ in <i>E. coli</i>	37
3.14	Regolazione della trascrizione da sistemi di trasduzione del segnale a due componenti	37
3.15	Regolazione dell'iniziazione della trascrizione ed allungamento negli eucarioti	37
3.16	Il ruolo delle cascate di segnalazione nella regolazione della trascrizione	37
3.17	Silenziamento genico attraverso imprinting genomico	37
4	Processamento dell'RNA	38
5	RNA regolatori	39
6	Traduzione	40
7	Modifica e targeting delle proteine	41
8	DNA mobile	42
9	Strumenti e tecniche della biologia molecolare	43

Capitolo 1

Struttura e funzione dei cromosomi

1.1 Organizzazione dei cromosomi

L'informazione genetica è impacchettata in almeno una molecola di DNA molto lunga, un cromosoma. Ogni cromosoma contiene una molecola di DNA a doppio filamento con molti geni e regioni di DNA non codificante. Si dicono intergeniche le regioni tra i geni. Se batteri ed archea possiedono cromosomi circolari gli eucarioti ne possiedono di lineari. La distribuzione dei geni varia tra gli organismi: quelli meno complessi tendono ad avere geni ordinati più densamente. La densità genica può variare anche sui diversi cromosomi degli organismi. Il numero dei cromosomi è caratteristico per una specie. Si possono fare incroci tra specie con numero di cromosomi diverso: in questo caso sono incapaci di accoppiarsi durante la prima parte della meiosi e l'incrocio risulta sterile.

1.1.1 Ploidia

Con ploidia si intende quanti cromosomi identici possiede un organismo:

- Aploidia: 1 cromosoma come nel lievito.
- Diploidia: 2 cromosomi come negli umani.
- Poliploidia: più di 2 cromosomi come nelle piante.
- Aneuploidia: un numero anormale di cromosomi, può avvenire in caso di sindromi genetiche o cancro.

La poliploidia viene sfruttata nei prodotti ortofrutticoli per aumentarne le dimensioni.

Aneuploidia e aborti

L'aneuploidia può essere sopportata in un certo numero dagli organismi: si nota per la trisomia del cromosoma 21 (sindrome di Down) e le poliploidie, ma può essere mortale e causare un aborto spontaneo.

Gametogenesi femminile

Si nota come con l'aumentare dell'età della donna aumenta il rischio di aneuploidia per i figli. Questo avviene in quanto ogni donna nasce con tutte le uova diploidi già presenti anche se immature. Queste maturano una alla volta dopo la pubertà una volta al mese. La continuazione della meiosi bloccata comincia il giorno prima dell'ovulazione a causa della gonadotropina. Un oocita primario può causare più errori durante la segregazione cromosomica nelle due fasi della meiosi rispetto a un uovo più giovane risultando in un uovo aploide con più o meno cromosomi.

1.1.2 Ulteriore DNA presente nelle cellule

Cellule eucariote

Le cellule eucariote possono avere DNA addizionale oltre il DNA cromosomale, in particolare in:

- Mitocondri: forniscono le cellule con ATP e sono organelli racchiusi da membrana con il proprio, solitamente circolare, cromosoma singolo.
- Cloroplasti: derivano l'energia dalla luce solare nelle piante, possiedono un proprio cromosoma.

Si pensa che questi organelli derivino da un batterio ancestrale assorbito e mantenuto da un altro organismo unicellulare.

Cellule batteriche

Le cellule batteriche possiedono DNA addizionale nelle proprie cellule: piccolo DNA circolare detto plasmide. Questi tipicamente codificano poche proteine che conferiscono un vantaggio selettivo come una resistenza ad un antibiotico.

Virus

I virus sono agenti infettivi che trasportano informazioni genetiche come piccoli cromosomi a DNA o RNA. I cromosoma virale può essere lineare o circolare, a doppio o singolo filamento.

1.2 Impacchettamento del DNA cromosomale

Si nota come per potersi adattare alle dimensioni del nucleo, delle cellule o di organelli intracellulari il DNA deve essere compattato. Compattare il genoma svolge anche una funzione di protezione, rendendolo meno accessibile da agenti esterni.

1.2.1 Il nucleoide

Nei procarioti, in assenza di nucleo il DNA si organizza in un nucleoide. È composto per l'80% di DNA e per il restante 20 di proteine di compattamento e RNA. Il cromosoma è pertanto composto da un grande complesso DNA proteine detto cromatina. Il nucleoide appare come una regione che esclude cromosomi, occupa $\frac{1}{3}$ del volume della cellula ed è ancorato all'origine di replicazione nella membrana cellulare. Forma "loops" o domini di circa 40kb grazie alla proteina *HIF* (integration host factor), una piccola proteina carica positivamente per bilanciare le cariche negative sul backbone. Il DNA è successivamente superavvolto da altre proteine che piegano il DNA. L'IHF è costituito da un dimero su cui si forma il loop. Il superavvolgimento è controllato da altri fattori come fattori di

trascrizione e l'attività di DNA ed RNA polimerasi che creano due superavvolgimenti con polarità opposta ai lati della bolla.

1.2.2 DNA eucariotico

Nel nucleo degli eucarioti i cromosomi subiscono cambi visibili durante il ciclo di divisione cellulare. Nelle cellule umane diploidi il DNA deve essere compattato 300 000-400 000 volte.

1.2.3 Il ciclo cellulare e la dinamicità dei cromosomi

Durante la fase G_2 del ciclo cellulare i cromosomi replicati si trovano in uno stato poco avvolto e si forma il centromero. Nella profase compaiono le fibre del fuso e i cromosomi si condensano. Nella prometafase le fibre si attaccano ai cromosomi che continuano a condensarsi. Nella metafase i cromosomi si allineano. Nell'anafase i centromeri si dividono e i cromatidi fratelli si muovono ai poli opposti. Durante la telofase si riforma la membrana nucleare, i cromosomi si decondensano e scompaiono le fibre del fuso.

1.3 Impacchettamento del DNA cromosomale degli eucarioti

1.3.1 Istoni

Negli eucarioti gli istoni sono proteine leganti il DNA. Si trovano quattro istoni del nucleo: nascono molto presto nell'evoluzione ed essendo cruciali per la sopravvivenza sono altamente conservati. Gli istoni sono basici in quanto ricchi in lisina e arginina cariche positivamente grazie all'gruppo ammino $+NH_3$ che stabilizzano le interazioni tra il DNA e gli istoni. 146bp si arrotolano 1.75 volte intorno a un complesso istonico in maniera sinistrorsa per formare un nucleosoma. Il complesso istonico prende il nome di ottamero istonico. Il superavvolgimento negativo facilita la separazione più facile, necessaria per la replicazione e la trascrizione. L'ottamero istonico ha due di ognuno dei quattro istoni del nucleo: $H2A$, $H2B$, $H3$, $H4$. Inizialmente due dimeri $H3-H4$ si associano con il DNA e reclutano poi due dimeri $H2A-H2B$ per la formazione dell'ottamero. Nonostante tutto il DNA eucariote sia impacchettato dagli istoni i nucleosomi si formano preferenzialmente a sequenze di DNA. Il DNA è generalmente piegato dolcemente intorno agli istoni ma presenta curve più acute ???????? La scanalatura minore deve diventare più stretta durante il piegamento, cosa più favorevole in regioni ricche di AT . Gli istoni fanno 13 interazioni con gli istoni del DNA nucleosomale: i due dimeri $H3-H4$ legano il centro e le terminazioni del DNA mentre $2(H2A-H2B)$ legano 30bp su un lato del nucleosoma. Il core istonico è composto dai domini di histone-fold composti da tre α -eliche.

Code istoniche

Il nucleo di una proteina istonica è legato a una lunga coda N-terminale che si estende verso l'esterno. Sono lunghe tra i 20 e i 39 amminoacidi e non hanno strutture. Interagiscono con altri nucleosomi per aiutare un ulteriore compattamento del DNA e strutture cromatiniche di livello superiore. La coda può essere modificata chimicamente in modo da modificare la struttura della cromatina e la sua funzione promuovendo o prevenendo il reclutamento di proteine che regolano la trascrizione. $H2A$ e $H2B$ presentano anche code C-terminali che regolano la trascrizione.

Varianti istoniche

Le varianti istoniche sono altre proteine con stabilità diverse, domini specialisti che cambiano la funzione del cromosoma, sequenze diverse alle terminazioni. Hanno amminoacidi diversi che possono essere diversamente modificati. Le varianti istoniche sono depositate da complessi di rimodellamento della cromatina dipendenti da ATP e possono essere dipendenti o indipendenti dalla replicazione.

Interazioni con il DNA

Gli istoni interagiscono con il DNA attraverso interazioni elettrostatiche tra i gruppi fosfato del legame fosfodiester e gli amminoacidi basici negli istoni e attraverso legami a idrogeno tra l'atomo di ossigeno nei gruppi fosfato e gli atomi di idrogeno nei gruppi ammino degli istoni. Modifiche chimiche delle basi o code istoniche attraverso enzimi modificano le cariche locali e le interazioni.

1.3.2 Livelli di compattazione

L'impacchettamento cromatinico ha diversi livelli di compattamento.

Primo livello

Il primo livello di compattazione è la fibra di 10nm, che nasce dall'associazione con il DNA dei nucleosomi con apparenza di perline su un filo.

Secondo livello

La fibra è ulteriormente compattata da una quinta proteina istonica *H1* in una fibra di 30nm nel secondo livello di compattamento, un ordinamento regolare che avvicina i nucleosomi. *H1* si lega al DNA linker tra due nucleosomi successivi diminuendo la lunghezza di 7 volte. Anche le code istoniche sono coinvolte nella formazione di questo secondo livello.

Terzo livello

Il terzo livello di compattamento, con diametro di 300nm si forma grazie a domini di loop radiali e al legame con la matrice nucleare nelle cellule in interfase. La matrice nucleare è composta dalla lamina nucleare composta da fibre proteiche della matrice interna e da proteine che legano ad essa i cromosomi. Le proteine attaccano la base di un loop di DNA alla fibra proteica grazie a sequenze specifiche *MAR* (matrix-attachment region) e *SAR* (scaffold attachment region). Si nota come ogni cromosoma occupa nel nucleo un territorio determinato

Quarto livello

I loop radiali diventano altamente compattati e rimangono ancorati alla matrice nucleare. Mentre la cellula entra in profase la membrana nucleare si dissolve e non si trova più una matrice nucleare: la compattazione aumenta drammaticamente nel quarto livello di compattazione o condensazione. Alla fine della profase i cromosomi sono interamente eterocromatici con un diametro di 700nm. Pertanto i cromosomi in metafase subiscono poca trascrizione e unicamente nel centromero. In questo momento i cromosomi hanno accesso al fuso mitotico.

Quinto livello

Il quinto livello di compattamento avviene con la formazione dei cromosomi visibili e grazie alla condensina. La condensina è una proteina che si sposta nel nucleo durante l'inizio della fase *M*, si lega ai cromosomi e compatta i loop radiali riducendo il loro diametro. Un'altra proteina coinvolta è la coesina caricata durante la fase *S* per tenere uniti i cromatidi fratelli.

1.4 Modifiche covalenti degli istoni

1.4.1 Epigenetica

Si intende per epigenetica l'ereditarietà di fenotipi non causati da cambi nella sequenza del DNA. È un fenomeno principalmente eucariote ed è causata da cambi strutturali nella composizione dei nucleosomi (varianti istoniche), modifiche chimiche della coda o nucleo istonico che altera lo stato di compattazione della cromatina e l'attività del nucleo del nucleosoma, metilazione del DNA alla citosina e dal legame di DNA o RNA con RNA non codificanti. Queste opzioni alterano l'espressione genetica. Cambi epigenetici sono trasferiti da cellula madre e figlia durante la replicazione del DNA e un numero di sindromi e cancro sono dovuti alla mal-regolazione di attività epigenetiche. Nei batteri la trascrizione dipende principalmente dall'RNA polimerasi e la sua regolazione allo stadio di iniziazione. Metilazione di adenosina e citosina intervengono nell'espressione genica, nella replicazione e riparazione del DNA e come difesa contro attacchi virali. Le modifiche chimiche più comuni sono alle code istoniche ma anche gli amminoacidi del nucleo globulare degli istoni possono essere modificati. Le modifiche sono principalmente acetilazione, metilazione, fosforilazione, ubiquitinazione e sumoilazione: "PUMAS". Tali modifiche vanno a colpire la struttura cromatinica e il reclutamento di proteine specifiche su di essa. Le modifiche epigenetiche sono molto veloci e reversibili attraverso enzimi e sono alla base di una veloce e precisa regolazione dell'attività genica.

1.4.2 Acetilazione

La maggior parte della cromatina possiede istoni acetilati, specialmente nelle code *H3* e *H4*. È associata con una trascrizione attiva: l'eucromatina è più acetilata. L'acetilazione di code e nuclei ha effetto sulla struttura cromatina:

- Direttamente: neutralizza le cariche positive sulla lisina sulla coda istonica riducendo le interazioni tra le code e il DNA rendendo la cromatina più accessibile da proteine leganti il DNA.
- Indirettamente: la lisina acetilata agisce come un sito di riconoscimento e legame per proteine contenenti bromodomini o lettori che possono reclutare altre proteine, componenti di grandi complessi che regolano la trascrizione come *HAT*, complessi di rimodellamento della cromatina e fattori di trascrizione che agiscono come *HAT*.

L'enzima responsabile per l'aggiunta di un gruppo acetile (mono-acetilazione) al gruppo ammino + NH_3 della lisina è l'istone acetiltrasferasi *HAT*, mentre l'istone deacetilasi *HDAC* lo rimuove. L'acetilazione della lisina pertanto neutralizza direttamente la carica positiva di essa riducendo l'attrazione tra DNA PO_4^- e lisina NH_3^+ . Inoltre diventa un sito di legame per proteine con bromodominio e rimodellatrici della cromatina aprendola e attivando la trascrizione. La deacetilasi agisce come repressione della trascrizione. La (de)acetilazione in regioni promotrici ha un ruolo nell'iniziazione della trascrizione. Altre acetilazioni sono presenti lungo sequenze codificanti, con

ruolo nell'allungamento della trascrizione. Se ne trovano ancora in enhancers o in varianti istoniche che presentano trascrizione attiva. Le proteine contenenti un bromodominio possono legarsi a una o più lisine acetilate attraverso il dominio e contengono altri domini come un dominio *PHD* che si lega a lisine metilate.

1.4.3 Metilazione

La metilazione sugli istoni avviene grazie a un istone metiltrasferasi *HMT* che può aggiungere 1, 2 o 3 gruppi metile sul gruppo amminico della lisina *K* o 1 o 2 gruppi metile sul gruppo amminico dell'arginina *A*. La metilazione della coda e del core istonico ha due effetti sulla struttura cromatinica:

- Diretto: mantiene la carica locale della lisina positiva compattando il legame tra istoni e DNA.
- Indiretto: proteine contenenti cromodomini (*HP1*, *Polycomb*) riconoscono e legano a specifiche lisine metilate e reclutano proteine che causano il silenziamento trascrizionale (togliendo spazio al legame con fattori di trascrizione) o la sua attivazione.

La metilazione è associata sia con attivazione che con repressione della trascrizione in base al residuo che è metilato:

- Mono-metilazione di *K9* nella coda *H3* causa una cromatina attiva trascrizionalmente.
- Mono- o tri-metilazione di *K4* nella coda *H3* causa una cromatina attiva trascrizionalmente.
- Di- o tri-metilazione di *K9* nella coda *H3* causa una cromatina silente trascrizionalmente.

1.4.4 Fosforilazione

I fosfati sono aggiunti da chinasi e rimossi da fosfatasi. La fosforilazione aggiunge una carica negativa alla coda istonica. Fosforilazione di *S10* nella coda *H3* permette la crescita cellulare e trascrizione promuovendo l'acetilazione di *K14* sulla coda *H3*. La fosforilazione di *S10* e *S27* nella coda *H3* è correlata con la condensazione dei cromosomi durante la mitose. È importante per la replicazione e riparazione del DNA e per l'apoptosi.

1.4.5 Ubiquitinazione e sumoilazione

L'ubiquitinazione delle lisine consiste nell'aggiunta di una proteina di 76 amminoacidi catalizzata dall'ubiquitina ligasi e rimossa dalla de-ubiquitinasi. Il suo ruolo non è compreso a fondo e avviene specialmente nelle code C-terminali di *H2A* e *H2B*. Regola la trascrizione reclutando rimodellatori e risposte al danno del DNA. Una mono-ubiquitinazione di *H2A* causa repressione trascrizionale mentre se avviene a *H2B* causa un'attivazione indiretta in quanto richiesta per la mono-metilazione di *H3K4* e di *H3K79*. La sumoilazione della lisina è una modifica simile all'ubiquitinazione e gioca un ruolo nella regolazione di trascrizione e riparazione di DNA.

1.4.6 Codice istonico

Le grandi possibili modifiche in aggiunta con le loro interazioni porta alla definizione di un codice istonico in cui modifiche uniche definiscono certi stati di cromatina e di espressione genica.

1.5 Complessi rimodellatori dei nucleosomi

La cromatina compattata rappresenta una barriera per le proteine che devono accedere al DNA e pertanto inibisce processi come trascrizione. La composizione del nucleosoma, la compattezza del suo legame con il DNA e la sua locazione possono essere fisicamente cambiati da complessi di rimodellamento dei nucleosomi dipendenti da ATP. Questi complessi possono introdurre loop nel DNA avvolto intorno a un nucleo istonico, far scivolare il DNA lungo l'ottamero istonico o rimuovere l'intero ottamero o 1-2 proteine istoniche e trasferirle da qualche altra parte. Possono attivare o reprimere la trascrizione ma non sono usati per la replicazione.

1.5.1 Ruoli dei rimodellatori della cromatina

I complessi di rimodellamento della cromatina hanno diversi ruoli nello stato cromatinico. Possono intervenire dopo la deposizione degli istoni durante la maturazione dei nucleosomi portando a una loro spaziatura regolare. Possono inoltre alterare lo stato cromatinico riposizionando i nucleosomi, espellendoli completamente o solo alcune loro subunità. Possono inoltre compiere installazioni o rimozioni di varianti istoniche.

1.5.2 Sottofamiglie

Esistono diverse classi di rimodellatori dei nucleosomi, ma tutte contengono dei domini chiave:

- Dominio motore ATPasi come *Dexx* e *HELICc*.
- Bromodominio o cromodominio.
- Dominio legante actina *HSA*.
- Dominio per il legame alla coda istonica *SANT* e *SLIDE*.

Switch/sucrose non-fermentable

Il complesso *SWI/SNF* facilita l'accesso alla cromatina: fa scivolare ed espelle i nucleosomi per l'attivazione o repressione genica.

Imitation switch

Il complesso *ISWI* assembla e spazia i nucleosomi principalmente per la repressione della trascrizione.

Cromodominio elicasi legante il DNA

Il complesso *CDH* è usato per l'assemblaggio dei nucleosomi e la loro spaziatura, per l'accesso ai geni esponendo i promotori e l'editing attraverso l'incorporazione di *H3.3*. Aiuta i repressori a legarsi alla cromatina e reprimere i geni attraverso *HDAC* associate.

Richiedenti inositolo

Il complesso *INO80* interagisce con *HAT* per attivare la trascrizione. Interviene anche nell'assemblaggio e spaziatura dei nucleosomi oltre a sostituire *H2A* con *H2A.Z* per la riparazione del DNA.

1.6 Variazione nella struttura cromatinica

I cromosomi subiscono varie fasi di compattazione diversa durante il ciclo cellulare. Durante l'interfase, quando i cromosomi sono relativamente poco condensati, i geni sono trascritti e il genoma è replicato si trova un gran numero di compattazione lungo il cromosoma. Della trascrizione può avvenire nelle regioni eterocromatiche, ma la traslocazione di un gene da una regione eucromatica a una eterocromatica può prevenire attivamente la sua trascrizione. Il livello di compattamento della cromatina non è uniforme e l'epigenetica rappresenta il suo ultimo livello di regolazione.

1.6.1 Eucromatina

Si dicono eucromatiniche le regioni dove le fibre di 30nm formano domini radical loop formando cromatina a 300nm. Questa zona è trascrizionalmente attiva.

1.6.2 Eterocromatina

Nell'eterocromatina i domini radical loop sono ulteriormente compattati attraverso metilazione della coda istonica a formare una cromatina a 700nm. L'eterocromatina si divide in costitutiva, o regioni sempre eterocromatiche permanentemente disattivate rispetto alla trascrizione o silenti e facoltativa, o regioni di cromatina che cambiano stato tra eucromatina ed eterocromatina. Alcune zone dei cromosomi sono altamente eterocromatiche:

- Telomeri: regioni di DNA alla terminazione dei cromosomi.
- Peri-centromeri.
- Regioni con sequenze di DNA altamente ripetute come l'rDNA nei nucleoli.

1.6.3 Effetti della cromatina

La cromatina ha effetto su trascrizione, replicazione, ricombinazione e trasmissione dei cromosomi. Riarrangiamenti che spostano un'origine di replicazione nell'eterocromatina causano una replicazione tardiva, arrivando fino a ritardare la divisione cellulare. La ricombinazione coinvolge rotture e riunioni di DNA di diverse molecole. Le regioni eterocromatiche ne subiscono di meno, proteggendo la regione contro tale modifica, cosa che avviene come nei geni di ripetizione di DNA ribosomiale. I cromosomi devono essere completamente compattati affinché avvenga la trasmissione e segregazione dei cromosomi.

1.6.4 Nucleolo

Il nucleolo è la parte del nucleo che contiene i geni di rDNA. Gli esseri umani possiedono cinque cluster di rDNA vicino la fine di cinque cromosomi. Si dice regione organizzatrice dei nucleoli i trascritti di rRNA prodotti dalle ripetizioni dall'rDNA. rDNA codifica per l'rRNA ribosomiale e molte cellule possiedono migliaia di ripetizioni di rDNA per riuscire a soddisfare la richiesta di rRNA e produzione di ribosomi. Il nucleolo non è separato da una membrana: sono le proteine e le RNA ad esso specifiche che gli conferiscono diversi pattern di colorazione. Un sottoinsieme di ripetizioni di rDNA sono silenti trascrizionalmente ed eterocromatiche in modo da aumentare la stabilità delle regioni ripetute.

1.7 Metilazione del DNA

Il DNA può essere modificato chimicamente attraverso la metilazione, che avviene in batteri ed eucarioti. I gruppi metile possono essere aggiunti a residui di citosina per creare la *5-metil citosina* attraverso DNA metiltransferasi o DNA metilasi. La modifica è reversibile grazie alla DNA demetilasi. La metilazione è rischiosa in quanto può alterare il DNA permanentemente. Le citosine metilate infatti possono subire una spontanea deamminazione idrolitica che cambia la citosina in timina con cambio mutagenico.

1.7.1 DNA metilasi

Le DNA metilasi utilizzano un base flipping per accedere alla citosina: una citosina è fatta uscire dalla doppia elica: un amminoacido dell'enzima è inserito temporaneamente al suo posto. La citosina viene poi metilata e reinserita nel DNA.

1.7.2 Effetti della metilazione

Nei procarioti

Nei procarioti la metilazione del DNA distingue il DNA appena sintetizzato nel processo di riparazione: appena dopo la replicazione solo il filamento genitore è metilato: questa regione si dice emi-metilata. Quando gli enzimi di riparazione del mismatch ne trovano uno leggono lo stato metilato per identificare correttamente il filamento parentale e riparare quello appena sintetizzato. La metilazione permette anche ai batteri di distinguere il DNA genomico da quello virale invasivo: enzimi di restrizione tagliano il DNA del fago a siti di riconoscimento specifici e durante il taglio il batterio protegge il proprio DNA metilando i siti di restrizione.

Negli eucarioti

La metilazione del DNA negli eucarioti silenzia la trascrizione. È pertanto un'altra forma di silenziamento epigenetico. Non cambia la carica della base e l'effetto repressivo è indiretto in quanto comporta il reclutamento di proteine lettrici che riconoscono e legano la base metilata. La metilazione avviene tipicamente a siti *CpG* o *CpXpG*, dove *p* è il legame fosfodiesterico e *X* una base qualsiasi. Circa il 60% delle *CpG* umane sono metilate. La metilazione può anche essere ereditata. Alcuni complessi si legano specificatamente a DNA metilato come enzimi di modifica istonica e complessi di rimodellamento della cromatina. Alcune proteine leganti istoni possono reclutare DNA metiltransferasi.

Isole *CpG* Le sequenze *CpG* non sono distribuite uniformemente nel genoma ma si trovano in lunghezze di 1-2kb dove il 60% del contenuto di DNA forma queste isole *CpG*. Sono studiate principalmente per la disattivazione del cromosoma *X*, si trovano in tutti i geni housekeeping, principalmente nella zona 5' nel promotore. Sono principalmente hypo-metilate, protette dalla metilazione e si correla con un'alta attività di trascrizione. *CpG* sono riconosciute da proteine *MBD* (metil-CpG-binding domain) con un dominio di legame di DNA e di un dominio repressore della trascrizione che possono reclutare complessi di rimodellazione della cromatina che disattivano la trascrizione. La metilazione può anche proibire il legame con fattori di trascrizione alle proprie sequenze di riconoscimento del DNA in un processo di mascheramento di *C*. La demetilazione avviene quando un gene deve essere trascritto. La metilazione di *CpG* è ereditata grazie all'enzima DNA metiltransferasi *DNMT1* che riconosce il sito emimetilato e lo rende completamente metilato.

1.7.3 La disattivazione del cromosoma X è un esempio di silenziamento epigenetico della metilazione di cromatina nei mammiferi

Un cromosoma X in ogni cellula è disattivato nelle femmine in modo che abbiano la stessa quantità di prodotto di gene X come nei maschi che ne possiedono uno solo. Il DNA è altamente metilato, $H2A$ è sostituito con *MacroH2A-Z*, gli istoni sono modificati come in eterocromatina ed avviene una regolazione basata su long non-coding RNA. La disattivazione del cromosoma X è casuale ed avviene alla gastrulazione nell'embrione, ognuna delle cellule possono scegliere individualmente quale dei due cromosomi X disattivare e la scelta viene ereditata. Uno dei due cromosomi si presenterà pertanto più denso, compatto e su un lato del nucleo. Circa il 15% dei geni legati a X non vengono disattivati completamente e la loro attività genica varia tra i cromosomi disattivati. La maggior parte di questi si trova nelle regioni pseudoautosomiali *PAR*, dove X e Y si accoppiano durante la meiosi.

Non corretta disattivazione di X durante la gastrulazione

Gatti calico La colorazione rossa del pelo dei gatti è dovuta a un gene nel cromosoma X , in cui l'allele rosso sintetizza un enzima che crea il pigmento arancio, mentre un altro non lo esprime e causa una colorazione nera. Nel caso in cui un gene X non sia disattivato e i maschi presentano XXY presentano una colorazione arancio e nera, oltre ad essere sterili.

1.7.4 Imprinting genetico

Una parte dell'attività genetica è controllata dall'imprinting genetico, che regola l'espressione di geni materni e paterni nell'embrione, casualmente in alcune cellule è silenziata la copia materna, in altre quella paterna. Se una delle copie di un gene è silenziata e l'altra è stata deleta non si trova espressione genica.

Disordini fisici e neurologici dovuti alla misregolazione dei geni soggetti a imprinting attraverso metilazione di citosina

Sindrome di Rett Questa sindrome è dovuta a una mutazione disattivante in un allele del *MECP2*. Avviene quando *MECP2* in un allele non è espresso a causa di metilazione. Uno di questi geni deve essere sempre espresso per la vitalità.

Sindrome di Prader-Willy e di Angelman In queste due sindromi sono colpiti gli stessi alleli del cromosoma 15. *PWS* avviene quando una regione paterna di 7 geni è eliminata. *AS* avviene quando è eliminata la regione materna. Le sindromi si manifestano quando l'altro allele parentale è espresso sub-ottimamente a causa dell'imprinting. Un insieme allelico parentale deve essere intatto per la sopravvivenza dell'embrione.

1.8 La separazione dei domini cromatinici da elementi barriera

1.8.1 Variegazione da effetto di posizione

Un effetto epigenetico è la variegazione da effetto di posizione. Un suo esempio è il colore dell'occhio di *Drosophila* in cui si presentano rossi grazie all'espressione del gene *white*⁺. In alcuni casi gli occhi

possono presentare sfaccettature bianche se il gene viene convertito in una regione eterocromatica in qualche cellula in cui risulta silenziato.

1.8.2 Elementi di barriera

Le cellule possiedono elementi di barriera che separano eu ed eterocromatina. Questi elementi possono prevenire la diffusione dell'eterocromatina. In *S. pombe* due elementi di barriera affiancano una regione di eterocromatina silente intorno al centromero. Gli *H3* negli elementi di barriera sono altamente metilati a *K9* silenziando la regione, mentre quelli fuori la barriera sono altamente metilati a *K4* attivando la regione. La rimozione di questi elementi permette la diffusione di metilazione *K9* e della zona silenziata. L'eterocromatina può infatti diffondersi attraverso modifiche di istoni successive come deacetilazione di *H3* la sua metilazione a *K9* e il legame della proteina di silenziamento *Swi6*. Gli elementi di barriera agiscono come barriere fisiche e possono essere sequenze specifiche a cui si legano proteine regolatrici delle modifiche istoniche o grandi loop di cromatina. Elementi di sequenze di barriera possono ancorare gli anelli nella lamina nucleare: l'eterocromatina si trova nelle regioni periferiche del nucleo in quanto *SAR/MAR* affiancano spesso elementi di sequenza di barriera.

1.9 Elementi richiesti per la funzione dei cromosomi

1.9.1 Origine di replicazione

Le origini di replicazione sono regioni del DNA con sequenze specifiche richieste per la replicazione in batteri ed eucarioti. Le *Ori* sono dove il dsDNA è svolto e separato per prepararsi all'attacco delle proteine di replicazione. La replicazione è bidirezionale:

- Nei batteri si trova un *Ori* per cromosoma e si indica con *ter* gli elementi di terminazione della replicazione.
- Negli eucarioti si trovano diverse *Ori* lungo il cromosoma in quanto si ha più cromosoma da replicare

1.9.2 Centromeri

I centromeri si trovano in tutti i cromosomi eucarioti. Sono sequenze che si trovano tipicamente al centro del cromosoma e sono necessarie per la segregazione dei cromosomi durante la divisione cellulare. La maggior parte delle specie ne possiedono 1 per cromosoma. Si trova in una regione eterocromatica e dopo la replicazione del DNA 2 cromatidi fratelli si formano e sono uniti dai complessi di anelli di coesina. Il centromero appare come una costrizione dovuta all'arricchimento locale di coesina che si trova in una quantità minore lungo l'intero paio di cromatidi. Il centromero ha dimensioni variabili a seconda della specie e si distingue in:

- Centromero puntiforme con sequenze definite di poche centinaia di basi.
- Centromero regionale con centinaia di kilobasi.

Alcuni organismi possiedono molti centromeri lungo il cromosoma e sono detti olocentrici. I microtubuli si attaccano su tutta la lunghezza del cromosoma. Durante la mitosi si possono segregare frammenti di cromosomi.

Il cinetocore

Il centromero recluta più di 100 proteine che formano il cinetocore che attacca i cromatidi fratelli ai microtubuli che si estendono da poli opposti del fuso, permettendo ad esso di separare i cromatidi attraverso la depolimerizzazione dei microtubuli. Questo meccanismo di segregazione è altamente conservato.

Esempi di centromeri

S. cerevisiae Il centromero è lungo 125bp e possiede tre regioni *CDEI*, *CDEII* e *CDEIII*, la prima e la terza possiedono sequenze altamente conservate e singole mutazioni possono rompere la funzione del centromero. *CDEII* invece è una regione ricca di *AT* e la sequenza esatta non è fondamentale.

S. pombe Ogni cromosoma di *S. pombe* presenta un cromosoma con una sequenza di centromero leggermente diversa con un nucleo unico di 5-6kb con lunghe sequenze di ripetizioni inverse che lo affiancano.

Esseri umani I centromeri sono lunghi 1Mb e sono fatte di sequenze ripetute dette ripetizioni α -satellite, lunghe 171bp ordinate in ripetizioni di ordine più alto. I nucleosomi centromerici possiedono varianti istoniche di *H3* *CENP-A* particolarmente nelle regioni ricche di *AT* che potrebbe riconoscere gli *i-motivi* e diadi. Il centromero marcato da *CENP-A* è dove il cinetocore si assembla. Una sovraespressione di *CENP-A* causa un legame del cinetocore con tutto il cromosoma e una sua rottura durante la segregazione.

1.9.3 Telomeri

I telomeri sono regioni alle terminazioni dei cromosomi lineari e funzionano come cappucci protettivi. Negli esseri umani sono formati da sequenze ripetute centinaia di migliaia di volte di *TTAGGGG*, marcano la terminazione del cromosoma definendolo e impedendo la fusione di cromosomi alle loro terminazioni. Infatti più un cromosoma è lungo più è propenso a subire rotture. Il DNA dei telomeri consiste di un filamento ricco di *G* e uno ricco di *C*. La lunghezza totale delle ripetizioni varia tra i 50 000 e i 30 000bp in base alla specie. La sequenza ricca di *G* si estende 5'-3' verso la terminazione del cromosoma dove termina in una regione corta a filamento singolo. Negli organismi con telomeri lunghi questa regione può essere processata in un rolled back T-loop formato dall'invasione e accoppiamento di basi del filamento singolo con la sequenza a doppio filamento a monte. Le ripetizioni sono un sito di legame per proteine che le marcano come terminazioni naturali distinguendoli dalle rotture del DNA. La DNA polimerasi non può copiare la terminazione di una molecola di DNA e pertanto interviene la telomerasi per mantenere le terminazioni dei cromosomi.

Telomerasi

Le proteine *TRF1* e *TRF2* (*TTAGGGG* repeat binding factor), *TIN2* e *RAP1* si legano ai telomeri e proteggono le loro terminazioni. La terminazione di ogni telomero forma un T-loop composto da una ripetizione *TTAGGGG* 3' a filamento singolo che lega una sequenza complementare in una sequenza a monte denaturata detta D-loop (displacement loop) e viene stabilizzata da copie multiple di *POT1* che vi si lega. La telomerasi è una speciale DNA polimerasi che possiede una proteina e una componente a RNA: forma un *RNP*. L'RNA della telomerasi fornisce un corto stampo che specifica la sequenza della ripetizione telomerica che deve essere aggiunta. La telomerasi pertanto sintetizza il DNA telomerico usando l'RNA come stampo. La lunghezza dei telomeri è mantenuta

1.9. ELEMENTI RICHIESTI PER LA FUNZIONE DEI CROMOSOMI

nelle cellule staminali e germinali, mentre nei tessuti maturi si trova una telomerasi insufficiente e avviene un accorciamento dei telomeri. Che limita il numero di divisioni cellulari che la cellula può avere. Una sovraattivazione della telomerasi è implicata in molti cancro e permette alle cellule di continuare a crescere e a dividersi.

Capitolo 2

Replicazione

2.1 Replicazione del DNA semi-conservativa

Durante la divisione cellulare l'informazione genetica deve essere copiata e distribuita equamente tra le cellule figlie. Dopo la scoperta della struttura a doppia elica del DNA si ragionò come i due filamenti complementari sono copiati e replicati.

2.1.1 Modelli di replicazione

- Replicazione conservativa: il DNA rimane intatto come un doppio filamento e agisce come stampo.
- Replicazione semi-conservativa: un filamento agisce come stampo per sintetizzare un nuovo filamento complementare.
- Replicazione dispersiva: il doppio filamento si rompe nella sua lunghezza e frammenti sovrapposti servono come stampi per la sintesi.

Determinazione del modello semi-conservativo

Per determinare il modello di replicazione Meselson e Stahl idearono un esperimento che utilizzava un gradiente di densità e ultra-centrifugazione.

Tecnica del gradiente di densità Gli scienziati presero un tubo di plastica in cui era presente un gradiente del sale cloruro di cesio. In questo modo aggiungendo componenti di diversa densità in cima al gradiente e centrifugandoli ad alta velocità la forza gravitazionale trasporta le componenti nel gradiente in modo che si fermino quando la loro densità è uguale alla densità locale della soluzione di *CsCl*. Le componenti a bassa densità sono posizionate più in alto nel gradiente, mentre quelle a densità più alta in basso. Quando le componenti si sono mosse attraverso il gradiente durante la centrifugazione si prendono campioni dal basso all'alto attraverso frazionamento.

Rendere il DNA più pesante Per rendere il DNA di nuova sintesi più pesante rispetto a quello originale viene utilizzato un isotopo dell'azoto ^{15}N in quanto è la massa dell'elemento più frequente e può essere sintetizzato in forma radioattiva.

L'esperimento Gli scienziati fecero crescere una coltura di *E. coli* per 4 divisioni cellulari in un medio minimale contenente glucosio e con $^{15}\text{NH}_4\text{Cl}$ come l'unica fonte di azoto. Il DNA alla fine pertanto conterrà ^{15}N nelle basi nucleotidiche. Prendendo un campione della coltura della quarta divisione cellulare e isolandolo. Successivamente si isola il resto dei batteri attraverso centrifugazione, li si lava e risospende nel medio minimale con $^{14}\text{NH}_4\text{Cl}$ come unica fonte di azoto. Si lascia crescere e dividere la coltura così ottenuta prendendo campioni ogni divisione. Si isola il DNA dai vari campioni e lo si carica su un tubo a gradiente di CsCl separato. Successivamente si ultracentrifugano tutti i tubi in parallelo, si fraziona i campioni e si fa correre il DNA su un gel di agarosio e li si trasferisce su membrana di nitrocellulosa. Infine si espone la membrana a un foto film.

Conclusioni Si nota come in base al modello si osserverebbero comportamenti diversi:

- Modello conservativo: il numero di batteri con ^{15}N rimarrebbe costante e aumenterebbe quella con ^{14}N , presentando pertanto due bande, una per l'isotopo e una per ^{14}N .
- Modello dispersivo: i batteri presenterebbero tutti del DNA ibrido contenente sia ^{15}N che ^{14}N , presentando pertanto una banda unica all'ibrido.
- Modello semi-conservativo: si troverebbe nella popolazione un numero di molecole contenenti uno strand con ^{15}N e l'altro ^{14}N , mentre il resto tutto a ^{14}N , pertanto si noterebbero due bande, una per ^{14}N e una per l'ibrido.

Si osserva che avviene il terzo caso, determinando che la replicazione è semi-conservativa.

2.2 Il modello dei repliconi

Il modello dei repliconi è stato proposto nel 1963. Si indica con replicone la parte del DNA che sta venendo replicata. La replicazione inizia a una particolare sequenza di origine o replicatore. Una proteina iniziatrice si lega al replicatore per iniziare il processo di replicazione.

2.2.1 Scoperta del modello

La scoperta del modello dei repliconi avviene grazie a Cairns nel 1963 attraverso un'analisi autoradiografica del genoma in replicazione di *E. coli*. Si cresce la cellula in un medium contenente glucosio e azoto. Si aggiunge ad essa $[^3\text{H}]\text{-timidina}$ e si fanno avvenire due replicazioni del DNA in modo che questa si incorpori due volte nel filamento di nuova sintesi. Si lisa la cellula e la si espone a un foto film per due mesi.

Osservazioni

Si nota come dopo una replicazione un filamento non è radioattivo mentre l'altro lo è. All'inizio della seconda replicazione si forma una sezione con entrambi i filamenti radioattivi, permettendo di visualizzare come il DNA si replica in maniera semi-conservativa nella cellula. Nel replicone o *Ori* il DNA si apre formando bolle tra due forcelle di replicazione. La bolla si estende in maniera bidirezionale. Lo si nota osservando la radioattività ai due estremi della bolla di replicazione.

2.2.2 Origini di replicazione

Origine di replicazione singola

In caso di una singola *Ori* in DNA circolare questo si comincia a svolgere in tale sequenza producendo una bolla di replicazione con una forcella ad ogni terminazione. Le forcelle procedono lungo il cerchio producendo il modello Θ . Successivamente interviene la topoisomerasi II *girasi* che separa le due molecole di nuova formazione.

Origini di replicazioni multiple

In caso di multiple *Ori* in DNA lineare si formano varie bolle di replicazione con forcelle ad ogni estremità. Le bolle mano a mano che ne incontrano altre si fondono tra di loro.

2.3 Identificazione delle origini di replicazione

Nei genomi di batteri, batteriofagi, virus e plasmidi si trova un *Ori* per molecola di DNA, nei primi in quanto il loro DNA si replica in maniera indipendente da quella dell'ospite. L'origine di replicazione in *E. coli* o *OriC* è stata trovata attraverso un esperimento.

2.3.1 Esperimento

Si prende una coltura di *E. coli* e la si trasforma con un plasmide contenente del DNA di *E. coli* ottenuto attraverso enzimi di restrizione e un gene che codifica la resistenza all'ampicillina. In questo modo ogni colonia che cresce in presenza dell'antibiotico contiene un *Ori* nel plasmide. Si continua a ridurre la lunghezza del frammento fino a che non si verifica più la resistenza. In questo modo si riesce a determinare la sequenza minima e specifica dell'*Ori*. In *E. coli* è lunga 245bp e la parte che si apre formata da $3 \times 13bp$ è ricca in *A* e *T* in quanto le basi formando solo due legami a idrogeno sono più facili da aprire. La sequenza contiene inoltre $5 \times 9bp$ siti di legame per 5 proteine iniziatrici *DnaA*.

2.3.2 Origini di replicazione negli eucarioti

Si nota come i lunghi cromosomi lineari degli eucarioti possiedono multiple origine di replicazione in modo da replicare il DNA in un tempo ragionevole. Attraverso autoradiografia si identificano diverse origine di replicazione attraverso le bolle con diversa dimensione in base al tempo di formazione: se precoce o tardiva. Le forcelle di replicazione si muovono comunque in maniera bidirezionale e si uniscono tra di loro quando si incontrano. La lunghezza di repliconi individuali è di 100bp in lievito e mosche e tra i 75 000 e 175 000 in cellule animali e umane. Il tasso di replicazione negli eucarioti è di $2000 \frac{bp}{min}$, molto più lento rispetto ai batteri. Si nota come dalla velocità di replicazione il genoma di un mammifero potrebbe essere replicato in un'ora. Nonostante questo la fase *S* dura più di 6 ore in una cellula somatica. Questo avviene in quanto non più del 15% dei repliconi sono attivi in un dato momento. Ci sono eccezioni come le divisioni degli embrioni di *Drosophila*, con fase *S* molto più breve.

Identificazione degli *ARS* nel lievito *S. cerevisiae*

Si intende con *ARS* la sequenza replicante autonomamente o *Ori*. L'identificazione avviene in maniera simile a quella dell'*Ori* di *E. coli*: frammenti di DNA ottenuti attraverso enzimi di restrizione

2.4. PANORAMICA DELLA REPLICAZIONE DEL DNA

vengono introdotti un un plasmide e in cellule del lievito incapaci di crescere in coltura priva di istidina. Le colture in grado di crescervi contenevano un'origine di replicazione. Si nota come si trova un *ARS* ogni 135 000bs.

Ruolo della struttura cromosomica nella replicazione

Negli eucarioti non tutte le *Ori* sono utilizzate durante la replicazione: la loro attivazione è regolata nella fase *S* da proteine che regolano il ciclo cellulare, l'ambiente locale di cromatina (effetto di posizionamento). *ARS1* si trova vicino al centromero e *ARS501* vicino a un telomero. Il cambio di posizione cambia il momento di firing dell'*ARS*. Nel lievito le *ARS* vicine al centromero si attivano precocemente. Fattori di replicazione, modificatori della cromatina e rimozioni degli istoni leggono il codice istonico aprendo e chiudendo la cromatina determinando domini di replicazione precoce e altri di replicazione tardiva.

Mappatura fisica di *Ori* attraverso elettroforesi su gel d'agarosio

Questa tecnica, detta anche ibridizzazione del Southern blot traccia le aperture e i movimenti di un *Ori* in un pezzo di DNA durante la replicazione. Per farlo si isola il DNA genomico da una cultura asincrona e lo si taglia con un enzima di restrizione specifico. Si traccia il DNA con una sonda con un Southern blot su gel 2D. Ogni cellula rappresenta uno stato intermedio di attività di replicazione locale. Sul medium a 2D con agarosio con presente *EtBr* più denso in modo che la forma del frammento influisca sulla sua posizione. Il DNA osservato può assumere diverse forme in base allo stato replicativo:

- Y semplice con un grafico ad arco Y che rappresenta un replicatore passivo.
- Y doppia con un grafico ad arco a doppia Y che rappresenta una terminazione.
- Bolla simmetrica con un arco a bolla che rappresenta un'iniziazione.
- bolla asimmetrica con un arco a bolla e transizione ad arco ad Y che rappresenta un'iniziazione con un'origine non centrata.

2.4 Panoramica della replicazione del DNA

Affinchè la cellula si divida deve avvenire la replicazione del DNA. Si intende per replicazione la completa e fedele copia del DNA nei cromosomi della cellula. La replicazione è semi-conservativa: ogni filamento della doppia elica parentale agisce come stampo per la sintesi di un nuovo filamento per la cellula figlia. In modo da copiare lo stampo la base dello stampo deve essere identificata e deve essere aggiunta la base complementare. Questo garantisce a meno di mutazione che ogni doppia elica figlia sia identica a quella parentale e che ogni cellula figlia riceva molecole di DNA identiche.

2.4.1 Le fasi della replicazione del DNA

Iniziazione

Durante l'iniziazione viene riconosciuta l'origine di replicazione da una proteina iniziatrice che apre la doppia elica localmente e recluta elicasi. Le DNA elicasi continuano a svolgere l'elica per esporre DNA a singolo filamento che è circondato da proteine leganti *ssDNA*. L'iniziazione è controllata in modo che avvenga una sola volta per ogni ciclo cellulare. La sintesi del DNA necessita di un primer

in quanto può aggiungere nucleotidi a una terminazione $3' - OH$ esistente. Il primer è un piccolo filamento di RNA sintetizzato da una DNA primasi.

Allungamento

Dopo la sintesi del DNA primer la pinza scorrevole simile ad un anello è reclutata all'ibrido a doppio filamento ssDNA RNA primer. La DNA polimerasi si lega al DNA attraverso la pinza e il macchinario di replicazione o replisoma si muove lungo il DNA copiando i filamenti. Ogni base nel DNA parentale è letta dalla DNA polimerasi che aggiunge basi complementari al filamento in crescita in una direzione $5' - 3'$.

Terminazione

La terminazione avviene quando due forcelle diverse si incontrano o quando questa raggiunge la terminazione del cromosoma lineare. Il complesso di replicazione è disassemblato, i primer a RNA sono rimossi e sostituiti con DNA e la DNA ligasi connette le sequenze di DNA di nuova sintesi.

2.5 Iniziazione

Le origini di replicazione sono i siti in cui il DNA è inizialmente svolto. Le proteine iniziatrici si legano alle origini a siti di legame degli iniziatori permettendo il legame con l'elicasi e continuando a svolgere il DNA. Alcuni organismi hanno specifiche sequenze come origine ma è l'abilità di legare la proteina iniziatrice che definisce un'origine. Le proteine iniziatrici sono proteine leganti $ATP AAA^+$. In *E. coli* si chiama *DnaA*. Negli eucarioti l'iniziatore è il complesso di riconoscimento dell'origine *ORC*, in *S. cerevisiae* ha 6 subunità e si chiama *Orc1-6*. L'*ATP* regola il legame dell'iniziatore. Il legame di *ATP* con *Orc1* è richiesto per il legame di *ORC* con il DNA. Spesso le origini di replicazione possiedono una sequenza di DNA definita con DNA unwinding element, regioni ricche di *AT* che facilitano lo svolgimento in quanto possiedono solo 2 legami a idrogeno. In queste regioni l'iniziatore separa i due filamenti quando si lega al DNA e recluta altre proteine.

2.5.1 Svolgimento dell'*Ori* nei procarioti - *E. coli*

In *E. coli* l'origine di replicazione *OriC* è una sequenza di 245bp con 5 *DnaA* box di 9bp, 3 con alta affinità e 2 con bassa che legano 15 molecole di *DnaA*. Nelle 3 ad alta affinità il *DnaA* è sempre legato ad essi, mentre in quelle a bassa affinità si lega *DnaA-ATP* solo quando la replicazione deve iniziare. Tutte le proteine *DnaA* legano *ATP* e si multimerizzano in un filamento a spirale. Il filamento distorce il DNA producendo un superavvolgimento positivo locale svolgendo a valle la regione ricca di *AT* di $3 \times 13bp$ che subisce invece un superavvolgimento negativo. Successivamente *DnaA-ATP* e 6 molecole di *DnaC-ATP* caricano l'anello omoesamerico formato dall'elicasi *DnaB* sui singoli filamenti dell'origine. *DnaC* successivamente lascia l'*OriC* a seguito dell'idrolisi dell'*ATP*. *DnaB* recluta la DNA primasi *DnaG* che sintetizzerà il RNA primer. La pinza scorrevole si lega alla sequenza ibrida ssDNA-RNA primer.

Regolazione dello svolgimento di *OriC*

Inattivazione regolata di *DnaA* (*RIDA*) La replicazione del DNA è un punto di non ritorno: si devono prevenire rireplicazioni alla stessa origine e deve essere integrata con le altre attività di divisione cellulare. In *E. coli* il discriminante è la presenza di *DnaA-ATP* contro *DnaA-ADP*: solo

il primo può multimerizzarsi e svolgere il DNA. Dopo l'iniziazione l'*ATPasi* AAA^+ *HDA* lega la pinza e stimola l'idrolisi in *DnaA-ADP* che si dissocia dall'*OriC* e non può riattivare la replicazione. *RIDA* è il principale meccanismo regolatorio che previene la ri-replicazione nei procarioti.

Metilazione di *OriC* L'iniziazione può essere prevenuta attraverso metilazione del DNA: la DNA adenine metilasi *DAM metilasi* metila i residui *A* nella sequenza *GATC* lungo il genoma di *E. coli*. 11 siti *GATC* in *OriC* sovrappongono i siti di legame di *DnaA*. Dopo la replicazione solo un filamento di DNA è metilato (emimetilazione). La proteina *SeqA* si lega ai siti *GATC* emimetilati e blocca il legame della *Dam metilasi*. In questo modo previene la metilazione di entrambi i filamenti e il legame di *DnaA* con le sue box. Il blocco è temporaneo: l'origine è completamente metilata sul nuovo filamento dopo 10 minuti causando la dissociazione di *SeqA* dalla doppio filamento. I cromosomi completamente metilati si segregano nelle cellule figlie e sono in grado di legare *DnaA*.

Sequestro di *DnaA* a *datA* Il *DnaA* può legarsi alla sequenza di DNA *datA* che si trova vicino all'*OriC*. La regione può legare 370 molecole di *DnaA*. Dopo la replicazione *DnaA* viene sequestrato alla regione *datA* e non è disponibile per il legame in *OriC*. All'inizio di un nuovo ciclo di replicazione cambi locali nella sequenza *datA* causano una dissociazione di *DnaA* che può legare *OriC*. Il *DnaA* è attivato in *DnaA-ATP*.

2.5.2 Svolgimento dell'*Ori* negli eucarioti

Le origini di *S. cerevisiae* sono simili a quelle di *E. coli*. Sono lunghe tra 100 e i 200bp, siti *A* e *B1* a cui si lega *ORC*. I siti *B2* e *B3* sono adiacenti ad esse e ricchi di *AT*. *ORC* si lega intorno ai siti *A* e *B1*. La struttura cromatinica è importante per le origini negli organismi multicellulari: la replicazione inizia in regioni cromosomiche specifiche, ma le sequenze di DNA non sono conservate. In *Drosophila* *ORC* si lega a sequenze *Ori* come le code istoniche sono iperacetilate, pertanto *HAT* potrebbero essere coinvolti nella regolazione delle origini di replicazione.

Regolazione dell'attivazione di *Ori* negli eucarioti

Gli eucarioti possiedono multiple origini di replicazione. Ognuna di esse deve attivarsi una volta per ciclo cellulare. Sono selezionate in G_1 e attivate in *S*. Le origini non possono essere riutilizzate fino alla rifezione nella fase G_1 successiva.

Selezione Durante la fase G_1 *ORC* si lega a un'origine. *ORC* è costituito da 6 subunità: *Orc1-6*. 5 di esse sono AAA^+ *ATPasi*. *ORC* sottostà alla formazione del complesso *pre-RC*, formato da *ORC*, *Cdt1*, *Cdc6* e l'elicasi *MCM2-7*.

Attivazione Durante la fase *S* nel *pre-RC* le elicasi *Mcm2-7* vengono fosforilate dalla chinasi *Dbf4 DDK*. Sono reclutate *Sld2* e *Sld3* per formare il complesso *SDS* al *pre-RC* e vengono fosforilate dalla chinasi dipendente dalla fase *S* o *S-Cdk*. Entrambe le fosforilazioni risultano in una completa attivazione del *pre-RC*. *Cdt1* e *Cdc6* lasciano il complesso. Avviene il reclutamento dei fattori di iniziazione *Cdc45* e dei complessi *GIN5* e *SDS* che permette l'apertura e svolgimento dell'origine. La replicazione deve essere completa prima che avvenga la segregazione. Se avviene uno stallo delle forcelle viene attivata la risposta al danno del DNA e l'entrata in mitosi è bloccata fino alla correzione degli errori.

Rif1 (Rap1 interacting factor 1) Questa proteina regola positivamente l'attivamento delle origini precoci e negativamente quello delle tardive. Previene il reclutamento del fattore di iniziazione *Cdc45* al *pre-RC*. È conservata nelle cellule umane ed è il regolatore chiave del programma di replicazione del DNA, ovvero dell'ordine temporale di attivazione delle *Ori*.

Regolazione dell'iniziazione della replicazione negli eucarioti multicellulari

Essendoci molte origini lungo un cromosoma lineare l'ordine di attivazione dipende da:

- *Rif1*.
- Lo stato di acetilazione della cromatina: *HAT* sono richieste per il reclutamento delle proteine del *pre-RC*.
- Lo stato di trascrizione: le *Ori* si trovano spesso vicino a siti di inizio della trascrizione e la forcella di trascrizione produce uno stato di superavvolgimento positivo a monte e negativo a valle, semplice da aprire per la replicazione.

In G_1 le origini selezionate che verranno attivate nella fase *S* sono determinate dal punto di decisione delle origini *ODP*. Specifici domini di replicazione sono regioni di DNA cromosomiale che contengono origini attivate nello stesso momento. Le proteine richieste per la replicazione sono concentrate in strutture nucleari dette replication factories, dove i domini di replicazione co-localizzano con 14 forcelle di replicazione per factory. La pinza scorrevole resa fosforescente *PCNA-GFP* rende visibile i foci di replicazione al microscopio. Le origini dormienti presentano un *pre-RC* assemblato ma non sono attive, ma possono essere attivate velocemente quando una forcella di replicazione vicina entra in stallo. Nei mammiferi l'eucromatina nell'interno del nucleo è replicata precocemente, mentre l'eterocromatina nella periferia è replicata tardivamente.

2.5.3 DNA elicasi

Dopo l'apertura e attivazione all'*Ori* il dsDNA deve essere ancora svolto. Questa operazione viene catalizzata dalla DNA elicasi, un esamero che si lega a un DNA a singolo filamento aperto alla sequenza *Ori* che si muove in basso verso il doppio filamento per svolgerlo alla forcella di replicazione in svolgimento. L'energia necessaria è fornita dall'*ATP*. L'elicasi è inoltre responsabile del reclutamento di altre proteine richieste per la replicazione: il complesso replisomico. Il filamento rimanente viene legato da una proteina. In *E. coli* l'elicasi è la *DnaB* e possiede 6 subunità identiche (omoesamero), negli eucarioti ed archea l'elicasi è complesso *MCM2-7* e comprende 6 subunità diverse (eteroesamero). Ognuna delle 6 subunità lega *ATP* a coppie causando un cambio di conformazione, mentre l'idrolisi e rilascio di *ADP* causa un ritorno alla conformazione iniziale. L'elicasi assume un comportamento pulsante: svolge il DNA e si spinge in avanti. L'elicasi batterica si muove lungo il filamento principale, mentre quella eucariotica in quello lagging.

Risoluzione delle strutture secondarie

Essendo che ssDNA può formare strutture secondarie che rende la sua copia difficoltosa e più sensibile a danno proteico *SSB* nei batteri e *RPA* (replication protein A) negli eucarioti si devono legarsi come omotetrameri al ssDNA in modo da tenerlo aperto. Sono fisicamente rimosse durante la polimerizzazione del DNA.

Risoluzione del superavvolgimento

Mentre il DNA viene svolto viene introdotto uno stress torsionale in quanto la separazione dei filamenti risulta in un superavvolgimento a valle di essa. Questo rende più difficile per l'elicasi proseguire nella separazione. Per questo devono intervenire topoisomerasi che risolvono il problema rompendo transientemente il DNA e permettendo il rilassamento del superavvolgimento. La ligasi successivamente chiude il DNA. Nei procarioti sono presenti solo le topoisomerasi II, negli eucarioti anche le I.

2.5.4 Sintesi di primer a RNA o RNA-DNA dalla DNA primasi

Dopo l'apertura dell'*Ori* la DNA primasi *DnaG* in *E. coli* è reclutata dalla DNA elicasi. Agisce esclusivamente alla forcella di replicazione producendo una corta sequenza RNA o RNA-DNA. Non richiede un esistente 3' – OH per la sintesi a differenza della polimerasi. La primasi batterica possiede due subunità e crea un primer di 10-30 basi, mentre quella eucariotica ne possiede 3 e crea un primer misto DNA-RNA. Due subunità funzionano come primasi mentre una come DNA polimerasi α aggiungendo un primer di DNA all'RNA. L'attività della primasi è unita a quella dell'elicasi e insieme formano il complesso primosoma. Dopo che la subunità DNA polimerasi α della primasi eucariotica crea una corta lunghezza di DNA la DNA polimerasi replicativa III per i batteri e δ o ϵ negli eucarioti la sostituisce e sintetizza il resto del DNA nel processo di polimerase switching. Questo avviene ogni volta che un frammento di Okazaki viene creato tra i primer. La polimerasi replicativa viene reclutata dalla pinza scorrevole e determina l'inizio dell'allungamento.

2.6 Allungamento

2.6.1 Pinza scorrevole

L'allungamento inizia con il reclutamento della pinza scorrevole che permette l'alta processività della DNA polimerasi mantenendola stabilmente legata al DNA. A una sequenza di ssDNA stampo viene caricata la pinza da una proteina di caricamento. Sia la pinza che il suo caricatore sono conservati in batteri, archea ed eucarioti:

- Pinza scorrevole: pinza β nei procarioti e *PCNA* (proliferating cell nuclear antigen) negli eucarioti.
- Proteina caricatrice della pinza: fattore di replicazione *C*: *RFC*.

La pinza scorrevole è un anello con un buco da 35Å che racchiude il primer ssDNA. È molto stabile e rimane associata con il DNA una volta caricata. Affinchè possa associarsi al DNA il caricatore della pinza, una struttura ad anello a 5 subunità deve aprirla. Specifiche subunità del caricatore sono AAA^+ *ATPasi* e quando lega *ATP* causano cambi conformazionali che guidano il legame con la pinza, la sua apertura e il reclutamento al DNA. La pinza scorrevole, una volta legata al primer recluta la DNA polimerasi attraverso un motivo a 8 amminoacidi. Il caricatore della pinza ha bassa affinità per la pinza scorrevole fino a che non è legato all'*ATP*. Quando lo lega il caricatore lega la pinza. Il complesso ha un alta affinità per il primer a ssDNA.

Legame con il primer

Il legame con il primer stimola l'attività *ATPasica* del caricatore che chiude la pinza e si rilascia. L'idrolisi dell'*ATP* riduce l'affinità di *RFC* per il DNA. La pinza rimane associata con il DNA e

recluta l'oloenzima DNA polimerasi permettendo l'inizio dell'allungamento. Il caricatore della pinza rilasciato può essere ricaricato con *ATP* per ripetere il processo a un altro primer. Si dice oloenzima un complesso multiproteico in cui un enzima centrale è associato con componenti addizionali che ne aumentano la funzione.

2.6.2 Sintesi del DNA

Materiali richiesti

La sintesi del DNA richiede il complesso stampo a ssDNA e primer. In vitro il primer è a DNA, mentre in vivo è a RNA o RNA-DNA. Lo stampo è il filamento di ssDNA che è letto dalla DNA polimerasi. Oltre al complesso sono richiesti deossiribonucleotidi: i monomeri *dNTP* come *dATP*, *dCTP*, *dGTP* e *dTTP*. La sintesi avviene sempre nella direzione 5'-3', mentre la lettura nella direzione inversa.

Polimerizzazione

La polimerizzazione del DNA consiste nella formazione dei legami fosfodiesterici. αP in *dNTP* si lega alla terminazione 3'*OH* del primer e viene rilasciato pirofosfato $\gamma P-\beta P$. L'energia che spinge la reazione in avanti deriva dall'idrolisi del pirofosfato da parte della pirofosfatasi che causa l'irreversibilità della reazione.

2.6.3 DNA polimerasi

Le principali DNA polimerasi processive coinvolte nella replicazione del DNA sono la DNA polimerasi III nei batteri e le DNA polimerasi γ e ϵ negli eucarioti. Le DNA polimerasi rimangono attaccate al DNA grazie alla pinza scorrevole per lunghe sequenze prima di dissociarsi rendendo la polimerasi processiva. Le polimerasi processive sono altamente conservate e contengono multipli domini e regioni con diverse funzioni tra cui la ricerca di errori. I tre domini della DNA polimerasi sono detti pollice, dita e palmo che insieme assomigliano a una mano destra. Il DNA in crescita a doppio filamento "il braccio" si trova nel palmo mentre il ssDNA passa attraverso le dita. I domini delle dita aiutano a posizionare il nucleotide in arrivo, il pollice mantiene il dsDNA allungato ma non contribuisce alla reazione di polimerizzazione.

Funzione

L'aggiunta corretta di un nucleotide al filamento in sintesi attiva la polimerasi attraverso un cambio conformazionale: la mano rilascia il DNA dopo aver aggiunto il nucleotide e si muove di una base per leggere lo stampo da 3' a 5'. Il sito attivo della DNA polimerasi possiede gruppi carbossilati di due residui di aspartato con due ioni $2Mg^{2+}$. I siti attivi catalizzano un trasferimento di fosforile unendo il 5'*P* del nucleotide in arrivo al 3'*OH* del DNA in crescita per formare un legame fosfodiesterico. La reazione consiste nell'attacco nucleofilo dal 3'*OH* al α -fosfato del *dNTP* in arrivo, rilasciando i fosfati β e γ come pirofosfato. Gli ioni magnesio sono critici: uno attiva il priming di 3'*OH* abbassando il suo *pKa* e favorendo l'attacco nucleofilo, l'altro interagisce con l'ossigeno negativo dei gruppi fosfato $\beta\gamma$ e posiziona α -fosfato vicino al priming 3'*OH*.

2.6.4 Fedeltà della polimerizzazione del DNA

La DNA polimerasi processiva fa un errore ogni 100 000 nucleotidi. L'identità dei monomeri è controllata durante e dopo l'aggiunta di un *dNTP* al filamento in crescita. La fedeltà è mantenuta

da “proofreading”. Durante l’addizione la polimerasi riconosce il nucleotide corretto grazie alla forma precisa nel palmo quando accoppiato con lo stampo. Nucleotidi scorretti hanno forme diverse e non entrano nel sito attivo con la stessa precisione. Non è richiesta alcuna energia per questo processo.

Riparazione dell’errore

Dopo l’aggiunta del nucleotide la DNA polimerasi si muove al nucleotide successivo nel filamento di stampo ma rallenta quando viene aggiunto un nucleotide scorretto. L’operazione di proofreading serve a riconoscere questi errori. Successivamente la terminazione scorretta è rotta nei legami a idrogeno e viene “flipped out” nel sito dell’esonucleasi che rimuove le basi mal-accoppiate attraverso attività di delezione 3’-5’ di esonucleasi: le funzioni di polimerasi ed esonucleasi sono spazialmente separate. La terminazione 3’OH nel nuovo filamento è successivamente riimmessa nel sito attivo e reinizia la sintesi. Per questo passaggio è richiesta energia.

Cause di errori

Tautomeria dei nucleotide L’inserzione di un nucleotide errato può essere dovuta a un riposizionamento transiente dei doppi legami delle quattro basi all’azoto che cambia la posizione dell’idrogeno legato al gruppo amminico o forme tautomeriche. Una volta accoppiate le basi tautomeriche possono riconvertirsi in posizione normale ma la coppia rimane mal-accoppiata.

- La timina si lega con la forma enolica della guanina.
- L’adenina si lega con la forma imminica della citosina.
- La guanina si lega con la forma enolica della timina.
- La citosina si lega con la forma imminica dell’adenina.

Un altro caso è la depurinazione delle purine, ovvero rimozione o perdita del nucleotide che può accadere a bassi pH e può causare transizioni e trasversioni.

- Transizioni: una purina è sostituita da un’altra purina.
- Transversioni: una purina è sostituita da una pirimidina.

I mal-accoppiamenti tautomerici sono riconosciuti dalla DNA polimerasi e corretti dall’attività dell’esonucleasi o da sistemi di riparazione post-replicazionali. Quando non corretti causano delle mutazioni.

- Transizioni (di entrambi i tipi) una purina è sostituita da un’altra purina ($A \leftrightarrow G$).
- Transversioni (del secondo tipo) una purina è sostituita da una pirimidina: ($A \leftrightarrow C \vee T$, $G \leftrightarrow C \vee T$).

Inclusione di un *NTP* Le concentrazioni intracellulari di *NTP* sono molto più alte rispetto a quelle di *dNTP*. Questi vengono discriminati in base alla posizione 2’OH. La DNA polimerasi contiene anche una tirosina nel dominio catalitico che si scontra con il 2’OH del ribosio prevenendo il loro posizionamento nel sito catalitico e la formazione del legame fosfodiesterico. La capacità di riconoscimento di questi errori è diversa per ogni DNA polimerasi. Viene comunque inserita la base corretta per l’accoppiamento e viene corretta attraverso un processo di riparazione post-replicatorio attraverso *RNasi H1* e *RNasi H2* che rimuovono il ribonucleotide dal filamento e una DNA polimerasi non processiva e una DNA ligasi chiudono il buco.

2.6.5 Sintesi del DNA discontinua

La DNA polimerasi è in grado di sintetizzare in direzione 5'-3' ma entrambe le forcelle si muovono bidirezionalmente. Pertanto per forcella un filamento può essere sintetizzato 5'-3' usando un primer a RNA nel mezzo della bolla e la polimerasi segue l'elicasi. Questo viene detto filamento continuo o guida. Il secondo filamento deve essere sintetizzato in maniera discontinua da 5' a 3' e viene detto filamento ritardato. Nessun primer a monte è disponibile sul DNA svolto mentre l'elicasi si muove. Viene detto "lagging" in quanto la sintesi è più lenta rispetto all'altro filamento in quanto dipende dalla produzione di molti primer. Corti primer a RNA o RNA-DNA sono posti alla forcella di replicazione da una DNA primasi legata alla DNA elicasi nel complesso primosoma. Nei procarioti i primer a RNA sono posti ogni 1000-2000 basi con superavvolgimenti a monte. Negli eucarioti invece gli intervalli sono di 150-200 basi in quanto lo svolgimento è più difficile della cromatina con i nucleosomi. La terminazione 3'OH del primer è il sito di inizio per la DNA polimerasi e la sintesi di un pezzo di DNA da un primer al successivo viene detto frammento di Okazaki.

Maturazione dei frammenti di Okazaki in *E. coli*

Il DNA viene sintetizzato come frammenti corti e discontinui e i frammenti sono uniti quando i primer a RNA sono rimossi. In *E. coli* la DNA polimerasi III sintetizza il DNA che si estende dal primer e si dissocia dal DNA quando raggiunge il primer successivo, mentre la pinza scorrevole rimane attaccata. La DNA polimerasi I rimuove il primer attraverso "nick translation" e riempie il buco lasciato dal primer con DNA, mentre il nick viene unito da una DNA ligasi.

DNA polimerasi I La DNA polimerasi I di *E. coli* è una polimerasi ad alta fedeltà e non processiva che viene usata per rimuovere i primer a RNA e per riparare i danni del DNA. Ha una struttura diversa rispetto a quella processiva. È una proteina con capacità di polimerasi 5'-3' ed esonucleasi in entrambe le direzioni. L'enzima può essere rotto con la proteasi subtilisina, e l'esonucleasi genera il frammento di Klenow con attività esonucleasica che viene usata per rimuovere overhang in 3'OH o riempirlo in 5' per rendere blunt le terminazioni di DNA trattato con enzimi di restrizione.

DNA ligasi La DNA ligasi è un enzima che catalizza la formazione del legame fosfodiesterico tra due molecole di DNA. Lo fa rimuovendo γ -fosfato in 5' e favorendo l'attacco nucleofilo grazie allo ione magnesio.

Maturazione dei frammenti di Okazaki negli eucarioti

La rimozione del primer e il riempimento dello spazio nel nuovo DNA è diversa negli eucarioti. La DNA polimerasi δ sintetizza il DNA partendo da un primer fino a che raggiunge il successivo. Successivamente sposta il primer a valle insieme al frammento di Okazaki creando un "flap". La Flap endonucleasi *Fen1* lo taglia (con omologo presente in archea). La DNA polimerasi δ successivamente si stacca dal DNA e la DNA ligasi unisce il nick. La DNA polimerasi δ viene reclutata da una pinza scorrevole alla terminazione 3'OH al primer successivo.

2.6.6 Attività del replisoma alla forcella di replicazione

La DNA polimerasi replicativa è parte di un grande complesso: il replisoma. Questo è composto da 3 DNA polimerasi, 3 pinze scorrevoli, 3 proteine Tau, un caricatore della pinza, una DNA elicasi, una DNA primasi, *SSB/RFC*, e solo nei procarioti da una DNA polimerasi I. La proteina Tau τ

lega la DNA polimerasi al caricatore della pinza. Si trova un replisoma per forcella di replicazione e sintetizza contemporaneamente sia il filamento principale che quello ritardato.

- Negli eucarioti la DNA polimerasi ϵ sintetizza il filamento guida, mentre le DNA polimerasi α e δ sintetizzano quello ritardato.
- Nei procarioti una DNA polimerasi III sintetizza il filamento guida, mentre due DNA polimerasi III sintetizzano quello ritardato.

Il modello a trombone

Le DNA polimerasi si muovono in direzioni diverse sui due filamenti, ma si muovono seguendo la forcella di replicazione e la stessa elicasi. Questo viene permesso dalla creazione di un loop da parte del filamento ritardato. Il loop deve essere rilasciato dalla polimerasi ogni paio di secondi. Si nota come le polimerasi sul filamento guida e ritardato si trovano insieme spazialmente e sono regolate insieme.

2.7 Terminazione

Quando la replicazione inizia il processo non si ferma fino a quando il replicone è replicato. Nei batteri le forcelle a destra e a sinistra si fondono su un plasmide o sul genoma batterico circolare, mentre negli eucarioti una bolla si fonde con un'altra in arrivo da destra e un'altra in arrivo da sinistra sul cromosoma lineare.

2.7.1 Terminazione nei batteri

Nei batteri una sequenza di terminazione *Ter* determina dove entrambe le forcelle si fondono. È composta da 10 ripetizioni di 23bp ed è localizzata diametricamente opposta a *OriC*. Le ripetizioni di *Ter* sono 5 in senso orario e 5 antiorario. La proteina terminatrice *Tus* si lega a ognuna delle 10 sequenze *Ter* fermando la forcella di replicazione, causando il disassemblaggio del complesso di replicazione e il corto pezzo di DNA non replicato è riempito dalla DNA polimerasi I e chiuso dalla DNA ligasi. *Tus* possiede due facce: una non permissiva che blocca la replicazione e una permissiva che permette la sua continuazione. Dopo la terminazione e ligazione i genomi sono incastrati e vengono liberati da topoisomerasi durante la decatenazione.

2.7.2 Terminazione negli eucarioti

In alcune regioni dei cromosomi eucariotici la replicazione è bloccata in una direzione per impedire scontri con il macchinario di trascrizione che arriva da un altro lato. Nel lievito tale sito è detto barriera della forcella di replicazione e si trova nel rDNA. La proteina *FOB1* si lega al sito e previene la progressione della forcella di replicazione.

Replicazione alla fine di un cromosoma lineare

Il meccanismo per la sintesi del filamento ritardato non può replicare la terminazione di un cromosoma lineare in quanto la rimozione dell'ultimo primer da parte di *RNAasi H* lascia un vuoto che non può essere riempito, pertanto una porzione del cromosoma (telomero) potrebbe essere accorciata durante diversi cicli di replicazione.

2.8 Replicazione dei telomeri

Il problema della replicazione delle terminazioni dei cromosomi è risolto dai virus *T4* unendo diverse copie del cromosoma lineare che poi si dissociano. Gli eucarioti invece possiedono una telomerasi per prevenire il problema della replicazione delle terminazioni. I telomeri variano tra i 100bp e i 20 000bp in base alla specie. Si accorciano ad ogni ciclo di replicazione e possono essere allungati attraverso la telomerasi che aggiunge nuove sequenze telomeriche alla terminazione. Nella *Drosophila* i *non-LTR* retrotrasposoni *HeT-A* e *TART* si traspongono ripetutamente nelle terminazioni cromosomiche per produrre una regione simile ai telomeri. Il mantenimento della lunghezza del telomero attraverso eventi di trasposizione addizionali.

2.8.1 Telomerasi

La telomerasi è una speciale DNA polimerasi, una ribonucleoproteina *RNP* formata da un'enzima e una molecola di RNA. L'enzima o *TERT*, telomerasi trascrittasi inversa è conservata negli eucarioti. L'RNA fornisce lo stampo per la sintesi delle ripetizioni della telomerasi. La telomerasi si lega all'overhang di DNA nel telomero a singolo filamento. L'overhang alla terminazione 3' si lega con l'RNA che viene usato come stampo per sintetizzare DNA. La sintesi viene ripetuta più volte.

2.8.2 Mantenimento della lunghezza

Il mantenimento della lunghezza dei telomeri viene raggiunto in due passi:

1. Allungamento dell'overhang 3': la telomerasi si lega all'overhang 3', sintetizza del DNA, lo allunga, si trasloca e riallunga il filamento.
2. Replicazione del filamento complementare: viene sintetizzato un primario da una primasi, il gap viene riempito da una polimerasi, il primer viene rimosso e una ligasi lega i filamenti.

Si noti come rimane comunque un overhang che è coinvolto nella formazione del D-loop.

2.8.3 Problemi nel mantenimento della lunghezza del telomero

Nelle cellule somatiche adulte il gene che codifica *hTERT* non è ben espresso a causa di repressione epigenetica, ma altamente espresso in cellule fetali e staminali. I telomeri si accorciano nelle cellule causando la risposta del danno al DNA e un arresto *p53* dipendente e morte cellulare. Inoltre la repressione epigenetica può impedire il legame della telomerasi con la cromatina del telomero (approfondire).

2.8.4 Il limite di Hayflick

La lunghezza dei telomeri alla nascita è di 15 000bp e per divisione cellulare il telomero si accorcia di 300-1000 coppie di basi. Ad un certo punto tra le 50 e le 60 divisioni si raggiunge il limite di Hayflick in cui una cellula smette di dividersi per evitare ulteriore erosione del cromosoma e raggiunge la senescenza. Molte cellule cancerogene possiedono una telomerasi sovraespressa e continuano a dividersi e proliferare o uno stato epigenetico alterato ai telomeri che porta alle stesse conseguenze.

2.9 Correzione degli errori post-replicativa

Il DNA mismatch repair *MMR* non è parte dell'attività della DNA polimerasi e riduce il tasso di errore di 100 volte. Un nucleotide mismatch causa una distorsione locale nella doppia elica riconosciuta dal dimero *MutS* che si lega ad essa e recluta il dimero *MutL* che stabilizza il legame con il DNA distorto. La riparazione avviene sul nucleotide di nuova sintesi, riconosciuto grazie alla metilazione sullo stampo di *GATC*. Il complesso *MutS-MutL* recluta l'endonucleasi *MutH* che si lega alla sequenza metilata sul filamento stampo più vicino al mismatch, taglia il filamento di nuova sintesi vicino alla sequenza, aiutata occasionalmente dall'elicasi *UvD*. L'esonucleasi 3'-5' rimuove il DNA fino al sito di errore e la DNA polimerasi III lo risintetizza. Quando si incontra il DNA dopo il taglio i filamenti sono riattaccati da una ligasi. L'efficienza di riparazione si abbassa allontanandosi da *GATC*.

2.10 Mantenimento delle modifiche istoniche

Si nota come i nucleosomi sono in parte rimossi dal replisoma in movimento. Le proteine istoniche devono pertanto essere reclutate ai posti corretti nel DNA appena reclutato e mantenere le modifiche epigenetiche (eredità epigenetica). I nucleosomi sono distrutti dalla forcella di replicazione e devono rilegarsi dopo che questa li ha superata e le modifiche epigenetiche sono aggiunte su essi. Questo avviene in quanto il nucleosoma parentale non è rimosso: 50% sono distribuiti equamente tra i filamenti singoli: avviene una segregazione degli istoni parentali. Di questi solo quelli marcati epigeneticamente 2(*H3-H4*) sono mantenuti, mentre gli altri sono rimossi anche se rimangono vicini. Quelli non marcati sono inclusi nei nuovi nucleosomi attraverso il chaperone *Caf1* insieme ai 2(*H2A-H2B*) parentali o nuovi attraverso il chaperone *NAP1*. Pertanto si trovano 4 tipi di nucleosomi nei filamenti sintetizzati. Infine il nucleosoma parentale originale serve come stampo per produrre lo stato epigenetico su tutti i nucleosomi che sono rimasti vicini.

2.11 DNA polimerasi specializzate

Tutte le DNA polimerasi catalizzano la stessa sintesi del DNA ma differiscono nella velocità e tasso di errore.

- DNA polimerasi replicative: sono processive e catalizzano la replicazione dei genomi durante la fase *S*, hanno un'alta fedeltà di polimerizzazione e un'attività esonucleasica 3'-5' altamente efficiente.
- DNA polimerasi non-processive o distributive: rappresentano la maggior parte di DNA polimerasi, polimerizzano corte lunghezze per mantenerlo in uno stato sano (riparazione), hanno bassa fedeltà, sintetizzano il DNA sul filamento danneggiato o sintesi di translesione *TLS*.

Certe DNA polimerasi hanno altre attività e tutte sono reclutate dalla pinza scorrevole sul DNA che permette loro di stare attaccate al DNA, altrimenti dopo ogni addizione di nucleotide potrebbero separarsi.

2.11.1 DNA polimerasi batteriche

Altamente conservate in tutti i procarioti *E. coli* possiede 5 DNA polimerasi:

- DNA polimerasi I: attiva nella riparazione del DNA ed estende e matura i frammenti di Okazaki.
- DNA polimerasi II: attiva nella riparazione del DNA.
- DNA polimerasi III: polimerasi replicativa che catalizza la polimerizzazione del DNA a grande lunghezza.
- DNA polimerasi IV e V: riparano danni al DNA che bloccano il replisoma.

2.11.2 DNA polimerasi specializzate

DNA polimerasi *TLS*

Esistono in tutti gli organismi e sono le DNA polimerasi per la sintesi di traslesione. Il DNA in una cellula subisce danno costante e deve essere riparato, in particolare il dimero di timidina nello stampo del filamento singolo causa un arresto della forcella, rottura ed il doppio filamento e morte cellulare. Le DNA polimerasi *TLS* promuovono la replicazione, ma hanno alto tasso di errore a causa dell'assenza dell'attività nucleasica. Quando il dimero di timidina causa uno stacco della DNA polimerasi III ad esso si lega la DNA polimerasi *TLS* che avendo un sito di legame più aperto riesce a superare il blocco e continuare per un breve tratto la sintesi con minore fedeltà prima di essere di nuovo sostituita dalla DNA polimerasi III.

Trascrittasi inversa

La trascrittasi inversa è una DNA polimerasi dipendente da RNA: copia l'RNA in ssDNA o cDNA. Usa un singolo filamento di RNA come stampo e necessita di un primer. Sono molto simili alle DNA polimerasi e sono codificate da virus e retrotrasposoni, elementi di DNA mobile negli eucarioti. La telomerasi è una trascrittasi inversa specializzata. I virus a RNA o retrovirus come *HIV-1* e influenza A hanno un genoma a RNA che codifica per una trascrittasi inversa nella cellula ospite. Il gema è trascritto in dsDNA che si integra nel genoma ospite.

Capitolo 3

Trascrizione

3.1 Panoramica della trascrizione

L'informazione conservata nel DNA è utilizzata per creare proteine o molecole di RNA funzionale. Si intende per trascrizione il processo di copia di un filamento di DNA in una molecola di RNA detta trascritto. Il processo di trascrizione viene svolto da una RNA polimerasi. Un cofattore esterno RNA polimerasi elicasi separa i filamenti di DNA e permette i ribonucleosidi trifosfato di accoppiarsi con il filamento stampo. Per produrre una proteina da una molecola di DNA la sequenza di RNA è letta dal ribosoma nella traduzione. L'RNA in questo processo viene detto RNA messaggero *mRNA*. Si nota come la RNA polimerasi non possieda attività esonucleasica e pertanto non possa correggere errori di mal-accoppiamento. Il tasso di errore è di 10^{-4} e il tasso di da 40 a 80 nucleotidi al secondo. Si nota come rispetto alla DNA polimerasi è più lenta, inefficiente e meno accurata.

3.1.1 Il processo di trascrizione

La trascrizione può essere divisa in iniziazione, allungamento e terminazione. Inizia quando la RNA polimerasi si lega a una sequenza di DNA che precede il gene: il promotore. Il sito di inizio di trascrizione *TSS* è la prima base ad essere trascritta ed è notata con +1. L'RNA viene trascritto nella direzione 5'-3' con il filamento letto in direzione opposta come nella sintesi del DNA. Pertanto si indica come basi a monte quelle 5' e a valle quelle 3'.

Iniziazione

Durante l'iniziazione la RNA polimerasi separa i filamenti di DNA per creare una bolla di trascrizione tra le 12 e le 14bp e inserisce i primi ribonucleoside trifosfati *NTP* mentre si trova al promotore. Quando il RNA è di lunghezza sufficiente la RNA polimerasi lascia il promotore "promoter clearance" e cambia conformazione per essere più stabilmente associata con il DNA permettendo l'allungamento del RNA.

Allungamento

L'allungamento inizia dopo la clearance del promotore e la RNA polimerasi si muove lungo il DNA aggiungendo ribonucleotidi e allungando il trascritto di RNA. La RNA polimerasi svolge il DNA a valle e close quello a monte mantenendo la dimensione della bolla di trascrizione di dimensione

costante per impedire la formazione di R-loop. Nella bolla una regione del trascritto di 8-10bp è accoppiata con il DNA mentre il resto è estruso dalla polimerasi.

Terminazione

L'allungamento continua fino a che la polimerasi incontra una sequenza di DNA detta terminatore che segnala la fine della sintesi di RNA. Il RNA è rilasciato e la RNA polimerasi si dissocia dal DNA.

3.1.2 Nomenclatura dei geni

In un gene viene indicato con +1 il sito di inizio della trascrizione *TSS*, con numeri negativi la zona del promotore e con *TTS* il sito di terminazione della trascrizione. Si dice con prossimale la zona più vicina al *TSS*, distale quella che si trova allontanandosi verso il *TTS*. Si intende per sequenza codificante *CDS* la sequenza del gene senza introni, mentre con open reading frame *ORF* la sequenza con gli introni.

3.1.3 Regolazione della trascrizione

La trascrizione è regolata per produrre il RNA richiesto al tempo corretto. La cromatina negli eucarioti presenta una sfida per la trascrizione in quanto i nucleosomi prevengono il legame e il movimento del macchinario di trascrizione attraverso la cromatina. Sono pertanto richiesti:

- Rimodellamento dei nucleosomi: per riposizionare gli istoni lontano dal DNA che deve essere trascritto.
- Chaperone degli istoni per riassemblare e disassemblare i dimeri nucleosoma-istone.
- Enzimi che modificano le proteine istoniche epigeneticamente per permettere o prevenire il legame di proteine che regolano la trascrizione.

3.2 L'enzima centrale della RNA polimerasi

Se i procarioti possiedono 1 RNA polimerasi gli eucarioti ne possiedono 3 principali:

- RNA polimerasi I *Pol I*: trascrive i geni di grandi RNA ribosomali *rDNA*. Agisce nei nucleoli 5-6 negli umani, 1 nel lievito.
- RNA polimerasi II *Pol II*: trascrive i geni di RNA messaggero *mRNA*, RNA non codificanti corti e lunghi, *miRNA* con ruolo nella regolazione dell'espressione genica attraverso interferenza e *snRNA* RNA piccoli nucleari con un ruolo nel processamento di *mRNA*. Agisce nel nucleoplasma.
- RNA polimerasi III *Pol III*: trascrive una varietà di RNA come tutti gli RNA transfer *tRNA*, il piccolo 5S RNA ribosomale, *snRNA U6* componente dello spliceosoma, *7SL RNA* o long non-coding RNA parte della particella di riconoscimento del segnale che regola la traduzione. Agisce sia nei nucleoli che nel citoplasma.

Le piante possiedono una quarta RNA polimerasi con ruolo di trascrizione degli RNA non codificanti con ruolo nell'espressione genica. Gli eucarioti possiedono anche RNA polimerasi mitocondriali. Si intende per nucleoli raggruppamenti non circondati da membrana di 6 ripetizioni di rDNA trascritti da *Pol I* e *Pol III*.

3.2.1 Struttura

Le RNA polimerasi sono formate da diverse subunità che formano il nucleo polimerasico.

RNA polimerasi batteriche

Le RNA polimerasi batteriche sono le più piccole con 5 subunità che si assemblano in un complesso nucleare con lobi simili a mascelle che formano una pinza. Le mascelle sono formate dalle subunità β e β' con le due α e ω alla base. La base della fessura è il sito attivo dell'enzima. Ogni subunità α possiede un dominio N terminale αNTD e un dominio C terminale αCTD unite da un collegatore flessibile.

Subunità e loro ruolo

- α (2): responsabile per l'assemblaggio del complesso.
- Domini N terminali: interagiscono con le subunità β e β' .
- Domini C terminali: si legano al DNA promotore: elemento a monte.
- β (1): contengono il sito catalitico, formano i legami fosfodiesteri legando Mg_2^{+} e svolgono la correzione degli errori.
- β' (1): mantiene l'enzima legato al filamento stampo, svolge e riavvolge il dsDNA.
- ω (1): è un chaperon: promuove la stabilità strutturale della RNA polimerasi.

RNA polimerasi di eucarioti ed archea

Tutte le RNA polimerasi possiedono le 5 subunità centrali viste prima altamente conservate, specialmente al sito attivo. Si trovano ulteriori subunità in RNA polimerasi di archea ed eucarioti ordinate intorno alle 5 centrali. L'attività catalitica pertanto rimane conservata in tutti e tre gli alberi della vita.

Ruoli della RNA polimerasi II Oltre a trascrivere il DNA la RNA polimerasi II accoppia la trascrizione con il processamento del trascritto di RNA: il dominio C terminale CTD di *Pol II* è cruciale per questa funzione: è la terminazione della subunità *Rpb1* e esiste come ripetizioni di una sequenza di 7 amminoacidi: *Tyr-Ser-Pro-Thr-Ser-Pro-Ser* con 26 ripetizioni nel lievito e 52 negli umani.

3.3 Riconoscimento dei promotori

Il nucleo della RNA polimerasi può sintetizzare il RNA ma non può riconoscere e legarsi alla sequenza promotrice di un gene. Si rendono pertanto necessarie ulteriori subunità che si legano direttamente al promotore. Si forma pertanto un oloenzima dalla loro unione con l'enzima nucleare.

3.3.1 Batteri

Nei batteri le subunità che si legano al promotore sono dette fattori σ . Ne esistono diverse che riconoscono promotori specifici per promuovere la trascrizione di geni specifici in base a particolari condizioni di crescita. I fattori σ si legano a sequenze che definiscono i promotori batterici composti tipicamente da due elementi: uno a -35 e uno a -10 detto box di Pribnow. Ogni fattore sigma ha una sequenza di legame specifico e una particolare spaziatura tra i due elementi. Questo in modo da regolare la trascrizione: più vicine le sequenze e la spaziatura a quella particolare del fattore σ più forte il legame e più alti tassi di trascrizione.

Fattori sigma

Nell'oloenzima della RNA polimerasi il fattore σ^{70} è costitutivo e 3 dei suoi 4 domini riconoscono specifici elementi dei promotori:

- Il dominio 2 si lega alla regione -10 e aiuta a separare il dsDNA “promoter melting”.
- Il dominio 3 riconosce le due basi della regione estesa -10 .
- Il dominio 4 riconosce l'elemento -35 , è attaccato a una parte flessibile del nucleo dell'enzima che permette diversa spaziatura tra -35 e -10 .

Alcuni fattori sigma sono regolati in risposta a condizioni ambientali o di sviluppo. Possono essere regolati sia a livello trascrizionale o traduzionale o alterando la stabilità del proprio mRNA. Sono inoltre regolati da:

- Fattori pro-sigma: proteine con domini inibitori che devono essere rotte prima che il fattore σ possa associarsi con l'enzima RNA polimerasi.
- Fattori anti-sigma: proteine che legano i fattori σ inibendo la loro funzione.

Fattori anti-sigma e regolazione dell'assemblaggio del flagello in *Salmonella typhimurium* Mentre le proteine che formano la base del flagello sono sintetizzate il fattore anti-sigma *FlgM* si lega a σ^F impedendo il suo legame con la RNA polimerasi- σ^{70} . σ^F promuove la trascrizione di geni necessari per il completamento dell'assemblaggio del flagello. Negli ultimi passi della sintesi delle proteine del flagello *FlgM* è esportata dalla cellula in modo che σ^F possa legarsi alla RNA polimerasi sostituendo σ^{70} e promuovere la trascrizione dei geni dell'assemblaggio permettendo la creazione del flagello.

3.3.2 Eucarioti

Anche i promotori eucariotici e di archea necessitano di ulteriori proteine per direzionare la RNA polimerasi ai promotori. Questi sono detti fattori di trascrizione generali *TF* seguito dal numero della *Pol*. Negli eucarioti *TFII* si assemblano al promotore e con l'enzima del nucleo formano il complesso di pre-iniziazione. *Pol I* e *Pol III* richiedono diverse proteine per formare il complesso di pre-iniziazione. I promotori per la RNA *Pol II* possiedono spesso una TATA box con sequenza di consenso *TATAA* tra le 25 e le 30bp a monte dell'inizio della trascrizione. Tutte le polimerasi eucariotiche necessitano della TATA binding protein *TBP* per iniziare la trascrizione che viene riconosciuta da *TBP* associated factors *TAF* formando il complesso *TFIID*. Alcuni altri elementi includono l'elemento di riconoscimento *TFIIB BRE*, l'elemento iniziatore *INR* e l'elemento promotore a valle *DPE*. Molti promotori non hanno questi elementi e la loro variabilità rende difficile la

3.3. RICONOSCIMENTO DEI PROMOTORI

loro predizione. Il complesso di pre-iniziazione contiene 32 proteine fattori di trascrizione generali e la RNA *Pol II* con 12 subunità. I promotori per *Pol II* possono dividersi in:

- Prossimali: distanti meno di 200bp da *TSS* nei lieviti unicellulari.
- Distali: distanti fino a 10kb dal *RSS* come negli eucarioti multicellulari.

- 3.3.3 Formazione del complesso di pre-inizio
- 3.4 Iniziazione della trascrizione e transizione a un complesso di allungamento
- 3.5 Allungamento della trascrizione
- 3.6 Terminazione della trascrizione
- 3.7 Principi della regolazione della trascrizione
- 3.8 Domini leganti il DNA in proteine che regolano la trascrizione
- 3.9 Meccanismi per regolare l'iniziazione della trascrizione nei batteri
- 3.10 L'operone *lac* in *E. coli*
- 3.11 L'operone triptofano *trp* in *E. coli*
- 3.12 Regolazione della trascrizione da parte di riboswitches trascritti
- 3.13 Regolazione dell'espressione genica del batteriofago λ in *E. coli*
- 3.14 Regolazione della trascrizione da sistemi di trasduzione del segnale a due componenti
- 3.15 Regolazione dell'iniziazione della trascrizione ed allungamento negli eucarioti
- 3.16 Il ruolo delle cascate di segnalazione nella regolazione della trascrizione
- 3.17 Silenziamento genico attraverso imprinting genomico

Capitolo 4

Processamento dell'RNA

Capitolo 5

RNA regolatori

Capitolo 6

Traduzione

Capitolo 7

Modifica e targeting delle proteine

Capitolo 8

DNA mobile

Capitolo 9

Strumenti e tecniche della biologia molecolare