Trabajo Práctico Bioinformática

Guia Tema 2.1 - Grupo 2

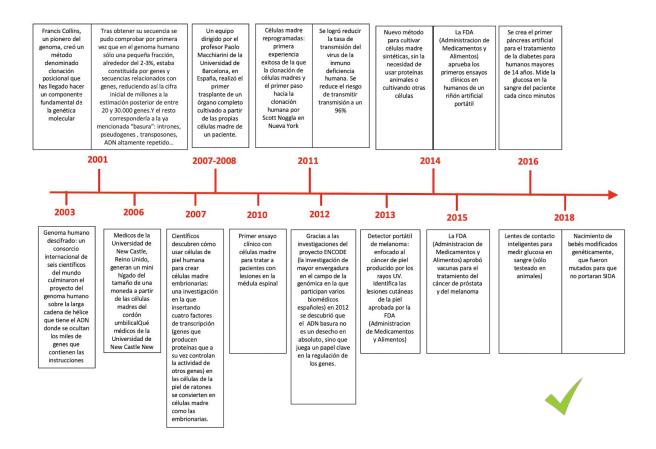
Profesor:

Danilo Ceschin

Integrantes:

Nazareth Chaile, Lucia Mansilla, Milagros Londero, Camila Cobresi, Ticiana

1. Link linea del tiempo: Descubrimientos



2. La bioinformática es una ciencia multidisciplinar que surge de la necesidad de comprender las funciones biológicas, más concretamente los genes. Es un área interdisciplinaria, que mezcla subáreas de biología, informática, tecnologías de la información, matemáticas y estadística. La bioinformática utiliza la experiencia de estos diferentes campos de conocimiento en el procesamiento, organización, análisis e interpretación de datos biológicos. Este trabajo in silico (hecho por computadora), comprende la redacción y ejecución de programas que utilizan algoritmos de teoría de grafos, inteligencia artificial, minería de datos, modelado estadístico, procesamiento de imágenes y simulación por computadora, entre otras cosas.

X

se aplican en diferentes áreas como la genética clínica, la biología celular y molecular y la genética evolutiva, y precisó que la importancia creciente de la bioinformática está relacionada con la necesidad de "procesar grandes cantidades de datos en el contexto de los marcos teóricos adecuados, con el objetivo de obtener resultados y predicciones más precisas y robustas en las diferentes áreas de aplicación

2.1

- 3. Desarrolle brevemente a que se refieren los siguientes bloques temáticos de la Bioinformática:
 - a. **Organización de la Información:** Es la creación y mantenimiento de bases de datos y bancos de datos.
 - b. **Acceso a la Información:** Consiste en la búsqueda de información en bases de datos y en la comparación de información con la de las bases de datos.
 - c. Algoritmos: Es una serie ordenada de instrucciones, pasos o procesos que llevan a la solución de un determinado problema.
 - d. Búsqueda de secuencias: Es una forma de representar y comparar dos o más secuencias o cadenas de ADN, ARN, o estructuras primarias proteicas para resaltar sus zonas de similitud, que podrían indicar relaciones funcionales o evolutivas entre los genes o proteínas consultados.
 - e. Genómica: Se enfoca en la estructura, función, evolución y mapeo de los genomas, tiene como objetivo la caracterización colectiva y la cuantificación de los genes, que dirigen la producción de proteínas con la ayuda de enzimas y moléculas mensajeras.
 - f. **Transcriptómica:** Es el estudio del conjunto de ARN (ARNr, ARNt, ARNm, ARNi, miARN) que existe en una célula, tejido u órgano.
 - g. Proteómica: Es el estudio a gran escala de las proteínas, en particular de su estructura y función.
 - h. . **Metabolómica:** Es el "estudio sistemático de las huellas únicas que dejan los procesos celulares específicos en su paso", es decir, el estudio del perfil de los metabolitos (moléculas pequeñas) de una muestra biológica.
 - i. Análisis de Imágenes y Visualización de Imágenes: Consiste en la extracción de información derivada de sensores y representada gráficamente en formato de dos o tres dimensiones, para lo cual se puede utilizar tanto análisis visual como digital (procesamiento digital de imágenes). Abarca la fotografía en blanco y negro y color, infrarroja, imágenes satelitales, de radar,

- radar de alta definición, ultrasonido, electrocardiogramas, electroencefalogramas, resonancia magnética, sismogramas y otros.
- j. "Data Mining": El objetivo general del proceso de Data Mining consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior aplicando métodos de inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos.
- k. Búsqueda de patrones: Es el acto de comprobación de una determinada secuencia de fichas para la presencia de los componentes de algún patrón. La coincidencia por lo general tiene que ser exacta y los patrones generalmente tienen la forma de secuencias o estructuras de árbol.
- Modelado: Es el uso de simulaciones por ordenador de subsistemas celulares (tales como redes de metabolitos y enzimas que comprenden el metabolismo, caminos de transducción de señales, y redes de regulación genética), tanto para analizar como para visualizar las complejas conexiones de estos procesos celulares
- m. **Simulación:** Es un intento de modelar situaciones de la vida real por medio de un programa de computadora, lo que requiere ser estudiado para ver cómo es que trabaja el sistema. Ya sea por cambio de variables, quizás predicciones hechas acerca del comportamiento del sistema.
- n. "System Biology": Es el campo de investigación interdisciplinaria de los procesos biológicos en el que las interacciones de los elementos, internos y externos, que influyen en el desarrollo del proceso se representan con un sistema matemático.
- 4. La bioinformática y la biología se relacionan en que ambas tienen como objetivo desarrollar nuevas tecnologías dor 42 se incluyen a los a 43 ces genómicos, todo ello en beneficio de la salud humana. desarrolla dispositivos capaces de contener grandes volúmenes de información, por lo que podríamos decir que estas disciplinas se desarrollan a la par en beneficio de la sociedad.
- 5. En bioinformática además de encontrar disciplinas como la biología y la informática también encontramos tales como la genética, las matemáticas aplicadas, la estadística, las ciencias computacionales, la inteligencia artificial, la química y la bioquímica. Todas estas áreas nutren a la bioinformática desde distintos ámbitos ya sea como la matemática para la resolución de problemas o las ciencias computacionales para el almacenamiento de datos.

Universidad Catolica de Cordoba Facultad de Ingeniería Córdoba, Argentina

Trabajo de Bioinformática

Integrantes:

Cabral, Camila

Caceres, Martin

Cobresi, Ticiana

Menel, Angelo

Nanfara, Abril

Paschini, Catalina

Introducción a la Bioinformática - Docente: Danilo Ceschin

Tema 1.3 – Introducción a la Biología. Trabajo Práctico

La envoltura nuclear encierra el ADN y define el compartimiento nuclear. Está formada por dos membranas concéntricas. La membrana nuclear interna contiene proteínas específicas que actúan como sitios de unión de la lámina nuclear que la sostiene, y que contacta los cromosomas y los ARNs nucleares. Esta membrana está rodeada por la membrana nuclear externa, similar a la membrana del retículo endoplásmico (RE), con el que se continúa. De hecho, la membrana nuclear externa puede ser considerada como una región especializada de la membrana del RE. Al igual que la membrana del RE rugoso, está generalmente tachonada externa comprometidos en la síntesis de proteínas. Las proteínas sintetizadas en esos ribosomas son transportadas al espacio entre las membranas nucleares interna y externa (el espacio perinuclear), el cual a su vez es continuo con el lumen del RE. El núcleo contiene muchas proteínas que ayudan a mediar sus funciones específicas. Esas proteínas, que incluyen histonas, ADN y ARN polimerasas, reguladores génicos, y proteínas procesadoras de ARN, son importadas desde el citosol, donde son sintetizadas. Deben pasar a través de las membranas nucleares externa e interna para llegar al interior del núcleo (el lumen nuclear). Este proceso de transporte es selectivo: muchas proteínas sintetizadas en el citosol son excluidas del núcleo. (Alberts, Bruce, Watson y col., 2007, "Molecular Biology of the cell", 5º edición, Editorial Garland Publishing) Discuta y conteste las siguientes preguntas en relación al párrafo del punto anterior.

1. Discuta y conteste las siguientes preguntas en relación al párrafo anterior.

1.1 ¿Por qué las células eucariotas presentan un núcleo? ¿Cuál es el sentido de una membrana nuclear?

Las células eucariotas poseen cromatina (cor ADN) englobada por una doble membrana, denominada envuelta nuclear, formando una estructura denomina núcleo.

El sentido de la membrana nuclear es una doble membrana que delimita el núcleo de la célula. Sirve para separar los cromosomas del resto del contenido celular.

1.2 Las funciones asociadas al material genético son la replicación, expresión, almacenaje y mutación. ¿Qué cree Ud. que significa cada uno de estos términos? Explique cómo el modelo de doble hélice de Watson y Crick se ajusta a esas cuatro funciones.

Replicación: La replicación del ADN es el proceso mediante el cual se duplica una molécula de ADN. Cuando una célula se divide, en primer lugar, debe duplicar su genoma para que cada célula hija contiene un juego completo de cromosomas.

<u>Expresión</u>: La expresión <u>ri</u>nica es el proceso que permite obtener proteínas a partir de genes. Los genes son secuencias de nucleótidos de ADN que codifican la información necesaria para la síntesis de proteínas. Esta síntesis tiene lugar en dos pasos: transcripción y traducción.

Almacenaje: El material genético se 72 nplea para guardar la información genética de una forma de vida orgánica y, en eucariotas, está almacenado en el núcleo de la célula. Para todos los organismos conocidos actualmente, el material genético es casi exclusivamente ácido desoxirribonucleico (ADN).

<u>Mutación</u>: En genética se denomina mutación genética, mutación molecular o mutación puntual a los cambios que alteran la secuencia de nucleótidos del ADN. Estas mutaciones en la secuencia del ADN pueden llevar a la sustitución de aminoácidos en las proteínas resultantes.

1.3 Describa las características del modelo de doble hélice de Watson y Crick. ¿Qué tipos de interacciones estabilizan la doble hélice?

Puentes de hidrógeno, Hidrofóbicas, acetilo ????????

Se multiplica siempre de 3'a 5'

La doble hélice es la descripción de la estructura de una molécula de ADN. Una molécula de ADN consiste en dos cadenas que serpentean una alrededor de la otra como una escalera de caracol. Cada cadena tiene una espina dorsal en la cual se alternan un azúcar (desoxirribosa) y un grupo fosfato. A cada azúcar se une una de las cuatro bases: adenina (A), citosina (C), guanina (G) o timina (T). Las dos cadenas se mantienen unidas por enlaces entre las bases nitrogenadas, adenina formando enlaces con la timina, y citosina con la quanina.

Las bases nitrogenadas (hidrofóbicas) se encuentran apiladas en el interior de la doble hélice, en planos perpendiculares a su eje. La parte exterior (grupos fosfato y azúcares) es hidrofílica.

Las bases de una de las cadenas o hebras están unidas mediante puentes de hidrógeno con las bases nitrogenadas de la otra cadena o hebra, uniendo ambas cadenas

-¿Qué es un gen? Es una Unidad de información básica.

1.4 Consulte la bibliografía y transcriba un concepto de gen. Discuta y analice su alcance. ¿Que son los intrones y los exon

Un gen es un segmento corto de ADN. Los genes le dicen al cuerpo cómo producir proteínas específicas. Hay aproximadamente 20,000 genes en



cada célula del cuerpo humano. Juntos <u>forman constituyen el material</u> hereditario para el cuerpo humano y la forma como funciona.

En la secuencia del DNA el contenido codificante de un gen no está distribuido de forma continua a lo largo de dicho gen, sino que tiene discontinuidades, llamadas intrones, cuya secuencia no codifica proteínas. Las partes de la secuencia que codifican proteínas son los exones. 82



- 2. Responda las siguientes preguntas.
 - 2.1 Si una molécula de una doble hélice de DNA tiene un contenido G + C del 56%, ¿cuáles son los porcentajes de las cuatro bases (A, T, G y C) de esta molécula? (proporción de las bases que hay)

2.2. Discuta si la siguiente proposición es verdadera o falsa: "Si una molécula de ADN contiene 45% de guanina, debe contener 55% de timina". Fundamente su respuesta.

Falso: las bases nitrogenadas siempre se presentan en pares con su respectiva base.

2.3. Cuando la molécula de ADN es sometida a la acción del calor, las dos hebras se separan (desnaturalización del ADN), y la temperatura a la cual la separación ocurre es dependiente del contenido de bases. Específicamente, al ADN con mayor proporción de pares G-C se desnaturaliza a mayor temperatura que las moléculas con contenido más alto en A-T. Explique por qué.

Esto es debido a que la unión A-T es más débil que la del G-C, debido a que el par G-C tiene 3 puentes hidrógenos mientras que el par A-T estan unidas por 2 puentes hidrógenos.

- **3.** El "Dogma Central" de la Biología: La información genética almacenada en el ADN es transcrita a moléculas de ARN mensajero que luego pueden ser traducidas en el proceso de síntesis de las proteínas.
 - 3.1 El Código genético se basa en la secuencia ordenada de bases nitrogenadas de los nucleótidos en el ADN, determinando la secuencia de

aminoácidos en las proteínas. Principales características del Código genético:

i) ¿Qué es un codón y cómo está determinado?

Un codón es un conjunto de 3 nucleótidos. Este conjunto forma la unidad básica del proceso de traducción del ARN. Existen 64 codones diferentes por las diferentes combinaciones de los cuatro nucleótidos. Un conjunto de codones definen un gen. (Corresponde a un aminoácido específico)

ii) Relación codón-aminoácido, secuencia de los codones en el ADN para su lectura (codones de iniciación y finalización), universalidad del Código genético.

Un anticodón es la secuencia que complementa los codones. Estos son imprescindibles para el reconocimiento de los codones del ARN mensajero y, en consecuencia, para el proceso de síntesis proteica en todas las células vivas. El ADN tiene ciertos codones que determinan el inicio o fin de la traducción del ARN mensajero.

- **3.2 El ARN** posee características estructurales que le permiten su síntesis a partir del ADN, la interacción con proteínas y tener ciertas actividades catalíticas para intermediar en la transmisión de la información genética.
- i) El ARN, ácido ribonucleico, también es un polímero de nucleótidos: Analiza su estructura, a partir de los 4 nucleótidos principales y del polímero. El ARN puede formar regiones internas de apareamiento de bases en su cadena lineal, que le sirven para la regulación de sus funciones.

El ARN es una cadena de una hebra (Monocatenaria), que se encuentra en el citoplasma. Los nucleótidos del ARN están formados por una azúcar pentosa, un grupo fosfato y una base nitrogenada (Adenina, Guanina, Citosina y Uracilo).

- ii) Existen diferentes tipos de ARN asociados a diferentes funciones. Respecto la síntesis proteica, que pueden mencionarse tres tipos mayoritarios: el ARN mensajero (ARNm), el ARN de transferencia (ARNt) y el ARN ribosomal (ARNr); estudia sus estructuras generales y a qué otras moléculas se asocian para desarrollar sus actividades.
- El **ARN mensajero** (ARNm) presenta una estructura lineal de una sola hebra, que puede formar horquillas en determinados tramos cuyas bases son complementarias. Su función es trasladar la información genética del ADN a los ribosomas, para la síntesis de proteínas. Cada molécula de ARNm es complementaria a un fragmento o gen de ADN, que sirve de molde para su síntesis durante la transcripción:
- ARNm monocistrónico: el ARNm contiene la información necesaria para la síntesis de una proteína (eucariotas).

 ARNm policistrónico: el ARNm contiene información para la síntesis de varias proteínas (procariotas).

El **ARN de transferencia** (ARNt) es el encargado de transportar los aminoácidos en el citoplasma para la síntesis de proteínas. Está formado por 70-90 nucleótidos, algunos de los cuales presentan bases poco frecuentes (distintas de A, C, G o U) y presenta fragmentos con estructura de doble hélice y otros en los que se forman bucle.

El **ARN ribosómico** (ARNr) es el más abundante (en torno al 80 % del ARN celular). Sus moléculas son largas y monocatenarias, aunque presenta fragmentos con estructura de doble cadena. Tiene función estructural ya que se encuentra asociado a proteínas formando los ribosomas, orgánulos encargados de la síntesis de proteínas.

iii) Analiza esquemáticamente la participación de estos 3 tipos de ARN en el proceso de síntesis de proteínas.

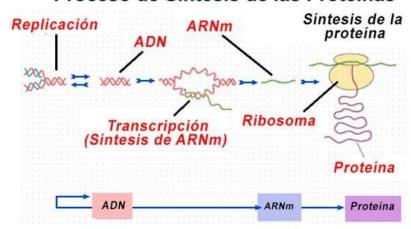
El ARN mensajero (ARNm) transmite la información genética almacenada en el ADN. Mediante el proceso conocido como transcripción, secuencias específicas de ADN son copiadas en forma de ARNm que transporta el mensaje contenido en el ADN a los sitios de síntesis proteica (los ribosomas).

Los aminoácidos (componentes de las proteínas) son unidos a los ARN de transferencia (ARNt) que los llevarán hasta el lugar de síntesis proteica, donde serán encadenados uno tras otro. La enzima aminoacil-ARNt-sintetasa se encarga de dicha unión, en un proceso que consume ATP.

Los ribosomas son los orgánulos citoplasmáticos encargados de la síntesis proteica; ellos son los encargados de la unión de los aminoácidos que transportan los ARNt siguiendo la secuencia de codones del ARNm según las equivalencias del código genético.

Es un ARN que forma parte de los ribosomas y es esencial para la síntesis proteica en todos los seres vivos. Los ARNr forman el armazón de los ribosomas y se asocian a proteínas específicas para formar las pre-subunidades ribosomales.

Proceso de Síntesis de las Proteínas



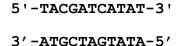
iv) Existen otras moléculas de ARN denominadas ARN no codificantes. Describa brevemente sus características y funciones.

Un ARN no codificante (ncRNA) es una molécula de ARN funcional, que a diferencia del RNA mensajero no se traduce en una proteína. La secuencia de ADN de la que un ARN no codificante se transcribe, a menudo se llama un gen de ARN no codificante.

Los genes de ARN no codificante incluyen funcionalidades abundantes y muy importantes como ARN de transferencia (tRNA) y ARN ribosomal (rRNA), así como también en ARN, tales como snoRNAs, microARNs, siRNAs y piRNAs y el ncRNA largo, que incluyen ejemplos tales como Xist y HOTAIR(HOX antisense intergenic RNA). El número de ncARNs (no codificantes de proteínas)en el genoma humano es desconocido; sin embargo, recientes estudios transcriptómicos y bioinformáticos sugieren la existencia de miles de ncARNs. La función de muchos de los ncARNs identificados recientemente no está confirmada, siendo posible, que muchos de estos, sean no funcionales.

Al igual que con las proteínas, las mutaciones o los desequilibrios en el repertorio ncRNA en el cuerpo pueden causar una variedad de enfermedades como cancer, síndrome de Prader-Willi, Alzheimer, etc.

4. Para la hebra de DNA 5'-TACGATCATAT-3' la hebra de DNA complementario correcta es:



5'-ATATGATCGTA-3'

5. a) Escribe la secuencia complementaria del siguiente nucleótido:

ATCGUTAGCTUAG

C=G

A=T

U=A

5' → 3'

ATCGUTAGCTUAG

3' → 5'

TAGCAATCGAATC

5' → 3'

CTAAGCTAACGAT

Las letras son A, C, G y T, que simbolizan las cuatro subunidades de nucleótidos de una banda ADN - adenina, citosina, guanina, timina, que son bases covalentemente ligadas a cadenas fosfóricas.

b) ¿Podría existir esta secuencia en la naturaleza? ¿Por qué?

No podria existir, ya que no puede existir en una misma linea timina y uracillo, el primero es del ADN y el segundo del ARN

- 6. Un ARN mensajero tiene 336 nucleótidos de longitud, incluyendo los codones de iniciación y de terminación. El número de aminoácidos de la proteína traducida a partir de este ARNm es: (Explique su respuesta)
 - a. 335

b. 112

- c. 168
- d. 111
- e. cualquier valor ≤ 111
- f. ninguna de estas opciones

PORQUE? 336 / 3 (códon) = 112

7. Se usa un ARNm sintético de secuencia repetitiva 5'-CACACACACACACACAC... en un sistema sintetizador de proteínas, en ausencia de células. Asumiendo que la síntesis de proteínas pueda comenzar sin la necesidad de un codón de iniciación, ¿qué producto o productos pueden esperarse tras la síntesis de proteínas?

Una proteína, con una secuencia de dos aminoácidos alternantes diferentes.

Dado que se ha supuesto que la iniciación puede proceder sin el codón AUG para la metionina, la traducción puede comenzar en cualquier lugar pero debe proceder en unidades de tres nucleótidos a partir del punto de inicio. Así, son codones posibles CAC, ACA, CAC, ACA independientemente del punto de iniciación. Dos codones pueden dar una proteína compuesta de dos aminoácidos alternantes

Poli(CA) codifica un polipéptido donde se alternan histidina y treonina.

8. Bajo condiciones donde la metionina debe ser el primer aminoácido:

a. ¿Qué proteína estará codificada por el siguiente ARNm?

/*El ARN mensajero (ARNm) es una molécula de ARN de cadena simple, complementaria a una de las cadenas de ADN de un gen. El ARNm es una versión del ARN del gen que sale del núcleo celular y se mueve al citoplasma donde se fabrican las proteínas. Durante la síntesis de proteínas, un orgánulo llamado ribosoma se mueve a lo largo del ARNm, lee su secuencia de bases, y utiliza el código genético para traducir cada triplete de tres bases o codón, en su aminoácido correspondiente.*/

Esta hay que corregirla...

- b. ¿Cuál es la secuencia de la hebra de ADN que codifica para este ARNm?
- 3'-GGAGTATACGCGGTAATATTCACTGTGTGT-5'
- c. ¿Cuál es la secuencia de la hebra de ADN molde sobre la que se sintetiza este ARNm?
- 5'-CCUCAUAUGCGCCAUUAUAAGUGACACACA-3'
- 3'-GGAGTATACGCGGTAATATTCACTGTGTGT-
- 5'-TGTGTGTCACTTATAATGGCGCATATGAGG-3'

9. ¿Qué entiende por epigenética? Explique brevemente.

13.1

La epigenética es una palabra de origen griego y significa literalmente por encima (epi) del genoma y es un campo emergente de la ciencia que estudia los cambios hereditarios causados por la activación y desactivación de los genes sin ningún cambio en la secuencia de ADN subyacente del organismo, es decir, se refiere al estudio de los cambios heredables en la expresión de los genes sin cambios en la secuencia del ADN: las marcas se producen en la cromatina –formada por ADN enrollado sobre proteínas y que contiene a los GENES antes de que sean interpretados-. Además existen moléculas capaces de regular a los ARN mensajeros, que son el producto de los genes una vez transcritos.

10. Explique brevemente cual es la importancia de la diferenciación celular.

La diferenciación celular es el proceso por el que las células adquieren una forma y una función determinada. La morfología de las células cambia notablemente durante la diferenciación, pero el material genético o genoma permanece inalterable, con algunas excepciones.

Por este mecanismo una célula no especializada se especializa en numerosos tipos celulares. Su importancia radica en que sin la diferenciación no habría tejidos ni órganos especializados en realizar una función, dado a que todas las células serían iguales.

Permite que la célula realice determinados procesos y cumpla ciertas funciones. De este modo también da lugar a que un tejido se encargue de determinadas funciones, diferentes a las de otros tejidos y los procesos a nivel sistémico sean más eficientes y regulables.

Cuando se pierde la diferenciación celular las células se dividen sin ningún control y se desarrolla cáncer, ya que son células que no pueden cumplir con su labor o han perdido su regulación.

Índice de comentarios

2.1	Faltan definiciones y fuentes de dónde fueron sacadas. Además de la selección de una de ellas.
4.1	¿este es un objetivo de la biología?
4.2	No sólo en salud humana
4.3	¿Quién desarrolla dispositivos?
6.1	envoltura
7.1	Hay partes del genoma que se transcriben y se produce la expresión de transcriptos que no necesariamente codifican para proteínas.
7.2	ن؟
7.3	ن؟
7.4	¿Qué se multiplica? Por otro lado, si se quiere referia a la lectura de la hebra de ADN, esta SIEMPRE es de 5' a 3'
7.5	¿Le dicen al cuerpo? ¿Le dicen cómo?
8.1	definición no adecuada
8.2	En realidad los exones codifican para cadenas de aminoácidos. Luego, la combinación de la expresión de los exones contituye una proteína.
8.3	Т
13.1	La redacción del párrafo es poco clara