

2022

# Genética

GUÍA DE SEMINARIOS

—Dime una cosa, Glaucón: veo que en tu casa hay perros cazadores y gran cantidad de aves de raza. ¿Acaso, por Zeus, no prestas atención a los apareamientos y crías de estos animales?

—¿Cómo? —preguntó.

—En primer lugar, ¿no hay entre ellos, aunque todos sean de buena raza, algunos que son o resultan mejores que los demás?

—Los hay.

—¿Y tú te procuras crías de todos indistintamente o te preocupas de que, en lo posible, nazcan de los mejores?

—De los mejores.

Platón, *La República*.  
Genética aplicada hace 2.400 años.

**Horarios de cursada**

*Turno mañana:*

Comisión 1 y 2 los martes de 9:00 a 13:00 h.

*Turno tarde:*

Comisión 3 los martes de 14:00 a 18:00 h.

**PERSONAL DOCENTE Y HORARIOS DE CONSULTA**

<b>Día</b>	<b>Horario</b>	<b>Docente</b>
Lunes	14:00 a 16:00	Dr. Aníbal Lodeiro
Miércoles	10:00 a 12:00	Dra. Ma. Silvia Tacaliti
Jueves	9:00 a 11:00	Ing. Ftal. Florencia Bongiorno
Jueves	14:00 a 16:00	Dra. Luciana Saldúa
Viernes	10:00 a 12:00	Dra. Érica Tocho

**Bibliografía**

Alberts, B., Lewis J., Johnson A. (2003) *Biología Molecular de la Célula*, 4ª ed. Editorial Omega. Disponible en la Biblioteca Central de la Facultad.

Falconer DS, Mackay TFC (2001) *Introducción a la Genética Cuantitativa*. Acribia. Disponible en la Biblioteca Central de la Facultad.

Griffiths A.J.F., Wessler S.R., Lewontin R.C, Carroll S.B. (2008) *Genética*, 9a. ed. McGraw Hill-Interamericana. Disponible en la Biblioteca Central de la Facultad.

Klug W.S., Cummings M.R. (2013) *Conceptos de Genética*, 10a ed. Pearson. Disponible en la Biblioteca Central de la Facultad.

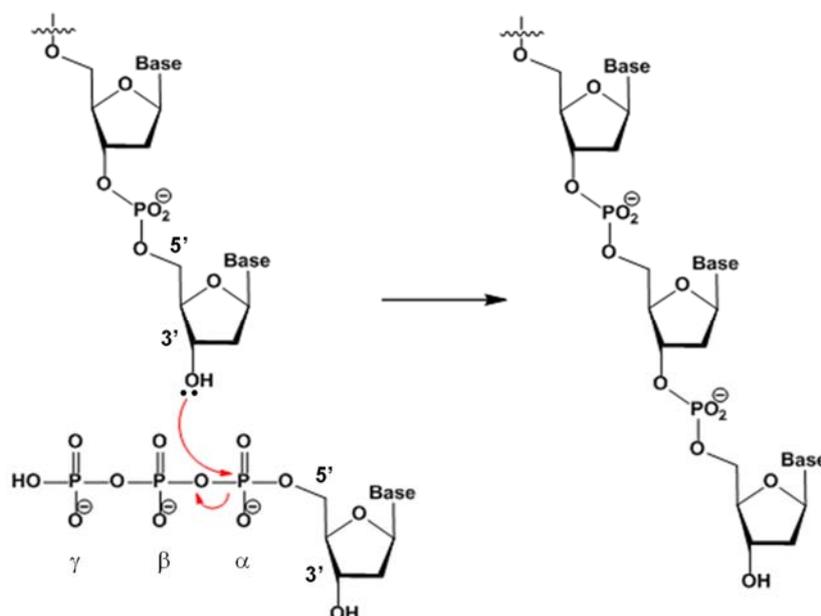
Lodish H., Berk A., Zipusky S.L., Matzudaira P., J., Baltimore D. & Darnell J. (2002) *Biología Celular y Molecular*, 4a ed. Editorial Médica Panamericana. Disponible en la Biblioteca Central de la Facultad.

## 1. Naturaleza y replicación del material genético.

Sin dudas, la capacidad de replicarse en forma autónoma a partir de sustancias simples y energía del entorno es una de las características sobresalientes de los organismos vivos. Esta capacidad de una célula para producir otra célula reside en el hecho de que las "instrucciones" para generar una célula se encuentran en su ADN, y éste puede copiarse. De este modo, una célula se divide en dos células hijas iguales, cada una de las cuales se divide a su vez produciendo dos células hijas iguales, y así sucesivamente. Por lo tanto, a partir de una célula inicial tendremos dos, a partir de estas dos:  $2 \times 2 = 4$ , a partir de estas cuatro:  $4 \times 2 = 8$ , a partir de las ocho:  $8 \times 2 = 16 \dots$  es decir  $2 \times 2 \times 2 \times 2 = 2^4$ . O bien, para decirlo en forma general, a partir de una sola célula obtendremos  $2^n$  células en  $n$  generaciones. ¿Qué pasaría si en vez de partir de una sola célula partiéramos de más células, por ejemplo 3? Se cumpliría lo mismo: en la primera generación tendríamos  $3 \times 2 = 6$  células, en la segunda  $6 \times 2 = 12$  células, en la tercera  $12 \times 2 = 24$  células... es decir:  $3 \times 2$ ;  $3 \times 2 \times 2$ ;  $3 \times 2 \times 2 \times 2 \dots$  o sea  $3 \times 2^1$ ;  $3 \times 2^2$ ;  $3 \times 2^3 \dots$  y generalizando, podemos decir que a partir de un número inicial de células igual a  $N_i$ , en  $n$  generaciones obtendremos un número final  $N_f = N_i \times 2^n$ .

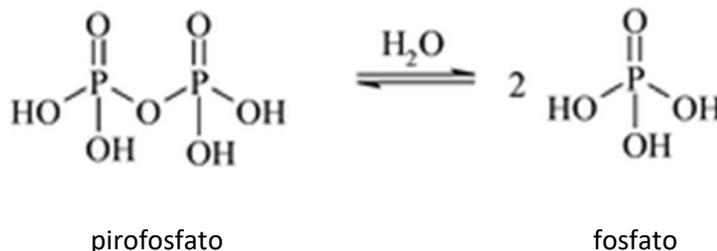
De la misma forma, a partir de una doble cadena de ADN obtendremos  $2^n$  dobles cadenas en  $n$  generaciones y a partir de  $N_i$  cadenas,  $N_f = N_i \times 2^n$  dobles cadenas en  $n$  generaciones. Este proceso de división para dar dos descendientes iguales puede durar indefinidamente, mientras haya alimento y energía. ¿Cuál es el "alimento" de las dobles cadenas de ADN? Son los desoxirribonucleótidos trifosfato, que llamaremos genéricamente dNTP, que serán los que se vayan integrando a la doble cadena a medida que ésta va siendo polimerizada. ¿Y la energía? Justamente, también está en los dNTP y aquí reside quizás uno de los hechos más asombrosos: este proceso es **espontáneo**.

¿Qué quiere decir que sea espontáneo y por qué es asombroso? Vayamos por partes: existen unos principios fisicoquímicos, llamados principios termodinámicos que nos dicen cuándo un proceso es espontáneo y cuándo no. Hay que aclarar que la idea de espontaneidad no tiene que ver con la cinética: que un proceso sea espontáneo no quiere decir que sea rápido, simplemente quiere decir que el estado final que nosotros postulamos es más estable que el estado inicial, y por eso el *sistema tiende a moverse desde el estado inicial al final*. Por ejemplo, si nos preparamos un café con leche, primero tenemos el café y la leche por separado, mezclarlos es fácil, pero una vez que los mezclamos ya no los podemos volver a separar a menos que utilicemos algún procedimiento químico bastante sofisticado. ¿Por qué? Ambas sustancias [el café y la leche] difundieron una en otra durante la mezcla, y el estado final en donde ambas sustancias están repartidas homogéneamente en la taza es más estable que si en media taza hubiera leche y en la otra media hubiera café. El estado final "café con leche" es, entonces, más desordenado que el estado "café por un lado y leche por el otro". El Segundo Principio de la Termodinámica capta este hecho con una magnitud llamada **entropía**, que hace referencia al "desorden" de un sistema. Todos los sistemas tienden al aumento de la entropía, lo que significa que tienden al desorden. Para ser más precisos, todos los sistemas poseen un cierto nivel de energía, pero dentro de ese nivel, cada componente del sistema puede tener una configuración diferente. Cuantas más configuraciones haya en un sistema, más desordenado va a estar. Esto es lo que pasa cuando un dNTP se une a una cadena de ADN creciente. Veamos el mecanismo de la reacción en la Fig. 1.1:

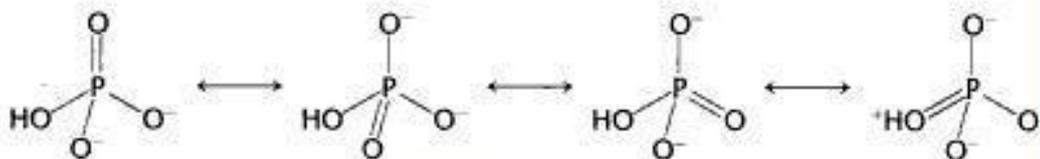


**Fig. 1.1** Esquema del mecanismo de incorporación de un dNTP a una cadena de ADN. Se muestra la ubicación de los fosfatos  $\alpha$ ,  $\beta$  y  $\gamma$  en el dNTP entrante. Las flechas rojas indican el movimiento de los pares de electrones. Por simplicidad, se muestra solo una de las dos hebras.

Es decir, un par de electrones libres del  $\text{-OH}$  que se encuentra en la posición 3' de la cadena de ADN ataca al enlace éster (entre dos ácidos) que se encuentra uniendo el fosfato  $\alpha$  (más próximo a la ribosa) con el fosfato  $\beta$  (el que está en el medio) en el extremo 5' del dNTP entrante. La consecuencia es que este par de electrones se "lleva" el desoxirribonucleótido monofosfato (dNMP) a la cadena de ADN, quedando libres los dos fosfatos que estaban en el extremo 5', dando origen a una molécula de *pirofosfato*:  $(\text{dNMP})_n + \text{dNTP} \rightarrow (\text{dNMP})_{n+1} + \text{pirofosfato}$ . Este pirofosfato luego se hidroliza a dos fosfatos:



Los dos fosfatos que se generan pueden adoptar distintas configuraciones, por *resonancia* del doble enlace entre el átomo de P y uno de los átomos de O. Es decir, el doble enlace que en el diagrama superior está escrito para arriba podría escribirse para los costados o para abajo, y seguiría siendo la misma molécula:



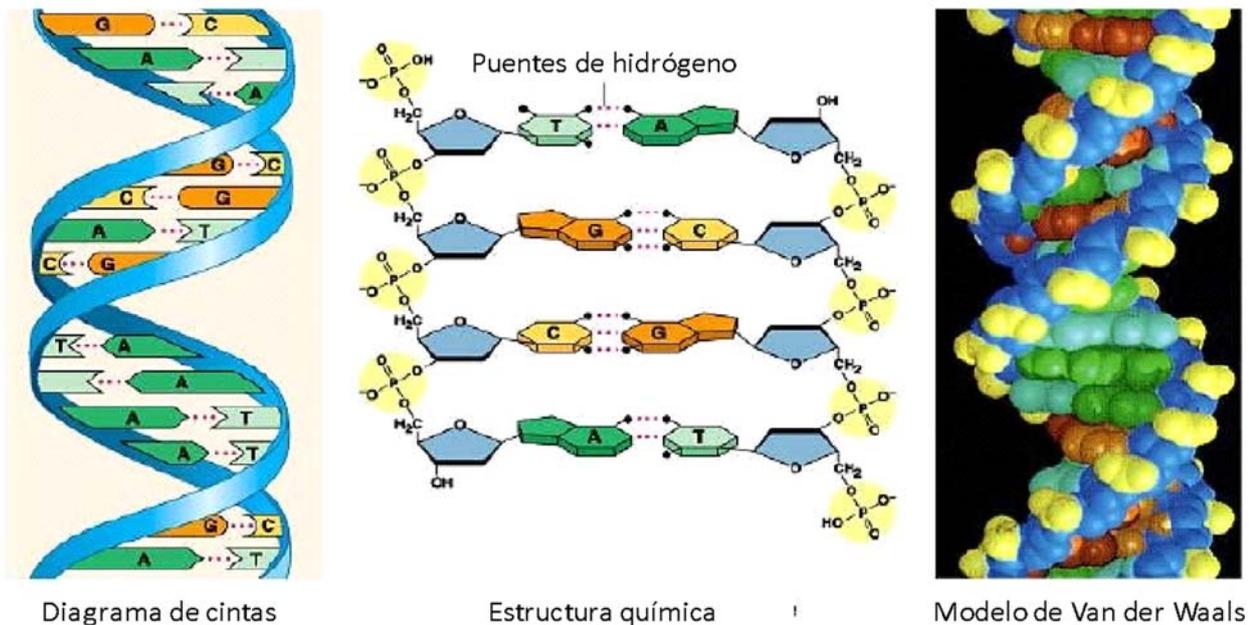
Y aquí hay una clave importante: como cada molécula de fosfato puede encontrarse en cuatro configuraciones distintas, podemos preguntarnos cuántas configuraciones obtendremos con las dos moléculas de fosfato **simultáneamente**. Por ejemplo, si una de las moléculas puede estar en la configuración que se ve a la izquierda en el esquema, la otra podría estar en cualquiera de las cuatro configuraciones, y por lo tanto en este caso tendremos cuatro combinaciones entre las dos ( $1 \times 4$ ). Si agregamos que la primera molécula esté en la siguiente configuración, también aquí tendríamos cuatro posibilidades con la segunda molécula, es decir que ya vamos en  $2 \times 4 = 8$  combinaciones de configuraciones distintas. Y si consideramos el total, con la primera en cualquiera de las cuatro combinaciones y la segunda también en cualquiera de las cuatro combinaciones, para el conjunto tendremos  $4 \times 4 = 16$  posibilidades, es decir,  $4^2$ . Si aplicamos el mismo razonamiento a la molécula de pirofosfato que les dio origen, veremos que, si en el resto fosfato de la izquierda la doble ligadura está "para arriba" como en el esquema, en el resto fosfato de la derecha puede estar en cualquiera de *tres* posiciones, ya que la doble ligadura no puede ubicarse en el enlace éster, porque nos quedaría un átomo de O con tres enlaces (hacé la prueba). Esto significa que, en el pirofosfato tendremos solamente  $3^2 = 9$  posibles configuraciones. Entonces, el sistema con los dos fosfatos resultantes de la reacción es más desordenado que el sistema donde el pirofosfato aún no había reaccionado, y por eso el proceso es espontáneo en la dirección de la hidrólisis del pirofosfato. Esto es lo que impulsa la síntesis de ADN: cuanto más pirofosfato se consume, más dNTP tiende a unirse al ADN por la ley de acción de masas.

Esto nos explica por qué la síntesis del ADN es en el **sentido 5'  $\rightarrow$  3'**: debe haber un  $\text{-OH}$  libre (en el extremo 3' de la molécula) capaz de atacar al enlace éster entre los fosfatos  $\alpha$  y  $\beta$  del dNTP entrante (que están en su extremo 5'), que una vez unido a la cadena de ADN en crecimiento deja un extremo 3'  $\text{-OH}$  libre, que continúa el proceso.

Sin embargo, como dijimos antes, la espontaneidad del proceso no nos dice nada acerca de su *velocidad*. Si bien las reacciones descritas arriba son espontáneas, son muuuuuuy lentas, a menos que se agregue un catalizador. En la célula, este catalizador es una enzima llamada ADN polimerasa. Pero esta enzima hace algo más que acelerar la reacción. Como recordarás, el ADN es una doble cadena antiparalela, estabilizada por puentes de hidrógeno entre las bases adenina (A) y timina (T) y entre guanina (G) y citosina (C). En realidad, hay dos puentes de hidrógeno entre cada par A y T y tres puentes de hidrógeno entre cada par G y C, con lo cual estos últimos pares son más estables. Así, cada cadena sirve de *molde* para la otra: una cadena 5'ACGT3' será molde de (se apareará con) otra 3'TGCA5' (antes de seguir leyendo, escribí ambas cadenas, una arriba de la otra, y representá con dos líneas los puentes de hidrógeno que

unen A con T y con tres líneas los puentes de hidrógeno que unen G con C). Observá que ambas cadenas son *antiparalelas*.

Las bases, a su vez, se dividen en purinas y pirimidinas, siendo las purinas más grandes que las pirimidinas. Tanto A como G son purinas, mientras que C y T son pirimidinas. Es decir que con el apareamiento de bases que mencionamos, siempre una purina se une con una pirimidina. Dado que las bases están hacia el interior de la doble hélice de ADN, este apareamiento de bases determina que el *diámetro* de la molécula sea constante (Fig. 1.2). Si, por el contrario, se unieran purinas con purinas y pirimidinas con pirimidinas, el contorno de la estructura “ondularía” y se parecería más bien a un chupetín.



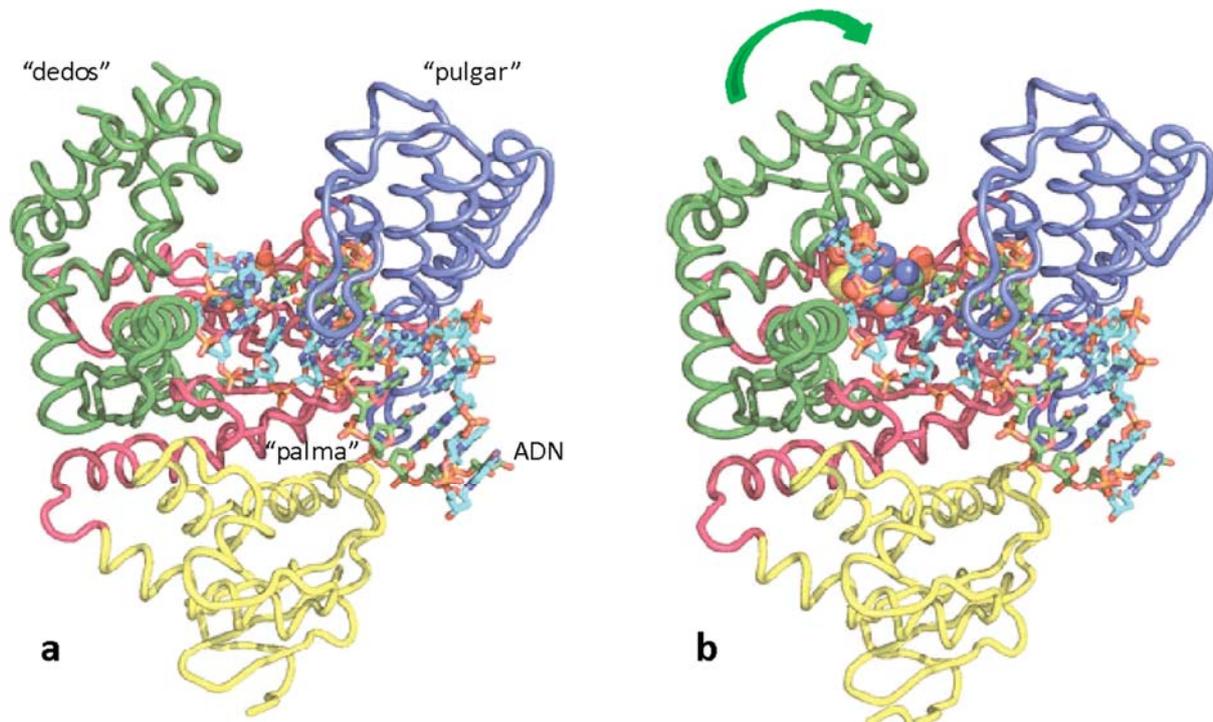
**Fig. 1.2** Estructura de la doble hélice del ADN. A la izquierda se muestra un “diagrama de cintas”, que representa solo las cadenas de azúcar-fosfato (en celeste) y los planos de las bases complementarias. Al centro se representa la estructura química simplificada de los fosfatos (amarillo), las desoxirribosas (celeste) y las bases purínicas (A y G, con dos anillos) y pirimidínicas (C y T, con un anillo). Observá la polaridad de cada cadena por la orientación de los anillos de desoxirribosa. A la derecha se muestra la doble hélice con los radios de Van der Waals representados como esferas, con los mismos colores que en la representación química del panel central.

Esto es detectado por la ADN polimerasa, porque para que ocurra la reacción, la doble cadena de ADN debe pasar por una especie de canal que hay en el interior de la enzima, de forma tal que, si se deforma el cilindro, se “traba” y no puede continuar. El sitio activo de la ADN polimerasa tiene una forma que recuerda una mano derecha con la palma hacia arriba y los dedos casi cerrados, como si sostuviera una botella en posición horizontal. El canal que mencionamos está formado por la cavidad entre los “dedos”, el “pulgár” y la “palma”, como se muestra en la Fig. 1.3a. Cuando el ADN entra en el canal, se apoya en la “palma” y el próximo dNTP entra por la zona de los “dedos”. Si el dNTP es el complementario al que está en el molde, se producen una serie de interacciones por puentes de hidrógeno entre el dNTP, el molde y los aminoácidos del sitio activo, de tal forma que los “dedos” se cierran (Fig. 1.3b). Eso posibilita que otros aminoácidos de la proteína, que se encuentran entre los “dedos” y la “palma”, y que poseen carga tales como lisina (carga positiva) y aspartato (carga negativa), se posicionen cerca del -OH 3’ del extremo del molde y el trifosfato 5’ del dNTP entrante y de esta forma faciliten el ataque nucleofílico “empujando” los electrones del -OH sobre el enlace éster entre los fosfatos  $\alpha$  y  $\beta$ , que se mostró en la Fig. 1.1. Sin embargo, si el dNTP que se posicionó en el sitio activo no es el complementario al molde, los “dedos” no pueden cerrarse y el dNTP es expulsado de allí.

De esta forma, la estructura de la enzima garantiza que se forme el par de bases correcto sea cual sea la secuencia de bases y al mismo tiempo, cataliza (es decir, acelera) la reacción espontánea de incorporación del dNTP. Esto es la clave que permite que una doble hélice de ADN genere otra igual, y finalmente, que una célula genere otra igual. Es decir, es el sentido de la vida:  $5' \rightarrow 3'$ .

Ahora podríamos preguntarnos si se podría generar una doble cadena de ADN mezclando en un tubo de ensayo (es decir, *in vitro*) un ADN molde, los cuatro dNTP y la ADN polimerasa. La respuesta es no. Entonces, ¿qué más hace falta?

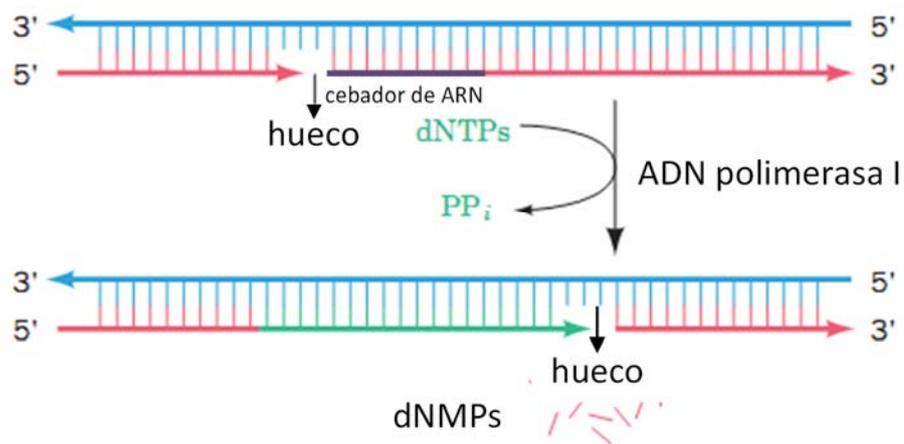
Una cosa importante es que, para “ofrecer” la cadena molde, las dos hebras de la doble hélice de ADN deben separarse, es decir, deben romperse los puentes de hidrógeno entre ellas. Esto requiere energía: por ejemplo, cuando hervís agua en una pava, el vapor de agua que se levanta consiste en moléculas de agua sueltas, porque el calor de la hornalla rompió los puentes de hidrógeno. Podríamos romper los puentes de hidrógeno con calor, cerca del punto de ebullición y como veremos, es lo que se hace en ciertas reacciones. Pero la célula no pasa por periodos de ebullición cada vez que tiene que replicar su ADN. En vez de ello, intervienen unas enzimas llamadas *helicosas*, que usan la energía del ATP para romper los puentes de hidrógeno. Sin embargo, con esto no alcanza. Los puentes de hidrógeno son estables, con lo cual, para el sistema es más estable tenerlos que no tenerlos. Es decir, si se rompieron, el sistema tenderá espontáneamente a formarlos de nuevo y eso puede ocurrir muy rápidamente, antes de que pase la ADN polimerasa. Para evitarlo, hay unas proteínas pequeñas, llamadas *proteínas de unión a simple cadena* o *SSB* (por su sigla en inglés) que inmediatamente se unen a las simples hebras, estabilizándolas antes de que pase la ADN polimerasa.



**Fig. 1.3** Estructura del sitio activo de la ADN polimerasa con una sección de doble hélice unida. Todas las moléculas se representan con un diagrama de “cintas” similar al de la Fig. 1.2. La región amino-terminal de la proteína se representa en amarillo, la región de los “dedos” en verde, la región del “pulgar” en azul y la región de la “palma” en fucsia. La cadena de ADN está apoyada sobre la “palma”. Los fosfatos se ven en naranja, los anillos de ribosa de la cadena molde en celeste, los de la cadena en crecimiento en verde, y los anillos de las bases en azul, al interior de la doble hélice. En **a**, no se ha incorporado el nuevo nucleótido y los “dedos” están abiertos. En **b** se está incorporando el nuevo nucleótido (dCTP, representado con esferas de radios de Van der Waals, como en la Fig. 1.2, panel derecho) y los “dedos” se cierran, posibilitando la catálisis del ataque nucleofílico de los electrones del 3'-OH de la cadena en crecimiento sobre el enlace fosfodiéster entre los fosfatos  $\alpha$  y  $\beta$  del dCTP (ver Fig. 1.1). Colocá vos los extremos 5' y 3' en las cadenas de ADN.

Otra necesidad, algo menos evidente, es que la ADN polimerasa no puede iniciar la síntesis de ADN “posándose” en una simple cadena. Si mirás de nuevo la Fig. 1.3, notarás que la “palma” requiere que por allí pase una doble cadena. Para generar la primera región de doble cadena se necesita que otra proteína produzca un pequeño fragmento de ácido nucleico que, una vez unido al molde, genera el fragmento de doble cadena que necesita la ADN polimerasa para empezar a catalizar la síntesis del ADN. Estos fragmentos son formados también en forma complementaria al molde, pero interviene otra enzima, del tipo ARN polimerasa, que no necesita de un fragmento de doble cadena para empezar, porque su sitio activo es algo diferente. Las ARN polimerasas son extremadamente importantes para el proceso de *transcripción* de la información genética y las abordaremos más tarde. La clave es que estas enzimas catalizan la síntesis de una pequeña cadena de *ácido ribonucleico* o ARN que es complementaria a una zona del ADN molde y así se genera el fragmento de doble cadena desde el cual la ADN polimerasa puede empezar a copiar. Dada esta función, a esos pequeños fragmentos de ARN se los denominó *cebadores* o *primers* (en inglés), y en un alarde de imaginación, a la enzima que cataliza la formación de los *primers* se la llamó *primasa*.

Pero esto nos genera otra necesidad: como la cadena de ADN se va sintetizando por partes, hay varios sitios a lo largo de la cadena adonde la ADN polimerasa inició procesos de síntesis, cada uno de los cuales comienza con un corto tramo de ARN cebador. Por lo tanto, si los cebadores o *primers* quedan a cada rato en la doble cadena de ADN, al cabo de un tiempo, va a estar toda “sucia” de ARN. Entonces, hay que remover los cebadores después de usarlos. De eso se encarga una ADN polimerasa en una versión algo diferente. Esta enzima posee una actividad llamada *exonucleasa* cuyo sitio activo se encuentra en un lugar alejado del sitio polimerasa, que cataliza la biosíntesis del ADN. La actividad exonucleasa cataliza la hidrólisis de los cebadores, nucleótido por nucleótido a medida que esta ADN polimerasa avanza, y simultáneamente, la actividad polimerasa va rellendo el sitio donde estaba el cebador con ADN, a partir del extremo 3’ del tramo anterior. De esta manera, siempre va quedando un “hueco” entre la base que le sigue a la que sacó y la que colocó, donde no hay enlace fosfodiéster. Este “hueco” se va corriendo a medida que la ADN polimerasa avanza (por eso en inglés se conoce este fenómeno como “nick translation”, o “traslado de la mella”) hasta que llega un punto en el que la ADN polimerasa no da más y se suelta del ADN. En ese momento, queda un “hueco” entre dos bases adyacentes, que no poseen enlace fosfodiéster ni hay trifosfato en el extremo 5’ de la próxima base, sino que allí queda un monofosfato (Fig. 1.4). Para poder rellenar este “hueco” actúa otra proteína llamada *ligasa*, que como no dispone de trifosfato en el extremo 5’, debe usar la energía del ATP para catalizar la formación del enlace fosfodiéster en el “hueco”.



**Fig. 1.4** Traslado de la mella o *nick translation*, catalizada por la ADN polimerasa I. La síntesis del ADN se produce de a tramos usando una cadena como molde (celeste), con lo cual, entre un tramo terminado en desoxirribonucleótidos (rosa) y el siguiente, que empieza con el cebador de ribonucleótidos (violeta), se genera un hueco o mella (parte superior). La ADN polimerasa I tiene actividad exonucleasa 5' → 3', con la cual cataliza la remoción de los ribonucleótidos del cebador, y además posee actividad polimerasa 5' → 3', con la cual cataliza la adición de desoxirribonucleótidos (dNTPs) a partir del extremo 3' que había quedado en el hueco. De esta manera, se van reemplazando los ribonucleótidos por desoxirribonucleótidos (verde), pero las actividades exonucleasa 5' → 3' y polimerasa 5' → 3' funcionando al mismo tiempo hacen que el hueco se vaya trasladando hacia el extremo 3' de la cadena. Una vez que la ADN polimerasa I se suelta de la cadena, que en ese momento ya es toda de ADN, permanece una mella adonde estaba el hueco, la cual debe sellarse mediante la acción de una ligasa, puesto que el nucleótido que queda en el extremo 5' de la mella no posee trifosfato (parte inferior).

Todas estas proteínas avanzan juntas durante la replicación, en una estructura compleja, llamada *horquilla de replicación*. El avance de la horquilla de replicación determina que se enrolle el ADN que está adelante, como cuando intentamos “abrir” una soga o una trenza simplemente tirando desde sus extremos. Para aliviar esa tensión, actúan otras proteínas llamadas *giratas*, que van cortando y rearmando enlaces por delante de la horquilla de replicación, también usando la energía del ATP para esa tarea.

Finalmente, debe observarse que, como el avance de la síntesis de ADN es en el sentido 5'→3', las dos hebras no pueden sintetizarse avanzando simultáneamente en el mismo sentido. La solución a esto es que, mientras una de las hebras avanza continuamente en el sentido 5'→3' la otra lo hace de a “saltitos” yendo “marcha atrás” sobre la otra hebra molde. Esos pedacitos de cadena luego se unen gracias a la actividad de la ADN ligasa. Por este modo particular de avance de la horquilla de replicación, a una de las hebras (la que avanza sin parar) se le llama *hebra continua* y a la otra (la que avanza de a “saltitos”) se le llama *hebra retrasada* o *discontinua*. Los fragmentitos de ADN de la hebra retrasada se conocen como *fragmentos de Okazaki* en honor al matrimonio de Reiji Okazaki y Tsuneko Okazaki, quienes fueron sus descubridores.

Ciertamente, toda esta explicación necesita un dibujo para ser mejor comprendida. Pero no lo vamos a poner, y en vez de ello, te sugerimos que, releyendo los párrafos anteriores, resuelvas los problemas 1.7 y 1.8, donde están las figuras que necesitas.

Además, existen muy buenas animaciones del proceso de replicación. Por ejemplo,

<https://www.youtube.com/watch?v=uEwyWgSvLc0>

<https://www.youtube.com/watch?v=YqjbmRQcyfM>

Después de leer este texto, ver los videos y resolver los problemas, te sugerimos que pienses en lo siguiente, para dimensionar el proceso.

1) Para dividirse en dos células nuevas, una bacteria tiene que replicar su ADN en unos 40 minutos. Si ese ADN contiene 4.000.000 de pares de bases ¿cuántos nucleótidos por segundo incorpora?

2) Sin embargo, dijimos que lo que se genera no es solo ADN, sino una **célula**, es decir que en ese tiempo tiene que fabricar también todos los componentes de la nueva célula.

3) Tratá de copiar esto lo más rápido que puedas, y fijate cuánto tardás:

ACGTTTCTACCGTCAGGGTGTGCATGTGTGCAAAATATCGACCGGTCAGCAGACCTCGATCGAGACGAGCAGATGCAGAGGGCT

4) Fijate cuántas veces te equivocaste.

5) Ahora compará con el tiempo que le lleva a la célula, y tené en cuenta que ella se equivoca menos de una vez en cada copia.

6) Y observá que la célula no solamente copió la secuencia, sino que mientras lo hacía iba fabricando el papel y la birome....

7) Recomendación final: **leé el libro. Esto no lo reemplaza.**

### Problema 1.1

Uno de los procesos más importantes para la agricultura sustentable es la fijación biológica de  $N_2$ , llevada a cabo por un grupo de bacterias denominadas diazótrofos. Uno de estos diazótrofos es *Ensifer meliloti*, que realiza la fijación biológica de  $N_2$  en simbiosis con plantas de alfalfa, las cuales, gracias a la simbiosis, pueden nutrirse con el N del aire en vez de usar el N del suelo. Bajo condiciones óptimas para el crecimiento, una célula de *E. meliloti* se divide cada 2 horas, aproximadamente. Supongamos que ninguna célula se muriera en estas condiciones.

- ¿Cuánto tiempo le tomaría a una sola célula de *E. meliloti* en un frasco de cultivo de 10 litros llegar a una densidad de células máxima de  $10^{10}$  células por ml (cultivo saturado)?
- ¿Cuántas células se obtendrían a partir de una célula luego de una semana de crecimiento?
- Sabiendo que una célula de *E. meliloti* pesa 0,95 pg ¿cuánto pesarían las células obtenidas en el inciso b)?
- ¿Entonces...?

### Problema 1.2

En la Tabla 1.1 se detalla la composición porcentual de A, T, C y G del ADN de cuatro especies diferentes, pero se han omitido ciertos valores. Completá los datos faltantes en la tabla y calculá la relación  $(A+T)/(G+C)$

	A	T	C	G	$(A+T)/(G+C)$
Humanos	31		20	20	
Trigo	27	27			
Ovejas				21	

### Problema 1.3

- El ADN en solución se desnatura mediante el empleo de calor. ¿Por qué ocurre esto?
- ¿Cómo podrías medir el grado de desnaturación que posee una molécula de ADN a distintas temperaturas?
- La temperatura a la cual se alcanza la mitad de la desnaturación de una molécula de ADN ¿depende de su composición de bases?
- Hacé un gráfico que muestre el proceso de desnaturación a medida que aumenta la temperatura para una molécula de ADN rica en (A+T) comparada con otra rica en (G+C).
- Una vez que una molécula de ADN ha sido desnaturada por calor ¿puede renaturalizarse por enfriamiento? ¿Cómo harías el experimento que corrobore o no esta pregunta?

### Problema 1.4

- Bajo ciertas condiciones el ADN posee carga eléctrica. ¿Cuáles son esas condiciones y a qué grupos moleculares ionizables se debe esta carga? Explicalo con un esquema de la molécula.
- Esta propiedad se aprovecha en un método de separación de las moléculas del ADN llamado electroforesis en gel. Esquematizá el procedimiento y cómo se vería el resultado de uno de estos experimentos si se separaran dos moléculas de ADN de distinto tamaño molecular.
- Si el procedimiento se basa en la carga eléctrica de la molécula, ¿por qué las separa por tamaños?

### Problema 1.5

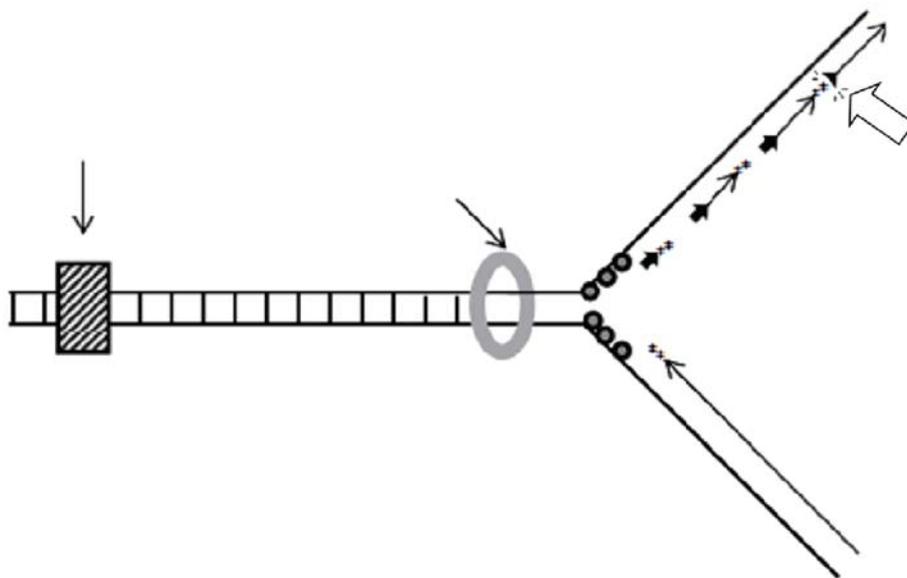
En función de la reacción de polimerización del ADN catalizada por las ADN-polimerasas:



- a) Completá la ecuación de la reacción, con los elementos necesarios para que la polimerización ocurra.
- b) ¿Cómo se aporta la energía necesaria para que se produzca la reacción?
- c) Esquematizá al menos el último nucleótido del  $(\text{dNMP})_n$  incluyendo el oxidrilo 3', el dNTP (incluyendo el extremo 5') y cómo ocurre el ataque y la formación del enlace fosfodiéster.
- d) Sobre la base de este esquema, explicá por qué esta reacción es termodinámicamente espontánea.

### Problema 1.6

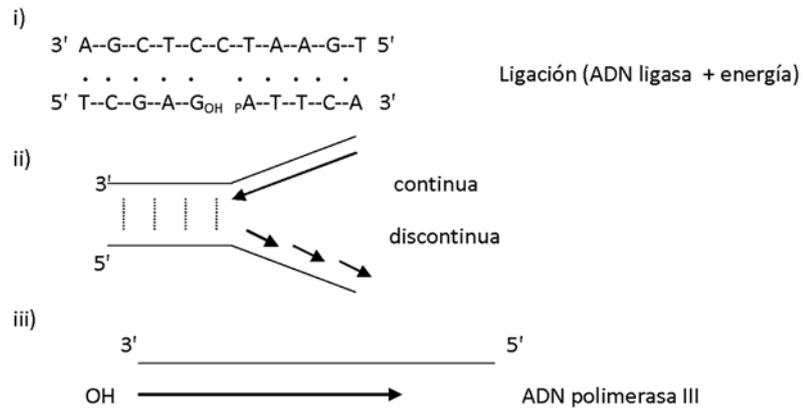
Considerá el esquema de la siguiente horquilla de replicación, donde la flecha indica el sentido de la reacción de adición de desoxiribonucleótidos por las ADN polimerasas, y el \* indica la incorporación en el ADN de dNTPs marcados radioactivamente:



- a) Marcá los extremos 5' y 3' de las cadenas del ADN. Identificá las enzimas representadas por un rectángulo y un óvalo gris, y describí las actividades que realizan. Agregá otras enzimas implicadas en el replicón. Mencioná cuáles necesitan ATP para realizar su actividad.
- b) Ubicá dónde debería encontrarse la primasa en el esquema. ¿Qué tipo de polimerasa es esta enzima?
- c) En la figura hay una flecha blanca que muestra que se ha acortado la molécula indicada con una flecha gruesa y corta (➡) ¿Cómo ocurre este proceso?
- d) ¿Por qué una de las hebras se sintetiza de manera discontinua? Explicá de qué forma se completa dicha hebra.

### Problema 1.7

Los siguientes esquemas corresponden a eventos replicativos en procariotas. Indicá si contienen errores, fundamentando brevemente tu respuesta:



## 2. Organización genética.

El ADN es una secuencia de desoxirribonucleótidos, pero ¿qué hay ahí? Muchas cosas, pero lo que más nos importa son los *genes*. A grandes rasgos, podemos decir que los genes son secuencias de desoxirribonucleótidos en el ADN. Es decir, secuencias dentro de la secuencia. Por ejemplo, la bacteria que mencionamos en la introducción del tema anterior tenía 4.000.000 de pares de bases en su ADN, y dentro de esa gran molécula de ADN hay alrededor de 4.500 genes. Es decir, hay unos 4.500 fragmentos de secuencia (genes) dentro de la secuencia principal, a la que llamamos *genoma*, cada uno de los cuales está encargado de codificar una **función** dentro de la célula. Es importante darse cuenta de que los genes siempre están en el mismo lugar en el genoma, de la misma forma que cada palabra de este texto ocupa un lugar determinado, y si se copiara el texto, cada palabra en cuestión debería ocupar el mismo lugar en todas las copias.

Por lo tanto, tenemos una primera diferencia entre genes y genomas. Cada especie tiene un solo genoma, que contiene todos sus genes; y el número de genes es muy variable de especie a especie: por ejemplo, mientras la bacteria que mencionamos tiene 4.500 genes, el ser humano tiene unos 30.000. Sin embargo, el genoma humano es proporcionalmente mucho mayor que el de la bacteria. Mientras el genoma bacteriano que pusimos como ejemplo tiene 4 millones de pares de bases, el genoma humano tiene 3.000 millones de pares de bases. Son tamaños tan grandes, que los referimos como “mega pares de bases” (Mpb): 4 Mpb para la bacteria y 3.000 Mpb para el humano.

Una simple cuenta nos habla también de la diferencia de organización entre un genoma y otro: si dividimos el tamaño del genoma por el número de genes, el resultado nos dice (más o menos) cada cuántos pares bases (pb) hay un gen. En el caso del genoma bacteriano,  $4.000.000/4.500 = 889$ , es decir un gen cada 889 pb, mientras que, en el genoma humano,  $3.000.000.000/30.000 = 100.000$ , es decir un gen cada 100.000 pb. Como los tamaños de los genes son similares (medidos en números de pb), podemos concluir de esto que el genoma bacteriano está mucho más “densamente poblado” de genes que el genoma humano. Lo mismo se extiende a los genomas de otras especies: los genomas procariontes son más “densos” que los genomas eucariotes. Este aspecto aún no se comprende bien, y hay muchas hipótesis que intentan explicar estas diferencias.

Pero volvamos a los genes. Supongamos la siguiente secuencia de ADN, que es una muy pequeña parte del genoma de una especie hipotética (no se escribe la cadena complementaria, porque queda sobreentendida):

5'-TCTAATACCCGTGGCAGGCTTCTGCAGGAGGTCCATGGATGTCTCCTACACTGTCATTCTGGGAGATGGCCCTATATAGCC-3'

Es decir, solo se representa una parte de la secuencia total, pero hay que entender que sigue habiendo bases a la izquierda del 5' y a la derecha del 3'. Pues bien: en este trozo de secuencia ¿dónde está el gen? ¿Cómo “sabe” la célula qué parte contiene un gen (o sea una instrucción para una función) y qué parte no? En una secuencia de ADN lo único que hay es eso: secuencia de ADN. No es como este texto, que tiene mayúsculas, espacios, signos de puntuación, que indican dónde empieza una oración, dónde empiezan y terminan las palabras, etc. Entonces, si en la secuencia de ADN tiene que haber señales, como por ejemplo una mayúscula, esas señales tendrán que ser también secuencias de ADN. En la secuencia de arriba hay tales señales, que iremos viendo más tarde. Por de pronto, la secuencia **TAATA** indica dónde empieza el gen; la “**A**” que está 9 pb después es la primera que va a aparecer en el ARN mensajero (ARNm) transcrito; la secuencia **AGGAGGT** indica dónde debe unirse el ribosoma al ARNm para iniciar el proceso de traducción; la secuencia **ATG** inmediata posterior indica el inicio de la traducción a proteína, y la secuencia **TAG** cerca del final indica la terminación de la traducción a proteína. Tratá de encontrar todas estas señales en la secuencia escrita más arriba.

Estas señales indican a los aparatos de transcripción y traducción dónde deben ejecutar sus tareas para que el gen se *exprese* en una proteína funcional. Entonces, podríamos decir que un gen debería tener todas estas señales, y la secuencia de aminoácidos de la proteína contenida entre ATG y TAG para ser considerado como tal.

Los genes se encuentran a lo largo del ADN en los cromosomas, y así como las distintas especies poseen distinto tamaño de genoma, también poseen distinto número de cromosomas. Por ejemplo, las bacterias en general tienen un solo cromosoma, pero es frecuente que posean *elementos de ADN extracromosomal* llamados *plásmidos*. Estos plásmidos en general son chicos (del orden de unos miles de pb) y poseen genes relacionados entre sí por su función. Por su parte, el ser humano tiene su genoma repartido en 23 cromosomas. Cada cromosoma está por duplicado en una célula somática, con lo cual ésta tiene 46 cromosomas. Los cromosomas pueden reconocerse por sus diferencias de tamaño y morfología (esencialmente por la ubicación del centrómero) y la clasificación de todos los cromosomas de una especie por tamaño y forma constituye el *cariotipo* de esa especie. Sin embargo, el ser humano también tiene

ADN extracromosomal, ubicado en las mitocondrias que pueblan el citoplasma. Las plantas, además de tener ADN extracromosomal en las mitocondrias, también lo tienen en los cloroplastos.

A lo largo de la evolución se fueron acumulando cambios (mutaciones) en el ADN, que se fueron seleccionando a favor y dieron origen así a la enorme biodiversidad que hoy habita nuestro planeta. Como los cambios a menudo son graduales, es posible reconocer el mismo gen en distintas especies comparando su secuencia de ADN. Más aún: puede *predecirse* la presencia de un gen en una especie solo por comparación de su ADN con el de especies donde ese gen haya sido caracterizado. Sin embargo, existe una cierta controversia sobre cuánto tienen que parecerse dos secuencias de ADN para concluir que se trata del mismo gen. Obviamente no es lo mismo comparar un gen humano con su homólogo en chimpancé que compararlo con su homólogo en levadura. De todos modos, mediante la comparación de una secuencia de ADN con las de genes homólogos en otras especies, es posible detectar las zonas que codifican las partes más importantes para la función de la proteína codificada por ese gen. Por ejemplo, es más probable que mutaciones que produzcan cambios en aminoácidos del sitio activo (como podría ser la zona entre los “dedos” y la “palma” de la ADN polimerasa esquematizada en la Fig. 1.3) den por resultado una forma inactiva de la proteína, que podría ser eliminada por la selección natural. Por lo tanto, es de esperarse que, comparando genes homólogos, encontremos menos cambios de nucleótidos en la zona que codifica el sitio activo que en otras zonas.

A continuación, daremos algunas definiciones útiles.

**Número x:** Número básico de cromosomas de una especie. En los individuos diploides, el número básico coincide con el número de cromosomas de un genoma ( $x = n$ ).

**Número n:** Número haploide o número gamético de un organismo.

**2n:** Número somático de cromosomas de un organismo.

**Haploide:** célula u organismo con una sola dotación cromosómica.

**Diploide:** célula u organismo que tiene dos dotaciones cromosómicas

**ADN:** ácido desoxiribonucleico. Es la molécula portadora de la información genética.

**Estructura Primaria:** en un polímero compuesto por nucleótidos, es la secuencia de nucleótidos. Cada nucleótido está compuesto por un azúcar pentosa (desoxirribosa en el caso del ADN, ribosa en el caso del ARN), una molécula de fosfato y una base nitrogenada, púrica (adenina o guanina) o pirimidínica (timina o citocina en el caso del ADN, o uracilo o citocina en el caso del ARN).

**Estructura secundaria:** basada en el modelo de Watson y Crick, el ADN está compuesto por dos cadenas **antiparalelas**, unidas por puentes de hidrógeno, helicoidales enrolladas alrededor de un eje imaginario. Este modelo requiere cantidades equimolares de las bases complementarias ( $A=T$  y  $G=C$ ) demostrado por Chargaff, siendo:

$\frac{(A+G)}{(T+C)} = 1$  universal y relación  $\frac{(A+T)}{(G+C)} \neq 1$  variando según la especie

En el caso del ARN, al ser monocatenario, los puentes de hidrógeno entre bases complementarias se forman en regiones donde la molécula se pliega sobre sí misma.

**Gen:** unidad física y funcional de la herencia, que se transmite de una generación a la siguiente. Constituido por ADN (en organismos celulares; los virus, por su parte, pueden tener genomas de ARN).

**Cromosomas:** son cuerpos filiformes constituidos por ADN y proteínas, su número es una característica constante de la especie. Luego de la duplicación o síntesis del ADN cada cromosoma estará constituido por dos **cromátidas hermanas** idénticas unidas por un centrómero para permitir la división celular. Dependiendo de la ubicación del centrómero, los cromosomas pueden clasificarse según su forma en metacéntricos, submetacéntricos, acrocéntricos o telocéntricos.

En organismos diploides ( $2n$ ), los cromosomas se encuentran de a pares. Cada miembro del par tiene el mismo tamaño, la misma posición del centrómero y los mismos genes, sin embargo, uno es de origen paterno y el otro materno. Este par de cromosomas constituye un **par de cromosomas homólogos**.

**Genoma o Genomio:** se denomina así al juego de cromosomas que un individuo recibe de uno de los progenitores diploides y por lo tanto está representado sólo uno de los cromosomas de cada par de homólogos. También se denomina genoma a la secuencia completa del ADN de la especie.

El número cromosómico es una característica de las células vegetales y animales. Dentro de cada especie existe una constancia numérica y estructural de los cromosomas, sin embargo, pueden ocurrir alteraciones en el número cromosómico normal. Hay dos clases principales de cambios cromosómicos numéricos: la euploidia y la aneuploidia.

Los cambios **euploides** afectan genomios enteros. Son aquellos organismos que tienen un número de cromosomas múltiplos del número **x**.

Los Monoploides ( $n$ ): llevan un juego de cromosoma o genomio, esto es, uno de cada cromosoma presente normalmente.

Los diploides ( $2n$ ) son los organismos que presentan dos juegos de cromosomas o genomios.

Los organismos con 3 o más genomios son llamados poliploides. La poliploidia es muy común en los vegetales, y para ciertas plantas constituye un factor importante en la evolución. Las dos clases principales de poliploides son los autopoliploides y alopoliploides, los cuales pueden distinguirse basándose en el origen de los cromosomas. Los autopoliploides aparecen cuando el mismo genomio se duplica. En cambio, los alopoliploides resultan cuando diferentes genomios se unen a través de la hibridación de plantas emparentadas o relacionadas entre sí.

Algunos ejemplos de poliploides:

Los triploides ( $2n=3X$ ): presentan 3 juegos de cromosomas o genomios.

Los tetraploides ( $2n= 4X$ ): presentan 4 juegos de cromosomas o genomios.

Los cambios **aneuploides** involucran a uno o más cromosomas aislados del complemento cromosómico. El número no es múltiplo del número básico  $X$ . En la naturaleza estas alteraciones son más graves que las euploides debido al desequilibrio génico que se produce en las divisiones celulares. Ciertas anomalías en la meiosis pueden llevar a producir gametas con  $n+1$  y  $n-1$  cromosomas.

Algunos ejemplos:

Monosómico  $2n -1$

Nulisómico  $2n - 2$

Trisómico  $2n+1$

Tetrasómico  $2n +2$

Doble Monosómico  $2n -1-1$

Doble Trisómico  $2n +1+1$

## Problema 2.1

Analizá e intentá responder las siguientes preguntas de concepto. Buscá la información que necesites.

- Explicá qué significan: gen, alelo, operón, exón, intrón, promotor, pseudogen, homólogo, ortólogo, parálogo, locus, elemento de secuencia, organela, plásmido.
- Considerando un genoma procarionta ¿todos los genes están codificados en la misma hebra de ADN? ¿Podrían dos genes distintos estar codificados en el mismo lugar, pero en hebras complementarias?
- Comentá el “dogma central de la genética molecular” y explicá por qué actualmente se lo considera una explicación insuficiente.
- ¿Por qué en los procariontas los genes tienden a agruparse en el genoma de acuerdo con su función?

## Problema 2.2

En la Fig. 2.1 se muestra un apilamiento de secuencias correspondientes a una región del gen que codifica el factor de transcripción RpoS en los siguientes microorganismos: 1) *Salmonella typhimurim*, 2) *Shigella flexneri*, 3) *Erwinia carotovora*, 4) *Yersinia enterocolitica*, 5) *Serratia entomophila*, 6) *Pseudomonas aeruginosa*, 7) *P. fluorescens*.

- ¿Con qué criterio se hizo el apilamiento y en qué suposiciones se basa?
- Explicá dónde están los extremos 5' y 3' de cada cadena de ADN mostrada en la Fig. 2.1, y qué significan los asteriscos y los guiones.
- ¿Por qué en las secuencias de la Fig. 2.1 hay zonas en las que se parecen más entre sí, y por qué ciertas posiciones son las mismas en todas las secuencias?
- ¿Se puede suponer que este gen está en el mismo lugar del cromosoma en las diferentes especies?

<i>S. typhimurium</i>	----AGAGCT--GTTATCGCAAGGGGCCACACA-----GCGTGTGTTGGAC
<i>S. flexneri</i>	----GGAAct--GTTATCGCAGGGAGCCACACA-----GCTTGTGTTGGAC
<i>E. carotovora</i>	----CGAGCT--GTTATCGCAAGGGTCCCACA-----GCGTGTCTTGAC
<i>Y. enterocolitica</i>	----AGAGCT--GTTGTCTCAAGCGTTACCCA-----ACGTGTATTGGAT
<i>S. entomophila</i>	----TGAGCT--GTTGGCGCAAGGTGTTAC-CA-----GCG-GTGCT-GAT
<i>P. aeruginosa</i>	TCCAAAAGCTCCGCTTCATTAAGCAACACAAATACATTGATTACACGCGTGCCTCGAT
<i>P. fluorescens</i>	GCCACCATTCTCTCTTCCAAACAACAAAGCACATCGACTACACGCGCGCTTGGAC
	* * * * *
<i>S. typhimurium</i>	GCGACTCAGCTTTACCTTGGTGAGATTGGGTATTCACTGTTAACAGCCGAAGAAGAA
<i>S. flexneri</i>	GCGACTCAGCTTTACCTTGGTAAGATTGGTTATTCACTGTTAACGCGCGAAGAAGAA
<i>E. carotovora</i>	GCAACACAGCTCTATTTAGGAGAGATCGGCTATTGCGCCTTTTAAACCGCAGAAGAAGAA
<i>Y. enterocolitica</i>	GCGACGCAGCTCTATCTGGGTGAGATCGGTTATTGCGCCTCTGCTTACCGCAGAGGAAGAG
<i>S. entomophila</i>	GCGACACAGCTCTATCTTGGTGAGATTGGTTATTGCGCGTGTGACCGCAGAAGAAGAG
<i>P. aeruginosa</i>	GCGACGCAGCTGTACCTCAATGAAATCGGCTTTTCTCCATTGCTCTCCCCGAAGAGGAA
<i>P. fluorescens</i>	GCAACGCAGCTGTATCTCAACGAAATCGGTTTCTCGCCCTGTTGACGCCGAAGAAGAA
	** ** *
<i>S. typhimurium</i>	GTCTATTTTGC GCGCTCGCGACCTGCGTGAGATGTCGCTTCTCGCCGTCGCATGATTGAG
<i>S. flexneri</i>	GTTTATTTTGC GCGCTCGCGACTGCGTGAGATGTCGCTTCTCGCCGCGGATGATCGAG
<i>E. carotovora</i>	GTCTATTTTGC CCGACGCGCGTGGTGGCGATGTCATCGCGCCGCGGATGATCGAG
<i>Y. enterocolitica</i>	GTCTATTTTGC GCGCGCTGCTCTGCGCGGTGACGTGCCGTCCCGCCGCGCATGATCGAA
<i>S. entomophila</i>	GTTTATTTTGC CCGCGCTGCTTGGTGGCGAGATGTCGCTTACGTCGCGGTATGATCGAA
<i>P. aeruginosa</i>	GTGCACTTTGC GCGGTTGTGCGAAAGTGGTGATCCGCGCGGCGCAAGCGCATGATTGAA
<i>P. fluorescens</i>	GTCCACTTCGCTGCTGCGCGAGAAGGGCGATCCCGCTGGTGGAAAGCGGATGATCGAG
	** *
<i>S. typhimurium</i>	AGTAACCTGCGTCTGGTGGTAAAAATGCCCCGCGTTATGGCAATCGTGGACTGGCGTTG
<i>S. flexneri</i>	AGTAACCTGCGTCTGGTGGTAAAAATGCCCCGCGTTATGGCAATCGTGGTCTCGCGTTG
<i>E. carotovora</i>	AGTAACCTGCGGTTGGTGGTAAAAATGCCCCGCGTTACAACAATCGTGGTCTGGCGCTG
<i>Y. enterocolitica</i>	AGCAACCTGCGGCTGGTGGTAAAAATGCCCCGCGCTATAGCAATCGCGGCTGGCGCTG
<i>S. entomophila</i>	AGTAACCTGCGGTTGGTAGTGAAGATTGCTCGCGCTTACAGTAATCGCGGTTTAGCGCTG
<i>P. aeruginosa</i>	AGCAACCTGCGGCTGGTGGTAAAAATGCCCCGCGCTATGTCAATCGTGGCCTGTGCGCTG
<i>P. fluorescens</i>	AGCAACCTGCGGTTGGTGGTGAAGATCGCCCCGCGCTATGTCAATCGCGGACTGTCCCTG
	** *

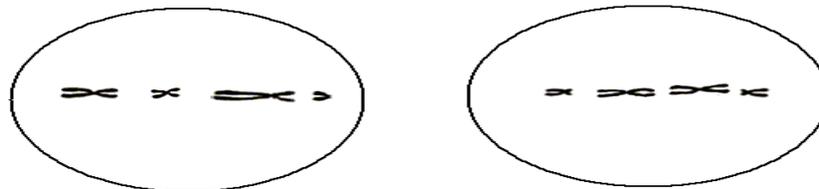
Fig. 2.1 (continúa en la página siguiente)

S. typhimurium	CTGGACCTGATTGAAGAGGGCAACCTGGGGCTTATCCGTGCAGTAGAGAAGTTTGACCCG
S. flexneri	CTGGACCTTATCGAAGAGGGCAACCTGGGGCTGATCCGCGCGGTAGAGAAGTTTCGACCCG
E. carotovora	CTGGACCTGATTGAAGAGGGGAACCTCGGCCTGATCCGTGCGGTTGAAAAATTCGATCCA
Y. enterocolitica	CTGGATCTGATTGAGGAGGGCAACCTCGGCCTGATCCGCGCGGTGAAAAAGTTTGACCCA
S. entomophila	CTGGATTTGATTGAAGAGGGTAACCTCGGTCTTATCCGTGCAGTGAAAAAGTTTGACCCA
P. aeruginosa	CTGGACCTGATCGAAGAGGGCAACCTCGGGCTGATCCGAGCGGTGGAGAAGTTTCGATCCG
P. fluorescens	CTCGACCTGATCGAGGAAGCAACCTAGGCCTGATCCGCGCGGTGGAGAAGTTTCGATCCG
	** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * **
S. typhimurium	GAACGCGGTTCCGCTTCTCAACATACGCAACCTGGTGGATTCCGACAGACAATCGAACGG
S. flexneri	GAACGTGGTTTTCCGCTTCTCAACATACGCAACCTGGTGGATTCCGACAGACGATTGAACGG
E. carotovora	GAAAGAGGATTCCGTTTTTCAACCTACGCAACCTGGTGGATTCCGACAAACGATAGAGCGG
Y. enterocolitica	GAACGCGGTTTTCCGTTTTCTCCACATACGCAACCTGGTGGATCCGACAGACGATCGAACGG
S. entomophila	GAACGCGGTTTTCCGTTTTCTCCACTTATGCCACATGGTGGATACGCCAGACAATTGAACGG
P. aeruginosa	GAACGCGGCTTCCGCTTCTCGACCTATGCCACCTGGTGGATTCCGTCAGACCATCGAGCGG
P. fluorescens	GAGCGGGATTCCGCTTCTCGACTACGCCACCTGGTGGATCCGCCAGACCATCGAGCGG
	** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * **
S. typhimurium	GCGATCATGAACCAAACCCGTACGATTGCGTTGCGGATTACATTGTTAAAGAGCTGAAC
S. flexneri	GCGATTATGAACCAAACCCGTACTATTGTTTTGCCGATTACATATCGTAAAGGAGCTGAAC
E. carotovora	GCGATCATGAATCAAACCCGTACCATTCCGTTTGCCGATTACATTGTCAAAGAACTCAAC
Y. enterocolitica	GCGATTATGAACCAAACCCGTACCATTGTTTTGCCATCCACATCGTCAAAGAACTGAAC
S. entomophila	GCAATAATGAACCAAACCCGTACCATTCCGCTCTGCCTATCCATATCGTAAAGAATTAAAC
P. aeruginosa	GCCATCATGAATCAGACCCGGACCATTCCGTTGCCGATCCATGTGGTCAAAGAACTCAAC
P. fluorescens	GCCATCATGAACCAGACCCGGACCATTGCGTTGCCGATCCATGTGGTCAAGGAGCTCAAC
	** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * **
S. typhimurium	GTATACCTGCGCACCGCACGTGAGTTGTTCGCATAAACTGGACCACGAACCGAGTGC GGAA
S. flexneri	GTTTACTTACGAACCGTACGTGAGTTGTCCATAAAGCTGGACTATGAACCAAGTGC GGTA
E. carotovora	GTTTATCTGCGCACCGCGCGGAACTGTCTCACAACCTGGATCACGAGCCGAGTGC GGAA
Y. enterocolitica	GTCTATCTGCGCACCGCGCGGAGCTTTCTCACAACCTGGATCATGAACCGAGCGCTGAA
S. entomophila	GTTTATTTGCGTACAGCGCGGAACTTTCTCATAAATTAGATCATGAACCGAGCGCAGAA
P. aeruginosa	GTCTACTTGC GCGCCCGGTGAGTTGACGCAAAACTCGATCATGAACCTTCCCCGAA
P. fluorescens	GTCTACCTGCGTGC GCGCGGAACTGACCCACAAGCTCGACCACGAACCTTCCCCGAA
	** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * **

Fig. 2.1 (viene de la página anterior) Apilamiento de secuencias de *rpoS* en los microorganismos indicados.

### Problema 2.3

¿Cuál de las células representadas a continuación es haploide y cuál es diploide? ¿Cuáles son sus cariotipos?



### Problema 2.4

El caballo tiene 64 cromosomas como número diploide.

a) ¿Cuántos **autosomas** habrá en un gameto?

b) ¿Cuántos **cromosomas sexuales** habrá en un gameto?

c) ¿Cuántos **cromosomas** recibirá un potrillo de cada progenitor?

d) ¿Cuántos cromosomas esperarías encontrar en los híbridos macho y hembra (mulas) resultantes del cruzamiento entre un caballo y un asno ( $2n=62$ )? Indica si los animales híbridos son fértiles o estériles.

### Problema 2.5

El esquema de la Fig. 2.2 representa una región del genoma de una bacteria donde se encuentran codificados cuatro genes, con funciones relacionadas entre sí y que por lo tanto se llaman *hid*.

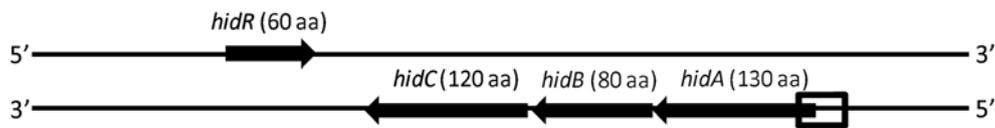


Fig. 2.2 Esquema de la ubicación de los genes *hid* en una región del genoma de una bacteria.

- ¿Qué representan las flechas y el rectángulo negro?
- ¿Qué significan los números seguidos de "aa"?
- ¿Por qué las flechas apuntan en sentidos opuestos?
- ¿Cuántos pares de bases (pb) tiene cada gen?

### Problema 2.6

En la Fig. 2.3 se representa la secuencia de ADN del genoma completo de la bacteria *Bradyrhizobium japonicum*, que realiza la simbiosis fijadora de  $N_2$  con plantas de soja. Las líneas de colores representan genes individuales, y el círculo celeste indica el porcentaje de (G+C) en cada región del genoma.

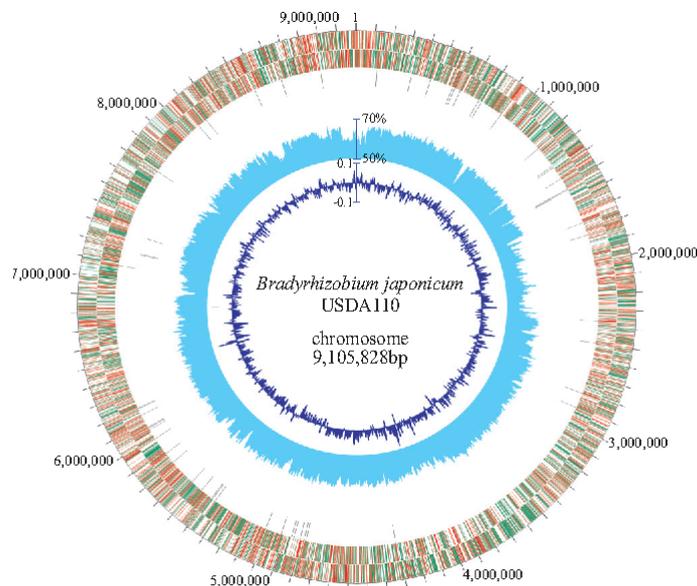


Fig. 2.3 Representación del genoma completo de *Bradyrhizobium japonicum*. El círculo externo representa la ubicación de los genes, y el círculo celeste más interno indica los porcentajes de (G+C) en cada región del genoma. La altura de cada línea celeste representa un porcentaje, según la escala indicada en la parte central, que va de 50% a 70%.

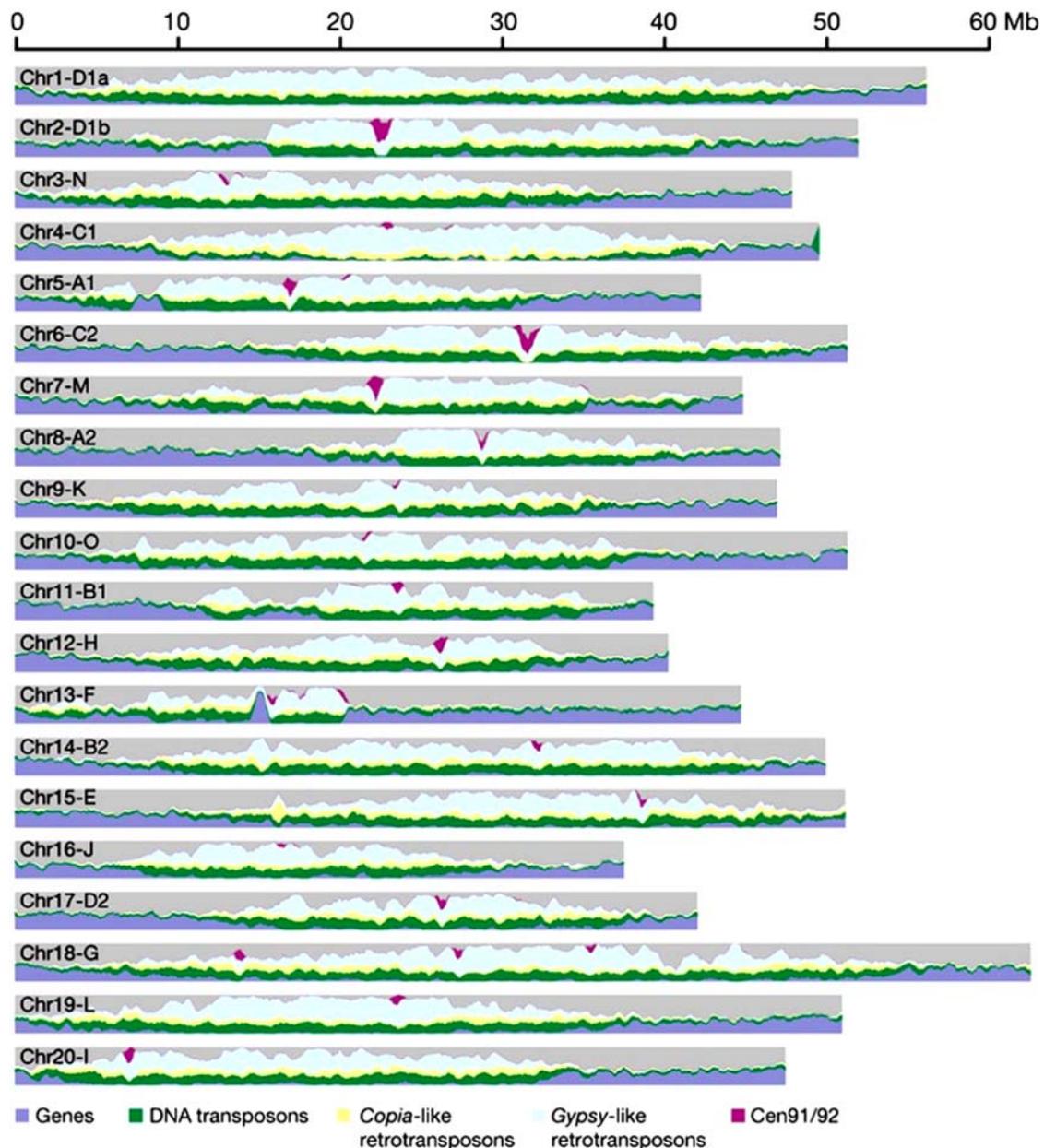
- ¿Qué relación tienen las líneas de color con las flechas del problema 2.5?
- Mirá atentamente la altura de las barras celestes del círculo interior. Fíjate que a la derecha son sensiblemente más bajas que en el resto del círculo. ¿Qué indica esto? ¿A qué puede deberse?

## Problema 2.7

En la Fig. 2.4 se representa la secuencia de ADN del genoma de la soja. Las zonas en gris oscuro indican los lugares de los cromosomas donde probablemente se encuentran los genes, mientras que las zonas en verde indican ADN proveniente de la inserción de elementos móviles como virus, transposones, etc. Finalmente, las zonas en violeta indican las secuencias de lo que probablemente son los centrómeros en cada cromosoma.

a) Compará este genoma con el de *B. japonicum* mostrado en la Fig. 2.3 y observá la densidad de ADN ocupado por genes en uno y otro caso. ¿Qué conclusiones podés sacar de la comparación de la distribución de genes en procariotas y eucariotas?

b) ¿Por qué en un genoma eucariota no puede especificarse inequívocamente dónde están los genes, y solamente puede indicarse una probabilidad, mientras que en los genomas procariotas como el de la Fig. 2.3 sí puede hacerse?



**Fig. 2.3** Representación del genoma completo de la soja (*Glycine max*). En la escala superior se indica el tamaño del ADN en megabases (Mb); 1 Mb = 1.000.000 bases. A la izquierda se indica el número de cromosoma (Chr). El código de color de la parte inferior de la figura indica las posiciones de los genes, elementos de inserción y centrómeros. Dado que en eucariotas no puede especificarse unívocamente la secuencia de los genes, los mismos se predicen con una probabilidad, que está dada por la altura de la línea serrada horizontal.

### Problema 2.8

El número cromosómico de la arveja (*Pisum sativum*) es  $2n = 14$ . Esquematizá cuál es el número cromosómico que puede esperarse en los siguientes individuos:

- a) Monosómicos
- b) Trisómicos
- c) Tetrasómicos
- d) Trisómicos dobles
- e) nulisómicos
- f) monoploides
- g) triploides
- h) autotetraploide
- i) alotetraploide

### Problema 2.9

Se sabe que distintas variedades de crisantemos tienen 18, 36, 54, 72 y 90 cromosomas; todos son múltiplos de una dotación básica de 9 cromosomas

- a) ¿Cómo describirías genéticamente a esas variedades?
- b) ¿Qué rasgo comparten los cariotipos de estas variedades?
- c) Se descubrió una variedad con 27 cromosomas, ¿será fértil o estéril? ¿Por qué?

### Problema 2.10

El origen del trigo ha sido profundamente estudiado. Se sabe que para dar origen al actual trigo pan (*Triticum aestivum*) han participado tres especies, dos de éstas diploides (las tres especies son  $x=7$ ). Una, cuyo genomio es A, se cruzó con otra cuyo genomio es B, dando origen a *Triticum tauschii*, (AABB), el trigo para fideos o trigo duro. Cuando ese híbrido natural se cruzó con la especie *Aegilops squarrosa* (genomio DD) aportó la calidad panadera y dio origen al alohexaploide que ha llegado hasta nuestros días.

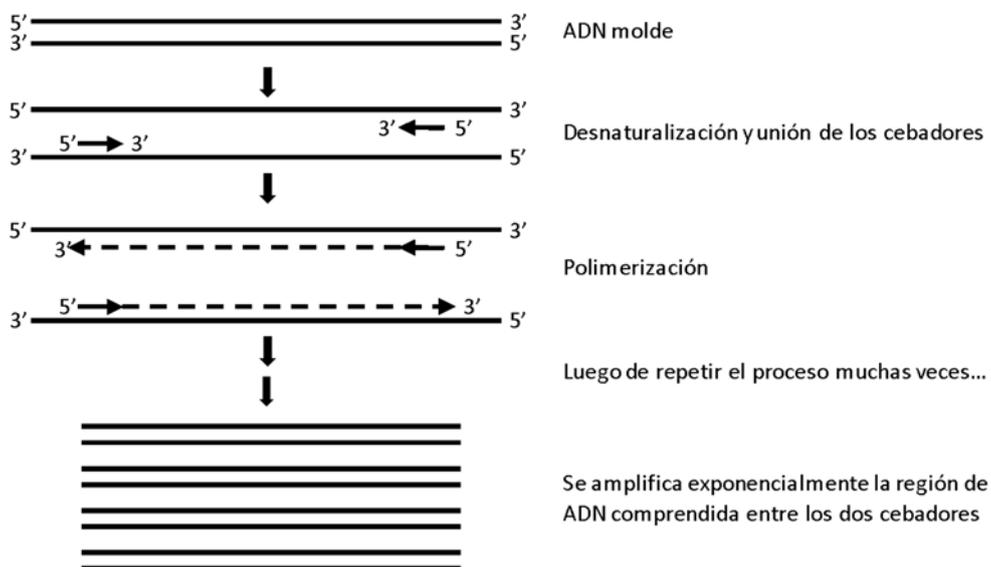
- a) Graficá los diferentes pasos que ha tenido el trigo hasta su situación actual, colocando en cada caso el valor  $2n$ ,  $n$ ,  $x$  (nivel de ploidía).
- b) Planteá los pasos para obtener el híbrido interespecífico *Triticale* resultante del cruzamiento entre *Triticum aestivum* ( $6X=42$ ) y *Secale cereale* ( $2n=28$ ), ambos con número básico = 7. Indicá en cada paso:  $2n$ ,  $n$ , y  $x$ .

### 3. Replicación del ADN *in vitro*.

La reacción en cadena de la polimerasa (PCR) es una reacción de replicación del ADN llevada a cabo *in vitro*, es decir en un sistema libre de células, en el cual se añaden solamente los componentes necesarios para que ocurra la reacción. En este caso, basta con utilizar el ADN molde, los cuatro desoxirribonucleótidos trifosfato (abreviados dNTP), la ADN polimerasa, dos cebadores (*primers*) de simple cadena, y un buffer apropiado que contenga  $Mg^{2+}$  (necesario para la actividad de las enzimas que usan nucleótidos trifosfato como sustratos).

La reacción se inicia con la desnaturalización del ADN molde por calor. Una vez que ambas cadenas están separadas, se disminuye la temperatura de forma tal de permitir que los cebadores de simple cadena se unan a sus secuencias complementarias en cada una de las dos cadenas de ADN molde. Para que esta unión sea específica de una sola secuencia de ADN, es necesario que los cebadores posean una longitud de al menos unos 20 nucleótidos, de forma de asegurar que esa secuencia no se va a encontrar repetida en otros lugares del genoma (a modo de ejemplo, una determinada secuencia de 20 bases podría encontrarse al azar cada  $4^{20}$  bases en el genoma). Una vez que los cebadores se unieron al ADN molde, se lleva la mezcla de reacción a la temperatura adecuada para que funcione la ADN polimerasa, la cual, a partir del extremo 3'-OH de cada cebador va a catalizar la polimerización de la cadena complementaria al ADN molde usando los dNTP que se encuentran en la mezcla de reacción.

Luego de varias repeticiones de este proceso (ciclos) puede obtenerse una amplificación exponencial de la región de ADN acotada entre los dos cebadores, como se muestra a continuación:



La enzima que se utiliza en la PCR es la Taq polimerasa, llamada así porque proviene de *Thermus aquaticus*, un microorganismo que habita aguas termales a muy alta temperatura. Una de las adaptaciones de este microorganismo a su ambiente consiste en que su ADN polimerasa (la Taq polimerasa) es resistente a temperaturas de 95 °C, y por lo tanto, en la PCR no es necesario reemplazarla luego de cada ciclo de calentamiento. Esto permite automatizar el proceso, que se lleva a cabo con un equipo denominado termociclador.

Una vez finalizada la PCR, los productos pueden separarse en una electroforesis en gel de agarosa. El principio de esta técnica consiste en hacer migrar las moléculas de ADN en un campo eléctrico hacia el polo positivo, dado que dichas moléculas tienen carga negativa a pH = 8, provista por los fosfatos que se encuentran en todos los enlaces fosfodiéster de las cadenas. Para separar las moléculas de acuerdo con su tamaño se introduce en la cuba electroforética un gel, normalmente de agarosa, cuya función consiste en generar una malla a través de la cual deben migrar las moléculas. Según lo ajustada que sea esta malla, las moléculas se verán tanto más impedidas de moverse cuanto mayor sea su tamaño, y eso podrá visualizarse como una serie de bandas cuyos desplazamientos (en cm) desde el punto de siembra serán directamente proporcionales al logaritmo de su peso molecular.

En YouTube hay disponibles varias animaciones que ilustran estos procesos. Por ejemplo:

<https://www.youtube.com/watch?v=vpAyYYArBOY> para la PCR

<https://www.youtube.com/watch?v=6vKLT5mQoBM> para la electroforesis en gel de agarosa.

### Problema 3.1

Una solución contiene fragmentos de ADN de doble cadena de 1000, 500, 250, 100 y 50 kb de tamaño. El presente diagrama representa un gel de agarosa. Ubicá en la primer celda o calle las bandas según el tamaño correspondiente. Indicá el tamaño de la banda de la segunda calle y explicá por qué esta banda es más ancha que las anteriores.



### Problema 3.2

Se quiere amplificar una secuencia de 144 pares de bases (bp) que solamente se encuentra en una cepa de levadura:

```
GGATCCTGCGGACGCAGCTTGTCTGTTCTCGTTACGATGCGCGTCTGCGTTAGCAGAGCATGGTGGGTGTGATCAGCGTGAATGGC  
ACCGACAAAGCCGTGGCGAGTCATACGAGGCTGCCCTGCGTTCTGCGACCCGCTCGCTT
```

a) ¿Qué significa esta sucesión de letras? ¿Representa ADN? ¿Una cadena? ¿Dos? ¿Dónde están los extremos 5' y 3'?

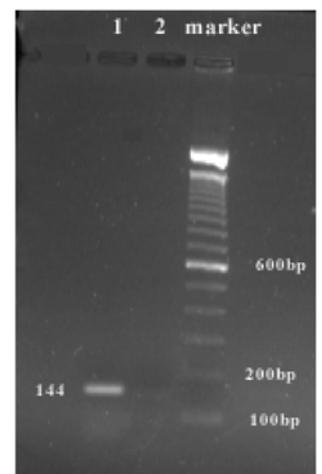
b) Escribí las secuencias de los cebadores empleados para amplificar el fragmento de 144 bp por PCR. Las secuencias **deben estar escritas en sentido 5'→3' con el extremo 5' a la izquierda**. Los cebadores deben ser de 20 nucleótidos cada uno.

c) Para llevar a cabo la reacción, además del molde, se añaden los cebadores o *primers* adecuados, una ADN polimerasa termo-resistente (Taq polimerasa) y el resto de los componentes necesarios para la reacción, de manera de completar un volumen final de 25 µl. Se emplean las siguientes condiciones de ciclado:

- 1) 92 °C durante 15 segundos
- 2) 56 °C durante 15 segundos
- 3) 72 °C durante 30 segundos

Escribí al lado de cada uno de los pasos anteriores, qué proceso ocurre. Comparalos con lo que ocurre en la replicación en las células. ¿Cuál es la fuente de energía para la separación de las hebras de ADN en uno y otro caso?

c) En la Fig. 3.1 se muestra el resultado de la electroforesis del producto de PCR obtenido con los cebadores que propusiste en b) y realizando la reacción como dijiste en c). En la calle 1 se utilizó como molde un extracto de ADN de la levadura, en la calle 2 se utilizó como molde agua destilada y en la calle "marker" se sembraron los marcadores de peso molecular. Explicá para qué sirve cada calle y si ellas nos permiten asegurar que la reacción de PCR fue exitosa.



**Fig. 3.1** Electroforesis en gel de agarosa del producto de la PCR realizada para amplificar un fragmento de 144 bp del ADN de una levadura. Calle 1: molde de ADN de la levadura, calle2: sin molde, marker: marcadores de tamaño molecular

### Problema 3.3

A continuación, se muestra la secuencia del gen que codifica una  $\beta$ -(1,6) glicosidasa en un hongo fitopatógeno. Las letras minúsculas indican zonas del ADN que no codifican la enzima, mientras que las letras mayúsculas indican la zona donde se codifica la enzima:

atcggtgtccaatcATGCCTGATTGTGGTCATCTTCATCTGTGGTCATTCATGTAACAATCAATGTAAGGTCATTGCTGCAGCATGCTG  
GCAGAATTGTCAGTCAAATGCAACACCAATGAGTCGTACAATGACAGGTCAATGAACGTCTGAATCTGTaagtcaaggtc.

- a) Escribí la secuencia de un par de cebadores de 20 bases cada uno, para amplificar por PCR la zona indicada en letras mayúsculas. Escribí las secuencias en el sentido 5'→3'.
- b) ¿Tiene que ser solamente esta región? ¿Entera tal cual, o se permite que peguen los cebadores más afuera? ¿Buscarías que terminen en CG?
- c) Explicá, sobre la base de la reacción de polimerización del ADN, por qué los cebadores deben orientarse en el sentido 5'→3'.

## 4. Marcadores moleculares.

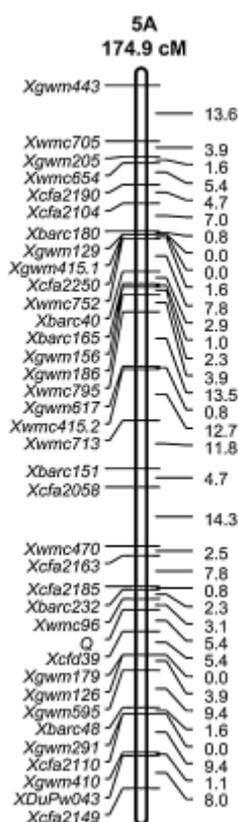
Los marcadores en general son aquellas señales, rastros o “marcas” que pueden ser de diferente naturaleza y que sirven para indicarnos la presencia cercana de un gen de interés en el genoma, identificar cepas, variedades, personas, etc. y establecer relaciones entre ellas. Existen diferentes tipos de marcadores: morfológicos, bioquímicos (isoenzimas y proteínas) y marcadores moleculares o genéticos (basados en ADN). Existen diversas técnicas de biología molecular disponibles para detectar variabilidad en la secuencia de ADN. La reacción en cadena de la polimerasa (PCR) es la más utilizada actualmente para poner de manifiesto a los marcadores moleculares.

Los marcadores moleculares que pueden ser detectados por PCR consisten en repeticiones cortas de secuencias, que de acuerdo al número de nucleótidos que tiene la secuencia que se repite, se califican como microsatélites o minisatélites. Sin embargo, el concepto de “satélite” no tiene que ver con estas secuencias. Simplemente se trata de repeticiones de secuencia que se encuentran en prácticamente todos los organismos vivos. Lo importante de estas secuencias repetidas es que, en una dada especie, se encuentran siempre en el mismo *locus*, y son polimórficas. Este último término hace referencia a que existen series alélicas muy grandes para estas secuencias.

En efecto, podemos considerar que un número dado de repeticiones de una cierta secuencia es un alelo, aun cuando no codifique ninguna proteína ni función biológica que pueda visualizarse como un carácter. Pero estos alelos en particular pueden identificarse y distinguirse si se los amplifica por PCR y se los separa por electroforesis en gel. El punto importante aquí es que los cebadores que se utilicen en estas PCR deben unirse al ADN molde **por fuera** de la zona repetida (el micro o el minisatélite) de forma tal de poderse amplificar el alelo entero.

Una vez que se los identifica, estos micro o minisatélites pueden ser mapeados en el genoma y también puede determinarse su ligamiento a genes de interés, utilizando para ello las mismas técnicas de cruzamientos que se verán en los TP 6 y 7.

A modo de ejemplo, en la Fig. 4.1 se muestra la ubicación de algunos marcadores moleculares en el cromosoma 5A del trigo. A la izquierda se indica la ubicación de los marcadores y a la derecha, la distancia (cM) entre dos marcadores contiguos.



**Fig. 4.1** Mapa de marcadores moleculares en el cromosoma 5A del trigo. A la izquierda se indica la ubicación de los marcadores y a la derecha, la distancia (cM) entre dos marcadores contiguos.

#### Problema 4.1

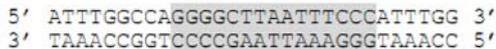
En el siguiente diagrama se muestra el polimorfismo de un marcador de tipo microsatélites (SSR) constituido por 3 alelos: A (GT)<sub>10</sub>, A1 (GT)<sub>8</sub> y A2 (GT)<sub>6</sub>. Mediante la técnica de PCR y utilizando cebadores diseñados en las regiones flanqueantes, los alelos pueden ser amplificados para luego ser separados mediante una corrida electroforética. Representará mediante un esquema el patrón de bandas que podría obtenerse en individuos diploides que lleven el mismo alelo en ambos cromosomas, o uno en un cromosoma y otro en el otro.



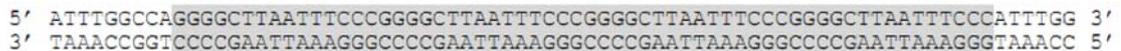
#### Problema 4.2

Se desea amplificar por PCR una secuencia repetida en tándem de número variable (VNTR) que permite identificar individuos según el número de repeticiones que posean en su ADN. Para el caso que nos ocupa, se trata de un VNTR que posee dos alelos: en uno de ellos hay una sola copia del VNTR, mientras que en el otro hay cuatro:

Una copia del VNTR:



Cuatro copias del VNTR:



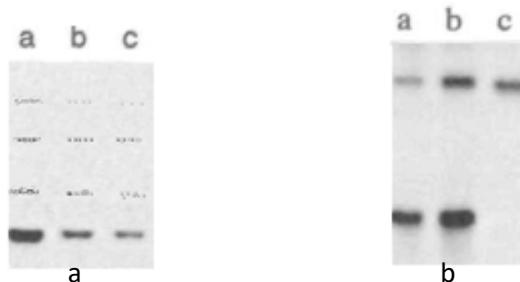
Se sabe que el individuo "a" es heterocigota para este VNTR, mientras que el individuo "c" es homocigota y posee sólo el alelo con cuatro copias del VNTR. El objetivo de esta PCR es identificar la composición alélica de un tercer individuo, al que llamaremos "b". Para ello, se utilizan en la PCR los siguientes cebadores (*primers*):

Forward: 5' pGGGCTT-OH 3'

Reverse: 5' pGGGAAA-OH 3'

**Nota:** Se representan cebadores de tamaño más corto que lo que deberían sólo para facilitar la lectura. **La solución al problema no tiene que ver con el tamaño de los cebadores.**

La reacción se realiza de la siguiente manera: Se obtiene el ADN molde de los tres individuos, y en paralelo se prepara la mezcla de reactivos. Luego, se divide la mezcla en tres partes iguales, y a cada parte se le agrega ADN molde del individuo "a", del "b" o del "c". La reacción se inicia incubando a 95 °C durante 2 minutos y después se realizan 30 ciclos de la siguiente manera: 95 °C 15 segundos, 53 °C 15 segundos y 72 °C 30 segundos, culminando con un ciclo a 72 °C 2 minutos. Luego, las muestras se corren en un gel de agarosa para observar los productos de amplificación, y como resultado, se observan las bandas que se muestra en la Fig. 4.1a. Sin embargo, este resultado no era el esperado, con lo cual los investigadores deciden cambiar de cebadores. Diseñan otros cebadores distintos, y utilizando el mismo protocolo para la PCR obtienen el resultado mostrado en la Fig. 4.1b, el cual finalmente les permite identificar el genotipo del individuo "b". Por lo tanto, elevan su informe, el cual es inmediatamente rechazado debido a la falta de un control.



**Fig. 4.1** Resultados de la PCR para cuantificar las copias de VNTR, utilizando los cebadores mostrados en la página anterior (a) o bien otros cebadores mejorados (b).

- a) ¿Qué utilidad tiene preparar la mezcla de reactivos de PCR antes de agregar el molde?
- b) ¿Cómo se dieron cuenta los investigadores de que el resultado de la Fig. 4.1a no correspondía a lo esperado?
- c) Explicá por qué los cebadores utilizados inicialmente no permitieron distinguir los tres individuos.
- d) Escribí la secuencia de los nuevos cebadores que permitieron obtener el resultado de la Fig. 4.1b.
- e) Indicá cuál es el control que faltaba y por qué es crítico para que el resultado de la PCR sea válido.

### Problema 4.3

Una de las producciones sustentables más difundidas en el Delta del Paraná es el cultivo forestal de Salicáceas: álamos (*Populus* spp.) y sauces (*Salix* spp.). Esta actividad provee materia prima para cuatro cadenas industriales: celulosa y papel, tableros de partículas, de bobinados y aserrío. Mediante el uso de marcadores moleculares microsatélites (del tipo repetición de secuencia única, abreviado SSR), se logró diferenciar a clones y líneas avanzadas del programa de mejoramiento de Salicáceas.

La Fig. 4.2 representa un gel de agarosa donde se observa la variación alélica de dos SSR (llamados PMGC 2217 y PMGC 2030). Las calles de 1-7 corresponden a clones de álamo y las calles 8-12 clones de sauce. M, corresponde al marcador de peso molecular.

- a) ¿Cuál habrá sido la secuencia de trabajo que los investigadores utilizaron en este estudio? Ordená del 1 al 5 los diferentes pasos del procedimiento:
  - Amplificación de los microsatélites mediante el método de reacción en cadena de la polimerasa (PCR).
  - Extracción de ADN de hojas de álamo y sauce.
  - Diseño de cebadores específicos para los marcadores microsatélites (búsqueda en bibliografía de la temática o en bases de datos en internet: ej NCBI).
  - Análisis de los geles, obtención de resultados y agrupamiento de clones.
  - Separación de los productos amplificados en geles de agarosa y visualización mediante tinción del ADN.



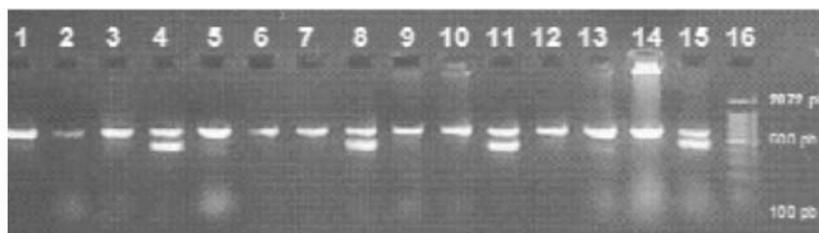
**Fig. 4.2** Electroforesis en gel de agarosa de los productos de PCR obtenidos con los cebadores de la Tabla 1. Calles 1-7: muestras de álamo. Calles 8-12: muestras de sauce. M: marcadores de peso molecular.

b) ¿Cuántos alelos podés identificar para cada microsatélite PMGC 2217 y PMGC 2030? ¿Considerás que fue posible diferenciar a todos los cultivares con el empleo de estos marcadores?

c) ¿Cuál de los dos marcadores es más polimórfico? ¿Podrías identificar cultivares homocigotas y heterocigotas para los marcadores? Da un ejemplo de cada caso.

#### Problema 4.4

Un productor avícola envía muestras de sangre de sus animales a un laboratorio para la determinación del sexo. El análisis se basa en la amplificación de un minisatélite llamado CHD que está ligado a los cromosomas sexuales (Z y W). En esta población de aves el minisatélite CHD posee dos alelos: CHDW con 10 repeticiones y CHDZ con 14 repeticiones. Teniendo en cuenta que en las aves los machos son homogaméticos (ZZ) y las hembras heterogaméticas (WZ) (es decir, tendrán el gen CHDZ/CHDZ y CHDW-CHDZ, respectivamente): indicá en la Fig. 4.3 qué muestras de las calles 1 a 16 podrían provenir de machos y cuáles de hembras.



**Fig. 4.3** Productos de PCR del minisatélite CHD de distintos individuos de una población de aves. Algunos de los individuos son machos y otras hembras, pero no se ha especificado el sexo de cada uno en las calles 1-16.

## 5. Expresión génica.

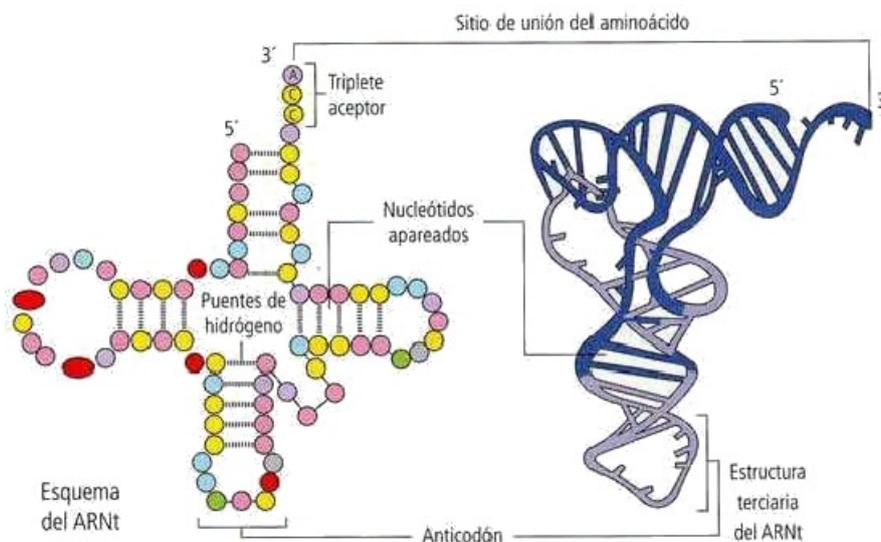
Anteriormente dijimos que los genes tienen “instrucciones” que determinan la función biológica, y que esas “instrucciones” se “expresan” en proteínas funcionales. Ahora podemos tratar de entender mejor estos procesos. Antes que nada, debemos recordar que una secuencia de ADN se transcribe en ARN y ésta en proteína, lo cual ha sido dado en llamar el “dogma central de la Biología Molecular”, aunque ahora sabemos que varios ARN no se traducen a proteína, y que la información de ciertos ARN puede también transcribirse en ADN (aquí se habla de *retrotranscripción*). Pero para simplificar, centrémonos en los procesos de transcripción y traducción.

Así como el ADN es una secuencia de desoxirribonucleótidos, el ARN es una secuencia de ribonucleótidos formada a partir de ribonucleótidos trifosfato (NTP). Si bien ambas moléculas son parecidas, cumplen funciones muy distintas, y en la célula el ARN no se replica, es decir que una cadena de ARN no puede servir de molde para producir otra. Esto es una suerte, porque de lo contrario, nuestro ARN aumentaría exponencialmente, nuestras células rebosarían de ARN, y nosotros explotaríamos chorreando ARN por todos lados.

Por su parte, las proteínas son secuencias de aminoácidos, y aquí sí hay algunas diferencias importantes. Una de ellas es que, mientras los ácidos nucleicos están formados por solo cuatro bases, las proteínas están formadas por 20 aminoácidos distintos. En segundo lugar, mientras que en los ácidos nucleicos solo podemos distinguir entre purinas y pirimidinas, en las proteínas podemos distinguir entre aminoácidos polares sin carga, polares con carga negativa, polares con carga positiva, no-polares, alifáticos, aromáticos, con grupos sulfhidrilo, y algunas combinaciones de todos ellos. Esto hace que la paleta de “letras” diferentes sea mucho mayor, y, por lo tanto, la diversidad de funciones también lo sea.

Tanto los ácidos nucleicos como las proteínas en solución pueden adquirir una serie de conformaciones tridimensionales, llamadas *estructura secundaria* y *terciaria*. Por ejemplo, la doble hélice es una estructura secundaria, que como vimos, se estabiliza por los puentes de hidrógeno entre las bases (Fig. 1.2). Si un ácido nucleico monocatenario (como el ARNm) se deja en solución, también va a formar puentes de hidrógeno, plegándose de tal manera que pueda formar las estructuras de doble hélice intracatenarias que maximicen la formación de puentes de hidrógeno dentro de la misma molécula. Un ejemplo conocido es el ARN de transferencia, que adopta la forma de una hoja de trébol (Fig. 5.1).

Las proteínas también tienen estructura secundaria y terciaria, como lo ejemplifica la ADN polimerasa de la Fig. 1.3. Los “resortes” que se ven allí son estructuras secundarias llamadas  $\alpha$ -hélice, y la combinación de  $\alpha$ -hélices da por resultado la estructura terciaria. La formación de un sitio activo como el que vimos entre los “dedos” y la “palma” en la Fig. 1.3 depende de que se forme correctamente la estructura terciaria.



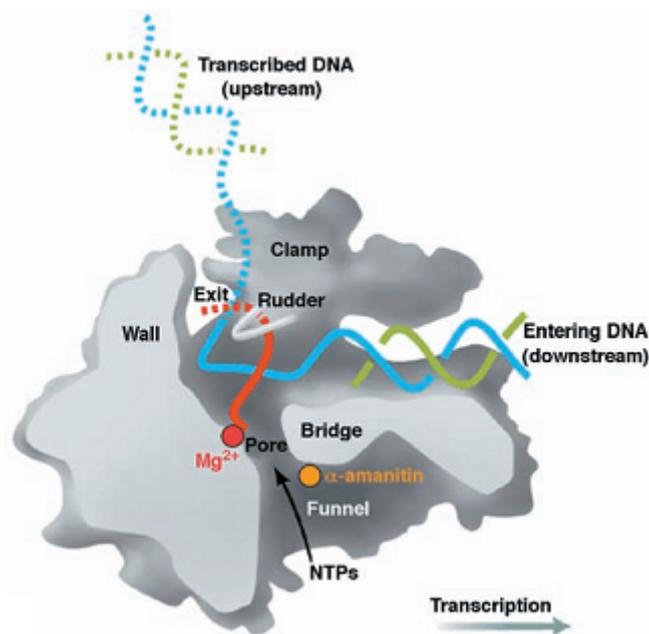
**Fig. 5.1** ARN de transferencia (ARNt) donde pueden apreciarse la estructura secundaria (izq.) la estructura terciaria (der.), los puentes de hidrógeno y la forma en que se pliega la molécula para maximizar la formación de puentes de hidrógeno. En las zonas donde hay doble hélice, es porque la secuencia de ribonucleótidos de una región de la cadena es complementaria con la de otra región (distante) de la cadena, con la cual se aparea. Por ejemplo, entre los extremos 5' y 3' hay una zona de apareamiento donde los nucleótidos rosas (G) son complementarios con los nucleótidos amarillos (C), quedando aparte algunos nucleótidos celestes (T) que no pueden aparearse.

La estructura terciaria se forma espontáneamente en solución, y un concepto de extrema importancia en la Biología Molecular es que la secuencia de aminoácidos de la proteína contiene la información necesaria para que se forme la estructura terciaria. Esto lo descubrió un bioquímico de origen noruego llamado Christian Anfinsen en 1961. Él tomó una solución de ribonucleasa activa (la enzima que cataliza la hidrólisis del ARN) y la sometió a desnaturalización con urea. Cuando la enzima estuvo desnaturalizada, es decir que había perdido toda su estructura terciaria, retiró paulatinamente la urea y al final del proceso, observó que la proteína había recuperado la estructura y la función. Desde entonces, existe la certeza de que conociendo la secuencia de aminoácidos puede predecirse la estructura y la función de cualquier proteína, y existen los métodos de cálculo para hacerlo. Sin embargo, el problema es tan complejo que ni con las computadoras más avanzadas ha podido ser resuelto hasta la actualidad.

El proceso de transcripción se inicia en el sitio donde en el Cap. 3 dijimos que “empieza el gen”. Ahora, con más propiedad, podemos decir que esta secuencia es una parte del *promotor*, que es la zona donde se une la ARN polimerasa para sintetizar ARN usando ADN como molde. Si recordás el Cap. 1, allí dijimos que la primasa es una enzima del tipo de la ARN polimerasa, y que no necesita cebador para iniciar la polimerización de la cadena de ARN. En vez de eso, cierta zona de la proteína reconoce la secuencia promotora en el ADN, se une allí, abre las dos cadenas y comienza a polimerizar una cadena de ARN “leyendo” la secuencia de nucleótidos de la cadena molde. Para hacer esto, el ADN pasa por un canal dentro de la ARN polimerasa, que además tiene un “embudo” (*funnel*) por donde entran los NTP que van a ser incorporados a la molécula. En ese sitio se confirma que el apareamiento de bases por puentes de hidrógeno es el correcto, y una especie de trinquete formado por el “puente” (*bridge*) impide que el complejo retroceda. Finalmente, una especie de “timón” (*rudder*) impulsa al ARN para salir de la ARN polimerasa por el otro extremo del canal (Fig. 5.2).

Una vez transcrito, el ARN que se va a traducir a proteínas se asocia a un complejo formado por los ribosomas y los ARNt. Los ribosomas catalizan la formación de los enlaces peptídicos entre aminoácidos, y los ARNt son los encargados de leer el código genético sobre el ARNm.

El código genético (Fig. 5.3) es la regla que permite vincular una “palabra” en “lenguaje ARN” con un símbolo en “lenguaje proteína”. Esto es precisamente traducir: pasar de un lenguaje a otro, para lo cual se necesita una regla de equivalencia. Esto es, un código. Un ejemplo que puede ser conocido es el código binario, como el que usan las computadoras. Un código binario está compuesto por ceros y unos, y solo con eso podemos representar todas las letras del abecedario. Sin embargo, hay que observar que cada letra deberá estar representada por más de un número. Por ejemplo, si representamos a la A con el 0 y a la B con el 1, nos quedamos sin nada para representar a la C, la D, la E, etc. Por lo tanto, deberíamos elegir *secuencias* de números para representar una sola letra.



**Fig. 5.2** Estructura de la ARN polimerasa catalizando la incorporación de un NTP a una cadena de ARN creciente (roja) mientras lee la secuencia del ADN molde (celeste). Las dos cadenas de ADN han sido separadas y la cadena codificante (es decir, la que tiene la misma secuencia que va a aparecer en el ARN) pasa por detrás de la enzima (verde). Los NTP entran por un embudo (*funnel*) y se incorporan al ARN mediante un mecanismo de reacción similar al descrito para la polimerización del ADN (Fig. 1.1). El retroceso del complejo es impedido por el puente (*bridge*) que hace de trinquete, y la salida de la molécula de ARN es impulsada por el timón (*rudder*). Todas estas sub-estructuras forman parte de la estructura terciaria de la ARN polimerasa y ejemplifican la flexibilidad de las proteínas.

La longitud de estas *palabras-código* de números para representar letras no puede ser de dos, es decir si representamos a la A con 00, a la B con 01, a la C con 10, etc., pronto también nos quedamos cortos de palabras-código. En realidad, necesitaríamos unos seis números por palabra-código para representar a las 28 letras (hacé la cuenta y explicá por qué). En el caso del código genético, como el “lenguaje ARN” tiene cuatro “símbolos” en vez de los dos que tiene el lenguaje binario, nos sobran 64 palabras-código de tres símbolos cada una para representar a los 20 aminoácidos. Cada una de estas palabras-código se denomina *codón* o *triplete*. Por lo tanto, el código genético es *universal*, es decir, sirve para todas las especies, y no está en ningún lugar de la célula, porque es una idealización.

Otra particularidad del código genético es que, al haber más palabras-código que aminoácidos, varios de ellos están representados por *codones sinónimos*. Esto implica que cambios en la secuencia de nucleótidos de un ADN pueden no reflejarse en cambios en la secuencia de aminoácidos de la proteína. Es decir, el código genético es estable. Dentro del código genético se encuentra el *codón de iniciación* de la traducción que vimos en la secuencia del Cap. 2, y tres codones de *terminación* de la traducción.

El mecanismo de traducción puede verse en:

<https://www.youtube.com/watch?v=9OZKLABLino>

Una diferencia importante de los procesos de transcripción y traducción entre procariontes y eucariotes es que en los primeros la transcripción y la traducción ocurren simultáneamente en el citoplasma, ya que no tienen núcleo, mientras que en los eucariotes la transcripción ocurre en el núcleo y la traducción en el citoplasma, esto es, son procesos separados en el tiempo y el espacio.

		Segunda base					
		U	C	A	G		
P r i m e r a b a s e	U	Phe UUU	Ser UCU	Tyr UAU	Cys UGU	U	T e r c e r a b a s e
		Phe UUC	Ser UCC	Tyr UAC	Cys UGC	C	
		Leu UUA	Ser UCA	Stop UAA	Stop UGA	A	
		Leu UUG	Ser UCG	Stop UAG	Trp UGG	G	
	C	Leu CUU	Pro CCU	His CAU	Arg CGU	U	
		Leu CUC	Pro CCC	His CAC	Arg CGC	C	
		Leu CUA	Pro CCA	Gln CAA	Arg CGA	A	
		Leu CUG	Pro CCG	Gln CAG	Arg CGG	G	
	A	Ile AUU	Thr ACU	Asn AAU	Ser AGU	U	
		Ile AUC	Thr ACC	Asn AAC	Ser AGC	C	
		Ile AUA	Thr ACA	Lys AAA	Arg AGA	A	
		Met AUG	Thr ACG	Lys AAG	Arg AGG	G	
G	Val GUU	Ala GCU	Asp GAU	Gly GGU	U		
	Val GUC	Ala GCC	Asp GAC	Gly GGC	C		
	Val GUA	Ala GCA	Glu GAA	Gly GGA	A		
	Val GUG	Ala GCG	Glu GAG	Gly GGG	G		

- indica el inicio de la traducción  
- indica la finalización de ésta

**20 aminoácidos:**  
Phe = fenilalanina  
Leu = leucina  
Ile = isoleucina  
Met = metionina  
Val = valina  
Ser = serina  
Pro = prolina  
Thr = treonina  
Ala = alanina  
Tyr = tirosina  
His = histidina  
Gln = glutamina  
Asn = asparagina  
Lys = lisina  
Asp = aspartato  
Glu = glutamato  
Cys = cisteína  
Trp = triptófano  
Arg = arginina  
Gly = glicina

Fig. 5.3 El código genético.

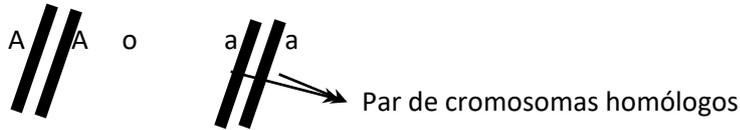
A continuación, daremos algunas definiciones importantes.

**Marco de lectura:** Cada una de las tres maneras en que pueden leerse tripletes una cadena de ADN. Por ejemplo, si se tiene la secuencia: ATACGTACATG..., se puede leer ATA CGT CAC ATG... si se empieza en el primer nucleótido, o bien TAC GTC ACA TG... si se empieza en el segundo nucleótido, o ACG TCA CAT G... si se empieza en el tercer nucleótido. De esta manera, la traducción a aminoácidos empleando el código genético daría Ile-Arg-His-Met... en el primer caso, Tyr-Val-Thr... en el segundo caso, o Thr-Ser-His... en el tercer caso. Es decir, codificaríamos tres proteínas diferentes. Notá que, si se empieza a leer desde el cuarto nucleótido, se repite la primera secuencia de aminoácidos, aunque con uno menos: CGT CAC ATG..., equivalente a Arg-His-Met... Es decir, hay solo tres marcos de lectura en una hebra, o bien seis en total si incluimos la hebra complementaria. No hay más ni menos.

**Marco de lectura abierto:** De entre los tres marcos de lectura posibles, uno se distingue si posee el codón de iniciación, ATG. En el ejemplo anterior, se iniciaría a partir del último triplete, pero siempre y cuando no haya un codón de “stop” o parada en el mismo marco de lectura a menos de 50 codones en la dirección 3’.

**Genotipo:** es la constitución genética del individuo. Puede ser:

**Homocigota:** un organismo diploide es homocigótico para un locus dado cuando en ambos cromosomas homólogos, dicho locus está ocupado por el mismo alelo. Se puede representar a dicho individuo de la siguiente manera:



**Heterocigota:** un organismo diploide es heterocigótico para un locus dado cuando en ambos cromosomas homólogos, dicho locus está ocupado por diferentes alelos. Se puede representar a dicho individuo de la siguiente manera:



**Fenotipo:** es la manifestación externa o aparente del genotipo; se refiere a múltiples apariencias biológicas, entre las que se incluyen características químicas, estructurales y de comportamiento que pueden ser observadas ya sea a simple vista o por otros métodos más complejos. Está influenciado por el ambiente:  $F = G + A$  (Fenotipo = Genotipo + Ambiente) por consiguiente, la expresión fenotípica de un individuo puede cambiar a lo largo de su vida.

**Dominancia completa:** Capacidad de un alelo para expresar su fenotipo a expensas de un alelo alternativo. Es la forma principal de interacción entre alelos (interacción intraalélica). El alelo dominante se expresará siempre que esté presente. En este caso los individuos con genotipo Aa y AA muestran el mismo fenotipo dominante.

**Variaciones de la dominancia:** no siempre el fenotipo de una progenie es igual al de uno de los padres, sino que puede ocurrir que presente un fenotipo diferente al de ambos padres. Puede ser:

**Dominancia incompleta:** Es el caso en el que el heterocigota produce un fenotipo intermedio entre los padres homocigotas. La notación génica es similar a la empleada en dominancia completa.

**Codominancia:** en el heterocigota aparece el efecto de ambos alelos. Cada uno de los alelos codominantes manifiesta independientemente su efecto. Para la notación génica, al no existir dominancia entre un alelo y el otro, se emplea las letras mayúsculas, pero con diferentes subíndices ( $R_1$  y  $R_2$ ). De esta forma no se indica si uno domina sobre el otro como usando mayúsculas y minúsculas.

Como siempre, **leé el libro. Esto no lo reemplaza.**

### Problema 5.1

Las ADN polimerasas son capaces de editar y corregir errores, pero las ARN polimerasas parecen no tener esta capacidad. Considerando que un error en la incorporación de una base puede producir el mismo error en la cadena polipeptídica codificada ya sea si se cometió en el ADN o en el ARN, ¿por qué existe esta diferencia tan grande en la capacidad de corregir errores entre la ADN polimerasa y la ARN polimerasa?

### Problema 5.2

Dada la siguiente secuencia de ADN:

```
5'  ACGTTACATGATATTGCTAATACTCTTTCATTCATTAAACTGGTA  3'
3'  TGCAATGTACTATAACGATTATGAGAAAGTAAGTAATTTGACCAT  5'
```

- ¿Qué productos obtendrías por transcripción?
- ¿Dónde deberían ubicarse los promotores para obtener los productos mencionados (a)?
- identificá la cadena molde y la codificante.
- Señalá todos los marcos de lectura abiertos que encuentres.
- ¿Qué productos obtendrías por traducción de los marcos de lectura abiertos mencionados en (c)?
- ¿Cuántos enlaces ~P de alta energía se hidrolizan en la síntesis de este ARN?

### Problema 5.3

La región codificante de un gen en eucariotas está formada por cuatro exones de 99, 75, 66 y 90 pb y tres intrones, intercalados entre los exones, de 45, 63 y 42 pb, respectivamente. Indicá:

- cuántos pb tendrá el ARNm precursor;
- cuántos nucleótidos tendrá la región que se traducirá del ARNm maduro.
- cuántos aminoácidos tendrá el péptido codificado.

### Problema 5.4

Explicá cuántos aminoácidos pueden ser codificados de la siguiente secuencia de ARNm:

- 5'-UUGCCUAUGGAUUGGAUG-3'
- 3'-AUGUAGGUAAGGUAGGCC-5'

- Suponiendo que comienza la lectura en el extremo 5' inmediatamente.
- Comenzando en la forma normal establecida por el código genético.

### Problema 5.5

¿Cuántas secuencias diferentes de ADN llevan información para la síntesis del polipéptido Met-Gly-Cys-Gly-Ala-Ser?

### Problema 5.6

Un plásmido lineal contiene solo dos genes, los cuales se transcriben en direcciones opuestas, cada uno desde el extremo hacia el centro del plásmido. Dibujá el plásmido y los transcritos, incluyendo:

- El ADN plasmídico, mostrando los extremos 5' y 3' de cada cadena.
- La cadena molde de cada gen.
- Las posiciones de los promotores y de los sitios de iniciación de la transcripción.

- d) Los transcriptos, indicando sus extremos 5' y 3'
- e) Las zonas donde deberían estar codificados los polipéptidos.
- f) Los lugares de terminación de la transcripción.

### Problema 5.7

El siguiente esquema representa una porción de la microscopía electrónica obtenida por Miller, mostrando la simultaneidad de los procesos de transcripción y traducción en *Escherichia coli*:

- a) Identificá qué representan 1, 2, 3 y 4
- b) Identificá los extremos 5' y 3' de la cadena molde del ADN y del ARNm.
- c) Agregá al esquema los polipéptidos en el proceso de síntesis indicando los tamaños relativos y los extremos N y C terminales del péptido más largo.
- d) Indicá con una flecha la dirección de movimiento de la ARN polimerasa
- e) ¿Qué diferencias presentará este diagrama en sistemas eucariotas?

### Problema 5.8

Suponé que estás trabajando con un mutágeno nuevo, y querés determinar cuál base es la que este mutágeno es capaz de alterar. Como dato previo, sabés que al aplicar este mutágeno se produjeron los siguientes cambios en la secuencia de aminoácidos de un dado producto génico:

Original:	Gln-His-Ile-Glu-Lys
Mutante:	Gln-His-Met-Glu-Lys
Original:	Ala-Val-Asn-Arg
Mutante:	Ala-Val-Ser-Arg
Original:	Arg-Ser-Leu
Mutante:	Arg-Ser-Leu-Trp-Lys-Thr-Phe

- a) ¿Cuál es la especificidad de cambio de base de este mutágeno?
- b) ¿Cuáles serán las consecuencias sobre la funcionalidad de la proteína en cada caso?
- c) ¿Qué será más perjudicial para el organismo: una mutación de corrimiento del marco de lectura o una mutación en el código genético? Explicá cómo podría ocurrir cada una desde el punto de vista molecular.