

L. Castaño, J.R. Bilbao

An Esp Pediatr 1996;45:315-320.

## Introducción a la biología molecular y aplicación a la pediatría (1): Conceptos básicos

### Introducción

Durante la última década, el mundo científico en general y la Medicina en particular han sido testigos de una de las revoluciones tecnológicas más espectaculares en las Ciencias Biológicas: la aparición de la tecnología de Biología Molecular. Su aplicación al desciframiento de las claves de la herencia humana han renovado en gran medida el concepto de diagnóstico, e incluso la terapéutica de un número cada vez mayor de enfermedades cuyos mecanismos etiopatogénicos y bioquímicos estamos descubriendo día tras día. Las publicaciones científicas bombardean constantemente con el descubrimiento de nuevos genes responsables de enfermedades, nuevos métodos de diagnóstico e incluso nuevas perspectivas de tratamiento. Poco a poco vamos conociendo las alteraciones moleculares que subyacen no sólo a enfermedades clásicamente catalogadas como hereditarias, caso del síndrome de Down, la fibrosis quística, etc., sino aquellas que son responsables de procesos mucho más complejos como el cáncer, la diabetes o la esquizofrenia. Conocer el defecto genético, nos permite llegar al diagnóstico definitivo de un proceso, o construir modelos animales de esas enfermedades sobre los que ensayar nuevas pautas de tratamiento. En la actualidad, el diagnóstico por tecnología molecular ocupa ya un lugar importante en la práctica médica habitual; además, estamos asistiendo al inicio de la puesta en marcha de tratamientos correctores del defecto a nivel genético, comenzando pues una segunda revolución en la medicina: la terapia génica.

A lo largo de los próximos capítulos desarrollaremos distintos aspectos generales de la Biología y Genética molecular, con el ánimo de introducir al especialista clínico en este campo fundamental en la Medicina actual. El objetivo de estos breves apuntes es que el lector conozca los conceptos básicos y las técnicas fundamentales de la Genética Molecular, para que le sirvan de referencia cuando profundice en el conocimiento de los avances moleculares de las enfermedades objeto de su estudio.

### 1.- Conceptos generales

#### 1.A.- El material genético:

Las células que constituyen el organismo humano, albergan

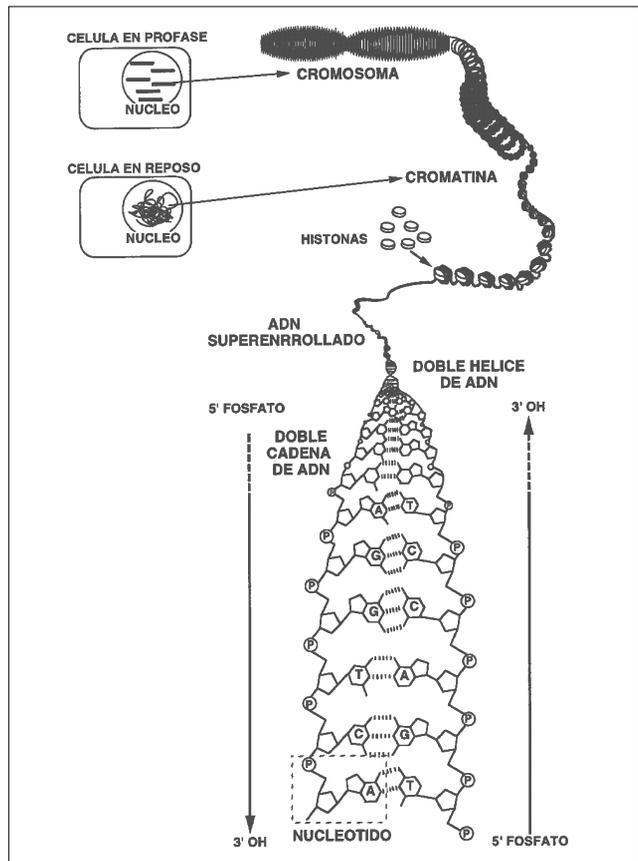


Figura 1. Estructura del ADN.

en su núcleo el material hereditario que contiene la totalidad de la información genética del individuo, a excepción del material hereditario mitocondrial. Esta información está "codificada" en una molécula, llamada **ADN** (del inglés Deoxyribonucleic Acid o ácido desoxirribonucleico), cuya estructura se detalla en la **Figura 1**. La totalidad del ADN de una célula se denomina **genoma**, y éste está organizado en unidades funcionales llamados **genes**.

El ADN es una molécula filamentos, enorme, que se enrolla sobre si misma, y se encadena con unas proteínas estructurales llamadas **histonas**, formando la molécula de **cromatina**. En la fase mitótica del ciclo celular llamada profase, la cromatina

Unidad de Investigación, Departamento de Pediatría, Hospital de Cruces, Barakaldo, Bizkaia.  
Correspondencia: Dr. Luis Castaño. Unidad de Investigación.  
2ª planta de Anatomía Patológica. Hospital de Cruces.  
Plaza de Cruces s/n. Barakaldo, Bizkaia 48903.

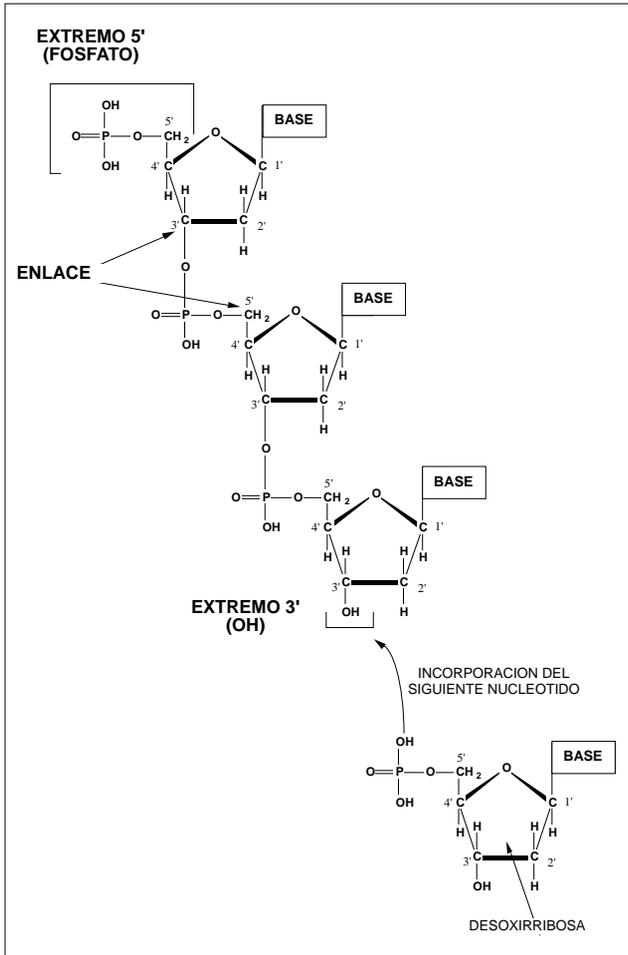


Figura 2. Dirección 5'-3' de una cadena de nucleótidos de ADN.

tina o complejo “ADN filamentos-histonas”, cambia de estructura formando los  **cromosomas** , estructuras visibles al microscopio óptico (Figura 1.). Cada cromosoma está formado por dos brazos (corto o “p” y largo o “q”) unidos por una estructura central o centrómero. Cada célula del organismo contiene 23 pares de cromosomas (22 autosomas y 2 cromosomas sexuales).

En su estructura íntima, el ADN está constituido por dos cadenas antiparalelas (corren en direcciones opuestas) formadas cada una por la concatenación de unas subunidades llamadas  **nucleótidos** , compuestos cada uno de ellos por un azúcar (desoxirribosa), un grupo fosfato y una de las cuatro bases nitrogenadas que forman el ADN [adenina (A), citosina (C), guanina (G) y timina (T)]. En cada una de las cadenas del ADN, el azúcar desoxirribosa de un nucleótido se une al fosfato del siguiente, en un sentido que se conoce como “ **dirección 5'-3'** ”, basándose en la numeración de los átomos de carbono del azúcar (Figura 2.). Para la formación de la doble cadena de ADN, una de ellas está unida a la otra por sus bases nitrogenadas, que quedan hacia el interior, enlazándose cada base con su correspondiente en la otra cadena mediante puentes de hidrógeno. Estos enlaces

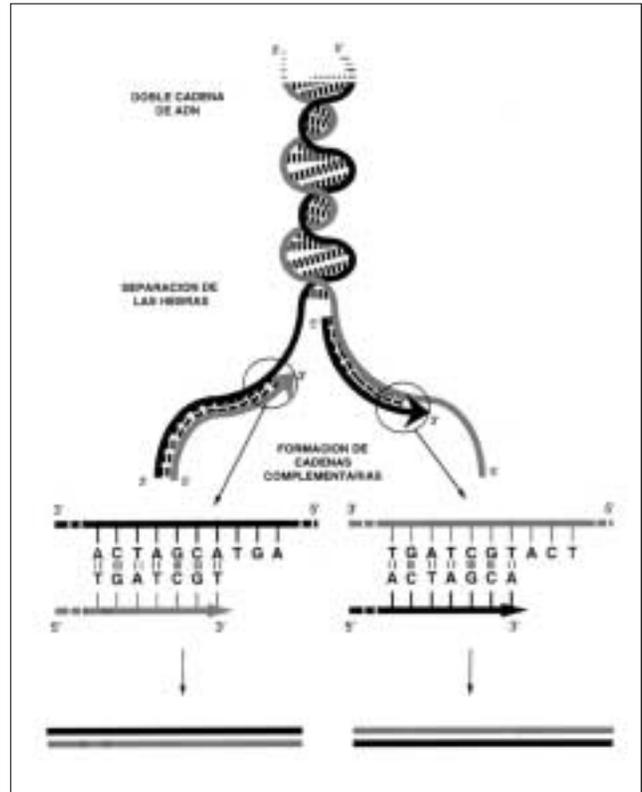


Figura 3. Replicación del ADN

se forman siempre entre la adenina y la timina (A-T), y la citosina y la guanina (C-G), con lo que la secuencia de bases de una de las cadenas es complementaria a la otra. Esta característica de doble cadena complementaria es fundamental para los procesos biológicos de la replicación (copia del material genético) y transcripción (síntesis de ARN mensajero), aspectos éstos que desarrollaremos posteriormente.

### I. B.- Replicación del ADN:

La división de una célula viene precedida por la duplicación de su material genético, el ADN, mediante un proceso que se denomina  **replicación**  (Figura 3.). Este proceso comienza con la ruptura de los puentes de hidrógeno y la separación de las dos cadenas que se abren en cremallera. Este proceso puede reproducirse en el laboratorio, sometiendo al ADN a calor o a álcalis, y se denomina  **desnaturalización del ADN** .

Cada una de las hebras separadas se comporta como un molde sobre el que se genera una cadena complementaria. La enzima  **ADN polimerasa**  se encarga de unir el grupo fosfato del nucleótido que se incorpora con la pentosa del nucleótido anterior, en dirección 5'-3'. El resultado de la replicación es la formación de dos moléculas bicatenarias de ADN idénticas (Figura 3.).

### I. C.- Función biológica del ADN: Transcripción y traducción

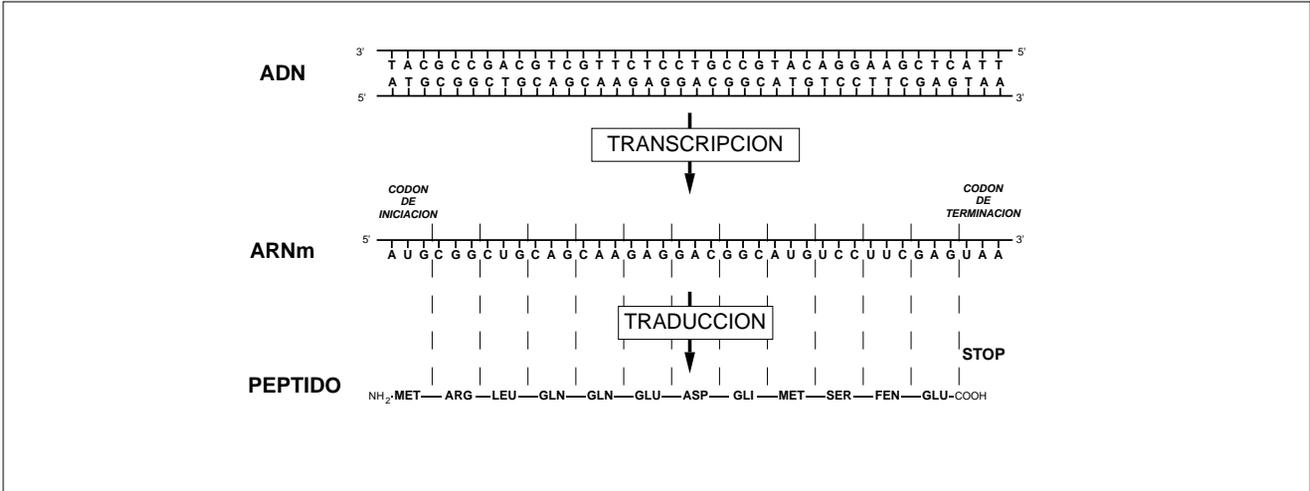


Figura 4. Transcripción y traducción. Correspondencia entre los tripletes de ADN, los codones de ARNm y los aminoácidos de un péptido.

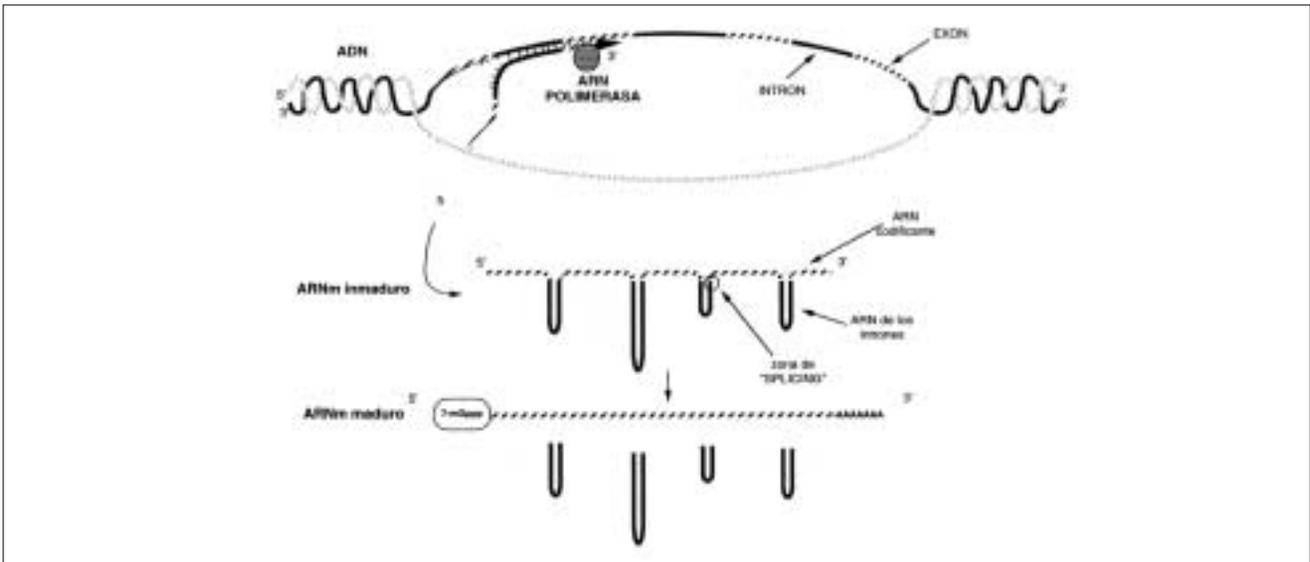


Figura 5. Transcripción del ADN y maduración del ARN mensajero.

La función de la molécula de ADN es la codificación de la información para la fabricación de todos los componentes proteicos de la célula, tanto estructurales como enzimáticos, y la transmisión de dicha información a la descendencia. El "lenguaje genético" en el que se codifica este mensaje viene determinado por el orden de las bases nitrogenadas de los nucleótidos en la secuencia del ADN. El elemento significativo de este lenguaje es la secuencia de tres bases de ADN o **triplete**. Como veremos más adelante, cada triplete de ADN se transcribe a un **codón** (unidad de codificación) o triplete de ácido ribonucleico mensajero (ARNm), que a su vez se traduce en un aminoácido de la proteína (traducción) (Figura 4). Este doble proceso, que constituye el dogma fundamental de la Biología, es el mecanismo por el que se lleva a cabo la síntesis de proteínas.

La información del ADN se estructura en genes, fragmentos de ADN que contienen la información para la síntesis de una proteína, que se encuentran dispersos por el genoma. No obstante, no todo el ADN porta información para la síntesis proteica; de hecho, sólo una parte muy pequeña del genoma está formado por ADN codificante (aquél que se llega a traducir a proteínas). Además, los propios genes contienen regiones codificantes, denominadas **exones** y regiones no codificantes (**intrones**) que se sitúan entre los exones, y que son eliminadas de la molécula de ARN mensajero antes de la traducción (Figura 5). La región de unión de un intrón con su exón inmediato se llama zona de unión o "**zona de Splicing**". Esta zona es importante porque aquí pueden ocurrir mutaciones de consecuencias graves, como veremos en otros capítulos.

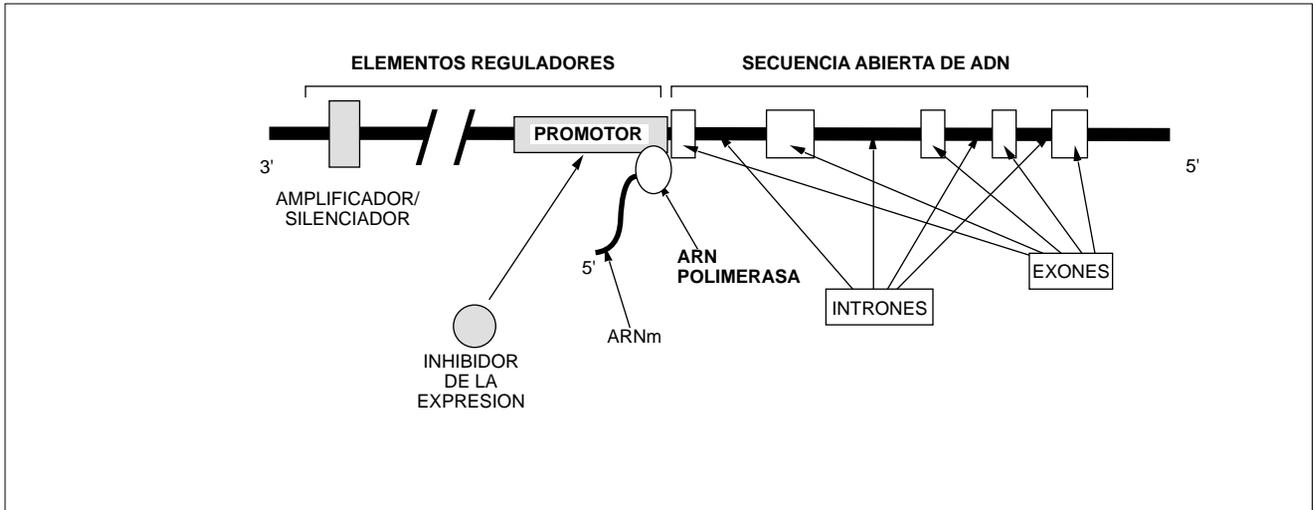


Figura 6. Estructura de un gen.

En Genética, el proceso de síntesis de proteínas se denomina **expresión génica**, haciendo referencia a la funcionalidad de los genes, y comporta dos procesos: transcripción y traducción.

La **transcripción** es el proceso de síntesis de ARNm a partir del molde de ADN. La molécula de **ARN (ácido ribonucleico)** es análoga al ADN, pero a diferencia de esta última, el ARN es una molécula monocatenaria, cuyos nucleótidos contienen ribosa en lugar de desoxirribosa, y uracilo (U) en lugar de timina (T). Además del **ARN mensajero (ARNm)** molécula encargada de llevar la información genética desde el núcleo a los ribosomas, existe también ARN de transferencia (ARNt) que transporta los aminoácidos al ribosoma para la formación de proteínas y ARN ribosómico (ARNr), componente fundamental de los ribosomas. La transcripción, se produce de forma similar a la replicación: las hebras de ADN se separan, y a partir de una de ellas que actúa como molde (hebra codificante) se sintetiza, por medio de la enzima **ARN Polimerasa**, una cadena complementaria de ARNm, en dirección 5'-3'. Posteriormente, tiene lugar la **maduración del ARNm**, que consiste en la eliminación de los intrones y la unión de señales para la traducción posterior. Estas señales son una cola de nucleótidos con adenina (cola poli-A) en el extremo 3' y un residuo de 7-metilguanosa (gorra m-7Gppp) en el extremo 5' (Figura 5).

La **traducción** es la síntesis del péptido en base al molde de ARNm maduro, conforme a las pautas del **código genético**, (Tabla I) según el cual, cada codón (tripleto de ARNm) se corresponde con un aminoácido. De los 64 tripletes posibles por la combinación de las cuatro bases (A, C, G y U), 61 de ellas codifican alguno de los 20 aminoácidos (hay diferentes tripletes que codifican para el mismo aminoácido), y los otros tres (UAA, UAG, UGA) **codifican señales de terminación** (final del péptido). Además, todos los péptidos sintetizados comienzan por el

Tabla I El código genético

PRIMERA POSICION	SEGUNDA POSICION				TERCERA POSICION
	U	C	A	G	
U	FEN	SER	TIR	CIS	U
	FEN	SER	TIR	CIS	C
	LEU	SER	STOP	STOP	A
	LEU	SER	STOP	TRP	G
C	LEU	PRO	HIS	ARG	U
	LEU	PRO	HIS	ARG	C
	LEU	PRO	GLN	ARG	A
	LEU	PRO	GLN	ARG	G
A	ILE	TRE	ASN	SER	U
	ILE	TRE	ASN	SER	C
	ILE	TRE	LIS	ARG	A
	<b>MET</b>	TRE	LIS	ARG	G
G	VAL	ALA	ASP	GLI	U
	VAL	ALA	ASP	GLI	C
	VAL	ALA	GLU	GLI	A
	VAL	ALA	GLU	GLI	G

aminoácido metionina, codificado por el codón AUG, llamado **codón de iniciación** (qué duda cabe que no todas las proteínas maduras tienen metionina como primer aminoácido, ya que éste puede perderse en procesos postraduccionales de maduración proteica. Esta característica es muy interesante en investigaciones que intentan aislar nuevos genes, ya que el hallazgo del tripleto ATG en una secuencia de ADN puede indicar el comienzo de una proteína, y por tanto de un gen. Las secuencias de ADN que comienzan por ATG, continúan con un número determinado de tripletes codificantes y concluyen con una señal de terminación se denominan **secuencias abiertas de ADN** (open reading frame), ya que pueden albergar un gen que se expresa.

En el proceso de traducción, que ocurre en los ribosomas citoplasmáticos, además del ARNm, interviene también el ARNt, encargado de llevar el aminoácido correspondiente a cada codón de ARNm, gracias a una secuencia complementaria a éste, el anticodón del ARNt.

#### **I.D.- Regulación de la expresión génica**

Es conocido que determinadas proteínas son características de ciertos tejidos e incluso de ciertos tipos celulares. Es necesario, por tanto, que existan mecanismos encargados de controlar la síntesis proteica o expresión génica en cada tipo celular. Este control tiene lugar fundamentalmente a nivel de la transcripción, y en él intervienen diferentes estructuras del propio gen.

Además de las secuencias abiertas de ADN, los genes contienen también elementos o secuencias reguladoras que modulan su expresión (Figura 6). Se trata de determinadas secuencias de ADN, generalmente localizadas previamente al propio gen, sobre las que interactúan moléculas (generalmente proteínas) capaces de activar o inactivar la expresión de esos genes, para que ésta ocurra en aquellos tejidos donde la proteína final es necesaria, y en el momento e intensidad adecuados. Entre las secuencias reguladoras destaca el **promotor**, indispensable para que se produzca la expresión, y que se encuentra antes del triplete ATG. El promotor contiene una secuencia universal, denominada **caja TATA** (por su secuencia T-A-T-A-T/A-A-T/A), lugar de unión de la ARN polimerasa necesaria para la transcripción del gen. A este promotor (o a secuencias próximas) pueden unirse otras moléculas que lo bloquean, convirtiéndose en inhibidores de la transcripción. Existen otras secuencias encargadas de regular la intensidad de la expresión génica, denominadas **amplificadores** o **silenciadores**, según su cometido. Estas regiones son muy variables y específicas de cada gen, y a ellas se unen, diferentes moléculas, como es el caso de receptores hormonales activados.

#### **Glosario**

**Acido desoxirribonucleico (ADN).**- Molécula portadora de la información genética, está formada por dos cadenas antiparalelas de nucleótidos compuestos por desoxirribosa, fosfato y una base nitrogenada (adenina, guanina, citosina o timina).

**ADN polimerasa.**- Enzima que une los nucleótidos de ADN durante el proceso de replicación.

**Acido ribonucleico (ARN).**- Polímero de nucleótidos formados por ribosa, fosfato y una base nitrogenada. A diferencia del ADN, es monocatenario, y presenta la base uracilo en lugar de timina. Los tres tipos de ARN: mensajero, ribosómico y de transferencia participan en la síntesis de proteínas en base al código genético.

**ARN mensajero (ARNm).**- Molécula de ARN que lleva la información genética desde el núcleo al ribosoma. Se transcribe a partir del ADN, y es el molde para la síntesis de proteínas en el ribosoma.

**ARN ribosómico (ARNr).**- Componente fundamental de los ribosomas.

**ARN de transferencia (ARNt).**- ARN que transporta los aminoácidos al ribosoma para la síntesis de proteínas.

**ARN polimerasa.**- Enzima que une los nucleótidos de ARN durante el proceso de transcripción.

**Caja TATA.** - Zona del promotor, en la región reguladora del gen, donde se une la ARN polimerasa para facilitar la transcripción.

**Codón.**- Conjunto de tres nucleótidos de ARN transcritos de un triplete de ADN, y que se traducen generalmente a un aminoácido.

**Codón de iniciación.**- Triplete de ARN compuesto de los nucleótidos A-U-G, que inician la síntesis de proteínas. Se traduce en metionina. En la fase de maduración de la proteína puede perderse esa metionina, motivo éste por el que muchas proteínas una vez maduras no comienzan con el aminoácido metionina.

**Codón de terminación.**- Tripletes U-A-A, U-A-G, U-G-A. Cuando se presentan en una secuencia de ADN significa que ese es el lugar de finalización de la secuencia de la proteína. Actúan pues como señal de "STOP".

**Cromatina.**- Estructura molecular formada por ADN e histonas, y se presenta en el núcleo celular. Contiene el material hereditario.

**Cromosoma.**- Cromatina estructuralmente organizada durante la profase del ciclo celular, y que se hace visible al microscopio. El genoma humano se organiza en 23 pares de cromosomas.

**Desnaturalización.**- Proceso de separación de dos cadenas complementarias de ADN, como paso previo para la replicación. Este proceso se realiza espontáneamente en el ciclo celular. En el laboratorio se puede reproducir sometiendo el ADN a calor o a álcalis.

**Expresión génica.**- Proceso de síntesis de proteínas, por funcionamiento de un gen.

**Exón.**- Regiones de ADN que se transcriben a ARN mensajero y consecuentemente codifican para un péptido. Se sitúan entre los intrones.

**Genes.**- Fragmentos de ADN que contienen la información para la síntesis de proteínas, y que se encuentran dispersos en el genoma. Está formado por una "secuencia abierta del gen", que codifica para el propio gen, y una región reguladora de su expresión, representada entre otros por el promotor.

**Genoma.**- Totalidad del ADN de una célula. Se organiza en genes

**Histona.**- Proteínas nucleares, que estructuran al ADN.

**Intrones.**- Regiones de ADN que no se transcriben a ARN mensajero, y consecuentemente no se traducen a proteínas. Se sitúan entre los exones. Su función no está clara aunque podría ser reguladora.

**Maduración de ARNm.**- Proceso de maduración por el que el ARNm pierde los intrones y se empalman los exones.

**Nucleótidos.**- Moléculas formadas por un fosfato, un azúcar pentosa (ribosa para los del ARN, y desoxirribosa para los del ADN), y una base nitrogenada (adenina, guanina, citosina o timina para el ADN, y adenina, guanina, citosina, o uracilo para el ARN).

**Promotor.**- Fragmento de ADN que se sitúa en la zona reguladora de cada gen, y cuya función es regular la síntesis o expresión proteica.

**Replicación.**- Proceso de duplicación de ADN.

**Secuencia abierta de ADN.**- Fragmento de ADN comprendido entre un codón de iniciación y otro de terminación, y que potencialmente puede representar el área codificante para una proteína. Del inglés "open reading frame".

**Traducción.**- Proceso de síntesis de los aminoácidos de una proteína a partir del ARN mensajero.

**Transcripción.**- Proceso de síntesis de ARN mensajero a partir de ADN.

**Triplete de ADN.**- Conjunto de tres nucleótidos, que forman la unidad del código genético, al transcribirse a un codón, y éste a su vez traducirse generalmente a un aminoácido.

**5'-3' (Dirección).**- Sentido de síntesis de una molécula de ADN o ARN. Los números se deben a que la unión entre nucleótidos se realiza de forma que el fosfato del primer nucleótido se une al carbono 5' del azúcar pentosa del siguiente nucleótido, y éste se une al fosfato del tercero por el carbono 3' de la misma pentosa.

**Zona de "Splicing" o unión.**- Secuencia de nucleótidos situada entre un intrón y su exón inmediato, por la que se rompe la cadena de ARN durante la separación intrón-exón que ocurre en el proceso de maduración del ARNm.

## Bibliografía

- 1 Watson J, Gilman M, Witkowski J, Zoller M. Recombinant DNA. Ed. Freeman & Co, 2nd ed. New York, 1992
- 2 Lewin B. Genes V, Ed. Oxford University Press, New York 1994.
- 3 Leon J. García JM. Manual de Genética Molecular, Ed. Síntesis, Madrid, 1ª Ed. 1990.
- 4 Ausubel F, Brent R, Kingston R, Moore D, Seidman J, Smith J, Struhl K. Current protocols in molecular Biology. Ed. J. Wiley & Sons, 1st Ed., 1994
- 5 Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A laboratory manual. De. Cold Spring Harbor Laboratory Press. 2nd Ed. New York 1989