



## FORMACIÓN CONTINUADA-ACTUALIZACIÓN EN MEDICINA DE FAMILIA

# Estructura y función del ADN y de los genes. I Tipos de alteraciones de la función del gen por mutaciones

M.L. Martínez-Frías<sup>a,b,c</sup>

<sup>a</sup>Departamento, de Farmacología, Facultad de Medicina, Universidad Complutense, Madrid, España

<sup>b</sup>Centro de Investigación sobre Anomalías Congénitas (CIAC ECEMC), Servicios de Información Telefónica (SITE y SITE), Instituto de Salud Carlos III (ISCIII), Madrid, España

<sup>c</sup>Grupo del CIAC en el CIBER de Enfermedades Raras (CIBERER), Madrid, España

Recibido el 7 de diciembre de 2009; aceptado el 7 de diciembre de 2009

Disponible en Internet el 5 de marzo de 2010

### PALABRAS CLAVE

Estructura del ácido desoxirribonucleico;  
Bases nitrogenadas;  
Código genético;  
Ácido ribonucleico mensajero;  
Transcripción alternativa;  
Tipos de mutaciones génicas

### KEYWORDS

DNA structure;  
Nitrogen bases;  
Genetic code;  
RNAm;  
Alternative transcription;

### Resumen

El ADN es «la molécula de la vida», y es la que lleva codificada la información genética característica de los diferentes seres vivos. Mediante ese código, regula el funcionamiento de cada tipo de célula; controla la transmisión de esa información, tanto en el tiempo como en el lugar de actuación de la misma; coordina la complejísima red de interacciones del funcionamiento celular y tisular; controla también su propia duplicación, reparación y autorregulación. Igualmente, controla y coordina los procesos de reproducción y mantenimiento de las características de cada especie. Todas estas actividades funcionales son reguladas y conducidas por un conjunto de instrucciones que constituyen el llamado *código genético*. El resultado se basa en un equilibrio entre la influencia del ambiente y esta compleja red funcional del ADN que muestra, además, un muy alto grado de plasticidad. Por ello, el genoma puede producir respuestas adecuadas a diferentes cambios del ambiente, manteniendo ese equilibrio. No obstante, a pesar de la importante capacidad homeostática del genoma, es susceptible de sufrir alteraciones por ciertos agentes que modifican el ambiente, dando lugar a efectos adversos y patológicos.

© 2009 Elsevier España, S.L. y SEMERGEN. Todos los derechos reservados.

### Structure and function of the DNA and genes. Types of alterations of the gene function through mutations

#### Abstract

The DNA is the “molecule of life,” and is that which carries the coded genetic information of the different human beings. Through this code, the functioning of each type of cell is regulated. It controls the transmission of this information, both in its time and place of action. It coordinates the very complex network of interactions of the cell and tissue function. It also controls its own replication, repair and self-regulation. Furthermore, it

Correo electrónico: mlmartinez.frias@isciii.es

## Types of genetic mutations

controls and coordinates the reproduction and maintenance processes of the characteristics of each species. All these functional activities are regulated and conducted through a combination of instructions that make up the so-called *genetic code*. The result is based on a balance between environmental influence and this complex functional network of the DNA that also shows a very high plasticity grade. Thus, the genome can produce adequate responses to different environmental changes, maintaining this balance. However, in spite of the important hemostatic capacity of the genome, it can be altered by some agents that modify the environment, giving rise to adverse effects and pathologic conditions.

© 2009 Elsevier España, S.L. and SEMERGEN. All rights reserved.

## Estructura del ADN

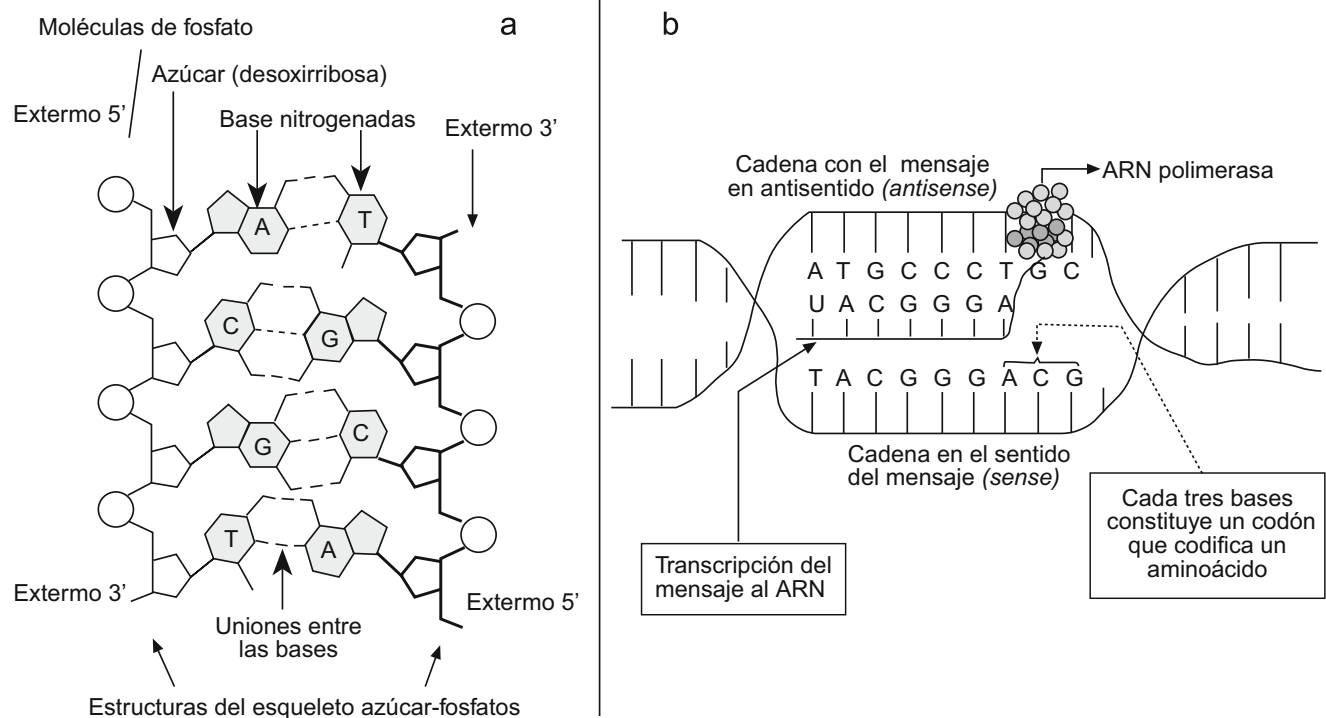
La molécula de ADN está constituida (fig. 1a) por una doble cadena en la que cada una de sus hebras está formada por uniones covalentes sucesivas entre un azúcar (desoxirribosa) y una molécula de fosfato. Cada azúcar de las dos cadenas está unido a una de las siguientes 4 bases nitrogenadas adenina (A), guanina (G), citosina (C) y timina (T). Estas 4 bases tienen distintas posibilidades de unión entre ellas a través de puentes de hidrógeno. Así, la A y la T, tienen 2 puentes de hidrógeno, mientras que la G y la C, tienen 3 puentes de hidrógeno. El número de puentes de hidrógeno establece una complementariedad específica entre las bases que determina sus uniones. Sin embargo, en la molécula del ácido ribonucleico (ARN), la T es substituida por el uracilo (U).

En el ADN, los dos extremos de los «esqueletos» de las dos cadenas complementarias de unidades «fosfato-desoxirribosa-base nitrogenada» (llamadas nucleótidos) terminan en un grupo

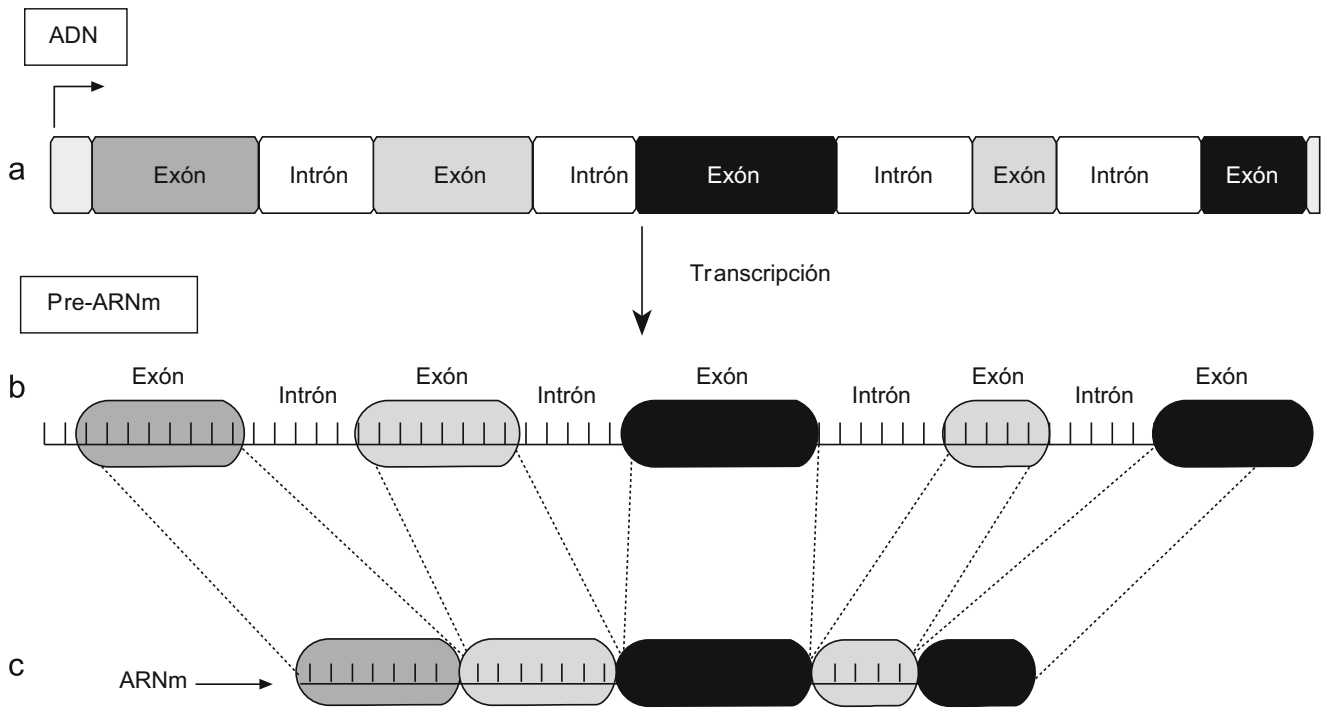
fosfato en uno de los extremos que se denomina extremo 5' (fig. 1a), y un hidroxilo del azúcar en el otro extremo, que se denomina 3'. Así, los dos «esqueletos» de desoxirribosa-fosfato se enfrentan en sentido contrario de manera que el extremo 5' se enfrenta siempre al 3' a través de las uniones complementarias de las bases, lo que confiere estabilidad a la doble cadena de ADN.

## Código genético

El ADN codifica la información genética mediante combinaciones de las bases, de forma que cada secuencia correlativa de 3 bases (triplete), que se denomina codón, codifica un aminoácido. Así, el codón ATG corresponde a la metionina (fig. 1b), y también es el que marca el sitio donde se inicia la lectura para el ARN mensajero (ARNm), que copiará el mensaje de los genes para trasladarlo al citoplasma, donde



**Figura 1** a) Estructura de la doble cadena de ADN. La complementariedad de las bases y los grupos terminales del esqueleto desoxirribosa-fosfato conforman la estructura estable de doble cadena del ADN. Los extremos de las cadenas terminan en un grupo fosfato llamado extremo 5', y en un hidroxilo del azúcar, que se denomina extremo 3'. b) Proceso de transcripción del mensaje genético al ARN mensajero. Las palabras en inglés se han escrito en cursiva.



**Figura 2** (a) Durante la transcripción se copia todo el ADN formando un pre-ARNm. (b) En este se produce la separación de los intrones y unión de los exones (*splicing*), (c) formando el ARN mensajero (ARNm), que dará la proteína.

se formará la proteína que codifica cada gen. Hay también 3 grupos de tripletes (TAA, TGA, TAG) que constituyen codones de parada.

Como las bases son 4, se pueden producir  $4^3 = 64$  combinaciones diferentes. Sin embargo, como los aminoácidos son 20, el código genético es redundante porque varias combinaciones de tripletes codifican un mismo aminoácido. Por ejemplo, la metionina solo la produce un único triplete, mientras que la glicina es codificada por 4, y la arginina por 6 codones.

### Función del código genético para la transcripción del mensaje

Cuando el ADN ha de trasladar el mensaje necesario para que en el citoplasma se forme una proteína, las 2 cadenas del ADN se separan en la zona que codifica el tipo y orden de los aminoácidos de esa proteína (el gen correspondiente). Ese mensaje va codificado en una de las dos cadenas, y ese mensaje se va a transcribir al ARNm, que lo trasladará al citoplasma para formar la proteína que codifica ese gen. Como se puede apreciar en la figura 1b, la cadena que lleva el sentido de los aminoácidos de la proteína (cadena *con sentido*) no es la que se copia sino la otra (cadena con el sentido contrario o *antisense*), manteniéndose así el orden adecuado (*el sentido*) de los aminoácidos en el ARNm para formar la proteína. Por ejemplo, según el esquema de la figura 1b, el código del segundo triplete del ADN de la cadena *con sentido* es el GGG (que corresponde al aminoácido glicina), sin embargo el ARNm, copiará el segundo triplete de la cadena *sin sentido*, que es CCC, y por tanto el ARNm llevará el código complementario GGG que codifica el aminoácido glicina, que es el que debe tener en segundo lugar la proteína que codifica ese gen. Por el

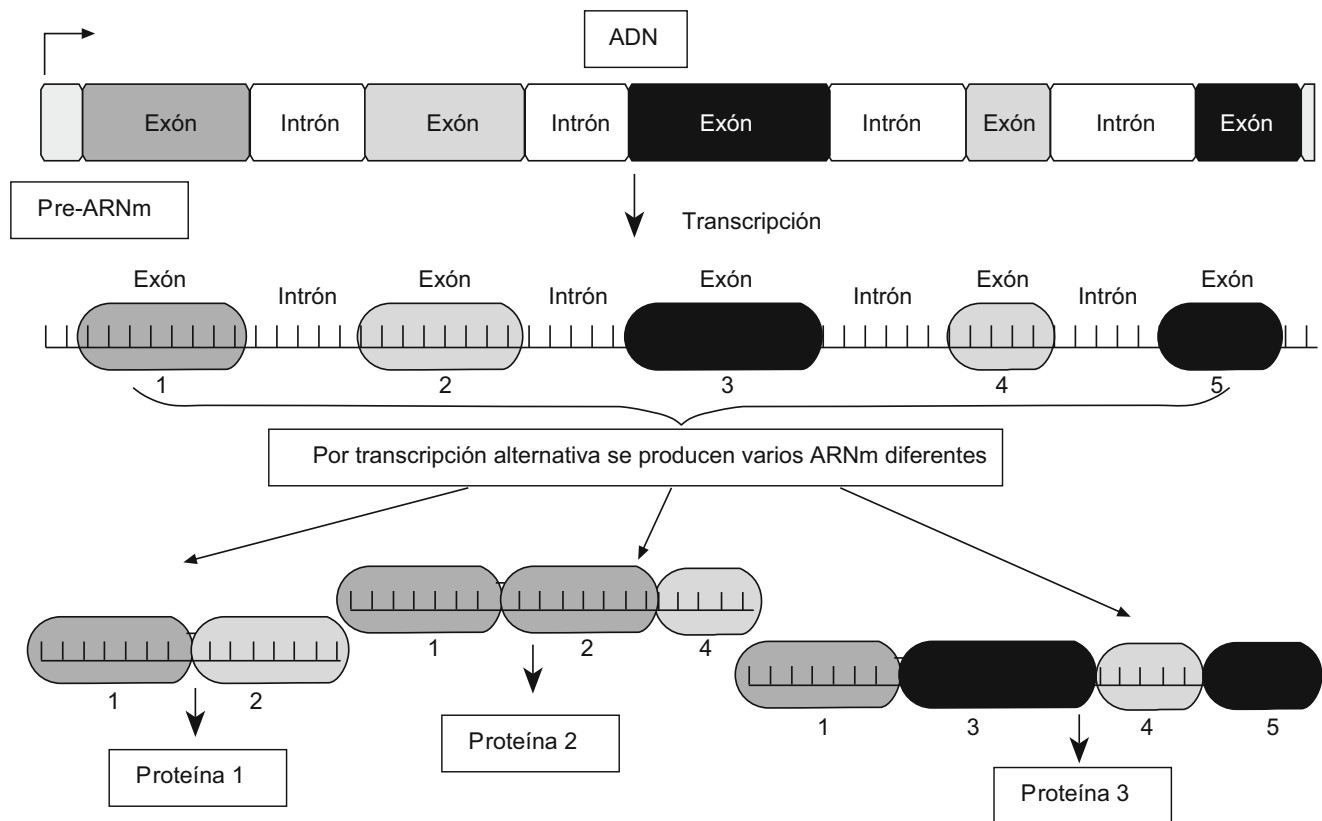
contrario, si el ARNm copiara la cadena *con sentido*, nunca se mantendría el mensaje. En nuestro ejemplo, si el ARNm hubiera copiado el segundo triplete de la cadena *con sentido*, el codón del ARNm habría sido CCC (prolina), por lo que el segundo aminoácido de la proteína no habría sido la glicina, sino la prolina, que no mantendría el orden y tipo de aminoácidos del código del gen.

### ¿Cuál es la estructura de los genes y cómo se forma del ARNm?

Los genes tienen la estructura que se muestra en la figura 2a, que consiste en segmentos denominados exones e intrones; siendo los exones los que llevan el mensaje para la formación de la proteína. Para ello, en la transcripción del mensaje genético, lo primero que ocurre es la copia completa del gen (con los exones e intrones) formando un pre-ARNm (fig. 2b). Posteriormente, se produce una separación de los intrones quedando sólo los exones formando así el ARNm (fig. 2c). Este proceso de separación de los exones y los intrones, y la posterior unión de los exones se denomina en inglés *splicing*, que significa unión. No obstante, en los últimos años, se está observando que los intrones también tienen importancia en estos procesos<sup>1,2</sup>.

### ¿Cómo es posible que se conozcan más de 100.000 proteínas distintas en los seres humanos cuando solo tienen alrededor de 25.000 genes?

Hasta finales del siglo pasado se consideraba que un gen codificaba una proteína y, como el número de proteínas



**Figura 3** Un mismo gen puede dar lugar a diferentes proteínas mediante el proceso de transcripción alternativa (*alternative splicing*). Es decir, formando distintos ARNm por distintas combinaciones de unión de los exones en cuanto al número y exón.

conocido era de unas 100.000, se pensaba que ese era el número de genes de los seres humanos. Sin embargo, los avances que se produjeron en esos últimos años del siglo XX en la tecnología molecular, que llevarían al desciframiento del genoma humano, el número de genes se estableció en unos 25.000.

Uno de los procesos conocidos que explican esa divergencia, es un mecanismo que se denomina transcripción alternativa o unión alternativa (*alternative splicing*). Este tipo de transcripción permite que un mismo gen pueda dar lugar a diferentes proteínas mediante la formación de distintos ARNm, a través de distintas combinaciones de los exones en la transcripción. En la *figura 3* se muestra un ejemplo en el que por transcripción alternativa de un mismo gen se producen 3 proteínas diferentes. De hecho, estudios muy recientes sobre lo que se denomina «transcriptoma» están mostrando que la transcripción alternativa es muy frecuente. Tanto que se ha llegado a considerar que un 75–95% de los genes que tienen muchos exones (*multi-exon genes*) realizan transcripción alternativa<sup>3</sup>.

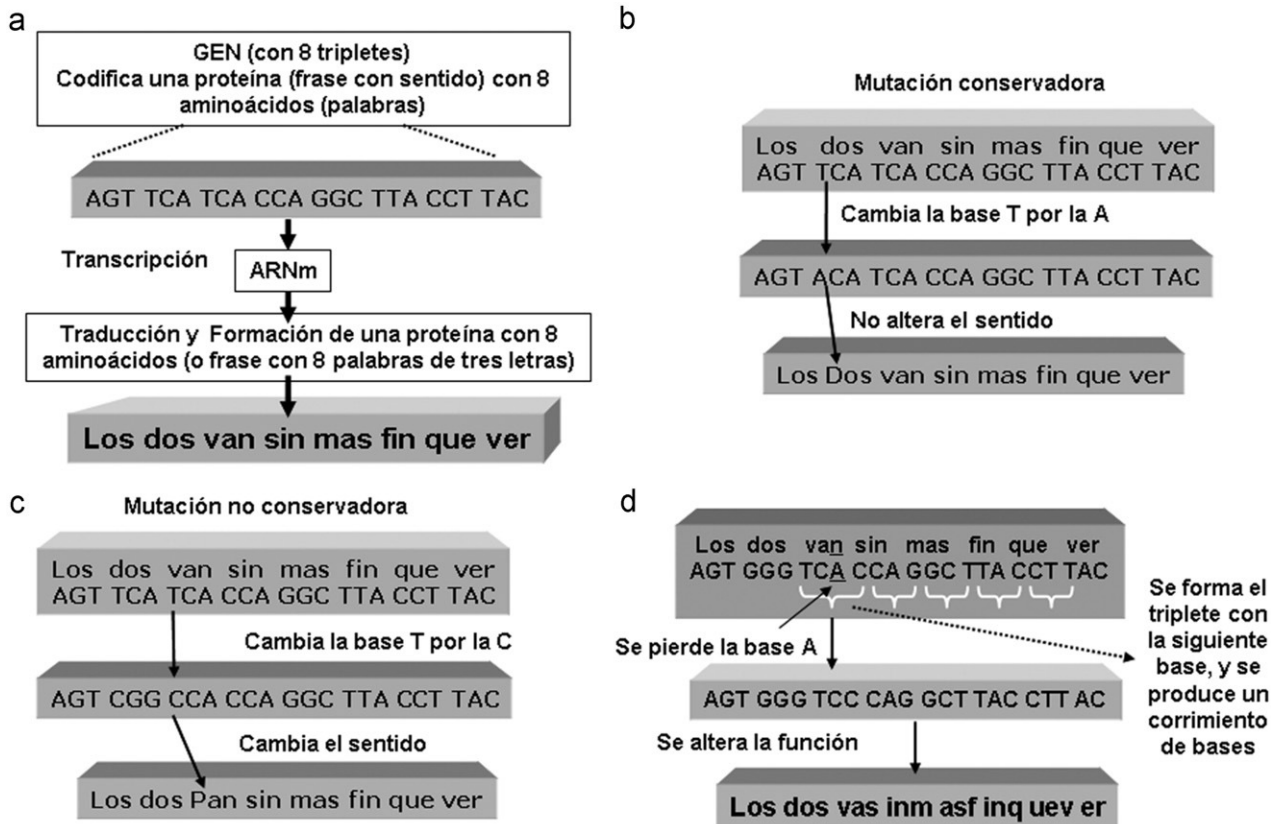
### Alteraciones del código genético por mutaciones

Como ya sabemos, combinaciones de 3 de las 4 bases A, G, T y C, codifican los distintos aminoácidos y, mediante diferentes combinaciones de aminoácidos de la cadena del ADN, se codifican las distintas proteínas. Cualquier cambio que se produzca en la combinación de las bases (o de los

tripletes), puede modificar el código del ADN y alterar la expresión de la proteína. Por ejemplo, si en un triplete TCA, que codifica el aminoácido serina, se produce un cambio de la T por una C, el aminoácido que ahora codifica el triplete CCA es una prolina, que es muy diferente de la serina, por lo que se altera la proteína. Estas alteraciones se consideran mutaciones génicas. Sin embargo, no todas las mutaciones van a ser patológicas, ni van a tener el mismo efecto en la proteína.

Teniendo en cuenta cómo los distintos cambios (mutaciones) afectan a la función de la proteína, esas mutaciones las podemos separar en 2 grupos: a) las que no alteran la función, que pueden ser de 2 tipos: mutaciones silenciosas y mutaciones conservadoras. b) Las que alteran la función, que pueden ser de varios tipos: mutaciones no conservadoras, mutaciones sin sentido y las que se producen por cambios (ausencia o ganancia) de un número de bases que no sea múltiplo de 3.

Con objeto de entender mejor estos aspectos, en la *figura 4* se muestran varios ejemplos. En primer lugar (*fig. 4a*) se esquematiza cómo el mensaje de un gen de ocho tripletes codifica una proteína de ocho aminoácidos, y el mensaje de esta proteína debe tener un sentido funcional. Ese mensaje, en nuestro ejemplo, daría lugar a una frase con ocho palabras de tres letras y con sentido, por ejemplo «los dos van sin más fin que ver». Sin embargo, si en el segundo triplete (TCA), que codifica la serina, se cambia la T por una A (ACA) el aminoácido resultante es la treonina (*fig. 4b*), que es muy similar a la serina, por lo que no cambia la función, y se considera una mutación conservadora. En



**Figura 4** a) Ejemplo de la transcripción del código de un gen. b) Ejemplo de mutación conservadora. c) Ejemplo de mutación no conservadora. d) Ejemplo de mutación «frame shift».

otras ocasiones, el cambio puede dar lugar al mismo aminoácido (por ejemplo TCA se cambia por TCC, que también codifica la serina), por lo que no cambia el sentido y se llama mutación silenciosa. Por el contrario, en la figura 4c se muestra que en el tercer triplete, el cambio de una T por una C, se traduce en el aminoácido prolina (CCA), que es muy diferente de la serina (TCA) por lo que altera la proteína (la frase ya no tiene sentido), siendo una mutación no conservadora. En el último ejemplo (fig. 4d), se pierde la base A del tercer triplete (que correspondería a la «n» de la palabra «van», que es la que codifica), por lo que el tercer triplete se formará ahora con la primera base del siguiente, este con la del otro, y así sucesivamente, lo que da lugar a aminoácidos diferentes y se trunca el sentido de la transcripción de la proteína, por lo que no se mantiene el mensaje. Lo mismo ocurre si se produce ganancia de alguna

base, o cualquier otro cambio de bases en números que no sean múltiplos de 3.

## Bibliografía

1. Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM, Armour CD, et al. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*. 2003;302:2141-4.
2. Thurman RE, Day N, Noble WS, Stamatoyannopoulos JA. Identification of higher-order functional domains in the human ENCODE regions. *Genome Res*. 2007;17:917-27.
3. Xing Y, Lee C. Relating alternative splicing to proteome complexity and genome evolution. *Adv Exp Med Biol*. 2007;623: 36-49 Review.