

BIOLOGÍA MODERNA Y CLONACIÓN

Francisco BOLIVAR ZAPATA *

SUMARIO: I. *Genes interrumpidos. Síntesis y procesamiento de RNA.* II. *El genoma y el proteoma del organismo vivo.* III. *El genoma y el proteoma humanos.* IV. *El uso e impacto de la información genómica en la salud. El inicio de la medicina molecular.*

I. GENES INTERRUMPIDOS. SÍNTESIS Y PROCESAMIENTO DE RNA

A principios de la década los años ochenta, en el siglo pasado, da inicio un esfuerzo muy importante encaminado al aislamiento y la caracterización de genes de organismos superiores, incluyendo el humano, con el objetivo general de entender en detalle la organización y la expresión de los genes. Así, utilizando diferentes estrategias y metodologías, se aísla un primer conjunto de DNA complementario a partir de copiar moléculas de RNA mensajeros específicos (figura 27), los cuales fueron, a su vez, utilizados para aislar los genes correspondientes a partir directamente del DNA cromosomal. Sorpresivamente, muchos de los genes cromosomales de organismos superiores resultaron tener un mayor número de nucleóti-

* Miembro de la H. Junta de Gobierno de la UNAM.

dos, es decir, un mayor tamaño, que los DNA complementarios correspondientes utilizados para aislarlos. Tras un análisis cuidadoso de estos resultados, que incluyó la determinación de la secuencia del material genético, fue posible concluir que, *de facto*, la mayor parte de los genes en los organismos superiores están constituidos por dos tipos de regiones de DNA. El primer tipo son regiones que codifican para la proteína, llamadas "exones", y el segundo tipo son regiones de DNA que no codifican para la proteína, a las cuales se les denominó "intrones". El DNA complementario, obtenido a partir de RNA mensajeros, sólo lleva o está constituido por las regiones de DNA que son los exones, y por eso siempre es de menor tamaño que el gen a nivel del DNA cromosomal.

Hoy sabemos, gracias al esfuerzo de muchos investigadores, entre ellos Chambon, que las células de organismos superiores como las nuestras han desarrollado mecanismos muy sofisticados para procesar el RNA que se produce durante la transcripción de un gen en el núcleo de la célula. La transcripción de RNA (figura 10) genera una molécula de RNA que es copia fiel del DNA transcrito. En el caso de las bacterias, esta molécula de RNA es directamente traducida en proteínas. Sin embargo, en el caso de los organismos eucariontes, la mayor parte de los transcritos de RNA son procesados para ser transformados en los RNA mensajeros que se traducen en proteína a nivel de los ribosomas.

En la figura 41 se muestra la estructura del DNA, a nivel del genoma, que codifica para la proteína B-globina, que es parte de la hemoglobina de la sangre en los seres humanos. Como puede verse en esta figura, este gen a nivel del cromosoma está constituido por alrededor de 1,600 pares de nucleótidos, de los cuales sólo 600 corresponden a tres secuencias de exones y los otros 1,000 pares de nucleótidos integran dos secuencias de intrones.

Al sintetizarse RNA a partir de este gen, se genera una primera molécula de "RNA premensajero" de aproximadamente 1,600 pares de nucleótidos. Este RNA es posteriormente "procesado" en el interior del núcleo de la célula. Este procesamiento consiste en remover fragmentos específicos de RNA, que en el caso de la B-globina corresponden a dos intrones, para generar así un RNA mensajero que sólo lleva la secuencia correspondiente a los tres exones. Al final de esta secuencia que codifica para los tres exones se localiza otra, a la que se le denomina "cola de poli A", que está presente en casi todos los RNA mensajeros de eucariontes. El RNA mensajero de aproximadamente 600 pares de nucleótidos es traducido en la proteína B-globina humana, que es una cadena polipeptídica de aproximadamente 150 residuos de aminoácidos (figura 16).

Es importante señalar que en términos generales, el procesamiento del RNA premensajero puede ser realizado de manera alternativa para el caso de varios genes, en cuanto a los fragmentos del RNA premensajero que son removidos. De hecho, se ha reportado que al menos el 35% de los genes humanos pudiera tener un procesamiento diferente y alternativo de su RNA premensajero, lo que implica que a partir de un RNA de este tipo se pudieran generar al menos dos RNA mensajeros diferentes (del mismo RNA premensajero) y de esta manera generar, a su vez, dos proteínas diferentes, tal y como se verá más adelante.

II. EL GENOMA Y EL PROTEOMA DEL ORGANISMO VIVO

Con toda esta información y gracias al mejoramiento y sofisticación continuos de las técnicas de DNA recombinante, en particular la aparición de técnicas poderosas de amplificación de DNA, tales como las técnicas de PCR

o reacción en cadena de polimerasa de DNA, y de secuenciación automática de DNA, hoy es posible analizar, inclusive sin clonar, genes de cualquier organismo, incluyendo al hombre. A través de ello se ha iniciado la etapa o la era del genoma, en la cual el esfuerzo ya no se concentra solamente en genes aislados, sino en el análisis del conjunto de todos los genes presentes en un organismo. En el caso de las bacterias, organismos unicelulares, su genoma lo conforman de 3,000 a 5,000 genes que se localizan en sus cromosomas, dependiendo de la especie (figura 42). Los humanos tenemos cerca de 40 mil genes localizados en las 46 cintas de DNA, los 46 cromosomas que conforman nuestro genoma, que por cierto es diploide (23 pares de cromosomas), lo cual quiere decir que tenemos la información genética por duplicado: una parte proveniente de nuestro padre y la otra de nuestra madre (figura 43).

En la era del genoma pretendemos conocer inicialmente cómo están organizados y localizados todos los genes en los cromosomas; es decir, todos los grupos que trabajamos en esta área tenemos como objetivo global contribuir a determinar los "mapas genéticos" de los seres vivos. Asimismo, pretendemos conocer cómo se regulan y en qué tipo de procesos celulares participan los genes y sus productos proteicos. En geografía, un mapa es la posición que guarda un país con respecto a los otros países en el planeta. En genética, un mapa es la posición que guarda un gen con respecto a los otros genes en las cintas de DNA que forman los cromosomas de un determinado organismo (figuras 2, 43 y 44). Si comparáramos uno de nuestros cromosomas e hiciéramos una analogía con un cassette o cinta musical, diríamos que de la misma manera en que en una cinta musical se encuentra almacenada información que se traduce en música, en una cinta de DNA se encuentra almacenada información ge-

nética que se traduce en proteínas. Hemos ya también señalado la analogía entre el DNA y una cinta musical, en donde sabemos que se encuentra almacenada información para diferentes canciones. Cada una de estas canciones en la cinta musical está a su vez almacenada en un segmento específico en esa cinta musical y de una manera lineal, es decir, primero se encuentra el segmento que codifica o guarda la información para la primera canción, luego el segmento para la segunda melodía, y así hasta la última canción; asimismo, estos segmentos de cinta son de diferentes tamaños y por ello las canciones también lo son. En el caso de una cinta genética de DNA, sabemos también que se encuentra almacenada información para cada una de estas proteínas o “canciones biológicas”, incluidas en un segmento específico de la cinta para cada una de ellas, que se llama gen. Estos segmentos de la cinta genética o genes se encuentran organizados también de manera lineal, es decir, uno después de otro, y que al igual que los segmentos de la cinta musical que codifican cada uno de ellos para una canción diferente y de diferente duración o tamaño, los genes codifican también cada uno de ellos para una proteína distinta y de diferente tamaño (figura 44).

El proteoma de un organismo es el conjunto de todas las proteínas que puede sintetizar ese organismo. Los humanos tenemos alrededor de 40 mil genes en cada célula de nuestro organismo, lo cual implica que podemos fabricar en nuestro organismo al menos 40 mil proteínas (ya que el procesamiento diferencial de ciertos RNA mensajeros de algunos genes humanos permitirá sintetizar más proteínas). En cambio, las bacterias que tienen entre 3 mil y 4 mil genes en su genoma (dependiendo de la especie) no tienen intrones, y por ello son capaces de sintetizar sólo entre 3 mil y 5 mil proteínas.

Hoy en día se ha logrado ya determinar la secuencia nucleotídica de todo el genoma de varias bacterias y también de varios organismos eucariontes, y con esto nace, *de facto*, la ciencia genómica, que nos permite el análisis comparativo de las secuencias de todos los genomas y de sus proteomas. El genoma de la bacteria *Escherichia coli* es uno de estos casos. Hoy se conoce la posición relativa en la cinta genética de su único cromosoma, de los 5,225 genes que constituyen el genoma de este microorganismo (figura 42). Gracias a esto, somos capaces de listar todas las 4,225 proteínas que puede sintetizar esta célula y que constituye su proteoma, cuya interacción hace posible los procesos de crecimiento y división de este organismo unicelular.

En el caso de los organismos eucariontes, que son aquellos que, a diferencia de los procariontes (bacterias), tienen todos sus cromosomas integrando el núcleo de la célula y rodeados por la membrana del núcleo, se han ya determinado las secuencias nucleotídicas completas de los genomas de varios de ellos, entre los que destacan el de la levadura, cuyo nombre científico es *Saccharomyces cerevisiae*, que es un microorganismo unicelular que se utiliza en la producción de alcohol; el de un gusano llamado *Caenorhabditis elegans*; el de la mosca de la fruta llamada *Drosophila melanogaster*; el de la planta *Arabidopsis thaliana* y, recientemente, el genoma de la especie humana.

La levadura es un organismo que tiene 16 pares de cromosomas, los cuales contienen entre todos ellos 6,241 genes. En el caso del gusano, este organismo tiene 6 pares de cromosomas y 19,099 genes; la mosca *Drosophila* tiene sólo 4 pares de cromosomas y en ellos se localizan 13,601 genes. La planta *Arabidopsis* tiene 25,498 genes en sus 5 pares de cromosomas, y en el caso de la especie humana se sabe que tenemos entre 30 mil y 40 mil ge-

nes. Estudios comparativos entre los genomas y los proteomas de estos organismos demuestran que hay un gran conjunto de genes y de proteínas muy parecidos que se encuentran en todos los organismos, incluyendo al hombre, y que son responsables de sus funciones biológicas primarias. De hecho, compartimos alrededor del 98.5% de nuestro DNA con el chimpancé, el 70% con el ratón y el 30% con la mosca.

Conocer las secuencias nucleotídicas de todos los genes de estos organismos, y a partir de ellas las secuencias de aminoácidos de todos sus productos proteicos, no significa, sin embargo, que de forma automática conoceremos todos los detalles del funcionamiento de estos seres vivos. *De facto*, estamos solamente al principio del entendimiento de cómo la regulación fina y sincronizada del genoma permite el desarrollo y funcionamiento del organismo vivo. Sin embargo, es indudable que nuestra capacidad para entender el funcionamiento y desarrollo de cualquier organismo será potenciada a través del conocimiento de su genoma, de sus instrucciones genéticas, de su proteoma y de la comparación de éstos con las presentes en otros organismos.

Por ejemplo, estudios recientes de comparación de genes entre los genomas secuenciados de los diferentes organismos eucariontes animales demuestran que, en el caso de la mosca, existen más del 65% de los genes que en los humanos son responsables de las enfermedades congénitas hasta ahora identificadas en la especie humana. En otras palabras, en la mosca se han encontrado al menos 177 genes que tienen equivalencia con genes humanos involucrados en enfermedades genéticas. Ejemplo de estos casos es el de una mutante de la mosca, que presenta una patología similar a la que se observa en pacientes con la enfermedad de Parkinson. Indudablemente, el análisis del funcionamiento de los genes y de los

productos proteicos en este insecto permitirá, en tiempos mucho más cortos, conocer con gran detalle las bases moleculares involucradas en muchas enfermedades genéticas humanas, a través de conocer lo que ocurre en otros organismos modelo, como la mosca, abriendo así posibilidades extraordinarias y novedosas para el tratamiento de este tipo de enfermedades en el humano.

Únicamente a través del conocimiento de cuándo y qué proteína aprecherà en cierta etapa del desarrollo del organismo vivo es que seremos capaces de entender en detalle la complejidad de toda interacción genética y proteica que subyace al más simple de los procesos de funcionamiento y desarrollo. Así pues, obtener el mapa genético y la secuencia del genoma y del proteoma del organismo vivo es elemento primario para el entendimiento del fenómeno de la vida y de su evolución.

III. EL GENOMA Y EL PROTEOMA HUMANOS

Hasta hace poco se tenían secuenciados y mapeados en nuestros cromosomas humanos (figura 43) más de 30 mil genes, utilizando para ello los DNA complementarios correspondientes. En los primeros meses de 2000 se publicó la noticia de que se había ya determinado la secuencia de todos los fragmentos de DNA que conforman el genoma humano, y que en un lapso de no más de dos años se tendría una secuencia del genoma humano con más del 95% de certeza.

Además, también a principios de 2000 se reportó la secuencia nucleotídica del cromosoma humano número 22, en el cual se localizaron 545 genes. Asimismo, se logró también la determinación de la secuencia nucleotídica del cromosoma número 21, el más pequeño de los 23 cromosomas de nuestra especie, y que tiene 33.5 millones de pares de bases y en el que se localizan 225

genes (figura 44). Estos números de genes, relativamente reducidos conforme a lo esperado por el tamaño de los cromosomas 21 y 22, hizo pensar que el número final de genes humanos no llegaría a los 80 mil o 100 mil originalmente estimados, sino tal vez serían entre 35 mil y 40 mil los genes de la especie humana.

Finalmente, en febrero de 2001, dos grupos reportaron simultánea pero independientemente en las revistas *Nature* y *Science* la secuencia del genoma humano. Ciertamente este resultado, alcanzado por cierto antes de lo previsto, es uno de los logros más importantes de la biología, comparable con el desciframiento de la estructura del DNA y con la teoría de la evolución, y va a permitir un avance extraordinario en el conocimiento del organismo humano. Entre los datos más relevantes del resultado de estos trabajos estarían los siguientes:

La secuencia que se reporta es un consenso de la porción de la "eucromatina" (región particularmente rica en genes), que corresponde alrededor del 95% del genoma haploide (es decir, 22 cromosomas somáticos y los dos cromosomas sexuales [figura 43]). Se reporta la secuencia de cerca de 3 mil millones (3×10^9) de pares de bases de los cromosomas humanos. Lo anterior significa que las 24 moléculas de DNA de nuestros 24 cromosomas que conforman el genoma humano miden un poco más de un metro, ya que cada par de bases está separado por 3.4 Å (un angstrom es 10^{-8} cm), así que al multiplicar $3 \times 10^9 \times 3.4 \times 10^{-8}$ cm el resultado es 102 cm.

Dos seres humanos somos genéticamente 99.9% idénticos. Lo anterior implica que, si bien entre dos individuos compartimos la mayor parte de nuestro material genético, hay, sin embargo, aproximadamente en promedio, tres millones (3×10^6) de nucleótidos diferentes entre dos miembros de la raza humana; encontrándose, entonces, en promedio entre dos individuos un nucleótido diferente

cada 1,200. Estas diferencias son, en lo general, el producto de mutaciones acumuladas en el genoma de la raza humana a través de los años. Aquellas que en particular son el resultado de una mutación de un solo par de bases en un gen (figura 14) reciben el nombre de “polimorfismos genéticos de un solo nucleótido” (SNP, por sus siglas en inglés).

En la población humana, un gen en particular ha sufrido, a lo largo del tiempo desde la aparición de la especie humana, diferentes mutaciones en distintos individuos, dando lugar así a estos polimorfismos genéticos (SNP). Al conjunto de todas las diferentes mutaciones en ese gen humano (presentes en toda la población humana) se le conoce con el nombre de “alelos” de ese gen en particular. Muchas de estas mutaciones no tienen efecto en la proteína para la cual codifica el gen; sin embargo, algunas otras mutaciones sí pueden cambiar ligera o profundamente la estructura de la proteína (figuras 14 y 16). En el caso de un cambio menos, el individuo que porta esta o estas mutaciones muy probablemente podrá vivir con este alelo sin problemas. Sin embargo, en el caso de mutaciones que cambian la estructura y por ende la función de la proteína de manera importante, los portadores pueden desarrollar enfermedades complejas e inclusive morir por estos cambios más severos. A la fecha se han reportado cerca de 2 millones de estos polimorfismos genéticos del tipo SNP.

Los polimorfismos genéticos son, en buena medida, la razón genética responsable de las diferencias físicas entre los seres humanos. Desde el punto de vista de la medicina, los SNP son marcadores genéticos importantes para entender no sólo las enfermedades, sino también las diferentes predisposiciones a las enfermedades que existen entre las razas y los individuos. Los polimorfismos genéticos son, en suma, los responsables de la identidad

genética individual de cada ser humano. La presencia e identificación de los SNP en cada individuo está dando lugar al desarrollo de una medicina molecular individualizada orientada al diagnóstico preventivo y al diseño de drogas específicas (farmacogenómica) a nivel del individuo. Ejemplos de lo anterior se señalan posteriormente.

Los dos artículos que reportan la secuencia del genoma humano señalan que no es posible fijar aún el número preciso de genes que conforman nuestro genoma, ya que hay secuencias que pudieran ser realmente genes, pero existen todavía definiciones que tienen que precisarse, sobre todo por el problema de la presencia en el genoma humano de "pseudogenes", material genético repetido que no funciona como un gen (véase más adelante). Sin embargo, ambos grupos concuerdan que el genoma humano tendrá entre 30 mil y 40 mil genes. Este número, que es importante por varias razones, es un número no mucho mayor al de los genes presentes en la planta *Arabidopsis thaliana* (alrededor de 26 mil genes) y en la mosca *Drosophila melanogaster* (alrededor de 13,600).

El material genético de los exones de estos 30 mil a 40 mil genes, que codifica para proteínas, sólo representa entre el 1% y el 2% de todo nuestro genoma. Los intrones reportados en los genes de nuestro genoma representan alrededor del 30% del total del genoma. Una diferencia importante con los otros organismos eucariotes secuenciados, a nivel de los exones y los intrones de los genes, es que en el caso del humano, los intrones son, en promedio, mayores en tamaño. Esta diferencia pudiera ser importante, ya que podría ser parte de la razón por la cual es posible obtener un mayor número de sitios de procesamiento alternativo de los RNA premensajeros que tienen mayores tamaños, para generar así más de una proteína a partir de cada transcrito de RNA premensajero. De hecho, se ha reportado recientemente

por varios grupos que existe una alta probabilidad de que ocurra este tipo de proceso, y que más del 35% de los RNA premensajeros humanos contienen sitios alternos de procesamiento.

Esta situación pudiera ser la responsable de explicar la aparente paradoja de porqué el genoma humano, con sólo 30 mil o 40 mil genes, de hecho codifica para un proteoma con más de 40 mil proteínas, resultantes de este tipo de procesamiento diferencial de los transcritos de RNA premensajero, y también de los procesos de modificación postraduccional de las proteínas. De acuerdo con lo anterior, no sería extraordinario que finalmente el proteoma humano estuviera constituido por un conjunto muy superior a 40 mil proteínas diferentes, y que algunos consideran que por la complejidad del organismo humano pudiera ser, incluso, mayor de 100 mil proteínas.

En cuanto al análisis del proteoma humano que se reporta por estos grupos, se señalarían algunas características relevantes. La primera es que una proporción más importante de lo que antes se pensaba de las proteínas humanas hasta ahora reconocidas está involucrada en procesos del metabolismo celular y en procesos de transcripción-traducción; lo anterior hace sentido en el marco de una actividad celular de transcripción y traducción muy intensa, consistente con la especulación de tener un proteoma funcional de mayor tamaño al del genoma. Otras funciones en las que se encuentra también involucrado un número elevado de proteínas se relacionan con los procesos de comunicación intra y extracelular, de transporte y de defensa e inmunidad.

Por otro lado, un resultado importante es que cerca del 1% de nuestras proteínas tiene probablemente un origen bacteriano, y por lo tanto los genes que las codifican debieron haberse adquirido mediante procesos de transfe-

rencia (infección) horizontal, ocurridos en diferentes etapas de la evolución de los vertebrados.

Es relevante mencionar también que se han identificado ya más de mil genes involucrados con enfermedades humanas, y también se han determinado genes que codifican para proteínas que pudieran ser blanco del desarrollo de nuevas drogas para el tratamiento de problemas de salud; ejemplos del uso de estos genes y sus proteínas se mencionan en la siguiente sección.

Otra característica muy importante de nuestro genoma es la presencia de lo que se conoce con el nombre de "material genético repetido". Al menos el 50% del total del material genético del genoma humano y probablemente más, son secuencias de bases que se repiten numerosas veces y de formas diferentes. Este tipo de DNA, en términos generales, se puede clasificar en cinco categorías, mencionadas por el Consorcio Internacional para la Secuencia del Genoma Humano: *a)* material repetido derivado de transposones; *b)* copias inactivas de genes, llamadas pseudogenes; *c)* secuencias de tamaño corto, repetidas varias veces; *d)* duplicaciones de regiones del genoma de 10-300 kb, que han sido copiadas e incorporadas en otra región del genoma, y *e)* bloques de secuencias repetidas en tandem, como los centrómeros y los telómeros de los cromosomas.

En estos artículos se señala que cerca del 90% de este material genético repetido en nuestro genoma (que equivale al 45% del total del DNA en el genoma) es del tipo del material repetido derivado de transposones. Los transposones son secuencias de DNA que se localizan en el genoma de todos los seres vivos y que tienen la propiedad de poder reubicar o translocar su posición en el genoma. Llevar a cabo esta función de relocalización pueden hacerlo, dependiendo del tipo de transposón que se trate, dejando copias perfectas o imperfectas del trans-

posón en el sitio o sitios nuevos de integración o relocalización. En los seres humanos hay cuatro clases de este tipo de DNA proveniente de transposones: a) los elementos tipo retrovirus; b) las secuencias LINE; c) las secuencias SINE, y d) los transposones de DNA.

Las secuencias repetidas del tipo de los retrovirus, de las cuales existen en nuestro genoma del orden de 450 mil copias, provienen de infecciones de virus que tienen RNA como material genético, que insertaron copias de su genoma como DNA en diferentes sitios del genoma humano a lo largo del tiempo. Muchas de estas secuencias contienen todavía un gen activo que codifica para la transcriptasa reversa. La transcriptasa reversa es una proteína con actividad de enzima que sintetiza DNA a partir de RNA, y es la enzima que usan el virus del SIDA y otros retrovirus para incorporar copias de su genoma en los cromosomas humanos. Varios miles de estos genomas virales están casi completos en su secuencia integrados en nuestro genoma; por suerte, la mayor parte de ellos son inertes por haber perdido alguna parte de un gen importante. Estos "retrovirus endógenos de humano", también conocidos con el nombre de *Hervs* (human endogenous retrovirus, por sus siglas en inglés), representan cerca del 8% de nuestro genoma.

La secuencia LINE son segmentos de DNA entre 6 mil y 8 mil pares de bases que se encuentran repetidas aproximadamente 850 mil veces, y que por ende representan el 21% del genoma humano. Son secuencias de DNA que contienen dos genes estructurales que codifican para dos proteínas. Al transcribirse el DNA de la secuencia LINE en RNA, estas dos proteínas se asocian al RNA transcrito y participan en su proceso de retrotranscripción, por la transcriptasa reversa, para copiar y luego integrar el material copiado de RNA en DNA en otra región del cromosoma. Las secuencias LINE son autónomas para realizar

este proceso y son también secuencias gregarias, ya que se presentan varias copias juntas en diferentes regiones de los cromosomas.

Las secuencias SINE son secuencias de DNA de 100 a 300 pares de bases que están repetidas cerca de 1.5 millones de veces en el genoma humano; lo cual representa aproximadamente el 13% de nuestro material genético. Las secuencias SINE, en particular las llamadas Alu, tienen semejanza con el gen que codifica para uno de los RNA que forman los ribosomas. Este gen que codifica para RNA ribosomal tiene un promotor interno que probablemente sea responsable de una transcripción especial que posibilita la formación de RNA pequeños que pueden, a su vez, transcribirse en DNA por la transcriptasa reversa y convertirse en estas secuencias Alu que sólo llevan una parte del gen que codifica para RNA ribosomal. Las secuencias Alu sólo se encuentran en primates.

Finalmente, las secuencias tipo transposones de DNA son del orden de 300 mil copias de secuencias parecidas a los transposones bacterianos, y representan el 3% de nuestro genoma. Varias decenas de nuestros genes se originaron muy probablemente a partir de este tipo de transposones de origen bacteriano.

Mucho se ha especulado con relación al posible origen y papel del material repetido en nuestro genoma. Ciertamente, parte importante de este material es el resultado de la transferencia horizontal de DNA a través de infecciones virales y bacterianas a lo largo de la evolución en los antecesores de nuestra especie. El material genético repetido ha sido denominado por algunos como "DNA basura o genes egoístas", por no tener una función biológica aparente.

Sin embargo, también hay otras opiniones y evidencias que consideran que los elementos de DNA repetidos de-

rivados de transposones pudieran participar en el rearrreglo de regiones específicas del genoma, lo que a su vez podría contribuir en ciertos procesos de adaptabilidad evolutiva del organismo. De los anterior es importante resaltar que el genoma ciertamente no es un ente estático, invariable, sino todo lo contrario. El genoma es un sistema dinámico, interactivo, que se rearrregla en cierta medida, lo que probablemente permite la adaptación, la evolución y la sobrevivencia del organismo ante cambios en los factores del medio ambiente.

El avance de la genética y de la ciencia genómica es vertiginoso. Lo anterior está contribuyendo al entendimiento profundo de los procesos finos involucrados en la organización, expresión y modificación de los genes, y a través de sus productos proteicos, del funcionamiento de la célula viva. Por otro lado, el comparar los procesos normales con procesos patológicos ha facilitado sustancialmente el entendimiento de enfermedades y problemáticas clínicas. Ciertamente, la determinación de la secuencia del genoma humano nos permitirá abordar problemáticas clínicas más complejas y enfermedades multifactoriales, en donde participan varios genes, de manera mucho más certera e individualizada, tal y como se menciona en la siguiente sección de este trabajo.

Finalmente, es fundamental señalar aquí que los factores ambientales juegan también un papel muy importante, e incluso pueden ser indispensables para desencadenar procesos normales y patológicos causados por uno o por la concurrencia de varios genes y sus productos proteicos. Así pues, nuestra vida y nuestras respuestas biológicas no es sólo cuestión de genes, sino de genes y de la interacción de sus productos proteicos con factores ambientales.

IV. EL USO E IMPACTO DE LA INFORMACIÓN GENÓMICA EN LA SALUD. EL INICIO DE LA MEDICINA MOLECULAR

Como hemos ya señalado, las técnicas de ingeniería genética han permitido, desde 1973, el aislamiento de muchos genes humanos y su utilización para la construcción de organismos transgénicos para la producción de proteínas humanas recombinantes, que hoy en día se utilizan en diferentes problemáticas clínicas y para el tratamiento y la prevención de enfermedades. Asimismo, con el avance de la ciencia genómica y proteómica —y particularmente con el desciframiento de la secuencia del genoma humano— tenemos una visión más avanzada de la forma en que están organizados los genes humanos, y también de las diferencias, los polimorfismos genéticos, que existen en todos y cada uno de los genes humanos y que son responsables de nuestra individualidad genética y por ello también de nuestra predisposición genética a enfermedades.

Esta nueva información genómica, aunada a las técnicas de ingeniería genética, permiten hacer un uso más sofisticado de los genes humanos, no sólo para producir proteínas específicas en organismos transgénicos, sino también para contender de manera más individualizada con aspectos fundamentales de la salud humana, entre los que señalaríamos el diagnóstico genético, la farmacogenómica y la terapia génica, tres áreas de gran importancia que se describen a continuación.

1. *El diagnóstico genético*

Las enfermedades genéticas humanas son el resultado de la presencia de mutaciones en uno o más genes humanos. Como resultado de lo anterior, en el individuo que

porta estos genes mutantes se producen procesos fisiológicos anormales que dan lugar a enfermedades genéticas, algunas de las cuales están listadas en la figura 45. Uno de los propósitos del diagnóstico genético es conocer la presencia de los genes mutantes en los individuos que los lleva, haciendo uso para ello de genes humanos normales y funcionales aislados en el laboratorio (figura 46).

El desciframiento del genoma humano, como hemos señalado, ha permitido identificar muchos polimorfismos genéticos, que son responsables de la individualidad genética de cada ser humano y también de su particular predisposición genética a contraer enfermedades. La detección y el diagnóstico temprano de estas diferencias representa un cambio cualitativo paradigmático en la práctica médica, ya que permitirá a cada individuo diseñar una estrategia de vida, incluyendo el posible tratamiento médico, más adecuada y más individual para contender con sus enfermedades genéticas presentes y futuras. Lo anterior es particularmente importante para aquellos individuos cuyas familias hayan mostrado una mayor susceptibilidad a contraer una cierta enfermedad, debido a la presencia de uno o varios polimorfismos genéticos específicos.

El extraordinario caudal de información que día a día emana del estudio del genoma humano nos ha incorporado cada día más a esta nueva etapa de la genética y de la medicina molecular moderna: la del diagnóstico genético preventivo e individualizado. Sin embargo, hoy en día, aun cuando hemos avanzado en forma increíble y extraordinaria en esta área de la genética humana —se conocen más de mil genes humanos involucrados en enfermedades genéticas— el diagnóstico genético todavía no representa una herramienta de uso cotidiano, en particular en países en vías de desarrollo como el nuestro. Es

indudable que, conforme se vayan conociendo y aislando, cada día más rápidamente, genes y polimorfismos genéticos asociados a enfermedades genéticas, particularmente aquellos que provean información acerca de las enfermedades humanas más comunes y de la predisposición a contraerlas, el diagnóstico genético individualizado será componente de toda estrategia médica en todos los sistemas de salud.

Reconociendo la inminencia del uso de esta poderosa tecnología de manera masiva, y sin dejar de resaltar el potencial de todo este nuevo paradigma médico, es sin embargo necesario destacar algunos aspectos éticos relevantes del uso de estas capacidades y de la información genética. Debemos distinguir muy claramente entre el diagnóstico orientado a los adultos, los niños, personas discapacitadas mentalmente e individuos no natos. Surge aquí el concepto y el aspecto fundamental de la privacidad genética y biológica, donde el concepto de “autorización de la obtención y uso de la información genética” debe aplicarse de manera distinta en estos grupos. Es obvio que se debe legislar para que el diagnóstico genético sólo se practique cuando exista una autorización por parte del individuo o su representante legal en el caso de infantes, cuyo DNA pretende examinarse. Sin embargo, es indudable también que aun cuando existan las leyes adecuadas, entraremos en problemáticas éticas, morales, filosóficas y jurídicas complejas cuando se otorguen autorizaciones por individuos ignorantes del tipo de consecuencias de los resultados que puedan obtenerse de este tipo de permiso. De forma realista, es posible prever, como lo han hecho notar muchos y en particular James Watson, descubridor de la estructura del DNA, que habrán muchos individuos que autoricen la realización de pruebas de diagnóstico, sin comprender todos los escenarios que los resultados pudieran tener en su vida futu-

ra. Así pues, resulta imperativo desarrollar programas de divulgación y también de educación sobre el DNA y la genómica que no sólo expliquen los aspectos fundamentales de esta disciplina, sino que también dejen claro el significado en particular de resultados positivos en pruebas de diagnóstico para enfermedades incurables, al menos por ahora. Hoy en día existen muchos individuos que conociendo tener el 50% de probabilidades de portar un gen mortal, como el asociado a la enfermedad de Huntington, no quieren conocer su destino final, prefiriendo vivir en la incertidumbre que tener que cargar con la idea de que a cierta edad desarrollarán un problema terrible.

2. *La farmacogenómica*

La industria químico-farmacéutica ha generado a lo largo de los años un conjunto muy importante de productos farmacéuticos que se utilizan en el tratamiento de diferentes problemas clínicos. Muchos de estos productos están dirigidos o tienen como blanco una proteína específica, en algún tejido de nuestro organismo. Estas proteínas tienen funciones particulares, dependiendo de su estructura (figura 16), tales como receptores de moléculas pequeñas o transductores de señales, y permiten el funcionamiento de la célula y del organismo. Ciertamente, el conocimiento del genoma y del proteoma humano facilitará diseñar nuevas drogas más potentes y específicas contra aquellos blancos proteicos ya conocidos. Sin embargo, este conocimiento también permitirá seleccionar un conjunto más amplio de genes y sus proteínas que pudieran ser blancos específicos de los actuales y de nuevos productos farmacéuticos para el tratamiento más efectivo de muchas enfermedades (figura 47).

Resulta importante resaltar que el grupo del Consorcio Internacional para la Secuencia del Genoma Humano re-

porta, en su artículo publicado en *Nature*, un primer conjunto de 18 nuevas proteínas con dominios receptores aparentes para diferentes moléculas, tales como la dopamina, y factores de crecimiento tipo insulina, entre otros, que pudieran ser blanco de éstos y otros productos. Indudablemente, el desarrollo de nuevas drogas, a través del enfoque y del apoyo de la ciencia genómica, y en particular de la farmacogenómica, abrirá una rica avenida para el desarrollo de nuevos medicamentos.

Por otro lado, como ya hemos señalado, el uso de la información de los polimorfismos genéticos, a nivel de cada individuo, permitirá también conocer las diferencias entre las variedades funcionales de cada uno de los genes humanos, que se conocen como los alelos de ese gen. Estas diferencias son las responsables, en muchas ocasiones, a través de la presencia en las células de proteínas ligeramente modificadas, de las diferentes susceptibilidades a las enfermedades y también de las diferencias en cuanto a la acción de los medicamentos.

A través de la identificación de estas diferencias presentes en cada individuo en las proteínas que funcionan como receptores de drogas o transductores de señales, podremos modular la concentración y/o afinidad de los medicamentos, inclusive al grado de diseñar drogas específicas para cada individuo, dependiendo del alelo particular que se tenga.

3. *La terapia génica*

Independientemente del buen uso de los sistemas de diagnóstico, que permitan orientar el estilo de vida, conforme al conocimiento de nuestros polimorfismos genéticos y por ello de nuestra predisposición a enfermedades, habrá siempre individuos que nazcan con padecimientos genéticos complicados y terribles. De esta realidad surge

la terapia génica, como una herramienta orientada a curar las deficiencias genéticas. Esta metodología implica el introducir una o varias copias de genes normales para sustituir la función de genes ausentes o mutantes en los cromosomas de las células de enfermos, para así curar la enfermedad (figura 48).

En el momento actual los esfuerzos correctivos de terapia génica están orientados a introducir genes normales o mensajeros antisentido en las células somáticas del organismo que presentan alguna deficiencia genética. Hace tan sólo unos años fue realizado un primer experimento en células de la médula espinal de un niño con una enfermedad de inmunodeficiencia; estas células aisladas y cultivadas *in vitro* fueron transformadas con genes normales para el defecto de la inmunodeficiencia y algunas de ellas incorporaron y expresaron el gen. Estas células modificadas genéticamente fueron posteriormente trasplantadas a la médula del propio niño enfermo, que después de este tratamiento se ha recuperado.

Éste es el primer experimento de terapia génica en humanos que abre una nueva área y una nueva era en la medicina y la biología modernas. Al día de hoy, son varios los experimentos de terapia génica realizados en humanos, algunos inicialmente exitosos, lo cual es extraordinario y alentador.

Por otro lado, no existen a la fecha, y al menos por ahora no deben permitirse, intentos para modificar genéticamente las células germinales de nuestros organismos, que son las células que son transferidas a las generaciones humanas posteriores. La decisión de no permitir efectuar experimentos en células germinales se sustenta principalmente en la convicción de que es inaceptable experimentar con la vida humana cuando no podemos conocer ni predecir por ahora el impacto global que pudiera tener

este tipo de intervención genética en un nuevo ser humano.

Si tratamos de juzgar la manera en que debemos proceder desde el punto de vista ético y moral con los procedimientos de la terapia génica y el diagnóstico genético, o con cualquier otro aspecto de la genética que pudiera tener consecuencias para la vida humana, diríamos, como ha señalado James Watson, que debemos orientar nuestros esfuerzos a propiciar el desarrollo de marcos jurídicos y de acciones sociales que impliquen o permitan los potenciales más altos para mejorar la calidad de la vida humana. Actuando de esta manera, sin embargo, se debe esperar controversia, ya que el nuevo conocimiento genómico modificará la percepción sobre nosotros mismos y nuestro lugar en el planeta, e indudablemente entrará en conflicto con valores tradicionales. Además, no debe olvidarse que, de cualquier manera, las herramientas del DNA recombinante y el conocimiento sobre los genomas y los proteomas están ya con nosotros, y que tenemos la obligación de usarlos no sólo para beneficio de la raza humana, sino de la vida misma.

Finalmente, señalaríamos que la Declaración de la Conferencia General de la UNESCO, de 1997, sobre "el genoma humano" resulta ser un avance fundamental en este sentido, ya que genera un marco moral y ético sobre los derechos y responsabilidades para el manejo de la información genética presente en el genoma de la raza humana, y permite simultáneamente abrir un espacio relevante para el análisis y el debate sobre el manejo de los genomas de otros organismos, con los que conformamos la biodiversidad de nuestro planeta.

El reto es ciertamente extraordinario y apasionante, y de resolverlo inteligentemente depende nuestra sobrevivencia.

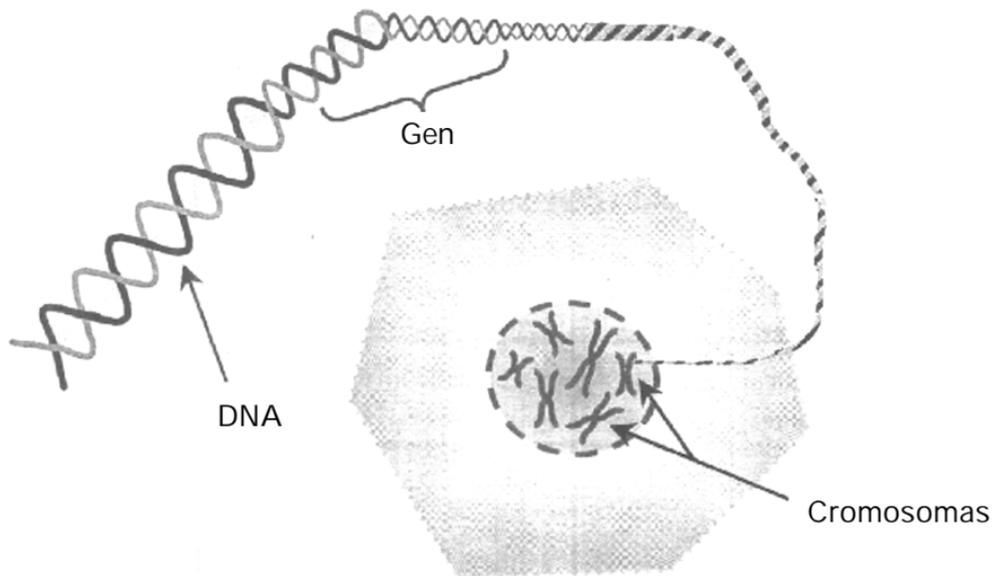


FIGURA 2. COMPOSICIÓN Y ORGANIZACIÓN DE LOS GENES EN LOS CROMOSOMAS

Los cromosomas son estructuras celulares que se encuentran localizados en el núcleo de las células y están formados por proteínas y ácido desoxirribonucleico (DNA). La información genética reside en el DNA y los genes son segmentos específicos de esta cinta genética llamada DNA. Es importante recalcar que cada ser vivo tiene un número específico y diferente de cromosomas con relación a los demás organismos vivos.

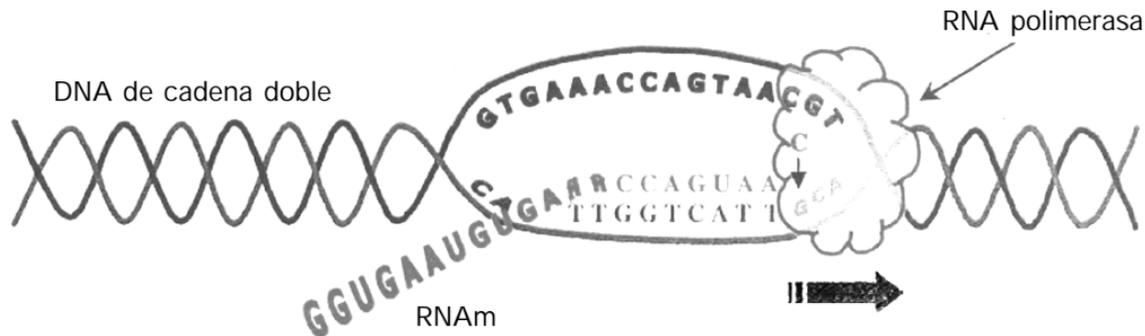


FIGURA 10. EL FENÓMENO DE LA TRANSCRIPCIÓN PERMITE LA SÍNTESIS DEL RNA MENSAJERO

Las moléculas del RNA mensajero (se muestra como cinta roja) son transcritas o producidas por la enzima RNA polimerasa, a partir de regiones específicas de material genético (que contienen uno o varios genes), utilizando como molde una de las dos cadenas del DNA. Las moléculas de RNA son polímeros lineales de cuatro tipos de nucleótidos: A, G, C y U, en donde la diferencia primaria con el DNA es que el Uracilo (U) es utilizado en lugar de la Timina (T), durante su síntesis. Estas moléculas de RNA mensajero, que son el primer intermediario en la síntesis de las proteínas, son exportadas del núcleo en el caso de las células eucariontes, y luego, en el citoplasma celular son leídas y su información es utilizada para sintetizar proteínas. Cada molécula de RNA mensajero es recipiente de la información para sintetizar una o varias proteínas específicas.

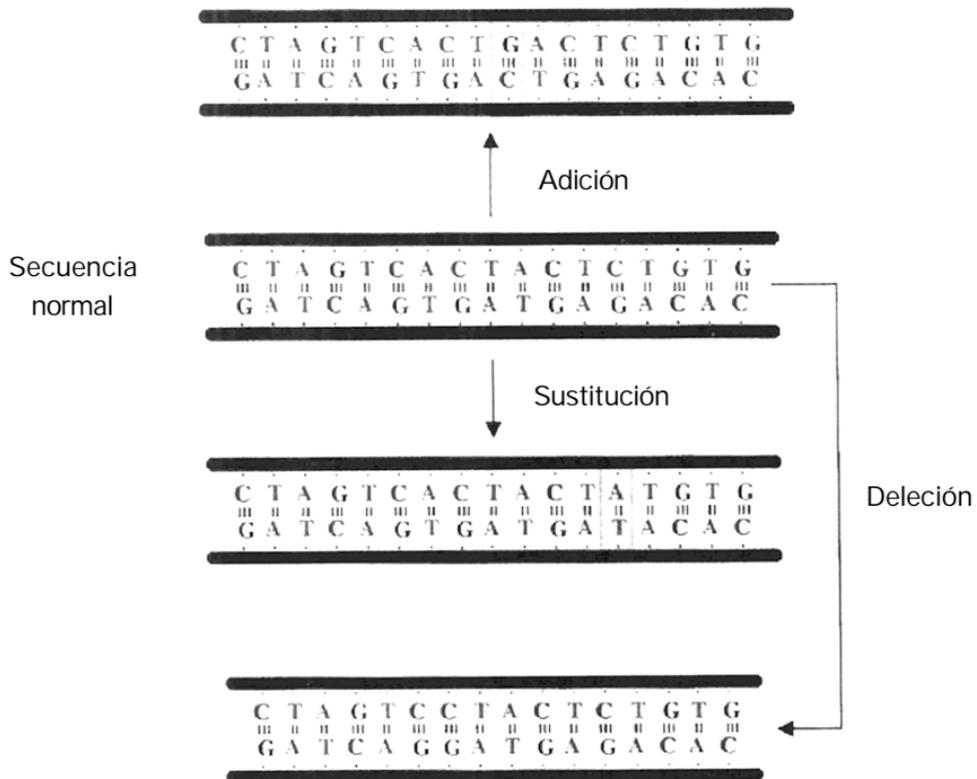


FIGURA 14. TIPOS DE MUTACIÓN

La figura muestra los diferentes tipos de mutaciones que pueden ocurrir para cambiar la secuencia original de nucleótidos en cualquier molécula de DNA: a) sustitución; b) adición, y c) deleción. En la figura se muestra el efecto de la sustitución, adición y deleción de un par de nucleótidos; sin embargo, las adiciones o deleciones pueden involucrar uno o muchos pares de nucleótidos.

Al cambiar la secuencia de nucleótidos de un gen, puede tener resultados diversos sobre el producto peptídico o proteína para la que codifica. Si el cambio es sólo de un par de nucleótidos, en muchos casos, no hay efecto en la actividad de la proteína. Sin embargo, en algunos casos este cambio sí afecta la función biológica.

Los cambios en los genes, debido a las mutaciones, son elemento fundamental en el proceso de la evolución de los seres vivos, ya que permiten el fenómeno de la variabilidad genética y biológica.

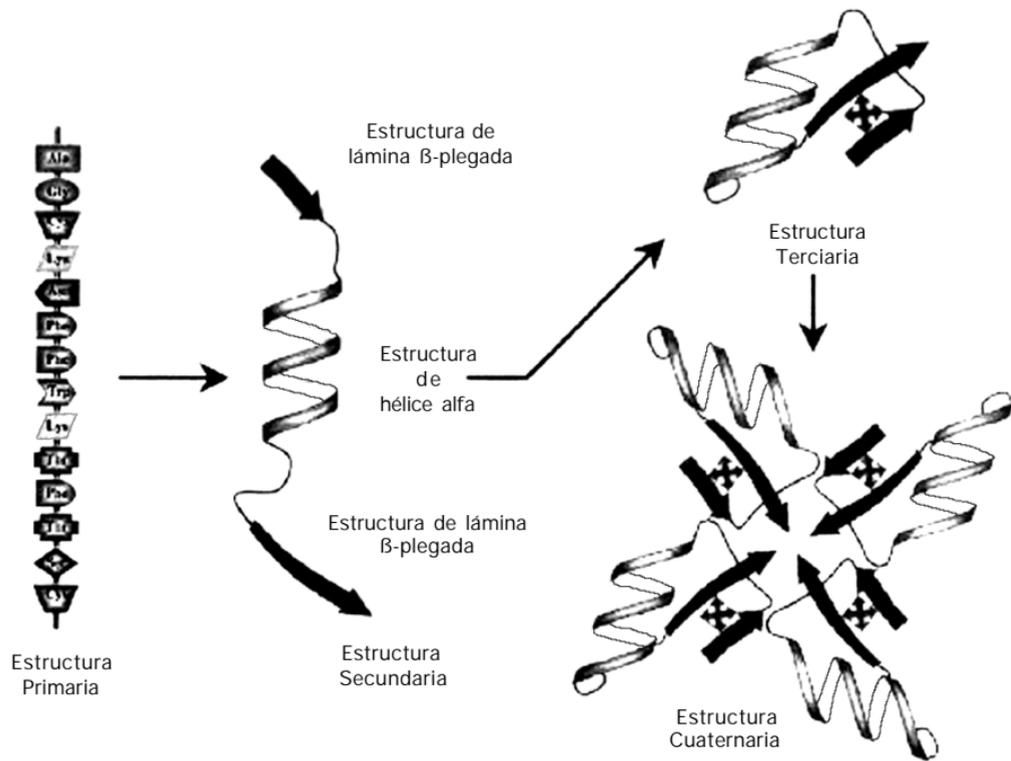


FIGURA 16. PROTEÍNAS: ESTRUCTURA Y FUNCIÓN

Las proteínas son las herramientas que tiene la célula viva para llevar a cabo la mayor parte de sus funciones. La función de cada proteína es específica y está determinada por la estructura tridimensional de la propia proteína. Las proteínas son polímeros de varias decenas o centenas de aminoácidos que son los monómeros de este tipo de polímero biológico. La estructura tridimensional (terciaria y cuaternaria) de cada proteína, está determinada por la estructura primaria, que es en realidad la secuencia de aminoácidos de la proteína.

En cada proteína existen regiones de este polímero que pueden organizarse, en lo que se conoce como estructuras secundarias, del tipo (alfa) hélice o β -plegada. A partir de este tipo de doblamientos, la proteína adquiere su estructura terciaria. Finalmente, la estructura cuaternaria de las proteínas es el resultado de asociar varias proteínas ya estructuradas a nivel terciario, lo que les permite un arreglo espacial multipolimérico para llevar a cabo sus funciones.

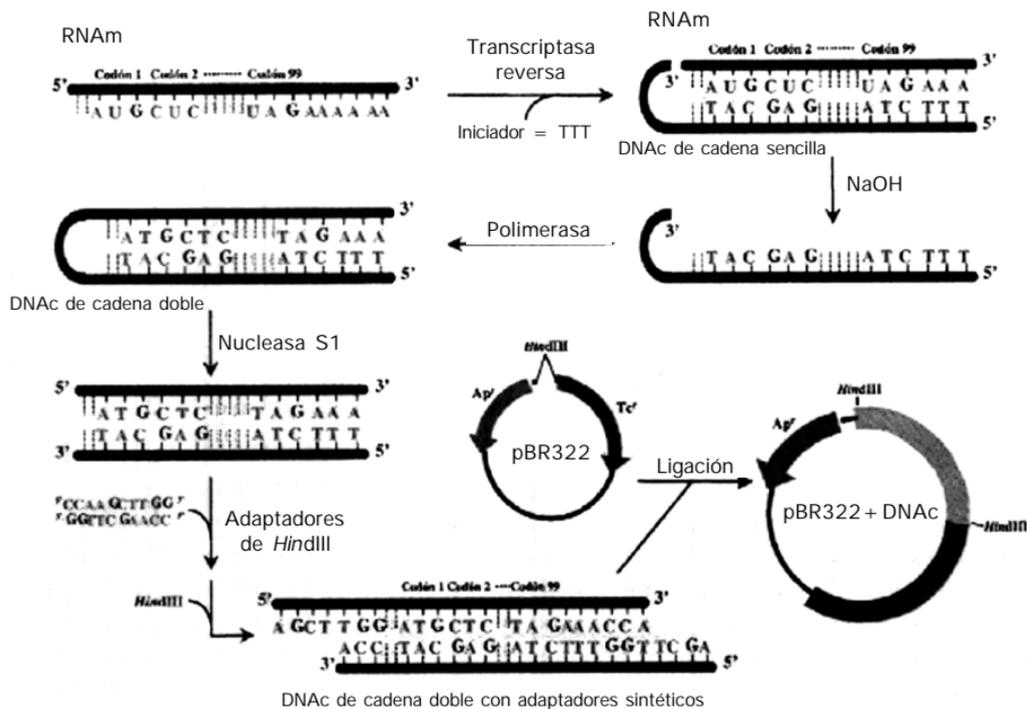


FIGURA 27. SÍNTESIS ENZIMÁTICA DE DNA A PARTIR DE RNA MENSAJERO

El copiado enzimático de una molécula de RNAm para producir una copia de DNA complementario (DNAc), comprende varios pasos. Primero, se copia el RNAm con la enzima transcriptasa reversa y el DNAc de hélice sencilla que se genera se libera del RNA al ser tratado con álcali (NaOH). Posteriormente, con una polimerasa, se obtiene la segunda cadena de DNA; finalmente, la estructura de asa terminal se digiere con una nucleasa que degrada hélice sencilla. Se genera así un DNAc de cadena doble con extremos rasurados que puede clonarse directamente o bien ligarse a moléculas adaptadoras de DNA. En el diagrama se muestra su unión con un adaptador de DNA sintetizado químicamente que contiene un sitio de reconocimiento para la endonucleasa *Hind*III. El DNA puede ser clonado directamente en un sitio de *Hind*III presente en un vehículo molecular, en este caso el plásmido pBR322.

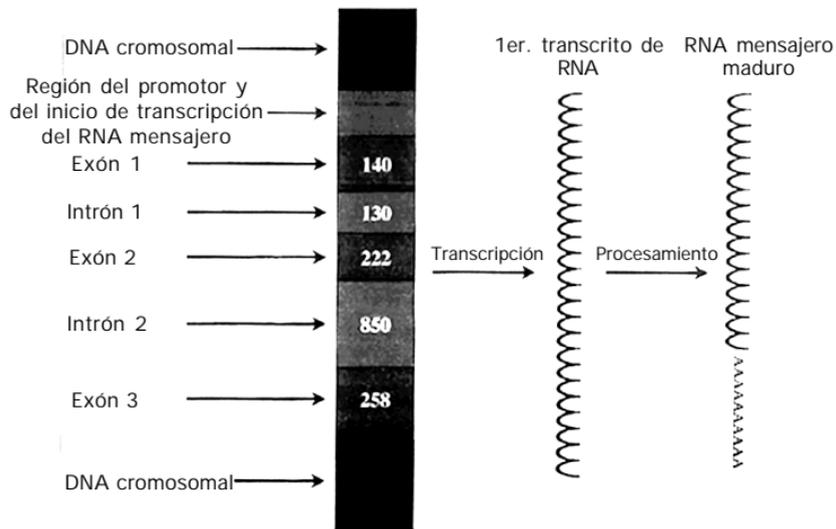


FIGURA 41. ORGANIZACIÓN DEL GEN DE LA PROTEÍNA β -GLOBINA

El gen que codifica para la proteína β -globina está formado por tres exones y por dos intrones, cuyos tamaños se señalan a nivel del DNA. Como resultado del proceso de transcripción de este gen, se produce una primera molécula de RNA que es copia fiel de todo el gen, incluyendo los exones y los intrones. Las regiones de RNA codificadas por los intrones en este RNA precursor son removidas a nivel del núcleo por enzimas específicas, generándose así la molécula de RNA mensajero, a la cual se le une una "cola de Poli A" formada por varios residuos de adenina.

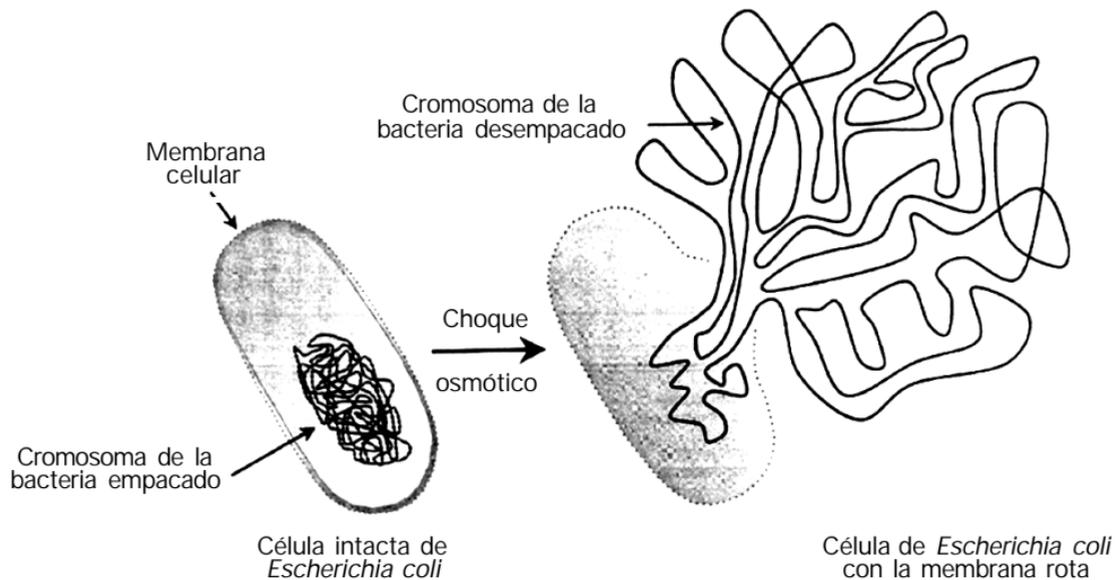


FIGURA 42. EL GENOMA DE LA BACTERIA *ESCHERICHIA COLI*

La figura muestra un diagrama de la bacteria *E. coli*, en la cual aparece el cromosoma bacteriano en forma empaquetada.

Al someter a la bacteria a un choque osmótico, ocurre un rompimiento de la membrana bacteriana, y el único cromosoma de esta bacteria, se desorganiza, se desempaca y eventualmente se fragmenta. En la figura se muestra el cromosoma desempaquetado, que mide un milímetro, y en el que se encuentran localizados 4262 genes que conforman el genoma de este organismo.

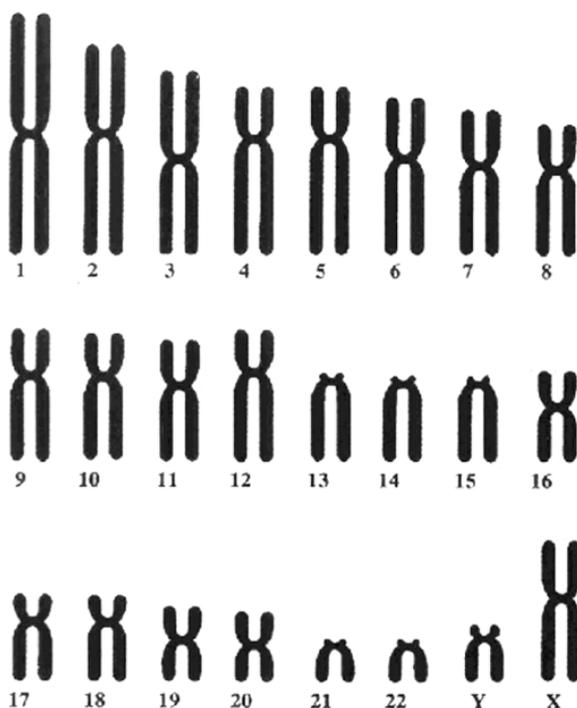
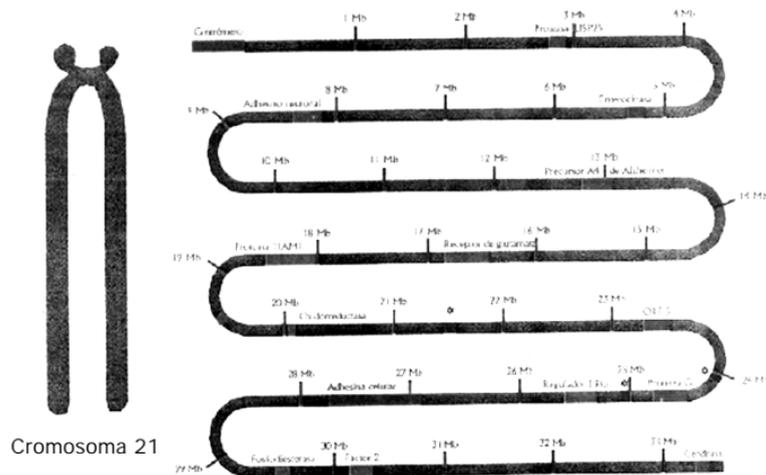


FIGURA 43. ESQUEMA DE LOS CROMOSOMAS HUMANOS

La figura muestra los 22 cromosomas somáticos o autosómicos y los dos cromosomas sexuales Y y X humanos. En ellos se localizan entre treinta y cuarenta mil genes y marcadores genéticos humanos.

Mapa genético del cromosoma humano número 21



Se presentan las posiciones aproximadas de algunos de los 255 genes localizados y reportados en este cromosoma. Este autosoma es el más pequeño de nuestros cromosomas y está compuesto por una molécula de DNA de aproximadamente 33.5×10^6 pares de bases. * Genes que codifican proteínas involucradas en el Síndrome de Down.

FIGURA 44. EL CROMOSOMA HUMANO # 21

Esquema del cromosoma humano # 21, cuya secuencia nucleotídica fue determinada en el año 2000. En la figura se muestra la posición de algunos de los 225 genes localizados en este cromosoma y entre ellos se señalan algunos de los involucrados en la enfermedad conocida con el nombre de Síndrome de Down, que ocurre cuando este cromosoma se encuentra presente en las células en tres copias, en vez de dos, que es lo normal.

ENFERMEDADES GENÉTICAS

Anemia falciforme
β-talasemias
Corea de Huntington
Diabetes
Distrofia de Duchenne
Enfermedad de Alzheimer
Fenilcetonuria
Fibrosis quística
Hemofilia
Hipoparatiroidismo
Neoplasia endócrina
Neurofibromatosis de von Recklinghausen
Renitis pigmentosa
Retinoblastoma
Síndrome de Lesch-Nyhan
Desórdenes inmunológicos

FIGURA 45. ENFERMEDADES GENÉTICAS

Listado que muestra algunas de las enfermedades genéticas más importantes. El elemento común en todas estas enfermedades es que en ellas ha habido mutaciones en uno o ambos de los dos genes que codifican para una proteína específica. Como resultado de este cambio, la proteína ya no es capaz de llevar a cabo sus funciones adecuadamente y se presenta la enfermedad en el individuo portador.

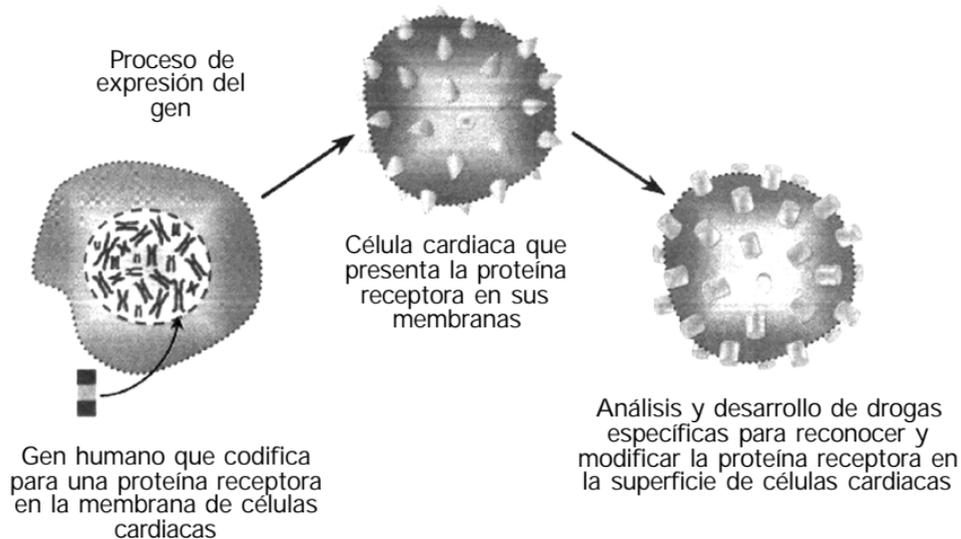


FIGURA 47. EL DISEÑO DE NUEVAS DROGAS; LA FARMACOGENÓMICA

La figura muestra un esquema mediante el cual es posible utilizar el gen que codifica para una proteína que funciona como receptora de algún ligando, en alguno de nuestros órganos (en este caso el corazón), para diseñar y desarrollar drogas específicas que se asocien y modifiquen la proteína receptora, cambiando su afinidad por el ligando. De esta manera es posible modular el funcionamiento de la proteína receptora.

Uso de genes humanos en medicina: terapia génica

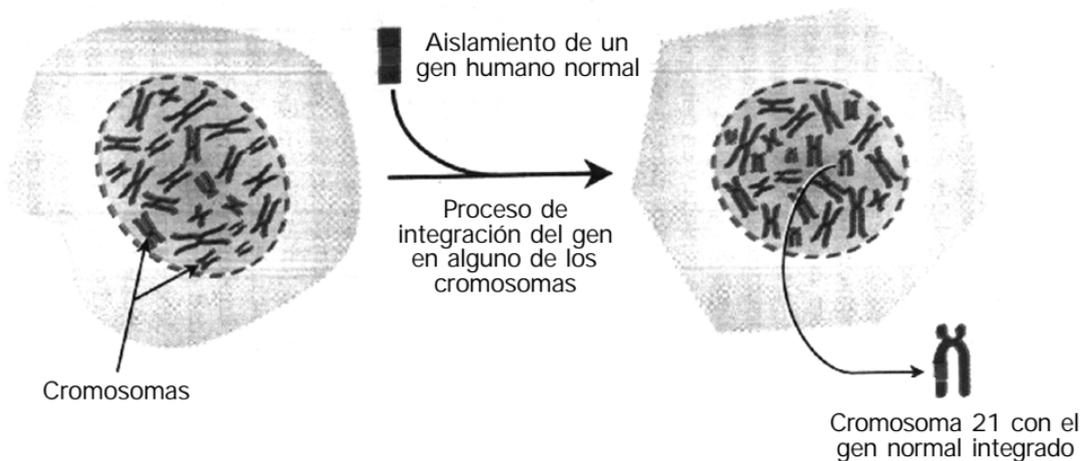


FIGURA 48. LA TERAPIA GÉNICA

En la figura se muestra cómo una célula que porta un gen mutante (en el lado izquierdo de la figura) puede incorporar la copia normal de este gen en alguno de sus cromosomas. Como resultado, esta célula se “transforma” en una célula que ahora tiene la capacidad de producir la proteína faltante. Esta estrategia se conoce como terapia génica.