

## El Proyecto del Genoma Humano y su Impacto en la Endocrinología \*

\* Conferencia dictada en la Jornada Anual de la Sociedad Chilena de Endocrinología y Metabolismo, La Serena, Chile 1 al 3 de noviembre de 2001.

Dr. Ronald Youlton R.  
Departamento de Pediatría, Clínica Las Condes

A mediados de los años ochenta un grupo de científicos consideró la posibilidad de analizar la secuencia del DNA humano y de localizar los genes, que en aquella época se estimaban en unos 100.000. El Proyecto del Genoma Humano fue oficialmente lanzado en 1990 como un esfuerzo de cooperación internacional (*International Human Genome Sequencing Consortium*), a 15 años plazo, para mapear nuestro genoma. A este esfuerzo se agregó posteriormente el de una empresa privada, *Celera Genomics*. En febrero de este año ambas entidades publicaron independientemente, un borrador con los resultados de la secuenciación del genoma (*Nature* 2001; 409: 860-921 y *Science* 2001; 291: 1304-1351). Ambas coinciden en que nuestro DNA consta de unos 3.200.000.000 pares de bases (3.200 Mbp), las que codifican para no menos de 26.000 pero no más de 38.000 genes (probablemente unos 33.000), cifra muy inferior a la considerada hace diez años atrás. Menos de 5% de nuestro DNA es codificante, lo que significa que los 33.000 genes estarían codificados por unos 160.000.000 de pares de base; el restante DNA es un mudo testigo de nuestra evolución.

Lo que se ha publicado este año es un borrador de la secuencia de los pares de bases que conforman nuestro DNA. Menos de la mitad del total de genes están identificados y de éstos, solo a una parte se le conoce su *locus* en los cromosomas. A modo de información, al 30 de julio de este año, la secuenciación se encontraba completada en un 47,1 %, en borrador en un 51,4% y no determinada en un 1,5%.

En forma paralela se ha estudiado el genoma de otras especies. La levadura *Saccharomyces Cerevisiae* tiene 6.034 genes, la mosca de la fruta *Drosophila Melanogaster* 13.061 genes, la pequeña lombriz *Caenorhabditis Elegans* 18.424 genes y la planta *Arabidopsis Thaliana*, 25.498 genes.

Se estima que 10% de los genes humanos tienen homología con genes de la mosca y del gusano y que la similitud del genoma humano con el del chimpancé alcanza al 99%, aunque la secuenciación del genoma de este último no está terminada. Se estima también que la similitud entre los distintos individuos de nuestra especie es superior al 99,9%.

Estas cifras son sorprendentes y echan por tierra nuestro concepto de que la complejidad biológica de una especie se relaciona con la cantidad de genes que posee. Surge entonces la pregunta: si tenemos algo menos que el doble de los genes del gusano ¿qué nos hace tan diferentes? La respuesta hay que buscarla en el producto de los genes, que son las proteínas.

Los factores de transcripción (FT) son proteínas que tienen la propiedad de unirse al

DNA y activar genes específicos, induciendo y regulando la transcripción y síntesis proteica. Se ha demostrado que existen familias y superfamilias de FT y que los miembros de cada familia aumentan de número en orden, desde la levadura al hombre. El genoma del gusano tiene unos 500 genes de FT, el de la mosca unos 700 y el del hombre, más de 2.000.

La complejidad biológica de las especies puede ser explicada entonces por el tamaño de la red de FT que posee, por la cantidad de genes que esta red regula y por la cantidad y diversidad de proteínas sintetizadas.

A su vez, las proteínas son mucho más complejas que el DNA. Un gen puede codificar para más de una proteína por procesamiento alternativo (splicing) del RNA. Un ejemplo es el de hormona de crecimiento (GH). Esta es una proteína de 191 aminoácidos cuyo peso molecular es 22 kd, cuyo gen estructural está en el cromosoma 17. Esta forma de GH constituye el 76% de la hormona circulante. Existe una variante de 20 kd que constituye el 16% de lo circulante y que se origina por procesamiento alternativo del RNA, de manera que los aminoácidos 32 al 46 del exon 2 quedan excluidos.

También hay que considerar que las proteínas pueden ser fosforiladas, glicosiladas, acetiladas, etc., cambios que pueden modificar sus funciones.

Más aún, algunas proteínas pueden ser subdivididas en fracciones menores después de sintetizadas, dando origen a polipéptidos que tienen acciones muy diversas. Tal es el caso de la proopiomelanocortina (POMC) cuyo gen estructural está en el cromosoma 2. Es una proteína de 241 aminoácidos que da origen a la adrenocorticotrofina (ACTH), la hormona melanofórica (MSH), la lipotrofina (LPH) y a la beta-endorfina.

También hay que tener en consideración que mutaciones diferentes en un mismo gen pueden generar proteínas con diferente nivel de actividad funcional, lo que puede manifestarse como fenotipos clínicos diferentes. Dos ejemplos pueden ilustrar este punto: las mutaciones en los genes CYP21 B (21 Hidroxilasa) y FGFR3 (Receptor 3 del Factor de Crecimiento de Fibroblastos).

Las mutaciones que ocurren en CYP21B provocan pérdida total o parcial de actividad de la enzima 21 Hidroxilasa, necesaria para la síntesis de cortisol en las glándulas suprarrenales; su déficit hace que los metabolitos previos al bloqueo se desvíen hacia una producción excesiva de andrógenos. Esta deficiencia es la causa más frecuente del síndrome de Hiperplasia Suprarrenal Congénita (>90% de los casos). Las mutaciones que anulan la actividad enzimática provocan las formas clásicas que se manifiestan en el período neonatal; aquellas que permiten algún grado de actividad residual causan las formas no clásicas de manifestación tardía. Estas últimas eran consideradas hace algunos años como formas adquiridas y no relacionadas con las formas clásicas.

El gen FGFR3, cuyo *locus* está en el cromosoma 4, tiene funciones relacionadas con el crecimiento. Ratones en los que se ha anulado (*knockout*) ambas copias del gen (ratón Fgfr3 *-/-*) tienen fémures muy largos, cuerpos vertebrales altos y una larga cola. Ello ha permitido concluir que la función de este gen es regular la osificación endocondral poniendo un freno al crecimiento. Mutaciones que inducen un aumento o ganancia de función del FGFR3 producen una frenación exagerada del crecimiento esquelético y causan la Acondroplasia. Otras mutaciones en este gen causan la Hipocondroplasia, una condición más leve y de expresión variable, así como también la Displasia Tanatofórica, que es un defecto letal.

Por todo lo anterior se estima que el Proteoma (conjunto de proteínas que produce una especie) es un orden de magnitud más complejo que el Genoma.

Las enfermedades genéticas se pueden separar en tres grandes categorías: cromosómicas, monogénicas o mendelianas y poligénicas o multifactoriales.

Las enfermedades cromosómicas se deben al exceso o ausencia de un cromosoma o parte de él. Ellas se observan en 65% de los abortos espontáneos del primer trimestre, en 6% de los mortinatos y en 0,6% de los recién nacidos vivos.

Los genes están contenidos en los cromosomas. Tenemos dos copias de cada uno (materna y paterna), con excepción de algunos genes del X y del Y que existen en copia única (ejemplos: receptor de andrógenos en el X, gen SRY determinante testicular en el Y).

A la fecha hay asignados solo 7.505 loci cromosómicos de los 33.000, pero su distribución no es homogénea. Así, el cromosoma 2 que es el segundo en tamaño, tiene asignados 462 loci y al cromosoma 19, que es menos de un tercio del tamaño que del anterior, se le han asignado 492. Una cantidad tan grande de genes, que pueden estar en exceso o déficit, explica las múltiples anomalías o la muerte embrionaria que ocurren en los casos de aberraciones cromosómicas.

Las enfermedades monogénicas son producidas por la mutación o delección de un gen. El libro *Mendelian Inheritance in Man* (MIM) editado por Victor McKusick en 1966 recopiló todas las enfermedades genéticas descritas en la época por los clínicos, los que se basaban en estudios familiares, así como también genes que no guardan relación con enfermedad, como grupos sanguíneos o variantes electroforéticas de enzimas. Este verdadero catálogo de enfermedades dominantes, recesivas, ligadas al X y al Y ha tenido sucesivas ediciones y está *on line* (OMIM), se actualiza diariamente y puede ser consultado libremente (<http://www.ncbi.nlm.nih.gov/entrez/omim>).

Su primera edición listaba 1.487 genes. Al día 3 de octubre de este año, OMIM registra 13.027 entradas, pero los genes relacionados con enfermedades suman 1.112 y las enfermedades mendelianas totalizan unas 1.500. La diferencia entre 1.112 y 1.500 se explica porque diferentes mutaciones en un mismo gen pueden producir diferentes enfermedades. Por ejemplo, mutaciones inactivantes y también algunas activantes del proto-oncogen RET causan la enfermedad de Hirschprung; otro tipo de mutaciones activantes de este gen causan el síndrome de neoplasia endocrina múltiple tipo 2 (MEN2).

Toda la tecnología que se ha desarrollado en la última década para el estudio del genoma humano, ha permitido la identificación y caracterización de genes cuya existencia era sugerida por evidencias indirectas.

Hace casi veinte años atrás tuvimos la oportunidad de estudiar una paciente que consultó por haber tenido abortos espontáneos a repetición. Era de baja estatura, con acortamiento mesomérico de las extremidades e incurvación distal de cúbito y radio. Radiológicamente tenía una deformidad de Madelung a nivel de ambas muñecas. Este conjunto de anomalías esqueléticas recibe el nombre de Discondrosteosis de Leri-Weill.

El estudio citogenético de la paciente demostró una translocación de la región heterocromática del cromosoma Y (que no transcribe) en la región distal del brazo corto de uno de sus cromosomas X, lo que debió ocurrir durante la meiosis paterna. Esta translocación era, con toda probabilidad, la responsable de los abortos al inducir errores en la meiosis y anomalías cromosómicas en los óvulos. En la búsqueda de la literatura nos encontramos con dos casos con características físicas similares. Publicamos este caso en la Revista Médica de Chile en 1985. La frase final de la discusión terminaba así: "... que la región Xp22pter codifique genes involucrados en el crecimiento y modelaje del esqueleto".

Este gen putativo fue reconocido en MIM en la edición siguiente con carácter de hipotético.

La región distal de los cromosomas X e Y está constituida por una secuencia de 2,6 Mb, idéntica en ambos. Estas regiones se recombinan durante la meiosis masculina, por lo que se las denomina regiones pseudoautosómicas (PAR1: *Pseudo Autosomic Region*). Los genes de esta región en el X escapan al proceso de inactivación en la mujer, de manera que ambas copias son activas en ambos sexos.

En 1997, Rao y cols identificaron y clonaron un nuevo gen en la región distal del brazo corto de los cromosomas X e Y al que denominaron SHOX (*Short stature Homeobox containing gene*).

El gen SHOX está contenido dentro de PAR1 y tiene una extensión de 170 Kb; transcribe dos proteínas, una de 292 aminoácidos y otra de 225, por procesamiento alternativo del RNA y se expresa en el tejido osteogénico y en el primer y segundo arcos faríngeos. Es un gen que contiene Homeobox, una secuencia de DNA de 180 bp altamente conservada, desde los invertebrados al hombre. Las proteínas Homeobox funcionan como FT, regulando la expresión de otros genes fundamentales en el desarrollo.

Los pacientes que tienen solo una copia de SHOX, como ocurre en el síndrome de Turner (45,X), en los casos que tienen una delección de la región terminal de Xp o Yp (como en nuestra paciente) o que tienen mutaciones intragénicas, son de baja estatura y presentan una o más anomalías esqueléticas, tales como mesomelia, cuartos metacarpianos cortos, cúbito valgo, deformidad de Madelung, genu valgum o extremidades inferiores relativamente cortas. También se han encontrado delecciones de SHOX en algunos pacientes diagnosticados como portadores de talla baja idiopática, sin anomalías esqueléticas. Pacientes que tienen tres copias de SHOX, como ocurre en el síndrome de Klinefelter (47,XXY), en los hombres XYY, en las mujeres 47,XXX y en otros raros casos de anomalías estructurales de uno de los cromosomas sexuales, tienen mayor estatura y metacarpianos y extremidades inferiores más largas.

Estas observaciones permiten concluir que SHOX funciona como un represor de la maduración del cartílago de crecimiento y de la fusión epifisiaria. Su deficiencia permite una maduración esquelética acelerada, que se manifiesta muy claramente durante la pubertad, particularmente en los segmentos distales de las extremidades, como se observa en la Discondrosteosis y en el síndrome de Turner; en el caso de este último se asocian los efectos del desbalance cromosómico. Las enfermedades comunes de los adultos (enfermedad coronaria, hipertensión arterial, algunos desórdenes psiquiátricos), así como las anomalías congénitas más frecuentes (labio leporino, displasia de caderas) y la mayor parte de las diferencias fenotípicas normales (estatura, peso, presión arterial) son de determinación multifactorial, es decir, producto de la interacción de diversos factores genéticos y no genéticos (medio ambiente, estilo de vida). La predisposición o susceptibilidad para tener una característica o desarrollar una enfermedad se hereda de ambos progenitores, pero su herencia no tiene los patrones mendelianos.

Si para las enfermedades monogénicas ha sido difícil individualizar el gen y ubicarlo en el mapa físico de los cromosomas, para las enfermedades o características de determinación multifactorial esta tarea lo será más.

Una manera de acercarse al problema es relacionar estas condiciones con polimorfismos del DNA, de los cuales hay varios tipos: RFLP (restriction fragment length polymorphisms), VNTR (variable number of tandem repeat.s), entre otros.

En el estudio de la secuenciación de nuestro genoma se ha encontrado que cada 1000 a 2000 bp hay un nucleótido que es distinto y que puede estar dentro de la secuencia de un gen, contiguo a él o en regiones no codificantes. Esta característica, que es hereditaria, se denomina polimorfismo de un nucleótido o SNP (single nucleotide polymorphism). Los SNP están presentes en todo nuestro genoma, se han identificado más de 1.400.000 de ellos y se los considera la fuente de variación entre los individuos de una misma especie, a la vez que pueden explicar la susceptibilidad a (o el efecto protector contra) diversas enfermedades.

Todo el nuevo conocimiento aportado por el estudio de nuestro genoma y por las técnicas moleculares van a introducir cambios significativos en nuestro modo de comprender, clasificar y tratar las enfermedades. Estos cambios se van a notar en diversos aspectos.

El diagnóstico molecular es posible en la actualidad para cerca de un centenar de enfermedades, número que está creciendo en forma exponencial. En nuestro medio, el análisis de mutaciones en casos de hiperplasia suprarrenal congénita, de neoplasia endocrina múltiple tipo 2 o de fibrosis quística son los más frecuentes.

El conocimiento preciso de la base genética de una enfermedad nos permitirá saber el

comportamiento biológico de ella.

Se ha desarrollado una nueva tecnología, la de las llamadas micro matrices de DNA (DNA *microarrays*). Consisten en una colección de cientos o miles de secuencias ordenadas en una fase sólida. De la muestra que se desea estudiar se extrae el RNA mensajero y a partir de éste, se sintetizan copias de DNA complementario utilizando nucleótidos que se marcan con fluorescencia; estos cDNA se hibridan en la micro matriz. La intensidad de la fluorescencia en cada punto de hibridación indica la relativa abundancia de los distintos mRNA. Ello permite el análisis de un gran número de genes en forma simultánea, generando patrones de expresión génica que pueden ser característicos de una enfermedad. Estas micro matrices también pueden ser utilizadas para la pesquisa de polimorfismos y mutaciones.

Una reciente publicación señala que el análisis de expresión génica permite distinguir las leucemias mieloide aguda y linfoblástica aguda, cuyos tratamientos son muy diferentes. De 6.800 genes estudiados, 1.100 se expresaron en forma diferente en ambas leucemias. De éstos hubo 50 genes que consistentemente permitieron diferenciar ambas leucemias con 85% de sensibilidad y 100% de especificidad. De esta forma, el médico y el paciente podrán optar por el tratamiento más apropiado para la enfermedad y para el enfermo.

El conocer la predisposición a desarrollar una enfermedad hará que las personas modifiquen algunos hábitos de vida o se sometan a procedimientos diagnósticos frecuentes y regulares o a tratamientos preventivos. Ejemplos son algunos cánceres de colon o mama o la neoplasia endocrina múltiple tipo 2. En la actualidad se puede prevenir o atenuar la masculinización de los fetos de sexo femenino afectados de hiperplasia suprarrenal congénita suministrando a la madre dexametasona a partir de la sexta semana de gestación; con ello se logra frenar las suprarrenales del feto, evitando la sobreproducción de andrógenos.

El riesgo de contar con información que indique la susceptibilidad o predisposición a patologías específicas, es que algunas personas podrían pasar su vida esperando una enfermedad que quizás nunca llegará, lo que podríamos llamar hipocondría genética.

La pesquisa neonatal de algunas enfermedades y su oportuno tratamiento permite que los afectados no desarrollen los síntomas de la enfermedad (prevención secundaria). En la actualidad ésta se hace para hipotiroidismo y fenilketonuria, pero en el futuro se extenderá a otras condiciones.

Los heterocigotos de mutaciones de enfermedades recesivas nunca enfermarán, pero pueden tener, con otro heterocigoto, descendencia afectada, a veces de enfermedades letales, como la enfermedad de Tay-Sachs. En ésta y otras enfermedades es posible el diagnóstico molecular a los portadores, de tal forma que ellos tomen sus decisiones reproductivas de manera informada.

La terapia génica consiste en reemplazar un gen mutado por uno funcional, lo que todavía puede considerarse como en etapa experimental.

La terapia basada en las características genéticas del paciente (farmacogenética) permitirá el uso racional de los medicamentos. La investigación de los polimorfismos de los citocromos P450 2C9 y 2D6 de una persona, que se relacionan con la velocidad de metabolización hepática de las drogas, hará posible que las dosis sean dadas a la medida del paciente, a la vez que se podrá predecir si éste está en riesgo de desarrollar alguna reacción adversa.

El día que se conozca cómo funcionan los genes y cuáles son los trastornos moleculares que ocurren en la enfermedad, se podrán diseñar drogas específicas para ese problema.

En otras circunstancias, en vez de sustituir el gen, se podrá sustituir la proteína (biosintética) que codifica o suministrar alguna molécula que interactúe e inhiba una proteína nociva. Un buen ejemplo es el de la leucemia mieloide crónica. En ella, la translocación cromosómica 9/22 (cromosoma Philadelphia) genera un gen quimérico y

una proteína anormal que induce una descontrolada multiplicación de la serie blanca. Recientemente se ha desarrollado una droga (STI 271) que se adhiere a esta proteína y bloquea su acción. Los ensayos clínicos han mostrado una clara mejoría y sin todos los efectos secundarios de la quimioterapia convencional.

Después de este análisis parcial y relativamente superficial del impacto del Proyecto del Genoma Humano sobre la medicina, podría quedar la sensación que la genética lo será todo, cayendo en el reduccionismo y en el determinismo. Nosotros somos el producto de nuestros genes y del ambiente en que fuimos concebidos, nacimos, nos criaron y de las circunstancias de la vida que nos ha tocado vivir. Como se ha dicho, los genes de Mozart por sí solos, no hacen el genio de Mozart. Finalmente es necesario recordar que las principales causas de enfermedad y muerte del homo sapiens son el hambre, las malas condiciones sanitarias, el consumo de tabaco, alcohol y drogas, los accidentes, la guerra, el sedentarismo y la polución ambiental.