

CÓMO IDENTIFICAR GENES QUE PROPICIAN UNA ENFERMEDAD, UTILIZANDO EL PROYECTO GENOMA HUMANO

Carlos Montoya Juárez*
Ivonne Guadalupe Fernández Carrasco*

RESUMEN

El proyecto genoma humano ha permitido conocer la secuencia de los 3 millardos de pares de bases del DNA humano, se presenta una descripción general del proyecto genoma humano, cómo se está desarrollando y las implicaciones que surgen con este nuevo conocimiento, las tendencias en medicina para la identificación de enfermedades genéticamente transmitidas, técnicas recientemente desarrolladas las cuales no serían posibles sin la información del genoma y las perspectivas en medicina como la terapia génica.

Palabras Clave: *Genoma humano, terapia génica, bioética.*

ABSTRACT

The human genome project has allowed to know the sequence the 3 billions of pairs of bases of the human DNA, in this article appears a general description of the human genome project, how this has being developed and the implications that arise with this new knowledge, the medicine tendencies for the identification of genetic diseases, approaches recently developed which will not be possible without the information of the genome and the medicine perspective like the genetic therapy.

Key Words: *Humangenome, genetic therapy, bioetical.*

ARTÍCULO RECIBIDO EL 29 DE MAYO DEL 2001 Y ACEPTADO EL 09 DE JULIO DEL 2001.

PROYECTO DEL GENOMA HUMANO

En febrero del año 2000 se publicó el número 6822 de la famosa revista Nature¹ en la cual el tema principal es el genoma humano, en este número se presentan los resultados de una titánica labor desarrollada por dos equipos:

- Consorcio internacional para la secuenciación del genoma humano.
- Celera Genomics (www.celera.com),

Los artículos de esta revista nos abren las puertas del futuro y las dejan abiertas para que lleguemos a una serie de nuevas tecnologías que pocos habían imaginado.

El Genoma Humano es un proyecto internacional, que inició en el año 1988. El objetivo principal es conocer la secuencia completa del genoma humano².

La idea de iniciar un estudio coordinado del genoma humano surgió de una serie de conferencias científicas celebradas entre 1985 y 1987.

Los objetivos del Proyecto son:

- Identificar los aproximadamente 100000 genes (actualmente se considera que sólo hay cerca de 30000) en el Ácido desoxirribonucleico (DNA) humano.
- Determinar la secuencia de 3 millardos de bases químicas que conforman el DNA.
- Acumular la información en bases de datos.
- Desarrollar de modo rápido y eficiente tecnologías de secuenciación.
- Desarrollar herramientas para análisis de datos.
- Dirigir las cuestiones éticas, legales y sociales que se derivan del proyecto.

¿QUÉ ES EL GENOMA?

Se llama genoma a la totalidad del material genético de un organismo. Se consideraba que el genoma humano tendría entre 50000 y 100000 genes distribuidos entre los 23 pares de cromosomas de la célula somática humana.

Cada cromosoma puede contener más de 250 millones de pares de bases de DNA, y se estima que la totalidad del genoma humano tiene 3000 millones de pares de bases.

* Unidad de Investigación Médica en Bioquímica, Hospital de Especialidades, Centro Médico Nacional SXXI, IMSS.

El genoma es el conjunto de instrucciones completas para construir un organismo. El genoma contiene datos que la célula emplea para el diseño de sus estructuras y el desempeño de sus actividades y las del organismo del cual forma parte la célula. Se localiza en el núcleo de las células. Consiste en cadenas de DNA estrechamente enrolladas en moléculas de proteína asociadas, organizadas en estructuras llamadas cromosomas. Si desenrollamos las cadenas y las adosamos medirían más de 1.5 metros, sin embargo su ancho sería ínfimo, cerca de 50 trillonesimos de pulgada.

La molécula de DNA consiste de dos cadenas enrolladas helicoidalmente, una alrededor de la otra como escaleras que giran sobre un eje, cuyos lados hechos de azúcar y moléculas de fosfato se conectan por uniones de nitrógeno llamadas bases³.

Cada cadena es un acomodamiento lineal de unidades similares repetidas llamadas nucleótidos, los que se componen de un azúcar, un fosfato y una base nitrogenada. La molécula de DNA contiene cuatro bases: Adenina (A), Timina (T), Citosina (C), Guanina (G).

El orden de estas bases es llamada secuencia de DNA, la cual especifica la instrucción genética requerida para crear un organismo particular con características que le son propias. La adenina y la guanina son bases púricas, en cambio la citosina y la timina son bases pirimídicas.

Las dos cadenas de DNA son mantenidas juntas por uniones entre bases que forman los pares de bases. El tamaño del genoma es usualmente expresado en el total de pares de bases (pb). En la especie humana, contiene aproximadamente 3 millardos de pares de bases. Otros organismos estudiados con motivo de este estudio fueron la bacteria *Escherichia coli*, la mosca de la fruta *Drosophila melanogaster*, y las ratas de laboratorio *Ratus noruegans*.

Cada vez que la célula se divide en células hijas, el genoma total se duplica. Durante la división, el DNA se desenrolla y rompe las uniones entre pares de base permitiendo a las cadenas separarse. Cada cadena dirige la síntesis de una nueva cadena complementaria con nucleótidos libres que coinciden con sus bases complementarias de cada cadena separada. Cada célula hija recibe una cadena vieja y una nueva. Cada molécula de DNA contiene muchos genes, la base física y funcional de la herencia. Un gen es una secuencia específica de nucleótidos o bases, los cuales llevan la información requerida para la construcción de proteínas que proveerán de los componentes estructurales a las células y tejidos como también a las enzimas para una esencial reacción bioquímica.

Sólo el 10% del genoma incluye la secuencia de codificación proteica de los genes. Entre mezclado con muchos genes hay secuencias sin función de codificación, de función desconocida hasta el momento.

EL núcleo de muchas células humanas contiene dos tipos de cromosomas, uno por cada padre. Cada conjunto, tiene 23 cromosomas simples, 22 de tipo autosómico y uno que puede ser X o Y que es el cromosoma sexual. Una mujer normal tendrá un par de cromosomas X (XX), y un hombre normal tendrá un cromosoma X y otro Y (XY).

Los cromosomas pueden ser observados mediante microscopio óptico y cuando son teñidos revelan patrones de luz y bandas oscuras con variaciones regionales. Las diferencias en tamaño y de patrón de bandas permiten que se distingan los 24 cromosomas uno de otro, el análisis se llama cariotipo.

Las anomalías cromosómicas mayores incluyen la pérdida o copias extra, o pérdidas importantes, fusiones, translocaciones detectables microscópicamente. Así, en el Síndrome de Down se detecta una copia adicional del par 21 o trisomía 21.

Otros cambios son tan sutiles que sólo pueden ser detectados por análisis molecular, se llaman mutaciones. Muchas mutaciones están involucradas en enfermedades como la fibrosis quística, anemia de células falciformes, predisposiciones a ciertos cánceres, o a enfermedades psiquiátricas mayores, entre otras.

Toda persona posee en sus cromosomas una copia de cada gen paterno y su correspondiente gen materno. Cuando ese par de genes materno-paterno son iguales, se dice que el individuo es homocigótico para tal rasgo, por el contrario se dice que es heterocigótico. Ejemplo: un gen que transmita el rasgo hereditario del color de ojos verde y el otro el color de ojos marrón. Se trata de heterocigotos para el rasgo color de ojos. Si a su vez, uno de esos genes domina en la expresión del rasgo al otro gen enfrentado, se dice que es un gen heredado dominante, de lo contrario se dice que es recesivo.

FLUJO DE INFORMACIÓN GENOMA-PROTEÍNA-ACCIÓN

Las instrucciones de los genes son transmitidas a través del ARN mensajero (RNAm), el cual es un intermediario transitorio. Para que la información de un gen sea expresada, un segmento de DNA que contiene un gen es copiado a un RNA. Este RNAm, se mueve desde el núcleo hasta el citoplasma celular, donde sirve como plantilla para la síntesis proteica.

La maquinaria celular que sintetiza proteínas traduce los códigos en cadenas de aminoácidos que constituyen la proteína molecular. En el laboratorio se puede aislar el RNAm y ser utilizado como plantilla para sintetizar un DNA complementario (DNAc), el cual puede ser usado para ubicar los genes correspondientes en el mapa cromosómico.

El mapa del genoma humano es una herramienta genética que permite estudiar la evolución del hombre y que cambiará drásticamente la medicina actual tal como la conocemos. Permitirá el tratamiento de enfermedades hasta ahora sin cura.

DESPUÉS DE CONOCERSE EL GENOMA HUMANO

El conocimiento del genoma permitirá que se creen nuevas drogas terapéuticas que desplazarán a las anteriores. Las nuevas drogas prometen tener menores efectos colaterales que las actuales.

Se podrá informar a una persona, que puede comer alimentos grasos porque carece de predisposición genética a la obesidad y a enfermedades cardíacas, pero que debe huir del alcohol porque es genéticamente propenso al alcoholismo.

Dentro de los llamados beneficios del Proyecto figuran a nivel de medicina molecular, la posibilidad de mejorar el diagnóstico de enfermedades, detección temprana de predisposiciones genéticas a ciertas enfermedades, el diseño racional de drogas, terapia génica, sistemas de control para drogas y farmacogenomas.

Se ha estudiado un gen que determina la producción de la proteína llamada SPARC, la que normalmente impide al organismo atacar y anular células cancerígenas. La terapia génica en estos casos actúa permitiendo que las células cancerosas sean atacadas por el organismo.

El estudio del polimorfismo del HLA permitirá estudiar la evolución y migración humana y tiene su utilidad en las mutaciones de linaje, migraciones de diferentes grupos poblacionales basados en el DNA mitocondrial, mutaciones del cromosoma, además de comparar los cambios evolutivos con eventos históricos.

También es útil en la identificación forense, para potenciales sospechosos en los cuales el DNA puede conducir a liberar a personas que fueran acusadas de crímenes injustamente, para identificar víctimas de catástrofes, paternidad y otras relaciones familiares, identificar y proteger especies en peligro, detectar bacterias que pueden contaminar agua, aire, alimentos, determinar compatibilidad de órganos donantes en programas de trasplante, determinar el pedigree en ganados y para autenticar productos de consumo como caviar, vinos.

En agricultura, ganadería y bioprocesamientos, se utiliza para mejorar la resistencia de cultivos ante insectos, sequías, para hacerlos más productivos y saludables, igualmente para producir animales más saludables y nutritivos, elaborar biopesticidas, vacunas comestibles y nueva limpieza del medio ambiente de plantas como tabaco.

El entender el genoma humano se considera tan importante tanto para la práctica clínica como para la investigación médica y se están diseñando cursos o módulos especiales para capacitarse en esta área, los cursos incluirían casos clínicos de los cuales se requiera acceder información en el Genbank, Online Mendelian Inheritance in Man (OMIM), y PubMed (bases de datos del Instituto Nacional de Salud NIH, donde se almacena y está a disposición del público la información

obtenida del genoma humano) y desarrollar procedimientos para el tratamiento⁴.

ESTUDIO DEL GENOMA HUMANO EN MÉXICO

En 1993 se celebró en México una reunión del Programa Latinoamericano del Genoma Humano (<http://www.unam.mx/genoma/>), en la que participaron muchos genetistas mexicanos y de otras naciones de Latinoamérica, así como algunos invitados norteamericanos y europeos. Fue un primer intento de información y comunicación.

Un año después, el 7 de Junio de 1994 la Universidad Nacional Autónoma de México (UNAM) llevó a cabo un taller sobre oportunidades y problemas de la investigación sobre el Genoma Humano, en el que participaron investigadores universitarios y profesores norteamericanos invitados. El 28 de noviembre de 1994 el Coordinador de la Investigación Científica de la UNAM, junto con un grupo de Directores de Facultades, Escuelas e Institutos de Investigación de la Universidad, autorizaron el inicio de labores de planeación para la participación organizada de la institución en este proyecto, que quedaron ubicadas en el Programa Universitario de Investigación en Salud.

Como resultado de este acuerdo se han llevado a cabo varias actividades, entre las que destacan:

La creación de este sitio en la World Wide Web (que incluye un inventario de los proyectos de investigación sobre genética molecular en la Universidad y un servicio de consulta de las principales revistas en este campo).

La elaboración de un proyecto multicolaborativo de investigación sobre el genoma humano, que pondrá las bases para el desarrollo de infraestructura y de nuevos proyectos a nivel nacional.

El objetivo de este proyecto de investigación es adquirir la capacidad para entender una enfermedad o un proceso biológico por medio de la identificación, a través del mapeo en el genoma, de los genes responsables. Se eligió, por consenso, a la diabetes mellitus no dependiente de insulina, tipo MODY ("maturity-onset diabetes of the young"), una enfermedad que por su frecuencia y sus peculiaridades en la población mexicana, así como por tener en estudio un número importante de familias multigeneracionales, constituye un modelo adecuado para desarrollar las habilidades, experiencia e infraestructura para efectuar investigación genómica de alta calidad. Por medio de esta investigación multicolaborativa, en la que participarán numerosas dependencias de la UNAM y a la que esperamos que se vayan sumando otras instituciones nacionales de educación superior y salud, a través del mapeo genético se buscarán los principales genes responsables de la diabetes MODY en la población mexicana.

TÉCNICAS USADAS EN LA INVESTIGACIÓN DE DEFECTOS GENÉTICOS

La investigación y la implementación de pruebas genéticas

lograron en 1970 una importante técnica para cartografiar los genes humanos o cariotipos. En el Instituto Karolinska de Suecia se descubrió un método para teñir los cromosomas humanos con colores fluorescentes, los que al ser iluminados con luz ultra violeta se hacen visibles como bastones a franjas claras y oscuras. Estos cariotipos son un instrumento muy útil para el diagnóstico de anomalías.

Para realizar una prueba en una persona adulta es suficiente con una sola gota de sangre, dado que el DNA se puede extraer de los leucocitos. También se puede extraer de las muestras de semen, algunos métodos permiten obtenerlo de la saliva e incluso, del cabello⁵.

Algunas de las enfermedades de las cuales ya existen pruebas disponibles son:

- Hemofilia (defecto en el control de las hemorragias).
- Fibrosis quística (acumulación de mucosidades en los pulmones, interfiere en la respiración).
- Mal de Alzheimer (enfermedad degenerativa neurológica marcada por una senilidad precoz).
- Anemia falciforme (anemia crónica hereditaria).

Esta técnica identifica la ausencia, duplicación o modificación de un gran segmento de DNA, y se puede observar a nivel microscópico.

EL FUTURO: TERAPIA GÉNICA

Consiste en la aportación de un gen funcional a las células que carecen de esta función, con el fin de corregir una alteración genética o enfermedad adquirida. La terapia génica se divide en dos categorías:

- La primera es la alteración de las células germinales lo que origina un cambio permanente de todo el organismo y generaciones posteriores. Esta terapia génica en la línea germinal no se considera en los seres humanos por razones éticas.
- El segundo tipo de terapia génica, terapia somática celular, es análoga a un trasplante de órganos. En este caso uno o más tejidos específicos son objetos, mediante tratamiento directo o extirpación del tejido, de la adición de un gen o genes terapéuticos en el laboratorio, junto a la reposición de las células tratadas en el paciente. Se han iniciado diversos ensayos clínicos de terapia genética somática celular, destinados al tratamiento de cánceres o enfermedades sanguíneas, hepáticas o pulmonares. Se realizan estudios para agregar un gen faltante o que supla la función del gen defectuoso, estos genes serían aplicados por vía intravenosa o inclusive oral. Sin embargo, estas técnicas están optimizándose y han sufrido no pocos descabros, se han tenido dificultades para encontrar un medio de transporte adecuado y problemas con el rechazo del DNA administrado.

CONSIDERACIONES ÉTICAS

El desarrollo científico, en lo que respecta al Proyecto Genoma Humano, abre las puertas a un sinnúmero de tratamientos que podrían ser beneficiosos para el hombre. Pero no se debe olvidar que esto implica manipular directamente los mecanismos que transmiten la vida y dirigen la evolución de las especies, incluyendo la nuestra.

Estos hechos desbordan por mucho nuestros conceptos de ética y humanidad, ya que nunca nos vimos enfrentados a la posibilidad de que la vida fuera manipulada de este modo.

Así surgen preguntas como: ¿Se debe prohibir o desaconsejar algún tipo de manipulación genética?, ¿A quién le corresponde la responsabilidad de discriminar entre lo permitido o no? Así, la UNESCO se compromete a promover y desarrollar la reflexión ética en los avances científicos en las áreas de la biología y la genética, proclamando los siguientes principios aprobando la declaración de éstos.

A. La dignidad humana y el genoma humano. Se refiere a la igualdad y dignidad de los individuos, cualesquiera que sean sus características genéticas; negando así la discriminación por características genéticas.

B. Derechos de las personas interesadas Se refiere a que toda investigación genética debe estar de acuerdo al país respectivo, y siempre con la previa información y aprobación del individuo. Si éste no está en condiciones de aprobarlo, sólo se llevará a cabo la investigación si ésta es indispensable para la salud del individuo.

C. Investigaciones sobre el genoma humano. Se refiere a que ninguna investigación podrá ir más allá de los derechos y dignidad humanas, y que todas las personas deben tener a su alcance los progresos biológicos y genéticos. A su vez éstas investigaciones deben estar orientadas a aliviar los males de la humanidad.

D. Condiciones de ejercicio de la actividad científica. Deben imponerse en los científicos responsabilidades especiales tanto en sus investigaciones como en los resultados de éstas. Los estados fijarán el marco de libre ejercicio de la investigación sobre el genoma humano, y éstos formarán comités que apreciarán los puntos éticos y jurídicos sobre estas investigaciones.

E. Solidaridad y cooperación internacional. Los estados deben promover investigaciones que prevengan y traten enfermedades genéticas o endémicas. Deberán fomentar la difusión internacional sobre esta investigación.

F. Fomento de los principios de la declaración. Se deberán fomentar estos principios a través de la educación y otros medios. Los estados garantizarán el respeto de éstos.

DIAGNÓSTICO MOLECULAR: DETECCIÓN DE MUTACIONES

El avance en el Proyecto genoma humano, en los mapas de secuencias ya disponibles impulsan la necesidad de disponer de técnicas capaces de rastrear DNA en busca de mutaciones asociadas con enfermedades humanas. En los próximos años se habrá identificado la mayor parte de los genes implicados en importantes enfermedades humanas. En las bases de datos se están introduciendo informaciones sobre mutaciones y sus implicaciones clínicas. Uno de los retos de la medicina será hacer uso de esa información para mejorar el diagnóstico y pronóstico de las enfermedades.

En la detección de mutaciones se ajusta perfectamente la expresión de “buscar una aguja en un pajar” puesto que se debe identificar un cambio en una base entre los 300 millones de ellas que puede ser la responsable de cambiar: el codón, el aminoácido, la proteína y hasta la acción de ésta. Por ejemplo, un cambio de un simple nucleótido Timina por Adenina en el primer exón de la β -globulina cambia el codón original de CTG a CAG, el ribosoma al sintetizar la proteína coloca una glutamina en lugar de una leucina, la proteína al tomar estructura cuaternaria, no adquiere la forma correcta de la proteína β -globulina y ésta se polimeriza, los eritrocitos producidos con esta β -globulina defectuosa son alargados y curvos en lugar de los convexos normales. Los eritrocitos defectuosos se agregan en la sangre venosa con baja tensión de oxígeno y una gran variedad de tejidos puede ser privado de oxígeno y sufrir daño, enfermedad llamada anemia falciforme, todo por un simple cambio en un nucleótido.

Para detectar mutaciones en sitios previamente determinados se han diseñado una serie de métodos que en general son la amplificación de un segmento específico de algunas centenas de pares de bases del DNA por medio de la técnica de reacción en cadena de la polimerasa (PCR) y la posterior identificación de los segmentos de DNA amplificados que portan la mutación. El método más común es el SSCP (Single strand conformational polymorphism) del cual aparecen cada día variantes para mejorar la separación de cadenas mutadas.

La detección de mutaciones en lugares sin especificar o cuando no se dispone de información de secuencia es una tarea muy ardua, aunque se han desarrollado algunas estrategias.

Sin embargo, lo más frecuente es la situación en que se trata de identificar mutaciones en múltiples sitios potenciales dentro de los segmentos de DNA de los que se conoce la secuencia normal.

De modo ideal, el procedimiento tendría que ser fiable, rápido, barato, y proporcionar información exacta sobre la posición y naturaleza de la mutación, así como requerir poco esfuerzo, ser automatizable y no usar reactivos peligrosos. Ninguno de los métodos actuales cumple este criterio. La secuenciación sería el método definitivo, pero sigue requiriendo mucho trabajo, y los métodos automatizados de alto rendimiento son caros.

Se está desarrollando un sistema que no usa electroforesis en gel (una de las fases más tediosas), llamado espectrometría de masas por láser de desorción/ionización asistida por matriz.

La división Applied Biosystems de Perkin Elmer (<http://home.appliedbiosystems.com/>) ha diseñado un análisis de hasta 32 mutaciones diferentes de un gen por amplificación simultánea de hasta 15 segmentos de DNA, seguida por ligado multiplex de oligonucleótidos específicos de alelos.

La misma empresa ha creado un ensayo de micro-PCR que puede realizar 24 ensayos de 2 μ cada uno en un chip, y cuyos resultados se pueden leer en tiempo real por medio de la estrategia de TaqMan.

Otro de los avances más prometedores son los “chips” de oligonucleótidos: arreglos de diferentes oligos unidos a soportes semisólidos. DNA marcado fluorescentemente puede entonces hibridarse a los oligos del chip, de modo que el patrón de hibridación se puede analizar por barrido de fluorescencia. Como se conoce la secuencia y posición de cada oligo en el chip, se puede determinar la posición y posiblemente la naturaleza de cada mutación.

Este enfoque de microchips con oligos ha sido llevado a un avance importante por la empresa Affymetrix (<http://www.affymetrix.com/>) usando técnicas fotolitográficas similares a las que emplea la fabricación de chips de silicio, se han podido obtener chips con 64.000 diferentes sondas de oligonucleótidos unidos a superficies planas de cristal. En 15 minutos el dispositivo puede detectar la hibridación de secuencias diana marcadas fluorescentemente a oligos del chip.

El Proyecto genoma humano hace uso de dos tipos de mapas para caracterizar el genoma: mapas genéticos de ligamiento, y mapas físicos.

MAPAS GENÉTICOS DE LIGAMIENTO

Los mapas genéticos se basan en el cálculo de la frecuencia a la que se co-heredan formas alternativas (alelos) de dos loci genéticos que están ligados formando parte de un mismo cromosoma.

A finales de los años 70 se recurre al análisis molecular de zonas de DNA no codificadoras y que son muy polimórficas: existen varios tipos de secuencias (algunas de ellas de naturaleza repetitiva, como los VNTR, los microsátelites, etc.). Entre las ventajas de los microsátelites se cuentan: contenido informativo muy alto, con lo que los análisis estadísticos mejoran en fiabilidad; distribución abundante y relativamente uniforme por todo el genoma; y que se pueden ensayar fácilmente mediante PCR. Además, estos loci genéticos sirven en genética clínica como marcadores útiles para localizar genes relacionados con enfermedades. Los polimorfismos moleculares han permitido que en la actualidad el Proyecto genoma humano haya generado detallados mapas genéticos

del genoma humano a un nivel de resolución en torno a 1 centimorgan (cM) o incluso menos.

MAPAS FÍSICOS E INTEGRACIÓN DE LOS MAPAS

Los mapas físicos especifican distancias físicas mensurables en pares de bases (pb). El mapa físico de mayor detalle es la propia secuencia del genoma, pero antes de llegar a obtenerla, hay que elaborar mapas físicos partiendo de resoluciones bajas y avanzando hacia las resoluciones cada vez mayores. Los mapas físicos de menor resolución son los propios cariotipos: la visualización microscópica de la dotación cromosómica haploide humana teñida con colorante de Giemsa nos muestra un patrón alternante de bandas claras y oscuras, en el que cada banda tiene una media de unos 7 millones de pares de bases. Actualmente existen novedosas herramientas de citogenética molecular (como las sondas fluorescentes *in situ* o FISH, la “pintura de cromosomas”, etc.) que permiten un mayor detalle y que, unidas a otras técnicas aumentan el arsenal de enfoques para el estudio de los genomas, de su dinámica y de sus alteraciones.

Los mapas físicos de mayor resolución se suelen elaborar a partir de bibliotecas de genes en las que el genoma a estudiar se encuentra fragmentado en multitud de trozos aleatorios y desordenados, cada uno de ellos clonado por separado en un vector adecuado: plásmido, cósmido, cromosomas artificiales de levadura (YAC), cromosomas artificiales de bacteria (BAC), etc. La idea para elaborar los mapas físicos es en cierto modo similar a la de ensamblar un rompecabezas: consiste en ordenar los fragmentos del genoma a base de buscar grupos de fragmentos que tienen alguna zona en común, es decir, ir hallando conjuntos de pares de fragmentos parcialmente solapados.

Las estrategias más recientes hacen uso de DNA humano en forma de unos 20.000 trozos independientes clonados en los llamados cromosomas artificiales de levadura (YAC, de sus iniciales inglesas), y buscando la “huella dactilar común” entre clones a base de la detección de determinadas secuencias repetitivas.

Lo último en metodología de mapas físicos ha sido el desarrollo de una especie de “marcadores físicos universales”, fácilmente generables, que permiten que los datos obtenidos en un laboratorio sean rápidamente compartidos y asumidos por toda la comunidad investigadora: se trata de los llamados “lugares etiquetados por su secuencia” (STS en inglés). Consisten en trechos cortos de DNA (de entre 100 y 1000 pb) cuya secuencia exacta se conoce y se sabe que es única en todo el genoma. Su facilidad de uso y su aceptación como “lenguaje común” estriba en que una vez que un investigador descubre una STS, cualquier otro puede obtenerla por sí mismo (ni siquiera hace falta el envío físico de muestras), simplemente fabricando *in vitro* los primeros correspondientes a sus extremos y amplificando la STS por reacción en cadena de la polimerasa (PCR). Los STS definen puntos concretos únicos del mapa físico, y constituyen magníficos marcadores o banderas fácilmente detectables.

Uno de los objetivos iniciales del Proyecto genoma humano era la obtención de mapas físicos con unos 30.000 marcadores repartidos de modo más o menos uniforme, de tal forma que cada dos marcadores consecutivos estén separados una media de 100 kb. Estos mapas de STS permiten la integración de los mapas genéticos y físicos, hacen accesible la fase de secuenciación y facilitan la clonación de genes implicados en enfermedades mediante la llamada estrategia del candidato posicional.

Una vez que se construyen los mapas, hay que refinarlos y purgarlos de posibles errores. Los errores suelen tener dos fuentes principales: algunos clones YACs son en realidad híbridos o quimeras producidas por artefactos durante el proceso de elaboración de la genoteca, y por lo tanto su mapa no refleja el orden genómico auténtico; y por otro lado, los programas de ensamblado de los mapas no son fiables al 100%. De ahí la importancia de confirmar y normalizar los datos mediante estrategias aceptadas por todos los investigadores.

MÉTODOS BÁSICOS DE SECUENCIACIÓN

- Método químico de Maxam y Gilbert (actualmente poco usado).
- Método enzimático de Sanger (terminación de cadena, o de los didesoxinucleótidos, ddNTP).
- Secuenciación automática según el método de Sanger.

Cada ddNTP se marca con un colorante fluorescente diferente, esto permite hacer la electroforesis en el mismo carril del gel. Las bandas de DNA son detectadas por su fluorescencia según pasan delante del detector. Si hacemos mover el detector en horizontal, podrá leer varias secuenciaciones al mismo tiempo y los datos pasan a un sistema computarizado.

NUEVOS MÉTODOS DE SECUENCIACIÓN

Microscopía de barrido (scanning) de efecto túnel (STM).

Una fina sonda se mantiene muy cerca del objeto (en este caso DNA), por medio de un sistema de control basado en la detección de una minúscula corriente inducida por el efecto túnel entre la punta de la sonda y el DNA. Una técnica muy parecida es la microscopía de fuerza atómica (AFM), en la que el control se debe a la medida de las fuerzas de van der Waals entre la sonda y la muestra. En cualquiera de los dos casos, la punta se mueve a lo largo del objeto, de modo que sus desplazamientos en vertical se miden y se registran, generando una imagen de la superficie del objeto. Aunque se han obtenido imágenes del esqueleto azúcar-fosfato de DNA de cadena sencilla y de cadena doble, falta determinar si se pueden “ver” las bases nitrogenadas. Si es así, la técnica podría lograr una secuenciación del orden de 1 Mb cada día.

Secuenciación por hibridación en chips con oligonucleótidos.

Se basa en sintetizar distintas sondas de oligonucleótidos, y unir las en disposiciones ordenadas (arrays) a una fina pastilla de nylon o vidrio. Este chip se prueba frente a un DNA marcado fluorescentemente, de modo que el patrón y cantidad de fluorescencia suministra información sobre la secuencia del

DNA en cuestión. La última generación de este enfoque es la combinación de técnicas folitográficas (como la de los chips de silicio para computadores) con síntesis química en fase sólida, que logra chips con ordenaciones de decenas e incluso centenares de miles de oligos distintos, que pueden usarse para identificar secuencias marcadas fluorescentemente en cuestión de minutos, por medio de un microscopio confocal de fluorescencia totalmente automatizado, que registra los datos.

Affimetrix ha logrado re-secuenciar por este método las 16 kb de DNA mitocondrial humano, con un dispositivo formado por 135.000 oligonucleótidos. Con la tecnología actual se pueden llegar a sintetizar en un día 400.000 oligos de 20 bases cada uno, dispuestos en un chip de 1,6 cm². Pero el objetivo final es lograr un chip con los 4 millones de sondas necesarias para secuenciar todo el genoma humano en una sola hibridación.

EL PAPEL DE LA INFORMÁTICA EN LOS PROYECTOS GENOMA

La informática ha sido uno de los objetivos esenciales del Proyecto genoma humano, debido a la gigantesca cantidad de datos que hay que recoger, analizar, comparar, interpretar y distribuir. La informática aplicada a la biología presenta dos subdisciplinas⁶: la *bioinformática*, se puede definir como el trabajo de investigación y desarrollo que se necesita como infraestructura de información de la actual biología; y la *biología computacional*, que es la investigación dependiente de computación dedicada a entender cuestiones biológicas básicas.

La adquisición de datos experimentales por métodos digitales está animando a la industria a diseñar y fabricar aparatos cada vez más sofisticados, que mejoran y aceleran la parte más rutinaria de la investigación. Para ello los aparatos incorporan sistemas computarizados de análisis y tratamiento de imagen visible.

BASES DE DATOS GENÉTICOS Y MOLECULARES

La gigantesca cantidad de datos generados en los proyectos genoma obliga a no utilizar medios escritos a la hora de publicar los datos: su difusión se hace por medios electrónicos, depositándolos en bases de datos públicos. El ritmo de acumulación de datos es vertiginoso y actualmente se duplican en menos de un año. Los biólogos del siglo XXI usarán esas bases de datos como un recurso indispensable de su trabajo cotidiano. En la actualidad funcionan principalmente dos tipos de bases de datos genómicos:

- El Consorcio Internacional de Bases de Datos de Secuencias está formado por GenBank, el Banco de datos de DNA de Japón (DDBJ) y el del EMBL, alberga los datos de secuencias. Las tres bases comparten y complementan la información. En España existe un nodo de EMBL residente en el Centro Nacional de Biotecnología (CNB) de Madrid.
- La Genome Data Base (GDB) se estableció para albergar los datos de mapas y relacionados (sondas, marcadores, etc.).

Actualmente la base de datos bibliográfica MEDLINE (mantenida por la NML estadounidense) está vinculada con las bases de datos genéticos y de secuencias. Por ejemplo, se puede hacer una búsqueda desde MEDLINE con palabras clave de una enfermedad, lo que da acceso a OMIM (Herencia mendeliana en línea), con referencias bibliográficas, y de ahí se puede saltar a los mapas genéticos, físicos y las secuencias, si están disponibles.

REFERENCIAS

1. Dennis K, Gallaher R, Campbell P. The Human Genome Everyone's genome. Nature 2001, 409(6822)813-958.
2. Klug WS, Cummings MR. Concepts of Genetics 5th Ed Upper Saddle River, New Jersey . Prentice-Hall, Inc. 1997: 477.
3. Karp Gerald. Cell and molecular biology concepts and experiments. New York. John Wiley & Sons, Inc. 1996:416.
4. Magee J, Gordon JI, Whelan A. Bringing the human genome and the revolution in bioinformatics to the medical school classroom: a case report from Washington university school of medicine. Acad Med 2001 Aug;76(8):852-5.
5. Semsarian C, Seidman CE. Molecular medicine in the 21st century. Intern Med J 2001 Jan-Feb;31(1):53-9.
6. Benton D. Bioinformatics-principles and potential of a new multidisciplinary tool. Trends Biothechnol 1996 14(8):261-72.

En las siguientes direcciones se puede encontrar información sobre proyecto genoma:

www.oml.gov/hgmis
www-ls.lanl.gov./masterhgp.html
www.hugo.dgb.org
www3.ncbi.nlm.nih.gov/Omim
www.nhg1.hih.gov