



EL DESARROLLO DE LA GENÓMICA EN EL SECTOR AGROALIMENTARIO

Ricardo Ramos Ruiz^(a) y Ana Ramírez de Molina^(b)

^aParque Científico de Madrid e ^bIMDEA Alimentación

Resumen

La secuenciación del genoma humano supuso en el año 2000 el origen de una nueva disciplina de la genética, conocida como genómica, que aplicada en alimentación se distingue como genómica nutricional. Los grandes avances tecnológicos desarrollados en esta área permiten realizar investigaciones y desarrollos hasta ahora impensables. Entre ellos, la secuenciación de genomas de especies de plantas de consumo, permitiendo optimizar sus propiedades en relación a procesos tan relevantes como la defensa frente a infecciones, mecanismos de maduración o su propia supervivencia. Pero más allá de la aplicación de los nuevos desarrollos tecnológicos a los procesos de producción, la genómica nutricional permite identificar los mecanismos por los que interaccionamos con los nutrientes y, por tanto, permite abordar la aplicación de la nutrición molecular para mejorar nuestra salud en una nueva forma de entender la alimentación. El desarrollo genómico proporciona las bases moleculares de la nutrición personalizada para la prevención y mejora de enfermedades crónicas, aportando nuevos objetivos para el desarrollo de una innovación alimentaria de alto impacto basada en la alimentación para la salud.

Abstract

The sequencing of the human genome in 2000 marked the beginning of a new discipline of genetics, known as genomics, which applied to food science is known as Nutritional Genomics. Great technological breakthroughs developed in this area allow research and developments in ways never before possible. Genome sequencing of varieties of agricultural plants allow optimize their properties in relation to processes as important as the defense against infections, mechanisms of maturation or their own survival. But beyond the application of new technology developments to production processes, Nutritional Genomics allow the identification of the molecular mechanisms by which we interact with nutrients and therefore can address the application of molecular nutrition to improve our health in a new way of understanding nutrition. Genomic development provides the molecular basis of personalized nutrition for the prevention and improvement of chronic diseases. Therefore, it provides new targets for the development of a high-impact innovation of food industry based on food for health.

1. El genoma

El genoma posee una estructura de sorprendente simplicidad. Al menos de una simplicidad aparente. Está formado por cuatro componentes esenciales sencillos –denominados bases o nucleótidos– que se ordenan en una larguísima secuencia de ácido desoxirribonucleico (ADN) formada en el caso de los humanos por unos tres mil millones de unidades. Esa larga cadena lineal contiene información compleja que está encriptada o codificada como si se tratara de un código informático al uso, con las cuatro letras de las bases A, C, G, T,¹ en lugar de los típicos ceros y unos.

Clásicamente se define como «gen» a la unidad de información localizada en el genoma (Lewin, 2008). Los genes son grupos de unos miles de bases ordenadas que contienen una información suficiente que las células son capaces de interpretar para fabricar proteínas, moléculas que desarrollan las funciones bioquímicas en las células. A su vez, las proteínas están

¹ Que se corresponden con las iniciales de las bases nitrogenadas adenina (A), citosina (C), guanina (G) y timina (T).

compuestas por una sucesión de componentes conocidos como aminoácidos que también se ordenan de una forma específica en una secuencia concreta y ordenada. En este caso los aminoácidos aportan una complejidad química superior, que capacita a las proteínas para adoptar una infinidad de estructuras tridimensionales con las que desarrollar las funciones biológicas «reales». La mayoría de las proteínas desarrollan su función como biocatalizadores capaces de crear un entorno químico adecuado para que las reacciones biológicas tengan lugar. De esta manera, el flujo clásico de la información genética es que una secuencia de bases en el ADN se comporta como una información codificada que se interpreta para definir una secuencia de aminoácidos en una proteína que una vez sintetizada es capaz de realizar una función.

El proceso de interpretación de la información desde una secuencia de nucleótidos hasta una secuencia determinada de aminoácidos fue un sorprendente hallazgo de la naciente disciplina de la *biología molecular* allá por los años 50 del siglo XX, que valió el premio Nobel a investigadores, entre otros, a Severo Ochoa (Ochoa, 2000).

La interpretación del código genético sigue un patrón de nuevo tan sencillo como que tres bases consecutivas en el ADN se interpretan como una señal para incorporar un aminoácido concreto durante la biosíntesis de una proteína. En este contexto hablamos de tripletes de bases como la unidad de información de la que se componen nuestros genes siguiendo la lógica de 3 bases $\rightarrow$ 1 aminoácido. De lo anterior se deduce que todos los genes tienen una constitución similar desde el punto de vista químico: una consecución de bases (A, C, G y T) de forma que unos genes solo se diferencian de otros en el orden de dichas bases, de igual forma que los genomas de unos organismos no se distinguen de otros más que en el número y orden de las bases de los que están compuestos.

En este sentido la forma de identificar inequívocamente un gen es «leer» el orden de bases que lo componen e interpretar dicha secuencia, que es lo que conocemos como *secuenciación del ADN*.

Probablemente el programa de colaboración científica más ambicioso que se ha desarrollado en el ámbito internacional, con participación de múltiples laboratorios y con gran repercusión mediática, ha sido la secuenciación del genoma humano en su integridad. Este proyecto iniciado en los años 1990 se prolongó durante más de una década y supuso una inversión de unos 3.000 millones de dólares y culminó con la famosa presentación del primer borrador de la secuencia del Genoma Humano en una rueda de prensa en la que participaron Bill Clinton y Tony Blair en junio del año 2000 (*The Human Genome*, 2001). Además de las innegables repercusiones científicas de dicho proyecto, sirvió también para hacer patente dos nuevos paradigmas de la biología molecular actual: la utilidad del trabajo en red en base a la composición de grandes consorcios internacionales trabajando de forma organizada detrás de un objetivo compartido, pero también la competencia entre consorcios públicos y privados. Toda una señal de las grandes repercusiones que iba a tener la publicación de una información tan esencial, donde se desvelan los secretos guardados por nuestros genes.

De hecho, la publicación de este primer esquema de nuestro genoma supuso en el año 2000 el pistoletazo de salida de una nueva disciplina de la genética, que conocemos como *genómica*, a la que se asocian un buen número de disciplinas entre las que cabe destacar, en nuestro ámbito, la *nutrigenómica*. En su conjunto, la genómica trata de interpretar cómo se rigen las señales que gobiernan el funcionamiento de nuestro material genético, la anotación funcional de los genomas y cómo se transforman las señales genéticas en actividad biológica real. La genómica utiliza sus propias herramientas, basadas en buena medida en los desarrollos tecnológicos que culminaron en la secuenciación del genoma humano.

Los resultados definieron mejor cómo es el *genoma humano*, compuesto por unos 20.000 genes diferentes –curiosamente, menos de los que se habían predicho– dispersos a lo largo de nuestros 23 pares de cromosomas, ocupando apenas un 5 % de nuestra dotación genética (Lewin B, 2008). Como decíamos al principio, esta organización es simple solo en apariencia ya que junto a las señales bien definidas que indican la presencia de secuencias génicas existe toda una serie de señales más «difusas» que modulan la forma en que se procesa esa información, definiendo cuándo, cómo, en qué proporción, durante cuánto tiempo, en respuesta a qué estímulos, en qué parte del organismo y con qué intensidad se expresa cada gen. Este proceso, conocido globalmente como *regulación de la expresión génica*, introduce a otra molécula de importancia muy relevante en los estudios de genómica, el ácido ribonucleico o ARN, que sirve como intermediario entre la información codificada localizada en el ADN y el proceso de síntesis de proteínas.

Podemos considerar entonces a nuestro genoma como una pista plagada de señales. Algunas de dichas señales son claras: inicio y finalización de genes, señales para comenzar la biosíntesis de las proteínas y señales para terminar dicha biosíntesis. Pero hay mucha más información, mucho más sutil, peor definida y más difícil de interpretar. Los genes se expresan (se transcriben en ARN) en base a señales de promotores y activadores o represores que son reconocidas por enzimas conocidas como factores de transcripción, y que a su vez se regulan por señales procedentes de la señalización intracelular. El proceso de expresión se transforma entonces en una compleja red en la que participan de forma coordinada muchas moléculas distintas, sometidas a distintas vías de regulación entre sí y por agentes externos. De esta manera los organismos reconocemos e interpretamos nuestro entorno y respondemos frente a él, mediante auténticas «conversaciones» que realizamos a nivel molecular con las herramientas (los genes) de las que disponemos en nuestro genoma. Complicando este modelo aún un paso más, las propias señales localizadas en el ADN no son completamente estables, sino que existen mecanismos de modificación genética conocidos de forma global como epigenéticos, que alteran transitoriamente, en ocasiones incluso de forma heredable, las propias señales de reconocimiento localizadas en el ADN.

2. Desarrollo tecnológico: la genómica

Las tecnologías de *secuenciación de ADN* se empezaron a desarrollar hace ya casi 50 años, en la década de los años 70 del siglo XX. Inicialmente se idearon dos métodos basados respectivamente en la degradación química del ADN (*secuenciación Maxam-Gilbert*) o en técnicas enzimáticas (*método de Sanger*, que finalmente fue el que acabó asentándose como método universal) (Sanger, 1977). Las enzimas responsables de este proceso son las conocidas como ADN polimerasas, cuya actividad incide justamente sobre la característica más relevante del ADN como material genético. El ADN debe copiarse de forma fidedigna para ser transmitido a las siguientes generaciones y, dentro de un organismo, como dotación a cada célula hija cada vez que una célula se duplica. Para generar una copia idéntica de sí misma el ADN aprovecha su estructura de doble hebra, donde una base localizada en una de las dos hebras solo tiene la opción de asociarse en la hebra opuesta a su base complementaria (A es la complementaria de T, C de G, y viceversa). Es decir, la secuencia de una hebra define de forma inequívoca cuál es la secuencia de la hebra opuesta; en otras palabras, determinar la secuencia de un ADN sintetizado supone directamente «leer» la secuencia del ADN utilizado de molde sobre el que se fabrica la secuencia complementaria.

En el sistema original (*secuenciación de Sanger*), se utilizaba un análogo radiactivo de un nucleótido para marcar la cadena naciente. A continuación se debía aplicar una técnica que permitiera resolver las moléculas en función de su tamaño, distinguiendo moléculas con una sola base de diferencia. Esta tecnología ha sido durante décadas la electroforesis en gel de poliacrilamida, y las moléculas se visualizaban posteriormente por radiografía de los geles con bandas radiactivas. Con esta tecnología, se conseguían unas tasas de productividad próximas a 1 kb (1.000 pares de bases) a la semana. Es decir, con un equipo de unas 10 personas trabajando a pleno rendimiento, se iban a necesitar décadas de trabajo para abordar un genoma entero del tamaño del humano (compuesto por 3,2 Gb, es decir 3.200 millones de pares de bases).

Partiendo de este escenario, se logró incrementar enormemente la productividad de esta tecnología. Una primera innovación importante de la secuenciación Sanger fue la sustitución de las técnicas de marcaje radiactivas por moléculas fluorescentes. Además de las evidentes ventajas en comodidad y seguridad, el marcaje fluorescente permite trabajar con las 4 bases en multiplex, es decir, marcando las 4 bases en el mismo tubo de ensayo, y separando cada base a posteriori simplemente mediante filtros ópticos (Zhou *et al.*, 2000). Y otro avance técnico aún más relevante en términos de productividad fue la sustitución de los geles de poliacrilamida por matrices de mayor resolución, basadas en micropartículas que se compactan sin necesidad de reacciones químicas como las de la acrilamida/bisacrilamida. Estas partículas se autocompactan en microcolumnas muy finas, por lo que dicha tecnología es conocida como electroforesis capilar. Trabajando con columnas capilares de mayor longitud se conseguía aumentar la longitud de lectura hasta cerca del millar de bases por cada columna capilar (Albarghouthi *et al.*, 2000). Si sumamos el desarrollo de equipos multicapilares corriendo en paralelo y a mucha mayor velocidad nos encontramos con un aumento de productividad de

hasta varios centenares de kilobases al día, es decir, un aumento de al menos dos órdenes de magnitud. Tales mejoras son las que permitieron imaginar que la secuenciación del genoma completo de un organismo superior era abordable y plantear y culminar el ambicioso proyecto de la secuenciación del genoma humano.

Aunque no haya alcanzado una proyección tan marcada, también se desarrolló una tecnología diferente de la de «4 nucleótidos a la vez», en la cual los nucleótidos se añaden uno por uno y se detecta su posible incorporación a la cadena molde. Esta tecnología es conocida como piro-secuenciación ya que la detección se basa en la medida de los grupo pirofosfato liberados cada vez que se incorpora un dNTP a una cadena de ADN en síntesis, y es especialmente útil para detectar y cuantificar poblaciones mixtas de ADN, por ejemplo la presencia de posiciones polimórficas, la detección de mutaciones en muestras mixtas (por ejemplo tejido tumoral mezclado con tejido sano) o la presencia de modificaciones sobre los nucleótidos (como caso típico, la metilación de citosinas).

2.1. La revolución de la genómica

En la última década se ha logrado desarrollar una nueva tecnología que ha vuelto a superar a este tipo de secuenciación capilar, que ahora ha pasado a considerarse como «tradicional», aumentando de nuevo la productividad en la escala de varios órdenes de magnitud. Se trata de la llamada secuenciación masiva inicialmente conocida como *Next Generation Sequencing*, un nombre que ya ha perdido su significado original puesto que es una tecnología perfectamente instaurada y madura en la comunidad científica. Normalmente nos referiremos a ella como *secuenciación masiva* o *secuenciación en paralelo*.

La primera tecnología de secuenciación en paralelo se describió en 2005 por Jonathan Rothberg y supone una muy ingeniosa automatización del proceso de piro-secuenciación, a la que siguió poco tiempo después una tecnología alternativa conocida como secuenciación por síntesis, semejante en su concepto a la secuenciación Sanger basada en 4 nucleótidos fluorescentes y basada en el seguimiento a tiempo real del proceso de secuenciación (Mardis, 2011).

La principal innovación de la secuenciación masiva es la alta productividad y la automatización del proceso de secuenciación. En los primeros prototipos se llegaban a alcanzar tasas de piro-secuenciación de casi 1 millón de reacciones en paralelo. En el segundo tipo de secuenciación por síntesis, mayoritario a día de hoy, podemos llegar a alcanzar rendimientos superiores a centenares de millones de moléculas secuenciadas en paralelo. Las longitudes de lectura que se alcanzan en cada reacción de secuenciación individual también han ido avanzando a lo largo de los últimos años, pasando de centenares de bases hasta los equipos actuales de última generación donde se pueden alcanzar kilobases de secuencia lineal en cada reacción (Metzker, 2010).

La secuenciación masiva supone distintas mejoras esenciales sobre las tecnologías desarrolladas apenas 10 años antes:

- Evita preparar cada fracción del genoma en un sistema biológico (plásmidos, virus, cromosomas artificiales...), todo el proceso se realiza *in vitro* y en un número reducido de etapas. Es una secuenciación acelerada.
- Permite trabajar en ausencia absoluta de información previa. Es una secuenciación en ausencia de premisas.
- Cada una de las secuencias obtenidas supone la lectura del ADN presente en una sola molécula o un único fragmento de un genoma. No hay por tanto ningún problema en trabajar con muestras de múltiples organismos combinados, un reto imposible de resolver con tecnologías pasadas. Es una secuenciación de moléculas individuales.
- Y por supuesto, la enorme cantidad de secuencia generada es lo que aporta toda su potencialidad. Un solo experimento de secuenciación genera suficiente información como para leer y repasar decenas de veces un genoma completo. Permite abordar la secuenciación de genomas tan complejos como los de muchos organismos vegetales y determinar la presencia de variantes genéticas en la escala del genoma completo. Es la tecnología de secuenciación de genomas.

La secuenciación en paralelo es en sí misma una tecnología en evolución y sigue desarrollando una constante búsqueda de mejoras. El primer salto cualitativo de la secuenciación masiva supone el análisis del material genético original sin modificaciones. Se trata de la secuenciación de molécula única mediante dispositivos de ‘smrt-cell’: *single-molecule real time sequencing* o basadas en nanoporos, que permiten alcanzar lecturas de longitudes larguísimas a partir de ADN que puede mantener todas las posibles modificaciones químicas que presentaba in vivo el genoma (Schadt *et al.*, 2010).

Y el segundo avance supone el análisis de genomas o transcriptomas a partir de muestras tan reducidas como puede ser una sola célula aislada, de forma que los tejidos no se analicen en global sino respetando y estudiando la individualidad de cada célula dentro de un sistema biológico (Shalek *et al.*, 2014). De nuevo, ambos objetivos eran inalcanzables para las técnicas clásicas, lo que permite afirmar que la nueva generación de la genómica nos está permitiendo comprender aspectos de nuestra biología completamente desconocidos y apasionantes.

A nivel de desarrollo comercial, los equipos de piro-secuenciación han tenido su lanzamiento gracias a las empresas Roche (Junior y FLX-454) y Life Technologies (Ion Torrent e Ion Proton), la secuenciación basada en síntesis e incorporación de nucleótidos (RTA; *real time analysis*) se ha desarrollado por parte de la empresa Illumina, y los nuevos equipos de molécula única están protagonizados por desarrollos de Pacific Biosciences (tecnología smrt-cell) y los basados en nanoporos por Oxford Nanopore y Genia.

El conocimiento de la secuencia genética completa constituye el primer paso para comprender las funciones y los mecanismos de actuación de los genes y de sus productos, así como modular estas actuaciones para su explotación en el sector agroalimentario.

No podemos dejar de considerar que la genómica no es una tecnología aislada y cerrada sobre sí misma. Existe de forma genérica todo un abanico de *tecnologías-ómicas* que engloban a todas aquellas disciplinas de investigación que estudian el conjunto o totalidad de un sistema biológico. Especialmente interrelacionada con la genómica podemos destacar a la transcriptómica (estudio de genes expresados en un organismo), la proteómica (estudio del conjunto de proteínas existentes en un sistema biológico) y la metabolómica (metabolitos implicados en las funciones biológicas).

En la Tabla 1 se resume el campo de aplicación de las principales ómicas con las que se relaciona la Genómica.

Tabla 1. Campo de aplicación de las principales ómicas de la genómica

Disciplina	Tecnología	Molécula de análisis	Resultado
Genómica	Secuenciación Sanger, secuenciación masiva	ADN	Identificación de genes y genomas
Transcriptómica funcional	Secuenciación masiva, <i>microarrays</i> , PCR a tiempo real	ARN, miRNAs smallRNA, lncRNAs...	Medida de la expresión génica
Transcriptómica descriptiva	Secuenciación masiva	ARN, <i>mirnas smallrna</i> , <i>lncrnas</i> ...	Identificación de genes, catálogos de genes
Metabolómica	Cromatografía, espectrometría (LC, MS, TOF)	Metabolitos, analitos	Cuantificación de metabolitos
Proteómica	Cromatografía, espectrometría (LC, MS, TOF)	Péptidos, proteínas	Mapas peptídicos, identificación de proteínas, cuantificación de proteínas
Metagenómica	Secuenciación masiva, PCR	ADN ambiental	Identificación de taxonomía de poblaciones

3. Programas de mejora genética en el sector alimentario

Como ya se ha comentado, la genómica y los avances tecnológicos desarrollados en esta área permiten investigar la dotación genética de un organismo y su interacción con el medio ambiente, así como los mecanismos moleculares responsables de sus posibles alteraciones. Derivado de ello, y aplicado al sector agroalimentario, permite la optimización de productos y procesos de interés.

3.1. La secuenciación de genomas en el ámbito agroalimentario

Dentro de las especies de interés agrícola, la secuenciación a escala genómica está empezando a abrir nuevas vías de conocimiento. Uno de los puntos de mayor interés es el del estudio de los procesos de domesticación y separación de especies y variedades, como se han

publicado en los últimos años para especies como la frambuesa, tomate, pepino, plátano y por supuesto para los cereales de mayor consumo (Doebly *et al.*, 2006; D'Hont *et al.*, 2012). El desarrollo de herramientas informáticas también está permitiendo una mejor fidelidad y rapidez en la anotación de genes y variantes. En conjunto, estos avances están optimizando a pasos agigantados la base molecular de las propiedades beneficiosas de las distintas variedades de las plantas de consumo en relación a procesos tan relevantes como la defensa frente a infecciones, los mecanismos de maduración o la propia supervivencia en función de las condiciones ambientales.

Desde un punto más práctico, los estudios actuales se están enfocando en la generación de mapas genéticos del máximo detalle, en especies como por ejemplo el trigo, la cebada o el arroz. La vía lógica de desarrollo es previsiblemente que en un futuro cercano, estos descubrimientos se transformen en marcadores que permitan seguir e influir de forma racional en la mejora agrícola basada en criterios racionales y seguros. El proceso supone generar datos genómicos de alta densidad que sirven para desarrollar marcadores de «calidad para cruzamiento», tanto más fidedignos cuanto mayor sea la calidad de los datos fenotípicos asociados. En algunas especies estos análisis se realizan actualmente mediante la tecnología de *microarrays* de alta densidad, que podrían ser sustituidas en la actualidad por técnicas de secuenciación en paralelo, de mayor capacidad y productividad.

Un ejemplo lo encontramos la reciente secuenciación del genoma de la remolacha azucarera, publicada en la prestigiosa revista *Nature* el año pasado, (Dohm *et al.*, 2014), en la que participaron investigadores del Centro de Regulación Genómica (CRG) de Barcelona junto con el Instituto Max Planck de Genética Molecular y la Universidad de Bielefeld. Gracias al desarrollo de las tecnologías de secuenciación se ha podido realizar este trabajo que aporta una nueva visión sobre cómo el genoma se ha ido forjando gracias a la selección artificial. La remolacha azucarera es el primer representante con el genoma secuenciado de plantas con flores cariofilales con más de 11.000 especies entre las que se incluyen plantas de importancia económica, como la espinaca o la quínoa. En este estudio han identificado genes hasta ahora desconocidos en regulación de la expresión génica posiblemente derivados de la selección artificial, información que puede ser de gran utilidad para la mejora del cultivo en relación con el rendimiento y calidad, y, por tanto, hacia su aplicación en el desarrollo de cultivos sostenibles.

3.2. Transgénesis

La genómica nos está permitiendo comprender como son los genes, cómo funcionan y en qué se basan sus propiedades; permite asociar genes a características y en ese sentido aporta información necesaria y suficiente para la mejora de las variedades alimentarias por selección de las mejor adaptadas a las necesidades de cada caso. Sin embargo, desde hace bastantes años, se ha venido desarrollado una tecnología capaz de incidir directamente sobre los organismos basado en una transformación radical: añadir un gen externo, que confiera una propiedad de la que se carecía.

Este proceso denominado *transgénesis* se ha aplicado ya sobre diversos organismos animales y vegetales. No se puede obviar que se ha generado un amplio debate que cuestiona sus logros, tanto en relación a su seguridad (una polémica que hasta ahora no ha conseguido demostrar ningún peligro para la salud) como por las repercusiones económicas de los monocultivos, e incluso por motivos éticos.

El diseño experimental de la transgénesis es teóricamente muy sencillo: se introduce la porción de ADN que codifica por el gen extra en la línea germinal y se deja que se integre en el genoma receptor por mecanismos de biología molecular. Posteriormente se analiza en la progenie qué individuos han sido capaces de integrar el gen exógeno y se estudia su efecto. Tradicionalmente este proceso se ha realizado mediante microinyección del cigoto recién fecundado, sistema que se ha refinado mediante técnicas de manipulación de células pluripotentes (López Guerrero, 2008). Así, las técnicas más actuales utilizan una forma muy elegante de conseguir la modificación genética. Se parte de células recogidas en el estado de blástula y se mantienen indiferenciadas el tiempo necesario para su manipulación genética. Una vez conseguidas dichas modificaciones, las células se reimplantan en otra blástula y se continúa con el proceso de desarrollo. Este diseño supone que no todas las células del organismo genéticamente modificado vayan a mostrar los cambios derivados del genotipo alterado. El razonamiento subyacente a estos procesos es el mismo que inspira la terapia génica. Tanto a nivel de individuo en el caso de la terapia génica, como a nivel de variedad en el caso de la mejora agroalimentaria, el objetivo es aportar una característica, o bien corregir un defecto o suplir una carencia utilizando herramientas que inciden directamente sobre el material hereditario.

Con un diseño así, se han buscado soluciones y mejoras en distintos ámbitos. Hay ejemplos que se han desarrollado dentro del ámbito de la ganadería, entre los que podemos destacar a los siguientes como algunos de los más relevantes:

- Vacas transgénicas que producen leche semejante a la humana, con una potencialidad alergénica muy reducida.
- Modelos animales para el estudio de enfermedades humanas o productores de anticuerpos «humanizados».
- Animales de granja de mayor tamaño para una mayor producción.

Sin embargo, es sin duda dentro del ámbito agrícola donde los avances de la biotecnología ha adquirido su mayor envergadura y avance con los desarrollos de plantas transgénicas. Como se ha comentado, la transgénesis necesita de la incorporación del material genético externo a las células de la línea germinal. En el caso de las plantas la transferencia del material genético se ve muy facilitado por la capacidad de regeneración de las plantas a partir de células maduras. Pero probablemente la causa principal del desarrollo de modelos transgénicos en plantas se deba a la importancia económica y social que tienen las plantas de interés agrícola como fuente básica de alimentación para la humanidad.

La mayoría de las modificaciones a las que se dirigen las plantas transgénicas tienen que ver con sus propiedades nutritivas o su mejora como especie de cultivo (López Guerrero, 2008). Las más conocidas sin duda son las siguientes:

- Plantas con capacidad de supervivencia en condiciones de cultivo más desfavorables, por ejemplo plantas con capacidad de resistir sequías más prolongadas o de producir sus cosechas más rápidamente.
- Plantas resistentes de forma natural a insectos, que eviten la necesidad del uso de pesticidas altamente tóxicos.
- Plantas con capacidades alimenticias adicionales, como es el caso del llamado arroz dorado, enriquecido mediante transgénesis de los genes productores de beta caroteno.

En 1992 se llevaron a cabo los primeros cultivos transgénicos de tabaco en China y en 1994 se aprobaron en Estados Unidos el consumo de organismos modificados genéticamente (OMG) para el consumo humano. En España se cultivó por primera vez en 1998 un maíz Bt resistente a dos especies de lepidópteros (taladros). Desde entonces ha habido un desarrollo espectacular en este ámbito, y actualmente hay una significativa producción mundial de cultivos transgénicos, siendo los más comunes la soja, el maíz, colza y el algodón. Frenando su desarrollo, la introducción de este tipo de alimentos en el mercado ha generado gran polémica tanto por las implicaciones económicas de los monocultivos, como referentes al miedo a que modificaciones genéticas impliquen riesgos para la salud o el medioambiente, aunque, como hemos mencionado, no se ha podido demostrar riesgo asociado hasta la fecha.

4. Aplicación del desarrollo de la genómica a la alimentación para la salud: la nutrición personalizada

A principios del siglo XXI el coste medio de secuenciación de un genoma humano rondaba los 20 millones de euros, constituyendo una barrera casi infranqueable para su aplicación en el sector agroalimentario. Sin embargo, el desarrollo tecnológico realizado en estos años ha resultado no solo en un vertiginoso avance en capacidad tecnológica y aplicaciones, sino también en una sorprendente reducción del coste, anunciándose a principios de 2014 por parte de Illumina el desarrollo que permitía «secuenciar un genoma personal por menos de 800 euros». Las consecuencias más importantes son su factible aplicación en áreas como la Biomedicina o las ciencias de la alimentación, permitiendo la conjunción de ambas en la nutrición molecular o la nueva nutrición para la salud: la *nutrición personalizada*.

4.1. Variaciones sobre un mismo tema: el genotipo

Pese a estar hablando del genoma humano como una sola entidad, no podemos considerarlo como una secuencia invariable, única e inmutable, absolutamente semejante para todos los individuos de la raza humana. Si bien nuestro genoma nos identifica, lo heredamos de nuestros padres y lo transmitimos a nuestros hijos, hay cierto margen de indeterminación. A lo largo de nuestra evolución se han ido produciendo sustituciones debidas en general a errores aleatorios producidos cuando el material genético debe copiarse acumulándose cambios sutiles pero persistentes. Con el devenir de las generaciones algunos de esos cambios accidentales han quedado fijados en el genoma, en muchos casos coexistiendo con las variantes originales. El resultado es que dentro de la larga secuencia de nuestro genoma se localizan posiciones que denominamos polimórficas, es decir coordenadas en las que podemos encontrar más de un nucleótido en una determinada posición. Recordemos además que poseemos dos copias de nuestro genoma, una heredada de cada uno de nuestros progenitores. Esto quiere decir que en esas posiciones polimórficas una persona puede poseer dos «variantes» diferentes, una en cada copia de su genoma.

La aparición de variaciones en nuestro genoma son el motor de la evolución y la causa de diferencias genéticas entre nosotros. Pensemos ahora en los genes que codifican las enzimas encargadas de catalizar una función bioquímica bien definida. De acuerdo con el código genético un cambio de un nucleótido por otro (llamémosle de forma genérica como «mutación») cambia en el código que define qué aminoácido se iba a incorporar en esa posición concreta. El cambio de un aminoácido por otro en la cadena nucleotídica puede resultar a veces irrelevante y resultar inocuo, pero también puede ocasionar un efecto devastador en la proteína, provocando su inactivación (por ejemplo, porque aporte una señal de fin de lectura) o puede quedarse en un efecto intermedio o más restringido. La función bioquímica de las proteínas como catalizadores de las reacciones biológicas se basa en la estructura tridimensional que adoptan en el espacio, facilitando la interacción de los metabolitos y estableciendo un entorno químico adecuado para el progreso de las reacciones. Un cambio de un aminoácido por otro por ejemplo en el centro activo de la proteína puede modificar la eficacia de estas interacciones, alterando la capacidad de interacción de los metabólicos entre sí, modificando constantes cinéticas, constantes de asociación, necesidad de cofactores, etc. afectando de esta manera en el proceso metabólico global (Crops, 2013).

Esta es la base racional de la interrelación entre genotipo y fenotipo. El genotipo es la dotación genética, la información en el ADN que eventualmente podría tener consecuencias sobre rasgos o características observables en el organismo, causando lo que conocemos como cambio fenotípico. Los cambios en el genoma sirven como marcadores, son sencillos de determinar y cuantificar con las actuales herramientas de la genómica y se pueden trasladar a cambios observables en las funciones celulares, son «observables» y cuantificables por técnicas bioquímicas o de cualquier otro tipo en función del efecto producido.

Frecuentemente nos encontraremos también que los distintos fenotipos responden a variaciones genotípicas aún más sutiles. Un caso típico es el de la capacidad de digestión de la lactosa. En nuestra infancia tenemos activas las enzimas que procesan la lactosa pero esta capacidad se pierde al pasar a la edad adulta, asociado a la transición de la alimentación materna a una alimentación diferente. Sin embargo, las poblaciones originarias del norte de Europa mantienen la capacidad de digerir la leche en la edad adulta, lo que se ha asociado evolutivamente a la posibilidad de alimentarse a base de productos lácteos por parte de poblaciones de pastores. Sorprendentemente, el marcador genético que podría ser responsable de dicha capacidad parece estar alejada unas 14.000 bases respecto al gen de la lactasa (LCT), responsable de la síntesis de la hidrolasa responsable de la función enzimática (*The food issue*, 2013). Esta variante identificada podría considerarse como un «marcador» de la tolerancia adulta a la leche que se ha ido extendiendo desde hace apenas unos 8.000 años y que coexiste en la población con las variantes originales.

Otro ejemplo que muestra las adaptaciones de nuestro genoma en relación con los alimentos es el del gen de la amilasa. En este caso, se ha asociado la presencia de múltiples copias del gen, normalmente asociado a una mayor tasa de actividad, en ciertas poblaciones cuya dieta está basada en un mayor consumo de carbohidratos (*The Food Issue*, 2013).

Finalmente, ¿cómo de variable es nuestro genoma? Es decir: ¿podemos esperar que haya muchos cambios genotípicos que puedan tener una repercusión en el fenotipo? Las técnicas actuales de secuenciación de genomas indican que al comparar dos individuos entre sí, o frente a la secuenciación del genoma humano «promedio» o «consenso», se pueden llegar a detectar millones de posiciones polimórficas (Alkan *et al.*, 2011). Pese a lo llamativo de este número, un sencillo cálculo nos permite tranquilizarnos pensando que aún mantenemos como mínimo un 99,9 % de homología con todos nuestros congéneres. Si realizamos este mismo cálculo centrando el análisis a las regiones directamente involucradas en la codificación de proteínas, la parte de nuestro genoma que conocemos como exoma, la variabilidad es mucho más reducida, habitualmente a algunos miles de nucleótidos. Esta subregión «esencial» de nuestro genoma corresponde a unas 40 Mb (40 millones de nucleótidos) lo cual nos señala que nuestra identidad se aproxima a un 99,99 % de nuestro genoma.

Sin embargo, esas diferencias aparentemente pequeñas son esenciales para interpretar las diferencias de comportamiento entre cada individuo, tanto en lo referente a la predisposición ante las enfermedades como frente a la distinta respuesta frente al consumo de determinados alimentos, actual objeto de intensa investigación en *genómica nutricional*, como herramienta útil para lograr ejercer una prevención eficaz y mejorar el tratamiento de enfermedades crónicas a través de la alimentación, mediante la nueva nutrición molecular o nutrición personalizada.

Las enfermedades crónicas como enfermedad cardiovascular, el cáncer o la diabetes, son las principales causas de mortalidad y pérdida de calidad de vida en nuestro país (<http://www.who.int/gho/countries/esp.pdf?ua=1>), causando más de 130.000 muertes en el año 2012, en gran parte de los casos después de años de intenso deterioro e importante coste sanitario.

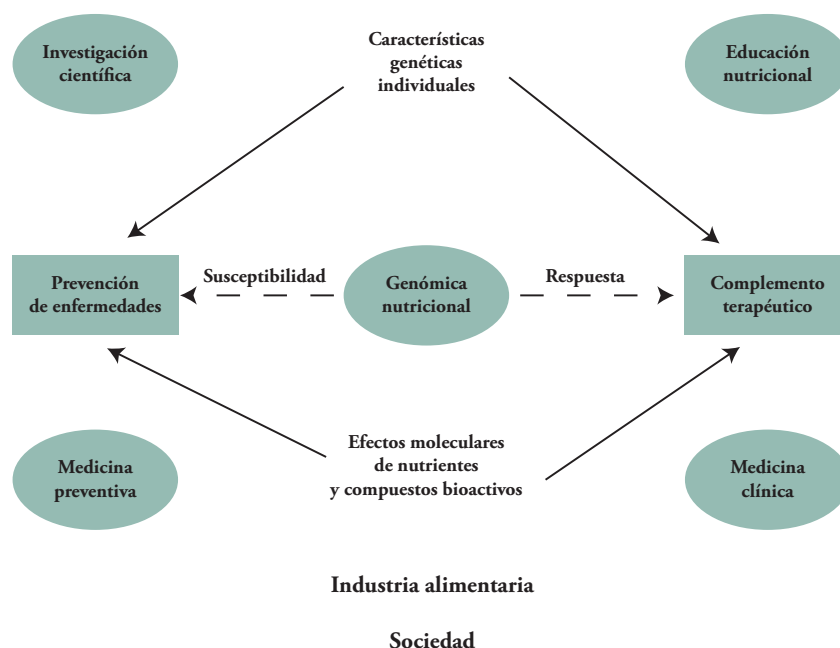
Según datos de la Organización Mundial de la Salud, la nutrición está en el primer plano como determinante importante en el desarrollo de estas enfermedades crónicas que puede ser modificado (OMS, Serie de Informes Técnicos, n.º 916). La genómica nutricional en el estudio de la interacción gen-nutriente constituye una herramienta clave para poder ejercer una nutrición para la salud de forma efectiva, tanto preventiva como a nivel de complemento terapéutico. Gracias al desarrollo tecnológico, el conocimiento de los mecanismos moleculares de acción de los nutrientes, así como el distinto efecto que son capaces de ejercer en función del genotipo de cada persona, permitirá realizar recomendaciones nutricionales personalizadas (o según subgrupos de población con variantes genéticas concretas) que puedan influir en el estado de salud de forma efectiva.

En los últimos años, el número de estudios que han empezado a describir interacciones entre genes y dieta ha experimentado un gran crecimiento. Un ejemplo ya muy bien establecido lo encontramos en el gen de la 5-10-metilentetrahidofolato reductasa (MTHFR), implicada en la formación de la forma primaria de folato sérico, cosustrato para la remetilación de homocisteína a metionina. Un aumento de homocisteína en sangre está asociada con un mayor riesgo de enfermedad cardiovascular (Nakai et al., 2001), proceso en el que esta enzima desempeña un papel clave. Se ha descrito una variante genética en el exón 4 de MTHFR que se encuentra con bastante frecuencia en la población, y que se traduce en un cambio en un aminoácido de esta proteína generando una enzima de menor actividad. Así, una baja ingesta de ácido fólico en personas con esa variante genética se traduce en mayores concentraciones séricas de homocisteína que les conferiría un mayor riesgo de enfermedad cardiovascular (Russo, 2003). Sin embargo, cuando la ingesta de ácido fólico en la dieta es más elevada, esta mayor cantidad de ácido fólico, compensaría el «defecto» genético disminuyendo el riesgo asociado. Así, la recomendación nutrigenética para personas con ese genotipo en MTHFR sería un consumo diario de alimentos ricos en ácido fólico, lo que tendría incidencia en sus niveles de homocisteína plasmática y en la prevención de riesgo cardiovascular.

Otro ejemplo lo podemos encontrar en el gen que codifica para PPARA (*peroxysome proliferator-activated receptor alpha*), importante regulador del metabolismo de lípidos relacionado con procesos de comunicación celular, que podría contener en su secuencia marcadores relacionados con el riesgo a enfermedad cardiovascular en relación a los niveles de ingesta de grasa saturada en la dieta. Recientemente se ha identificado una variante en este gen que se asocia a la disminución en factores de riesgo cardiovascular tras el consumo de grasa de leche. Así, personas con el genotipo TT en rs13554 de PPARA muestran un beneficio en su perfil lipídico e índices asociados de riesgo de enfermedad cardiovascular (ratios LDL, HDL y colesterol total) al consumir leche desnatada, aumentando significativamente dichos índices de riesgo con el consumo de leche no desnatada. Sin embargo, este efecto no se observa en personas con otros genotipos, que podrían evitar restricciones en estos productos ya que no se producen diferencias en su perfil lipídico al consumir distintos tipos de leche (Loria-Kohen et al., 2014). Así, el conocimiento de estas interacciones podría derivar en recomendaciones nutricionales personalizadas y en la introducción de nuevos productos para la salud en distintos subgrupos de población según sus características genéticas.

Sin embargo, los procesos no son simples ya que las interacciones genes-nutrientes ocurren a distintos niveles, al igual que los factores que influyen en el desarrollo de estas enfermedades son también complejos. Ahora se dispone de la tecnología necesaria para realizar análisis genéticos rápidos y fiables, pero hay que generar e integrar bien la información necesaria para hacer recomendaciones personalizadas de forma efectiva. Además, es esencial a su vez realizar un trabajo multidisciplinar para que la genómica nutricional tenga aplicación en nuestra sociedad, donde los profesionales de la salud juegan un papel esencial ya que es necesario que la nutrición molecular sea considerada como la herramienta de salud que es, integrándose nutricionistas especializados en la práctica clínica como un aspecto importante en atención primaria. También la industria alimentaria tiene que colaborar en este proceso, participando en el desarrollo y llegada al mercado de nuevos alimentos para la salud adecuados a distintos grupos poblacionales con características genéticas específicas, adecuación que será determinante para lograr su efectividad. Por último, una formación adecuada de la población para entender el impacto que puede tener en su salud y los beneficios de estos nuevos alimentos es también necesaria para que los beneficios derivados de los avances científicos en genómica nutricional y su enorme potencial llegue a la sociedad y, como consecuencia, se imponga de forma definitiva en el sector agroalimentario. En la Figura 1 se muestra como la industria alimentaria es una importante vía de conexión entre la medicina preventiva y la sociedad a través de la aplicación de la genómica nutricional, para la que se necesita la base de una investigación científica sólida y la educación nutricional necesaria para entender los avances científicos y nuestra responsabilidad en la mejora de la salud a través de la nueva nutrición.

Figura 1. Conexión de la industria alimentaria, la medicina preventiva y la sociedad a través de la genómica nutricional



5. Conclusiones y futuras perspectivas

La genómica es una ciencia que avanza de forma imparable. En apenas unos años se ha convertido en una tecnología revolucionaria que está cambiando la forma de relacionarnos con nuestro entorno, conociéndonos mejor como individuos desde el punto de vista de nuestros genes, conociendo mejor las bases moleculares del desarrollo de las enfermedades crónicas que constituyen pandemias en nuestra sociedad actual, y conociendo mejor cómo interaccionamos con los nutrientes y cómo mejorar nuestra salud.

Las tendencias actuales señalan hacia la consecución de los siguientes avances:

- Aplicaciones en alimentos transgénicos, diseñando modificaciones de forma mucho más precisa y dirigida, en respuesta a necesidades concretas.
- Diagnóstico y prevención de enfermedades de manera cada vez más detallada. Cabe esperar que la secuenciación genómica acabe siendo una práctica de rutina que sirva como una «prueba del talón universal» para detectar deficiencias o riesgos específicos asociados al genotipo. El impacto social de la secuenciación masiva será en este sentido, enorme. Sin duda, surgirán nuevos problemas –están surgiendo ya– en relación a la confidencialidad de los datos genéticos, el uso de información sensible por parte de terceros y las implicaciones éticas.
- Asesoramiento nutrigenético basado en la interacción gen-nutriente posibilitando directrices nutricionales personalizadas, con dos ámbitos de aplicación, prevención primaria de enfermedades, y también como complemento efectivo y personalizado del tratamiento.
- Finalmente, la industria alimentaria no va a ser ajena a este proceso y existirá un interés creciente en participar en el desarrollo de nuevos alimentos adaptados a las necesidades genéticas específicas de los diferentes grupos de individuos, forma de diferenciación y generación de valor añadido.

Como conclusión final, cabe reafirmar que la nutrición molecular es una realidad que se va a asentar porque integra conocimiento de múltiples vías que convergen: las bases moleculares de las propiedades nutritivas de los alimentos, las características de nuestros genes de forma individualizada y las interacciones entre todos ellos en relación con nuestro entorno. Una auténtica puerta abierta a la nutrición de precisión, herramienta de gran potencial social y económico.

Referencias bibliográficas

- (2001): «The human genome»; *Nature* 409(6822); pp. 745-964.
- (2001): «The human genome»; *Science* 291(5507); pp. 1145-1434.
- ALBARGHOUTHI, M. E. y BARRON, A. E. (2000): «Polymeric matrices for DNA sequencing by capillary electrophoresis»; *Electrophoresis* (21); pp. 4096-4111.
- ALKAN, C.; COE, B. P. y EICHLER, E. E. (2011): «Applications of next-generation sequencing. Genome structural variation discovery and genotyping»; *Nature Review Genetics* (12); pp. 363-376.
- BENJAMIN, L. (2008): *Genes (IX)*. Editorial Sudbury, MA, Jones and Bartlett Publishers.
- CROPS, G. M. (2013): «Promise and reality»; *Nature* (497).
- D'HONT, A. *et al.* (2012): «The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants»; *Nature* (488); pp. 213-219.
- DOEBLEY, J. F.; GAUT, B. S. y SMITH, B. D. (2006): «The molecular genetics of crop domestication»; *Cell* (127); pp. 1309-1321.
- DOHM, J. C. *et al.* (2014): «The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*)»; *Nature* (505); pp. 546-549.
- LÓPEZ GUERRERO, J. A. (2008): *¿Qué es un transgénico?* Editorial Sirius.
- LORIA-KOHEN (2000): «A genetic variant of PPARA modulates cardiovascular risk biomarkers after milk consumption»; *Nutrition* 2014 30(10); pp. 1144-50.
- MARDIS, E. R. (2011): «A decade's perspective on DNA sequencing technology»; *Nature* (470); pp. 198-203.
- METZKER, M. L. (2010): «Sequencing technologies - the next generation»; *Nature Reviews* (11).
- NAKAI, K.; ITOH, C.; NAKAI, K.; HABANO, W. y GURWITZ, D. (2001): «Correlation between C677T MTHFR gene polymorphism, plasma homocysteine levels and the incidence of CAD»; *Am J Cardiovasc Drugs* (1); pp. 353-61.
- OMS (2003-2013): «Dieta, nutrición y prevención de enfermedades crónicas»; *Serie de Informes Técnicos* (916). Informe de un grupo de estudio de la OMS. Ginebra, Organización Mundial de la Salud (OMS, S).
- RUSSO, G. T. *et al.* FRAMINGHAM OFFSPRING STUDY COHORT (2003): «Age and gender affect the relation between methylenetetrahydrofolate reductase C677T genotype and fasting plasma homocysteine concentrations in the Framingham Offspring Study Cohort»; *J Nutr* (133); pp. 3416-21.

- SANGER, F.; NICKLEN, S. y COULSON, A. R. (1977): «DNA sequencing with chain terminator inhibitor»; *Proceedings of the National Academy of Sciences USA* (74); pp. 5436-5437.
- SCHADT, E. E.; TURNER, S. y KASARKIS, A. (2010): «A window into third-generation sequencing»; *Human Molecular Genetics* (19); pp. R227-R240.
- SEVERO OCHOA (2000): «Base molecular del mensaje genético»; *Monografías* (20). Ed. CSIC.
- SHALEK *et al.* (2014): «Single-cell RNA-seq reveals dynamic paracrine control of cellular variation, Alex K»; *Nature* (510); pp. 363-369.
- THE FOOD ISSUE (2013): *Scientific American* 309(3).
- ZHOU, H.; MILLER, A. W.; SOSIC, Z.; BUCHHOLZ, B.; BARRON, A. E.; KOTLER, L. y KARGER, B. L. (2000): «DNA Sequencing up to 1300 bases in two hours by Capillary Electrophoresis with mixed replaceable linear polyacrylamide solutions»; *Anal. Chem.*, 72(5); pp. 1045-1052.