

Decodificando el Genoma del Maíz en Guanajuato.

Dr. Jean-Philippe Vielle Calzada¹

Qué curiosa realidad la de esta planta de porte mediano que todos los mexicanos orgullosamente nos comemos a diario. Por un lado, no tenemos dificultades en otorgarle un sitio preponderante como fundamento de nuestra cultura milenaria; el maíz (*Zea mays*) es simultáneamente el eje de la actividad económica productiva del medio rural, la base de la alimentación nacional, el organizador del tiempo y del espacio campesino, el elemento insustituible del arte culinario, la materia prima de innumerables artesanías, y el centro de un conocimiento popular acumulado durante milenios que se venera con refranes, leyendas, ritos y ceremonias. El centro de origen del maíz se encuentra en México. Lo anterior se encuentra fundamentado por tres evidencias categóricas: a) los restos arqueológicos del maíz más antiguo

han sido encontrados en cuevas de diversas partes del territorio nacional, b) el teosintle, el pariente más cercano y posible ancestro del maíz, solamente ha sido encontrado en forma natural en México y el norte de Centroamérica, y c) la diversidad genética más amplia del maíz se encuentra representada en las variedades criollas de nuestro país. Su domesticación, ocurrida hace más de 6,000 años, en los cauces del río Balsas, constituye el evento civilizatorio más importante previo a la llegada de Colón, y fué fundamental para el establecimiento de las múltiples culturas de Mesoamérica, desde los Mexicas hasta los Zunis. Y sin embargo, a pesar de esta extraordinaria riqueza cultural, la ciencia mexicana del siglo XX se interesó muy poco en conocer a fondo la biología de su alimento predilecto. Como es que el maíz forma mazorcas tan diferentes a cualquier otro cereal comestible o silvestre? Como es que forma hileras de hojas verticales que se mantienen en el mismo eje con respecto al sol? Como logran ciertas razas de maíz resistir a la sequía? De qué depende el tamaño de sus semillas y su contenido de proteínas esenciales para el ser humano? Si bien la respuesta a estas preguntas se encuentra estrechamente asociada con sus características

¹ Investigador Titular del LANGEBIO. Doctor en Ciencias (Ph.D) con Especialidad en Genética, Texas A&M University (1995); Posdoctorado en Genética Molecular del Desarrollo en Plantas en Cold Spring Harbor Laboratory (1999). Becario Internacional Howard Hughes 2001-2006. Investigador Nacional Nivel III. vielle@ira.cinvestav.mx

agronómicas más importantes, pocos son los genetistas mexicanos que se han dedicado a tratar de contestarlas. La trascendencia científica del maíz ha sido tan importante que algunos de los investigadores que se dedicaron al estudio de los procesos genéticos básicos que controlan su desarrollo fueron galardonados con el Premio Nobel. Desafortunadamente, ninguno lo obtuvo por descubrimientos realizados en México...

A principios del año 2001, un grupo de científicos mexicanos decidió que era urgente tratar de revertir esta tendencia lo más rápido posible (Tabla 1). Con el apoyo directo del CONACYT y de SAGARPA, y la participación decidida de la SEP y del Gobierno del Estado de Guanajuato, este grupo implementó una serie de iniciativas cuyo objetivo final es obtener la secuencia completa de todos los genes del maíz antes de finales de 2006. En los siguientes párrafos describimos las actividades del proyecto ZEA-2006 que se realiza en el Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO) del CINVESTAV – Campus Guanajuato, y analizamos su importancia a la luz de los avances más recientes en materia de desciframiento genómico a nivel internacional.

Tabla 1. Investigadores CINVESTAV participantes en el proyecto ZEA-2006

- Gustavo de la Riva
Postdoctorante del proyecto.
- Andrés Estrada-Luna y Cesar Alvarez Mejía
Auxiliares de Investigación. Nuevas estrategias de mutagénesis insercional.
- Araceli Fernández Cortés
Auxiliar de Investigación. Bioinformática y ensamblado genómico.
- Gustavo Hernández Guzmán
Postdoctorante. Generación de bibliotecas genómicas y alternativas de enriquecimiento en genes.
- Alfredo Herrera Estrella.
Investigador Titular. Generación de bibliotecas genómicas y substractivas de maíz.
- Luis Herrera Estrella
Investigador Titular y Director de LANGEBIO. Estudio de la respuesta a la asimilación de nutrientes.
- Enrique Ibarra Laclette.
Auxiliar de Investigación. Bioinformática y microarreglos.
- Beatriz Jiménez Moraila, Guillermo Corona Armenta
Coordinación del Laboratorio de Secuenciación.
- Octavio Martínez de la Vega.
Investigador Titular. Bioinformática y generación de bases de datos. Ensamblado genómico.
- Raymundo Méndez Canseco.
Auxiliar de Investigación. Secuenciación.
- June Simpson Williamson.
Investigadora Titular. Estudio de la respuesta del maíz a la sequía.
- Julio Vega
Postdoctorante. Generación de bibliotecas genómicas y alternativas de enriquecimiento en genes.
- Jean-Philippe Vielle Calzada
Investigador Titular. Estudio del tamaño del genoma y desarrollo reproductivo.

La Herencia Genética del Maíz

A pesar de los avances logrados en las últimas décadas por el mejoramiento genético tradicional, justo es mencionar que la mayoría de las características genéticas que hacen del maíz una planta excepcional fueron obtenidas por el esfuerzo intuitivo de las culturas prehispánicas. En términos de rendimiento en grano, las características esenciales de la mazorca existen desde tiempos milenarios, y se estima que solo el 5% de las 260 colecciones de maíz presentes en germoplasmas del mundo han sido utilizadas en programas de mejoramiento genético. Este porcentaje se reduce dramáticamente si tomamos en cuenta que el número real de criollos y variedades presentes en México sobrepasa con mucho cualquier colección de germoplasma mantenida por los centros de investigación internacionales.

Lo que resulta también innegable es que algunos de los descubrimientos más importantes en materia de genética clásica se obtuvieron gracias al estudio del maíz en los Estados Unidos de Norteamérica, cuando la mosca de la fruta (la famosa *Drosophila*) y el maíz competían estrechamente como los dos

modelos experimentales más influyentes en biología. A principios de 1930, el grupo de genetistas de maíz más importante de la historia se encontraba en la Universidad de Cornell (Nueva York), en donde George Beadle, Charles Burnham, Marcus Rhoades y Barbara McClintock – todos bajo la tutela del Dr. Rollins Emerson – comenzaron los primeros estudios clásicos de la planta (Figura 1). Con trabajos que se prolongaron por varias décadas, y antes de que se descubriera la estructura del ADN y su importancia clave como portador de la información genética, este grupo estudió fenómenos tan diversos como la esterilidad masculina, la estructura de los cromosomas, y la formación de semillas. El mayor logro científico del grupo de Cornell estuvo a cargo de Barbara McClintock (1902-1992) quien logró demostrar que ciertas propiedades de pigmentación del grano eran el resultado de la actividad de “elementos móviles” desconocidos, capaces de cambiar su localización dentro del material genético y alterarlo profundamente por mecanismos que violaban las más estrictas reglas de la genética tradicional; a dichos elementos los llamó “transposones”. Su descubrimiento le valió el Premio

Nobel de Medicina en 1983, y ha permitido estudiar la función de los genes de múltiples especies animales y vegetales al permitir el uso de los transposones para descomponer genes a voluntad y estudiar las consecuencias de dicha “descompostura” en el crecimiento de la planta. Antes, en 1958, George Beadle otro miembro de este grupo, recibía también el Nobel por haber descubierto que los genes actúan regulando eventos bioquímicos específicos. Durante los últimos 80 años, varias generaciones sucesivas de científicos derivados del pequeño grupo de Cornell han contribuido enormemente a enriquecer nuestros conocimientos de genética básica, de la biología de plantas, y del desarrollo del maíz.



Figura 1. Genetistas de maíz de Cornell en 1929. De izquierda a derecha: Charles Burnham, Marcus Rhoades, Rollins Emerson, Barbara McClintock y George Beadle (deteniendo al perro). Beadle y McClintock fueron galardonados con el premio Nobel en 1958 y 1983 respectivamente (cortesía del laboratorio Cold Spring Harbor).

¿Por qué Secuenciar el Genoma del Maíz?

Codificado en forma de ADN, el “genoma” es el conjunto global de información genética que dirige a las células para que formen un organismo vivo, lo hagan funcionar, lo distingan y lo identifiquen. Las principales unidades informativas del genoma son los genes, y se acepta generalmente que cada gen constituye el código de lectura de una proteína que debe cumplir una función en el organismo al que pertenece. La genómica estudia la organización y la función de los genomas con la finalidad práctica de transformar los métodos de mejoramiento tradicional y las estrategias biotecnológicas modernas.

Obtener la secuencia codificante de todos los genes del maíz no es ninguna frivolidad académica. A la entrada del nuevo milenio, 50 años después del descubrimiento de la estructura del ADN, la genética molecular ha dado lugar a una nueva revolución biotecnológica que transforma por completo las estrategias agrícolas y comerciales que imperan en el planeta. El componente esencial de esta revolución es el uso de herramientas moleculares, que ofrecen alternativas concretas para controlar

características vegetales con base en modificaciones del genoma vegetal. La implementación de este tipo de modificaciones depende del descubrimiento de genes cuya secuencia es obtenida - de manera rutinaria y sistemática - a partir de costosos proyectos de análisis genómico que se han desarrollado por lo general en países industrializados (Estados-Unidos, Japón, y la Comunidad Europea). Estos proyectos - financiados tanto por el sector público como por el privado - se proponen determinar la secuencia completa del genoma de un número creciente de especies cuyo interés científico y económico es incuestionable. Entre estas especies destacan, por su valor estratégico y comercial, los cereales comestibles. En 2004, la secuencia completa del genoma del arroz fue descifrada por un consorcio de laboratorios chinos, lo que constituyó el primer ejemplo a nivel mundial de secuenciación completa de un organismo vegetal por una sola Nación. Con esta iniciativa China le mostró al mundo de manera decidida que su apuesta científica y tecnológica contempla la genómica como punta de lanza de sus esfuerzos en materia biotecnológica. Como China, los países industrializados y las

grandes empresas agro-industriales confían en que el monto de sus inversiones actuales resulte insignificante frente a la abundancia monetaria que la revolución biotecnológica promete. El potencial monetario de la genómica depende en gran medida de la capacidad para proteger la propiedad intelectual correspondiente a la enorme cantidad de genes que se identifican a partir de estos proyectos y que serán usados para procesos biotecnológicos como para programas de mejoramiento genético, tanto por métodos convencionales como de ingeniería genética.

454: el Decodificador Genómico más Veloz del Mundo

No se pronuncia *cuatrocientos cincuenta y cuatro*, ni se trata del nuevo dispositivo táctico de Lavolpe para el balonpié nacional. Es “*four-five-four*” o “*cuatro-cinco-cuatro*”, el recientemente comercializado aparato de secuenciación que desde el año pasado ha revolucionado las tecnologías de desciframiento del ADN. Su tamaño no es mayor al de una fotocopidora convencional (Figura 2), y su manejo es más fácil

que el de una computadora portátil. Su ingenioso y eficaz sistema de lectura le permite alcanzar velocidades de secuenciación 100 veces superiores a la de sistemas de secuenciación convencionales que prevalecieron por más de 30 años. *454* es en realidad un conjunto tecnologías que permiten realizar lo que nunca se había logrado: descifrar más de 20 millones de letras (A,T,G o C, las letras del alfabeto del ADN) en menos de 5 horas. Para ello, la cadena completa del ADN de un organismo cualquiera se corta en pequeños pedazos de aproximadamente 700 letras. Cada uno de estos pedazos se liga a millones de micro-esferas de menos de 0.04 milímetros de diámetro. El aparato secuenciador *454* inmoviliza cada una de estas micro-esferas en una placa perforada que contiene 860,000 micro-celdas de 0.044 milímetros de diámetro en las cuales solo cabe una micro-esfera por micro-celda. Una reacción química emite luz cada vez que se lee una de las letras contenidas en el pedazo de ADN de las micro-esferas. Finalmente, una cámara digital de alta resolución permite detectar la señal luminosa y guardar toda la información en un banco de datos. Cada pedazo de secuencia leído forma una pieza del rompecabezas genómico que

poderosos sistemas de cómputo permiten volver a armar. El resultado final es la secuencia completa de todo el genoma del organismo elegido. *Consumatum est*: en menos de 2 años *454* convirtió a todos sus competidores en chatarra tecnológica costosa y caduca; así son las reglas actuales de la competencia genómica...



Figura 2. *454*, actualmente el secuenciador más veloz del mundo.

ZEA-2006: la Secuenciación Completa de los Genes del Maíz en México

A principios de 2001, con la ayuda de un donativo del Banco Mundial canalizado a través del Consejo Nacional de Ciencia y Tecnología y su programa de campos emergentes,

emprendimos la primera fase de una serie de proyectos encaminados a caracterizar la función del genoma del maíz mexicano. Conociendo el monto de la inversión norteamericana para proyectos genómicos de maíz – en aquel entonces casi 40 veces superior a la de nuestro país – decidimos enfocar nuestro esfuerzo en tres aspectos biológicos fundamentales para el cultivo del maíz en México: la identificación de genes importantes para la resistencia a la sequía, la identificación de genes involucrados en la asimilación de nutrientes (en particular fósforo), y la identificación de genes esenciales para la formación de semillas. Después de 5 años de trabajo conjunto que involucra a 5 grupos de investigación de la Unidad Irapuato del CINVESTAV, identificamos más de 12,000 genes de maíz que participan en alguno de estos procesos. Esta información es actualmente objeto de un análisis bioinformático continuo que nos permite seleccionar aquellos cuya función es esencial o importante para estrategias de mejoramiento genético. La información se encuentra organizada en una serie de bases de datos que pueden ser accedidas por usuarios nacionales con previa autorización.

El éxito de esta primera fase nos permitió establecer las bases del proyecto ZEA-2006 cuyo objetivo es obtener por primera vez a nivel internacional la secuencia completa de todos los genes del maíz, antes de finales de 2006. Se calcula que el genoma del maíz consta de aproximadamente 2,000 millones de nucleótidos (las letras del ADN), lo que representa aproximadamente 4 veces el tamaño del genoma del arroz (*Oryza sativa*) y 60% del tamaño del genoma humano. Se calcula también que el genoma del maíz contiene cerca de 55,000 genes, casi 2 veces más que el número de genes que se encontró en el genoma humano (es otra de las reglas de la genómica: no por tener un genoma más chico se tienen menos genes...). El apoyo entusiasta de SAGARPA como fuente financiadora de ZEA-2006 nos permitió adquirir un secuenciador 454 que agiliza sustancialmente la tarea (ver sección anterior). Además, gracias a la enorme diversidad genética que prevalece en México, en 2005 pudimos identificar una raza de maíz *palomero* originaria del centro del país que tiene un tamaño de genoma 25% inferior al de B73, la variedad que constituye la referencia de los estudios genómicos que prevalecen en el extranjero. Además de

su importancia económica (se calcula que tan solo los ciudadanos norteamericanos consumen más 17,000 millones de litros de palomitas de maíz al año), estos estudios nos han permitido determinar que el maíz *palomero* presenta un genoma más compacto, y con una densidad de genes mayor a la normal, lo que confirma que las razas de maíz mexicano presentan una estructura genómica distinta a la de los maíces de regiones templadas. Por lo mismo, la posibilidad de identificar genes cuyas funciones sean diferentes y relevantes para el agro mexicano realza la trascendencia del presente proyecto.

La elucidación completa de los genes del maíz constituye el primer proyecto de secuenciación genómica masiva que se realiza en México. El proyecto se lleva a cabo actualmente en el Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO), cuyas instalaciones están siendo construidas en el Campus del CINVESTAV de Irapuato Guanajuato. Con esta iniciativa el CINVESTAV asume el liderazgo que le corresponde en materia de investigación genómica, un liderazgo que se justifica en el marco del estudio y la preservación de la extraordinaria

biodiversidad genética que existe en nuestro país.