

ARTÍCULO

EL SECRETO DEL CHOCOLATE

Alejandro Garciarrubio

El secreto del chocolate

Resumen

Como resultado de un enorme esfuerzo multi-nacional, recientemente se publicó el primer borrador de la secuencia genómica de *Theobroma cacao*, planta de cuyas semillas se obtiene el chocolate. Además de ser una aportación significativa al conocimiento biológico, la secuencia constituye una valiosa herramienta para el mejoramiento racional del cacao. La intención del presente artículo es dar a conocer a un público no especializado los métodos, resultados y alcances de ese trabajo.

Palabras clave: cacao, chocolate, genoma, genómica, secuenciamiento.

The secret of chocolate

Abstract

Ensuing from a broad international effort, a recent publication reported the first draft of the complete genome sequence of *Theobroma cacao*, plant from whose fruit chocolate is obtained. Besides being a significant contribution to basic biological knowledge, the sequence provides a valuable tool for rational plant breeding. The present review attempts to bring to the non-specialized audience some understanding of the methods, results and implications of that work.

Keywords: cacao, chocolate, genome, genomics, sequencing

Introducción

Un secreto milenario

¡Nada endulza sonrisas, levanta enfermos, resuelve querellas y reconcilia amores como el chocolate! ¡Bendita droga permitida! Mientras que el alcohol embrutece, y el café desata ansiedades, el chocolate nos torna en seres mejores: joviales, generosos, apacibles, compasivos..., en suma, nos acerca a los ángeles. ¿Cuál es el secreto? ¿Qué afortunada mezcla de alcaloides, qué soborno a los sentidos, al gusto, al olfato, al tacto, a la vista..., a todos al mismo tiempo, qué despertar en la memoria de momentos felices y cuentos entrañables, explicará, cuando al fin la entendamos, la magia del chocolate?

Quizás un paso importante para llegar a la respuesta haya sido la reciente develación de otro secreto. La historia del chocolate está salpicada de secretos. Pero no me refiero al misterio de cómo incubar una cereza en licor dentro de un vientre de chocolate, ni cómo albinizar al oscuro chocolate, ni cómo lograr la cristalización perfecta, que permite que una barra sólida se derrita con el insinuado vaho de nuestra boca. Los secretos de la industria del chocolate son muchos, pero yo hablo de uno mucho más profundo: cómo hacer el cacao. Entiéndase bien: no cómo

sembrarlo, cuidarlo y cosecharlo, o cómo fermentarlo, procesarlo y fraccionarlo. Hablo de cómo hacer el cacao: cómo hacer el embrión, el tallo, el follaje..., cómo hacer que se nutra de fotones y aire, cómo crear un árbol que cargue sus semillas de lípidos y aromas deliciosos. Este es un secreto mucho mayor y más valioso que cualquier otro que la industria chocolatera esconda. Es tan valioso que se ha conservado en millones de copias, reproduciéndolo generación tras generación, para que no se pierda. Es tan detallado y complejo que apenas podría escribirse en diez tomos del tamaño del directorio telefónico de la ciudad de México. Sus elementos especifican una red tan robusta y complicada de interacciones, que la red de Facebook parece juego de niños. Este secreto es, pues, un tesoro tan codiciable, que no es sorprendente que se invirtieran millones de dólares en una operación de espionaje para obtenerlo.

El gran rompecabezas

Pero esta historia no está plagada de engaños, sobornos, extorsiones ni asesinatos. Manipulación, mentiras y traiciones no fueron las armas de estos detectives. Eso sí, tuvieron en sus manos tecnología tan, tan sofisticada que haría al arquetípico James Bond palidecer de envidia. El secreto, sobra decirlo, estaba en el genoma del cacao, copiado en el núcleo de cada célula de cada planta de *Theobroma cacao* (nombre completo, del cual “cacao” o “chocolate” son simples apodos cariñosos). Como todos los genomas (ya sea el de la mosca o el nuestro), el genoma del cacao está escrito en hebras de ADN (la famosa doble hélice, la “molécula de la vida”) que representan un texto, pasmosamente extenso, en un alfabeto de tan sólo cuatro letras (adenina, guanina, citosina y timina, o A, C, G y T, para ser más breves). El genoma de cacao está contenido en diez grandes tomos (cromosomas) de los cuales cada núcleo guarda dos copias. Los detectives de genomas tienen una forma curiosa de investigarlos. Primero obtienen muchas copias de ellos (una sola hoja de cacao tiene varios cientos de millones de núcleos), purifican el ADN, y luego estallan esa mezcla de copias del genoma en millones de diminutos fragmentos. Lo que resulta es un precioso rompecabezas desarmado.

La razón de esta rara costumbre es que, a pesar de sus sofisticados métodos, estos detectives son incapaces de leer más que unas cuantas letras, lo que hacen a partir de los extremos de cada molécula de ADN. Por eso, en el caso del cacao tuvieron que romper los diez cromosomas (moléculas gigantes de ADN que contienen decenas de millones de letras) en muchísimos pedazos más pequeños, para tener millones de extremos a los que leerles de 40 a 800 letras. Como es fácil imaginar, si uno comprara el tiraje completo del *New York Times* (unos cuantos miles de camiones), los hiciera explotar con dinamita y luego recogiera los millones de fragmentos generados, resultaría muy difícil armarlo para recuperar todas las notas y noticias que se publicaron ese día. Sobre todo, sería casi imposible reconstruir alguno de los ejemplares en particular. Más bien, uno iría pegando con cinta adhesiva los fragmentos que pensara que se traslapan, pues

—por venir de la misma edición— en la región compartida contendrían idéntica información. Un mejor símil de lo que hace un detective de genomas es tomar al azar uno de los millones de fragmentitos de periódico, leerlo, y pasar lo leído a una computadora. Repetir el proceso, leyendo millones de fragmentitos de periódico, hasta que la estadística diga que casi no hay ninguna parte del texto que no haya sido leída al menos unas cuantas veces (algunas partes, por puro azar, habrán sido leídas cientos de veces), y pedirle a un software especializado que, con base en las coincidencias, una las pequeñas lecturas en cadenas de texto más grandes. Procediendo así, cada vez se tendrán fragmentos más grandes: uno de ellos podría incluir $\frac{3}{4}$ partes de la primera página, otro se extendería desde la cuarta página de sociales hasta la mitad de la quinta, etcétera. Lo mismo sucede con los genomas: se logran armar fragmentos de texto muy grandes, aunque, casi por regla general, no se logran reconstruir cromosomas completos.

En el caso particular del cacao, cuyo genoma es un texto de unas 430 millones de letras (la suma de los diez cromosomas), los detectives leyeron 26 mil millones de letras a partir de los extremos de los fragmentos producidos. En el lenguaje secreto de estos espías, a esto se le llama leer un genoma “54X”, lo cual quiere decir que, en promedio, cada parte del genoma se leyó 54 veces (ahora ya no es un secreto). En vez de diez cromosomas, lograron armar 4800 grandes fragmentos, de diversos tamaños, pero con una longitud media de 470 mil letras. La suma total de los fragmentos dio poco menos que 327 millones de letras, lo que implica que en la parte del rompecabezas que no se logró armar queda aún el 24% del genoma del cacao. Esto no fue una gran desilusión: en todo rompecabezas hay regiones que son muy difíciles de armar, sobre todo aquellas donde las piezas carecen de personalidad, pues se parecen mucho entre ellas (como las zonas de follaje en un rompecabezas que represente una escena en un parque). Por suerte, esas regiones suelen ser las menos importantes, las menos informativas (el follaje no se armó, pero la pareja de novios en la banca y el niño con el perro se reconstruyeron a la perfección).

Fragmentos notables

De los genomas, las partes más importantes son los genes, es decir, los capítulos que definen cómo construir las piezas básicas de las que se ha de formar cada célula. Estas piezas, a las que los detectives de genomas llaman proteínas, han sido perfeccionadas a lo largo de millones de años, a tal grado, que parecieran tener vida propia: se reconocen unas a otras, perciben el entorno y actúan de acuerdo con éste, se trasladan de un sitio a otro, fabrican otras proteínas e incluso construyen la materia prima de la que éstas, o cualquier otro componente celular, se forman. En fin, es su preciso actuar lo que sustenta la armonía, dinamismo y estabilidad de las células. El 98% de los genes del cacao está en ese 76% del genoma que se pudo armar en textos largos, y casi ninguno se esconde en las regiones que no se armaron.

Para entender cómo es que se puede estar tan seguro de esto, es necesario hablar de otras moléculas de las células: los ARN. Cada ARN es como una fotocopia de un gen. Si usted fuera el presidente de una empresa con millones de empleados (como lo es una célula), y tuviera una única copia del manual de construcción (el genoma), seguramente haría fotocopias de algunos capítulos y no de otros, según el momento lo requiera. Lo fotocopiaría por dos razones. En primer lugar, su único manual original es demasiado valioso como para andarlo exponiendo a la intensa, diversa y, en ocasiones, peligrosa actividad de la empresa. En segundo lugar, dado que algunas actividades son mucho más frecuentes que otras, una sola copia de esas instrucciones sería penosamente insuficiente. Por eso, la célula hace fotocopias de sus genes en ARN. Para los detectives de genomas eso es muy conveniente: ya que los ARN son copias fieles de porciones del genoma. Si uno sabe el texto de un genoma y el de un ARN, podrá decir con certeza qué región del genoma está contenido en ese ARN, y que ésta constituye un capítulo (un gen) que define a una proteína.

Cabe aclarar que, hasta la fecha, el lenguaje del genoma (su vocabulario, puntuación y gramática) nos es completamente desconocido; un inconveniente mayúsculo para cualquier labor detectivesca. Agravando la situación, los textos verdaderamente importantes del genoma están inmersos en un océano de textos mucho menos importantes, textos olvidados por burocracias anteriores a lo largo de millones de años. Por ello, en el caso del tabaco, los detectives obtuvieron 700 mil "fotocopias" de ARN que, después de haber sido leídos y armados, por un método análogo a como se lee un genoma, resultaron tener 29 mil contenidos diferentes. Intencionalmente los ARN se obtuvieron de distintos tejidos, pues si, por ejemplo, sólo se hubieran obtenido de raíz, faltarían allí muchos ARNs que son importantes para la construcción de hojas, o tallos, u otro tipo de tejido. Al buscar los 29 mil contenidos en la parte armada del rompecabezas del genoma, se vio que prácticamente todos estaban allí, comprobando que la parte que fue imposible armar estaba casi desprovista de genes.

Deja vu

Tener el texto completo de un genoma y localizar todos los sub-textos que representan genes, es apenas el primer paso para entender cómo se construye y funciona un organismo. Es como el manual de esos aparatos que se compran desarmados: en la primera página suele venir una imagen con las distintas piezas numeradas. Eso es lo que se tiene ahora: una lista de las piezas con las que, si supiéramos hacerlo, podríamos armar un cacao. En primera instancia, no se sabe para qué sirve cada pieza, ni cuál interacciona con qué otras. Por supuesto, ayuda si alguna pieza es reconocible (esto parece un tornillo y aquello una tuerca), lo cual sólo es posible si ya se la ha visto con anterioridad en otro lado. En el caso del cacao, ese "otro lado" serían los manuales de construcción de otras especies vegetales. Los detectives de esta historia buscaron piezas que

fueran comunes con los genomas de uva, soja, álamo y Arabidopsis, un pariente de la mostaza de poco interés comercial, pero que es la planta más investigada, por las ventajas que da para su estudio en laboratorio. Quizás no sea el procedimiento más sutil, pero para investigar qué función tiene la pieza (proteína) descrita en cierto gen, no tenemos mejor manera que destruir el gen y observar qué propiedades de la planta se alteran o pierden. El experimento es arduo si se intenta con plantas como el cacao, que son árboles, pero es un poco más sencillo en Arabidopsis, que llega a adulto reproductivo en tan solo dos semanas, y cuyo diminuto tamaño permite cultivar miles de individuos en una cajita no mayor que un plato de café.

Volviendo a la página uno del manual de construcción, es obvio que los tornillos de media no pueden sustituir a los de una pulgada, o los de rosca derecha a los de rosca izquierda. Aun así, podemos generalizar que tienen funciones parecidas. Lo mismo sucede con las piezas (proteínas) de una célula: cuando se parecen lo suficiente podemos asumir que tienen funciones relacionadas (que forman “familias”, según el argot de los detectives de genomas). Interesantemente, esto se puede hacer también para las piezas cuya función se desconozca. No es necesario saber qué utilidad tienen las rondanas para intuir que todos los aros aplanados, en sus distintos tamaños, sirven para más o menos lo mismo. Los 28 mil genes del cacao fueron agrupables en 7100 familias de piezas, de las cuales 6400 ya habían sido vistas en Arabidopsis, o en alguna de las otras tres plantas comparadas. Muchas de estas familias sustentan funciones metabólicas y celulares fundamentales; es decir, pertenecen a procesos esenciales para la vida de las plantas. Otras 700 de estas familias sólo están en el cacao y, de momento, casi nada se puede decir sobre ellas.

En busca de El Dorado

El texto recién leído de un genoma es un vasto territorio sin explorar; imposible de estudiar en toda su extensión. Por eso es mucho mejor centrarse en unos cuantos aspectos de él. Los detectives que iniciaron este esfuerzo tenían una idea específica de dónde buscar el oro: en los genes de resistencia a patógenos, y en los que afectan la textura, sabor y aroma del chocolate.

La relación (¿enamoramiento?) de nuestra especie por el cacao es sumamente reciente: se cree que inició en Mesoamérica hace unos tres mil años, cuando algún antepasado visionario comenzó su domesticación. Mucho, muchísimo antes, millones de años atrás, otras especies ya habían desarrollado un enorme apetito por el cacao. Entre ellas se encuentran hongos, oomicetos e insectos. De la lógica más elemental se desprende que estos parásitos del cacao deban ser considerados por nosotros, los humanos, como nuestros enemigos naturales, a los cuales debemos dar guerra sin cuartel. Y sin duda la mejor forma de hacer esto es ayudar al cacao a que se defienda él mismo. Por eso, uno y cada uno de los proyectos de fito-mejoramiento del cacao

está dirigido a obtener variedades resistentes. Por supuesto, hasta cierto punto, todas las plantas se saben defender de sus patógenos. La pregunta es cómo hacer para que lo hagan mejor, y la respuesta a esa pregunta comienza por ver cómo lo han hecho hasta ahora. Muchas de las familias de genes que resultaron comunes entre *Arabidopsis* y cacao, se sabe que en la primera son parte de los mecanismos de defensa. Estas familias pertenecen a dos grandes clases (cuyos nombres, por no ofuscar al lector, omitiré). De la primera clase, cacao tiene 253 genes, y de la segunda, 297. Se cree que algunos de estos genes participan en respuestas a la agresión de patógenos muy particulares, mientras otros son parte de respuestas generales. En resumen, ahora los detectives tienen 550 interesantes genes que investigar, y los fito-mejoradores 550 cartas a las que apostar.

“Lípidos de almacenamiento, flavonoides, proantocianinas, teobromina, terpenos aromáticos”... La sola mención de estos nombres nos debería hacer salivar. Son las moléculas que le dan al chocolate sus cualidades incomparables. Claro que para la planta tienen un significado muy distinto. Los lípidos, que forman el 50% del peso de la semilla, son las reservas que alimentarán al embrión cuando la semilla germine. Los flavonoides le sirven de defensa contra insectos, microbios y otros patógenos, y le protegen de la luz ultravioleta y del poder oxidante de los radicales de oxígeno. Las proantocianinas son cadenas de flavonoides en reserva. Los terpenoides resinosos son hormonas y pigmentos, y también participan en la defensa. Los monoterpenos y sesquiterpenos, que son los terpenoides más volátiles, son aromas (como lo son en el perfume de las flores) y sirven para atraer polinizadores y otros animales benéficos. La síntesis de todas estas maravillosas moléculas es llevada a cabo por las proteínas que se describen en el genoma del cacao. Cuando éstas se comparan, familia a familia, con las de *Arabidopsis*, se entienden muchas de las propiedades del chocolate. Por ejemplo, 84 de los genes de cacao participan en la síntesis de lípidos y el genoma codifica un exceso de dos proteínas, que desvían los recursos de la célula hacia la producción de determinados ácidos grasos, que constituyen sus triglicéridos. Esto explica porqué los lípidos de chocolate tienen una temperatura de fusión alta (se derriten alrededor de los 36° centígrados, mientras los lípidos de otras semillas son líquidos a temperatura ambiente; como el aceite de girasol, por mencionar alguno). Otro ejemplo, la semilla del cacao es extremadamente rica en dos flavonoides particulares (catequina y epicatequina), y para cierto paso de la “cadena de producción” el cacao tiene 18 copias del gen, en tanto que *Arabidopsis* tiene sólo una. No hace falta que mencione más ejemplos; todos tienen la misma lógica, aunque las moléculas sean diferentes. Baste decir que son cientos los genes a los que debemos el placer de saborear un delicioso chocolate.

Los Dioses “hackeados”

Según una leyenda azteca, el cacao era alimento exclusivo de los dioses, pero Quetzalcóatl

regaló unas semillas a los toltecas y les enseñó a preparar chocolate. El pueblo tolteca prosperó en riqueza y sabiduría pero, un día, los dioses se dieron cuenta de que su hermano Quetzalcóatl les había robado esas semillas para dárselas a ellos, y que por eso ahora los toltecas bebían de ese elixir que era privilegio de los dioses. Decidieron vengarse. Un dios se disfrazó y convenció a Quetzalcóatl de que probara una nueva bebida que le quitaría todas sus penas: el pulque (tlachihuitli). Tras una enorme guarapeta, al día siguiente Quetzalcóatl se sintió tan mal por la vergüenza (y por la cruda, imagino yo), que decidió marcharse del pueblo, llevándose consigo las semillas. Por fortuna, en el camino unas poquitas de ellas se le cayeron de la mano (al parecer, por la región de Tabasco) y todo el cacao que ahora tenemos descende de ellas. ¿Qué de esta leyenda es cierto?... ¡Ya el genoma nos lo dirá!

Nota

La secuencia genómica del cacao fue obtenida por una colaboración de instituciones públicas y privadas (Chocolates Hershey's,) de varios países, incluido los Estados Unidos de América, Francia, Brasil, Venezuela, Corea y Costa de Marfil. La publicación original tiene esta referencia:

Bibliografía

“The genome of *Theobroma cacao*”, Xavier Argout et al., Nature Genetics, Vol. 43, Pags:101-108 (2011).