

DE LAS TRINCHERAS DEFENSIVAS BACTERIANAS A LA EDICIÓN EFICIENTE DE GENOMAS*

Juan Rafael Riesgo Escovar

Instituto de Neurobiología, Universidad Nacional Autónoma de México,
correo E: riesgo@unam.mx

Desde hace muchos años, la posibilidad de editar genomas a voluntad ha sido uno de los objetivos más acariciados por los investigadores en biología y medicina, particularmente en sistemas en donde la generación y caracterización de mutantes es difícil, o casi imposible. Aún en los sistemas en donde típicamente se generan mutaciones en cantidades suficientes en un tiempo corto, éstas generalmente se inducen de manera aleatoria, por lo que después de generar las cepas o líneas, queda el trabajo de reconocer, aislar y caracterizar a la mutante o mutantes de interés, de manera que aún en estos sistemas genéticos la operación es tardada, y normalmente requiere de mucho esfuerzo, especialmente en sistemas pluricelulares. Por otra parte, las mutaciones no se editan sino que se inducen de manera razonablemente aleatoria, y queda después el trabajo de establecer su naturaleza, y sus efectos. Durante mucho tiempo este fue el único modo de forzar cambios genéticos, y es usado también para, sin prejuicio, generar mutaciones en todos los genes susceptibles de mutar a un mismo fenotipo, o, dicho en otras palabras, aislar a la mayoría de los genes que intervienen en el mismo proceso de interés.

Se han ideado varios sistemas para generar cambios en los genomas que se basan en cortes de doble cadena de DNA en la secuencia diana, con la posterior inducción de la recombinación homóloga o de otros mecanismos de reparación de daño al DNA en la región del corte, para poder editar el genoma celular. Se sabe que generar cortes de doble cadena en la región que se quiere editar, proveyendo al mismo tiempo de una abundancia de copias con una secuencia editada, que sirve de templado durante la reparación del corte del DNA, funciona como método general de edición, siempre y cuando se pueda después identificar y aislar a las células con el cambio deseado.

En el sistema clásico de "knock-out" en ratones, las células troncales que albergan el cambio

deseado se incorporan a embriones tempranos, con la esperanza de que estas células modificadas formen parte de la masa interna (o el embrión propiamente dicho), y sean precursoras de las células germinales, para que las modificaciones se hereden a la descendencia. Sin embargo, y a pesar de sus incuestionables méritos, este proceso es complicado, tardado y caro, por lo que sus aplicaciones han sido exitosas en casos en donde de antemano se desea estudiar cambios particulares en gen o genes donde se tienen ya datos o se tienen evidencias que apuntan hacia la función o funciones que el gen desempeña. No se puede aplicar para estudiar y cuantificar el número de genes que se requieren para un proceso, por ejemplo, algo que si se puede cuando se generan mutaciones al azar y se aíslan cepas o líneas con fenotipos mutantes comunes.

Generar cortes de doble cadena secuencia-específicos en el DNA genómico es el paso crítico, pero es a la vez algo bastante difícil de lograr. Recientemente, se han ideado varios sistemas para poder generar esto: por ejemplo utilizar la especificidad de los dominios de dedos de zinc (que reconocen secuencias específicas en el DNA) acopladas a nucleasas inespecíficas, pero en todos los casos, los procedimientos son complicados, requieren de un esfuerzo considerable, despliegue técnico alto y tardan bastante tiempo.

Es en este escenario que se inserta el reciente (1) método de edición genómica conocido como CRISPR/Cas9. Es el más sencillo, barato y flexible de todos los métodos hasta hoy disponibles (2, 3). Muy recientemente se ha ampliado este método para poder generar cortes en moléculas de RNA, con lo que, teóricamente, se podrían generar individuos en donde el DNA está intacto, pero se perturba la expresión de uno o varios RNAs derivados del locus (4). Es decir, los métodos de CRISPR-Cas pueden servir para editar no sólo el DNA, sino también moléculas de RNA, en una suerte de equivalente

del método de RNAi. Falta aún saber cuál es la eficiencia del método aplicado al RNA, y que tan general es su aplicación, porque implica expresar una proteína, la C2C2, que tiene actividad de RNA-sa, junto con un RNA de tamaño pequeño que sirve de guía. Aún así, las posibilidades son muchas.

Sin embargo, ¿de dónde vienen estos sistemas de modificación de DNA (y RNA)? ¿Cuál es la función que normalmente llevan a cabo en los organismos? Las bacterias y otros microorganismos se han adaptado para contender con la invasión frecuente de ácidos nucleicos foráneos, y ataques a la integridad de su genoma. Para ello disponen de un arsenal de mecanismos de defensa, entre los cuales se cuentan las estrategias de enzimas de restricción y metilasas y los sistemas de inmunidad CRISPR/Cas.

Una de las estrategias mejor descritas es la que se basa en las enzimas de restricción, que son endonucleasas secuencia-específicas que cortan el DNA foráneo (el DNA propio que contiene la misma secuencia está protegido por metilación). De ahí el nombre de endonucleasas de restricción, porque restringen la invasión de DNA foráneo, que en muchos casos es debido a la incorporación del material genético de bacteriófagos, DNA invasor normalmente suficientemente largo como para que existan sitios de restricción en el mismo.

Además de esta estrategia, también tienen un sistema inmune, que los protege de reiteradas invasiones de las mismas secuencias de ácidos nucleicos. Estos sistemas constituyen sistemas de defensa inmune, pues guardan memoria de anteriores invasiones de ácido nucleico foráneo, y son capaces de montar una respuesta a estos retos.

Desde hace bastante tiempo se han descrito *loci* constituidos por repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas, o CRISPR, por su acrónimo en inglés, en los genomas de muchas bacterias. Son *loci* transcritos del genoma. En ellos las secuencias cortas o espaciadores son pequeñas secuencias insertadas entre las secuencias palindrómicas, y están formados a partir de fragmentos de los ácidos nucleicos invasores de eventos previos. Se piensa que las enzimas Cas1 y Cas2, presentes en todos estos sistemas, son las encargadas de procesar esos fragmentos de los ácidos nucleicos invasores, duplicar la secuencia palindrómica, e insertar la secuencia procesada del ácido nucleico foráneo en el *locus* CRISPR de la bacteria. En un estudio sobre estos sistemas, se examinaron 703 genomas de Archaea y Eubacterias, y de ellos el 44% tiene sistemas CRISPR/Cas reconocibles (5). En otras palabras, es un sistema muy frecuente y ampliamente difundido en genomas procarióticos.

La manera en que se lleva a cabo la defensa consta de tres pasos: en el primero, conocido como la adaptación, se procesa uno (o en algunos casos, más de uno), de los fragmentos del ácido nucleico invasor, y se inserta en un *locus* CRISPR. En el segundo paso, conocido como expresión, se transcribe un RNA largo de todo el *locus* CRISPR, y se procesa para generar fragmentos de RNA correspondientes a los distintos espaciadores (los espaciadores son secuencias derivadas de ácido nucleico invasor) y secuencia palindrómica. Finalmente, en el tercer paso, conocido como interferencia, se usa ese RNA procesado como guía para reconocer por complementariedad al ácido nucleico invasor, y por medio de la acción de una o varias nucleasas, cortarlo. De esa manera se neutraliza la invasión. Este sistema parece haber evolucionado mucho, ya que actualmente se acepta que dentro de este sistema general hay tres variantes básicas, y dentro de estas, varias subclases. Todas, sin embargo, tienen un *locus* CRISPR y genes Cas, como los genes Cas1 y Cas2.

El método básico de edición genómica CRISPR/Cas9 se basa en la adaptación del sistema homónimo de defensa o inmunidad bacteriano, ideado por las Dras. Doudna y Charpentier. La parte medular del sistemas CRISPR/Cas de edición, parte, como está dicho arriba, en la propiedad de utilizar fragmentos cortos del DNA diana, de manera que estas secuencias cortas sirvan de guía para posicionar nucleasas (Cas9) en secuencias del DNA diana complementarias a estos RNA guía, para generar cortes de doble cadena en el DNA diana.

En los sistemas bacterianos hay varios mecanismos CRISPR/Cas, pero el II es el más sencillo y es el que se ha adaptado a la edición de genomas. Originalmente constaba de tres moléculas: la nucleasa Cas9, que se posicionaba en secuencias específicas del DNA por medio del RNA adaptador, y el tracrRNA, un RNA corto parcialmente complementario a la parte palindrómica del crRNA (RNA guía). Este tracrRNA se encuentra codificado vecino a los genes Cas en otro *locus*.

Los crRNA son los que en el método reconocen por complementariedad al DNA diana. Seguido de la secuencia de reconocimiento entre el crRNA y el DNA debe de existir una secuencia en el DNA diana, denominada PAM (por motivo adyacente protoespaciador, por sus siglas en inglés), que para el caso de Cas9 tiene la secuencia NGG, donde N es cualquier nucleótido, seguido de un par de guaninas. Esta secuencia es crítica para que Cas9 pueda cortar el DNA, de modo que lo que se necesita para dirigir el corte de doble cadena a alguna región del genoma es una secuencia de 22 nucleótidos, con un nucleótido entre la secuencia de reconocimiento

y la PAM. Esto se traduce a una combinación que surgiría de manera aleatoria aproximadamente una vez cada 230,000 combinaciones. Dado que en la secuencia de reconocimiento se permiten algunos errores ("mismatches", en inglés; hasta cinco), esto da como resultado que en el método pueda haber "off-targets" (secuencias que son parcialmente complementarias y que pueden sufrir cortes también); en otras palabras otras dianas que pueden ser cortadas también por Cas9 además de la secuencia diana original. De todas maneras, el método es generalmente muy específico, y si se eligen con cuidado la secuencia o secuencias para Cas9, la edición puede ser muy eficiente. La Dra. Doudna recientemente ha mostrado que parte de esto se debe a que la proteína Cas9 recorre de manera muy eficiente los cromosomas del núcleo de células buscando el ancla del tracrRNA, por medio de microscopía de fluorescencia con Cas9 acoplado a un fluoróforo.

Una de las mejoras introducidas en el sistema fue fusionar el crRNA (o RNA guía) con el tracrRNA, de manera que una sola molécula de RNA sirve de anclaje para Cas9 y de sitio de reconocimiento para el DNA diana. Se ha demostrado que este RNA mixto es tan eficiente como las dos especies de RNA originales, y se le denomina también RNA guía. Si además se construye un vector que contenga la parte del RNAguía codificando para el tracrRNA junto con un sitio para clonar crRNA, se obtiene un sistema muy flexible y sencillo de usar. Sólo se necesita sintetizar el crDNA para el gen diana (generalmente se sintetizan éstos como "primers" o cebadores con sitios de clonaje a los extremos), éstos se clonan en el vector, y se usa este vector para transformar las células o el organismo en donde se quiere editar el genoma, junto con una fuente de Cas9, en la edición de genomas (Fig. 1).

El sistema original utiliza un solo RNA guía para dirigir a Cas9 al sitio diana y generar ahí cortes de doble cadena que serán reparados por los sistemas de reparación de las células, introduciendo mutaciones en la secuencia en un alto porcentaje de las veces (se ha reportado hasta 70% de eficiencia). Sin embargo, más recientemente se han introducido variantes que hacen al sistema más flexible y dirigido: si se elaboran dos RNA guía en el mismo gen diana, se pueden generar deleciones, y si se provee también de un templado en abundancia (que puede ser otro plásmido) con las mutaciones deseadas, se pueden generar reparaciones de la deleción usando como molde al DNA complementario con los cambios deseados. Esta estrategia es bastante eficiente, y se han reportado altas eficiencias de CRISP/Cas9.

Es un sistema universal, ya que se ha utilizado con éxito en muy variados sistemas: células

eucariontes en cultivo, levaduras, y animales modelo como *Danio rerio*, el pez cebra, plantas, el nematodo *Caenorhabditis elegans* y la mosca de la fruta, *Drosophila melanogaster*, entre otros (6, 7). También se ha modificado a Cas9 para que realice cortes solamente en una cadena del DNA. Con dos RNA guías se puede generar un corte en cada cadena de la doble hélice, pero cada corte a una cierta distancia del otro, no en el mismo par de bases del DNA y favorecer así la recombinación homóloga, o bien, usar una Cas9 modificada sin actividad de nucleasa, pero fusionada a GFP (GFP son las siglas en inglés de la proteína verde fluorescente), por ejemplo, para marcar sitios en el DNA genómico, de manera similar a como se hace con FISH (FISH son las siglas de hibridación *in situ* con marcadores fluorescentes en inglés), pero realizado de manera más sencilla.

Recientemente, la Dra. Doudna, entre otros científicos, ha llamado a una moratoria en la aplicación de este método en humanos, propiciado esto último, sin duda alguna, por el reporte de un grupo chino que publicó los primeros resultados de edición génica en embriones humanos triploides, no viables (8). Si bien es cierto que el método ha sido aplicado exitosamente ya en una gran variedad de organismos, y que, en principio, parece ser de aplicación universal, existen todavía problemas como los "off-targets", inherentes a todos los métodos de edición génica descritos hasta hoy. Estos "off-targets" pueden generar cambios en otras regiones del genoma (un solo corte de doble cadena es suficiente), y complicar la interpretación de los resultados. En el caso de embriones humanos, esto puede conllevar a efectos indeseados (algo así como daño colateral), y mientras no se tenga una manera eficiente y clara de separar estos efectos de los deseados, se debe de ejercer cuidado al aplicar estas tecnologías, además de otras consideraciones éticas.

Sin embargo, la aplicación de esta tecnología permite no sólo generar cambios dirigidos en prácticamente cualquier sistema vivo (básicamente en cualquier organismo susceptible de ser transformado), sino además, dada la sencillez y facilidad del método, de generar mutagénesis con muchos RNA guía, yendo más allá del estudio de un gen a la vez (9). En el caso de la mosca de la fruta, se tienen ya descritas varias cepas que expresan Cas9 en las células germinales y que, aunado a la transformación con un RNA guía dirigido a las mismas células germinales, permite generar con relativa facilidad los cambios deseados en el DNA genómico de las células germinales. Los cambios se pueden caracterizar fácilmente amplificando por PCR (reacción en cadena de polimerasa, por sus

CRISPR/Cas9 endógeno

CRISPR/Cas9 para edición genómica

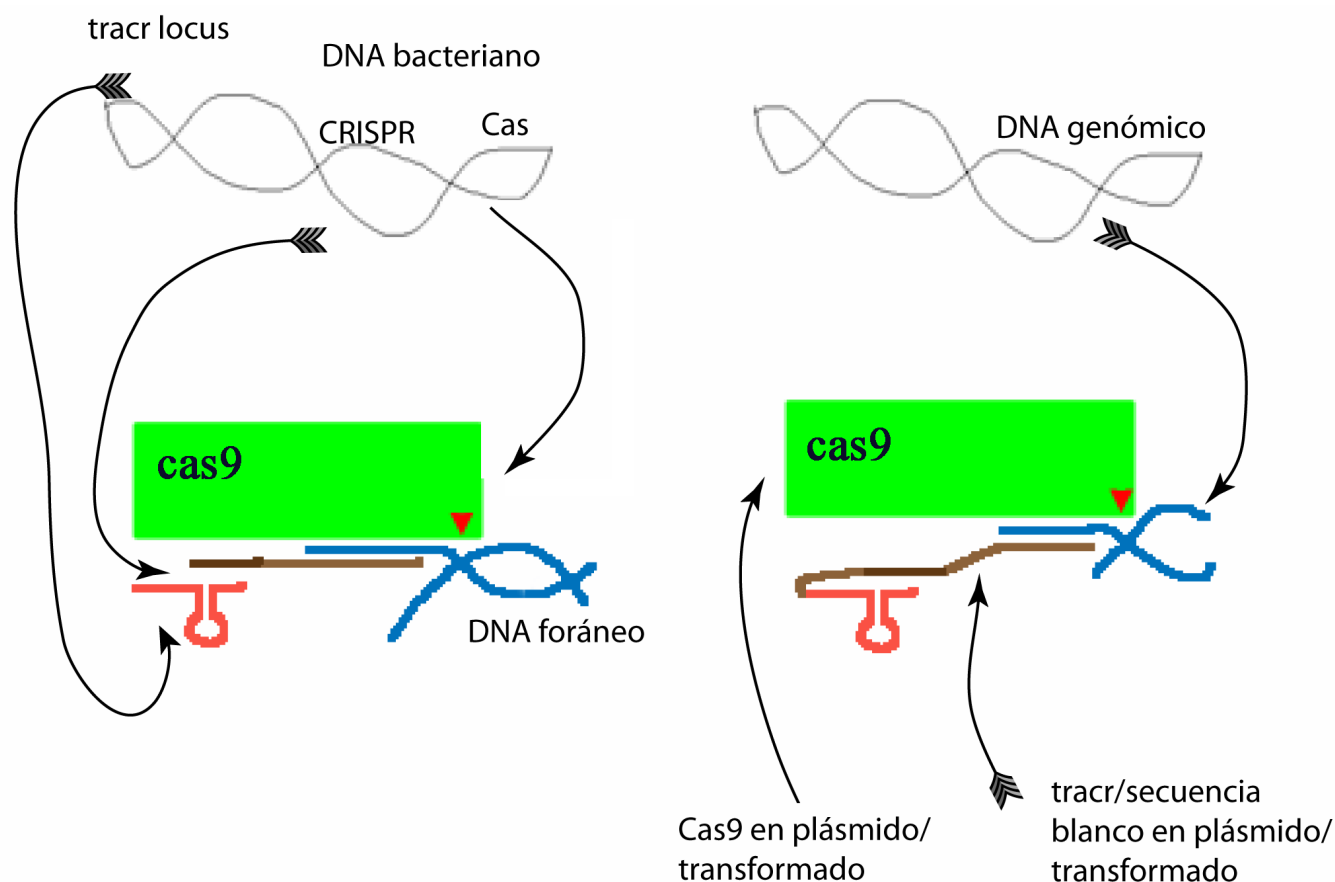



Figura 1. Comparación entre el sistema CRISPR/Cas9 de bacterias, y su adaptación para la edición de genomas. En la ilustración de la izquierda se muestran esquemáticamente los componentes del sistema tipo II de CRISPR/Cas (el sistema CRISPR/Cas9). Consta de tres loci presentes en el genoma bacteriano: el locus CRISPR, que contiene las secuencias palindrómicas y las secuencias espaciadoras derivadas de DNA foráneo; al lado las secuencias de los genes Cas (en este caso Cas1, Cas2 –no mostrados– y Cas9), y el locus tracr, que codifica para un RNA pequeño parcialmente complementario con los fragmentos procesados del locus CRISPR. El DNA foráneo es identificado por complementariedad y cortado en la secuencia PAM al lado de la secuencia CRISPR por Cas9 (triángulo rojo). En la ilustración de la derecha se muestra el sistema tipo para la edición de genomas: Cas9 y un gen quimérico compuesto del locus tracr y la secuencia complementaria a la secuencia genómica a editar son proveídos por transgénesis o por transformación a las células o tejidos blanco. El RNA quimérico tracr/sequencia blanco reconoce por complementariedad a la diana en el DNA genómico, y guía a Cas9 para realizar el corte de doble cadena en la secuencia PAM (triángulo rojo).

siglas en inglés) la región diana, y secuenciando los amplicones obtenidos.

En conclusión, y a pesar de algunos problemas con la especificidad –inherentes a todo método de edición genómica conocido– el sistema CRISPR/Cas9 representa la mejor opción, y en muchos casos, la única opción viable para generar edición

genómica. En consecuencia, la aplicación de este método está abriendo una nueva era en la investigación genética. Han pasado escasos cuatro años desde que se describió el método originalmente, y ya hay una gran cantidad de reportes utilizando esta tecnología. 

REFERENCIAS

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, and Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337: 816-821.
2. Charpentier E, and Doudna JA (2013) Biotechnology: Rewriting a genome. *Nature* 495: 50-51.
3. Doudna JA, and Charpentier E (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346: 1258096.
4. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DB, Shmakov S, Makarova KS, Semenova E, Minakhin L, Severinov K, Regev A, Lander ES, Koonin EV, Zhang F (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* doi: 10.1126/science.aaf5573.
5. Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9: 467-477.
6. Horii T, Hatada I (2016) Production of genome-edited pluripotent stem cells and mice by CRISPR/Cas [Review]. *Endocr J* 63: 213-219.
7. Kondo S, Ueda R (2013) Highly improved gene targeting by germline-specific Cas9 expression in *Drosophila*. *Genetics* 195: 715-721.
8. Liang P, Xu Y, Zhang X, Ding C, Huang R, Zhang Z, Lv J, Xie X, Chen Y, Li Y, Sun Y, Bai Y, Songyang Z, Ma W, Zhou C, Huang J (2015) CRISPR/Cas9-mediated gene editing in human triploid zygotes. *Protein Cell* 6: 363-372.
9. Wade M (2015) High-Throughput Silencing Using the CRISPR-Cas9 System: A Review of the Benefits and Challenges. *J Biomol Screen* 20: 1027-1039.