



**FACULTAD DE FARMACIA**  
**UNIVERSIDAD COMPLUTENSE**

**TRABAJO FIN DE GRADO**

**TÍTULO: EL SISTEMA INMUNE  
BACTERIANO (CRISPR), UNA HERRAMIENTA  
UNIVERSAL EN LA EDICIÓN DE GENOMAS**

Autor: San Félix García-Obregón, Ana

Tutor: Jesus Pla Alonso

Convocatoria: JUNIO 2017.

## ÍNDICE

<b>OBJETIVOS Y METODOLOGÍA DEL TRABAJO.....</b>	<b>3</b>
<b>OBJETIVOS.....</b>	<b>3</b>
<b>METODOLOGÍA.....</b>	<b>3</b>
<b>ABSTRACT.....</b>	<b>3</b>
<b>RESUMEN .....</b>	<b>4</b>
<b>DESARROLLO HISTÓRICO DEL CRISPR .....</b>	<b>4</b>
<b>MECANISMO MOLECULAR DEL SISTEMA CRISPR .....</b>	<b>7</b>
<b>APLICACIONES.....</b>	<b>11</b>
<b>EDICIÓN GENÓMICA .....</b>	<b>11</b>
Edición genómica mediada por Cas9 nativa .....	12
Edición genómica mediada por Cas9 nickasa (nCas9).....	14
Factores influyentes en la aplicación de CRISPR/Cas en células humanas .....	14
<b>MODULACIÓN TRANSCRIPCIONAL BASADO EN Cas9 INACTIVADA.....</b>	<b>14</b>
<b>EDICIÓN EPIGENÉTICA MEDIADA POR Cas9 INACTIVADA .....</b>	<b>15</b>
<b>RASTREO GENÉTICO DE AMPLIO ESPECTRO EMPLEANDO Cas9.....</b>	<b>16</b>
<b>ESCANEEO DEL GENOMA .....</b>	<b>17</b>
<b>CONTROL DE POBLACIONES.....</b>	<b>17</b>
<b>CONCLUSIONES .....</b>	<b>20</b>
<b>BIBLIOGRAFÍA .....</b>	<b>21</b>

## **OBJETIVOS Y METODOLOGÍA DEL TRABAJO**

### ***OBJETIVOS***

Este trabajo se ha realizado con el fin de dar a conocer una de las técnicas pioneras en edición genómica que ha evolucionado enormemente en los últimos años, introduciendo su historia, mecanismo molecular y posibles aplicaciones. Esta técnica conocida comúnmente como CRISPR (*Clustered Regulatory Interspaced Short Palindromic Repeats*) consiste en un sistema inmunitario adaptativo empleado por microorganismos para defenderse de la infección de virus o plásmidos al reconocer y cortar secuencias de DNA procedentes de la invasión vírica.

### ***METODOLOGÍA***

Para realizar el presente trabajo se han empleado portales de búsqueda de documentos científicos como son PubMed, Bucea o Google Scholar donde se han seleccionado los diferentes artículos utilizando las siguientes palabras clave: CRISPR, genome edition, Gene Drive, Cas9, dCas9, crRNA, tracrRNA, immunitary system, applications. Además, también se ha obtenido información en las páginas oficiales de la BBC y la revista Nature, en las cuales se han buscado las noticias más recientes relacionadas con CRISPR. Tras recopilar toda la información, se han destacado las partes más importantes de los diferentes artículos y se han resumido siguiendo el índice presentado.

### **ABSTRACT**

In order to be able to understand the genome and get its easy manipulation, scientists have been studying different ways of Genome Edition for so many years. Nevertheless, it has been the discovery of the CRISPR/Cas system which has broken with all previous known standards of genome edition. Originally, this system derived from the prokaryote's adaptive immune system mediated by clustered regularly interspaced short palindromic repeats (CRISPR). This technology allows to manipulate endogenous genes of the cellular genome using a RNA-guided nuclease, such as Cas9, with a high accuracy and in an easy and rapid way. Nowadays, CRISPR system has been used in a wide variety of organisms that have traditionally been challenging to manipulate genetically. Although all its features must be fully defined, CRISPR system is a powerful technic that enable a broad range of applications from basic biology to biotechnology and medicine.

## RESUMEN

Desde hace muchos años se llevan estudiando técnicas de edición genómica que hagan posible llevar a cabo un estudio exhaustivo del genoma celular. Sin embargo, ha sido el descubrimiento del sistema CRISPR/Cas el que ha supuesto un gran avance en el mundo de la edición genómica. Originalmente, constituye el sistema inmunitario adaptativo de los organismos procariotas mediado por un grupo de repeticiones palindrómicas cortas agrupadas e interespaciadas (CRISPR). Esta técnica permite manipular el genoma celular de formas nunca vistas hasta el momento empleando nucleasas guiadas por RNA, como es el caso de la nucleasa Cas9; de una forma sencilla, rápida y específica pudiendo actuar sobre genes endógenos concretos de múltiples organismos diferentes que hasta el momento habían sido imposibles de manipular genéticamente. Aunque todavía quedan por definir todas sus características, el sistema CRISPR posee una amplia gama de aplicaciones, como pueden ser el diagnóstico y tratamiento de enfermedades genéticas e infecciosas, el control de plagas o en el rastreo y el control de la expresión génica empleando variantes de Cas9.

## DESARROLLO HISTÓRICO DEL CRISPR

La primera descripción conocida del sistema CRISPR fue en 1987 tras la secuenciación del gen *iap* de *Escherichia coli*, donde se descubrieron una serie de repeticiones presentes en el genoma<sup>1,2</sup>. Más tarde el científico español Francisco Mojica describió la existencia de una serie de secuencias palindrómicas repetidas con estructura similar en diferentes especies procariotas y planteó que podrían tener una función importante en la resistencia a fagos de estos microorganismos. Mojica denominó a estas repeticiones “Short Regulatory Spaced Repeats” (SRSRs), más tarde conocidas como “Clustered Regulatory Interspaced Palindromic Repeats” (CRISPR)<sup>3</sup>.

La primera vez que observó dicho fenómeno fue en 1993, durante sus estudios con *Haloferax mediterranei*, una arquea con gran tolerancia a la sal. Se comprobó que la concentración salina del medio de crecimiento del microorganismo afecta la forma de corte de sus enzimas de restricción en el genoma microbiano. Mojica estudió los fragmentos de DNA alterados en los cuales descubrió múltiples copias de una secuencia palindrómica de 30 bases, separadas unas de otras por espaciadores de unas 36 bases<sup>3</sup>.

En 2003, Mojica vio que el loci CRISPR de *E. coli* poseía un espaciador con la secuencia del fago P1 y que todas las cepas con ese espaciador eran resistentes a dicho fago. Fue entonces

cuando llego a la hipótesis de que el loci CRISPR estaba implicado en un sistema inmunitario adaptativo en organismos procariotas<sup>3</sup>. Dicha evidencia fue corroborada por Philippe Horvarth de la universidad de Strasbourg tras sus estudios con cepas de *Streptococcus thermophilus* sensibles a un fago genéticamente caracterizado. Se aislaron las cepas resistentes al fago y se observó que habían adquirido secuencias derivadas del fago en su loci CRISPR. Además, la incorporación de múltiples espaciadores daba lugar a una mayor resistencia<sup>2-4</sup>.

Horvarth también estudio el papel de los genes cas, en concreto de cas 7 y cas 9. Comprobó que cas 7 era necesario para que la bacteria fuese resistente; sin embargo, aquellas especies que contenían un espaciador derivado de un fago no necesitaban el gen para ser resistentes lo que llevo a la conclusión de que cas 7 permite la generación de nuevos espaciadores y repeticiones. Por el contrario, cas 9 posee codificadas en su secuencia nucleasas necesarias para adquirir la resistencia al fago. Por tanto, cas 9 supone un componente imprescindible en el sistema inmunitario bacteriano<sup>3</sup>.

Más tarde, en 2008, se estudió la programación de CRISPR para ello se introdujo el sistema CRISPR de *E. coli* en otra cepa de *E. coli* que carecía del mismo. Esto permitió caracterizar un sistema de 5 proteínas cas denominado Cascade, el cual es imprescindible para cortar un precursor de RNA transcrito del locus CRISPR, en RNAs de 61 nucleótidos de longitud denominados crRNAs. Todos ellos comienzan con las 8 últimas bases de la secuencia repetida y continúan con la secuencia completa del espaciador y con el principio de la siguiente región repetida. Esta secuencia palindrómica es lo que daría lugar a una estructura secundaria en el crRNA. Para demostrar que dichas secuencias de crRNA eran la base del sistema CRISPR se diseñaron los primeros sistemas CRISPR artificiales programados contra genes del fago lambda ( $\lambda$ ). Efectivamente, las cepas que contenían la nueva secuencia de CRISPR eran resistentes al fago  $\lambda$ <sup>3</sup>.

Actualmente se sabe que la diana del sistema CRISPR es el DNA. Este hallazgo fue demostrado por Luciano Marraffini y Erik Sontheimer. Marraffini observó que una cepa de *Staphylococcus epidermidis* tenía un espaciador que coincidía con una región del gen nickase (nes) de plásmidos presentes en *Staphylococcus aureus* resistentes a antibióticos<sup>1,3</sup>. Junto con Sontheimer, modificaron este gen introduciendo un intrón en la mitad de la secuencia sin obtener respuesta ante la infección con el plásmido puesto que la secuencia intrónica no había sido eliminada, por lo que el espaciador no correspondía con la secuencia del gen nes modificado. Por tanto, quedaba demostrado que CRIPR actuaba frente al DNA<sup>3</sup>.

A pesar de todos los estudios realizados sobre el sistema CRIPR-Cas 9 no fue hasta 2010 cuando se descubrió la función de una pequeña secuencia de RNA que se denominaría CRISPR RNA trans-activado (tracrRNA). Fueron Emmanuelle Charpentier y Jörg Vogel quienes lo descubrieron durante sus estudios con *Streptococcus pyogenes*. Ambos observaron que un pequeño RNA era transcrito con una elevada frecuencia desde una secuencia adyacente al locus CRISPR y que poseía 25 bases prácticamente complementarias con las repeticiones de CRISPR. Esta complementariedad permite al tracrRNA hibridar con el precursor de crRNAs para que sea procesado por la RNasaIII dando lugar a los crRNAs maduros. Además, también se comprobó que este tracrRNA es necesario para que la nucleasa Cas9 realice el corte en el DNA diana<sup>3</sup>.

En 2012 se purificó el sistema Cas9-crRNA y se puso en contacto *in vitro* con DNA, observando que el complejo era capaz de actuar cortando la doble hebra de DNA. Además, se demostró que se podía reprogramar Cas9 con espaciadores diseñados para cortar en sitios de elección *in vitro*<sup>3</sup>.

En este mismo año, Charpentier y Jennifer Doudna demostraron que los RNAs tracrRNA y crRNA pueden funcionar *in vitro* cuando son fusionados en un pequeño RNA guía (sgRNA). Tras varias modificaciones para aumentar su eficacia *in vivo*, el sgRNA ha sido ampliamente utilizado en edición genómica<sup>3</sup>.

En 2011, Feng Zhang creó una versión del Cas9 de *S. thermophilus* y otra de *S. pyogenes* para usarlo en células humanas. Encontró que expresando Cas9 y un sistema CRISPR RNA diseñado para actuar sobre un plásmido que contenía el gen de la luciferasa, podría disminuir la intensidad de la luminiscencia en células humanas embrionarias de riñón. Observó que las células de mamífero, a pesar de no poseer la RNasaIII microbiana, eran capaces de procesar el crRNA. Ya en 2012, había diseñado un sistema de tres componentes formado por Cas9 de *S. pyogenes* o de *S. thermophilus*, tracrRNA y un locus CRISPR. Empleando dicho sistema en diferentes áreas del genoma humano y de ratón, observó que era posible mutar genes con una elevada eficacia y precisión, causando, tras el corte de Cas9, deleciones o inserciones mediante la unión de extremos no homólogos (NHEJ) o insertando nuevas secuencias corregidas mediante recombinación homóloga (HDR). Además, era posible editar varios genes a la vez programando sistemas CRISPR con un sgRNA diferente para cada gen. Sin embargo, no obtuvo buenos resultados *in vivo*, aunque realizando una fusión completa de toda la longitud de los dos RNAs que constituía una horquilla 3' del tracrRNA, se resolvía el problema<sup>3</sup>.

Zhang demostró la versatilidad del CRISPR siendo posible emplearlo para diseñar modelos en ratones de enfermedades hereditarias y de cáncer permitiendo realizar rastreos para encontrar los genes esenciales en el proceso biológico<sup>3</sup>.

Tras todos estos descubrimientos los trabajos empleando el sistema CRISPR siguen aumentando y cada vez son más los organismos en los cuales se ha editado su genoma empleando CRISPR y sus aplicaciones en biología, terapéutica y agricultura.

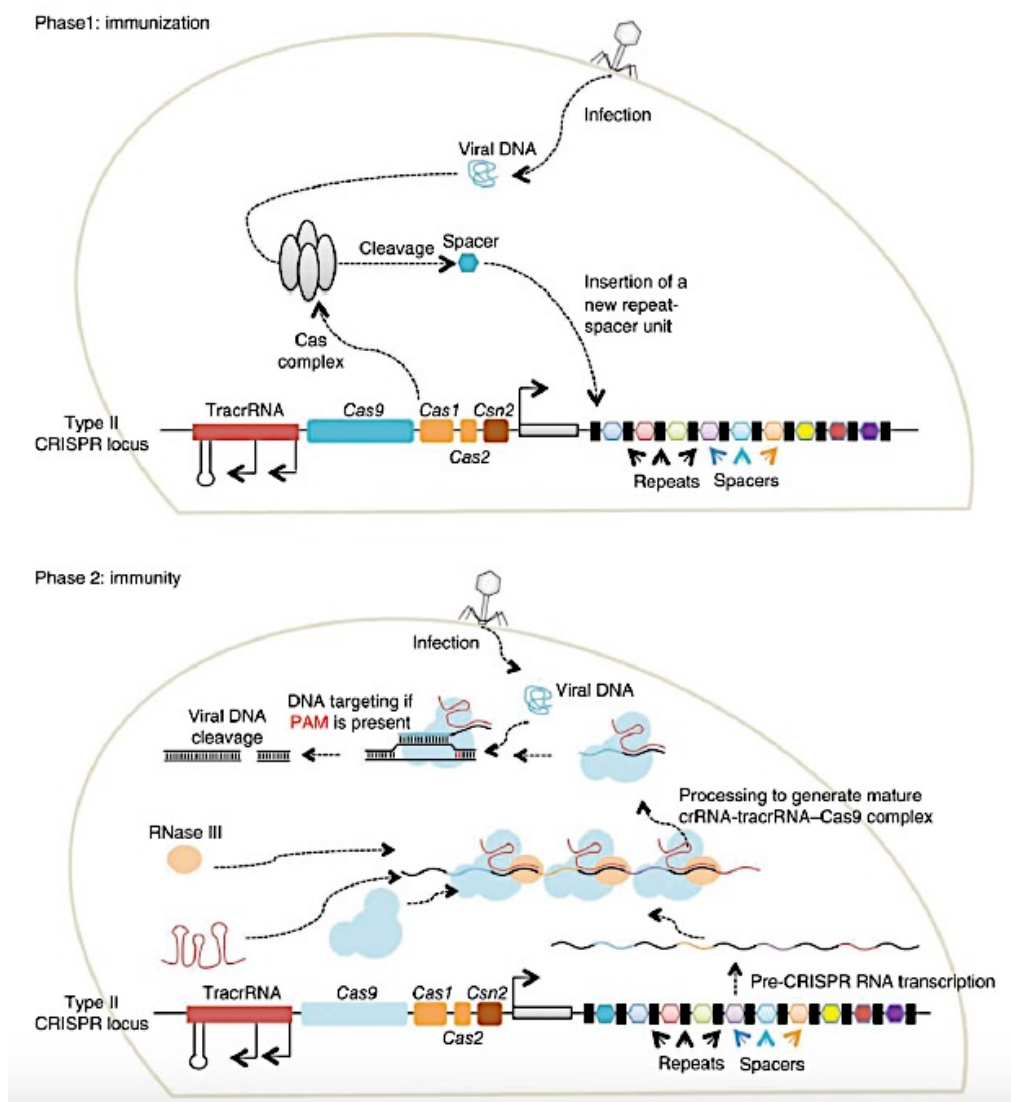
## MECANISMO MOLECULAR DEL SISTEMA CRISPR

El locus CRISPR está compuesto por los genes Cas y una serie de secuencias palindrómicas repetidas interespaciadas por pequeñas secuencias espaciadoras que poseen un origen plasmídico o vírico y que suponen el centro de la inmunidad específica a los fagos y plásmidos que posean una secuencia complementaria en su genoma<sup>2</sup>. Todo este conjunto constituye el sistema inmunitario adaptativo de organismos procariotas. Dependiendo del mecanismo molecular, se diferencian tres tipos de sistemas CRISPR. En el caso de los tipos I y III, las secuencias repetidas forman una estructura secundaria en forma de lazo en la base de la cual se produce el corte para obtener los crRNAs. Mientras que en el tipo II, las secuencias repetidas hibridan con un tracrRNA para después producirse el corte de la RNAsaIII.

La inmunidad adaptativa proporcionada por el sistema CRISPR consta de dos fases principales<sup>2,4</sup>:

1. **Fase de adaptación (Ilus.1 phase 1):** consiste en la adquisición de los espaciadores tras la primera infección por el plásmido o virus. Secuencias del genoma viral son introducidos en el locus CRISPR.
2. **Fase de inmunidad (Ilus.1 phase 2):**
  - a. **Biogénesis de la guía de RNA:** la región CRISPR es transcrita y procesada para obtener secuencias cortas de RNA (crRNAs) que contienen la secuencia de un espaciador complementaria a la del DNA invasor.
  - b. **Fase de corte:** estos crRNAs sirven de guía para dirigir el corte en el genoma viral por las endonucleasas tipo Cas.

El problema es que los fagos pueden escapar del sistema CRISPR a través de mutaciones en la región diana del genoma de manera que ya no se reconozca el sitio de corte. Sin embargo, la rápida adquisición de espaciadores por CRISPR permite responder rápidamente ante fagos que han conseguido evadir la respuesta inmunitaria<sup>2</sup>.



**Ilustración 1. Fases de la inmunidad CRISPR-Cas.** El locus CRISPR está constituido por un conjunto de repeticiones palindrómicas cortas de DNA (cajas negras) separadas por espaciadores provenientes de un fago o plásmido (cajas coloreadas). Este conjunto de repeticiones-espaciadores está flanqueado por un operón de genes asociados a CRISPR (genes Cas: Cas9 en azul, Cas1 y Cas2 en naranja) que codifican la maquinaria de inmunización. El locus CRISPR está precedido por un promotor o secuencia líder (caja gris) que contiene el promotor para la expresión. **Fase 1 de inmunización:** las secuencias espaciadoras similares a las del DNA del fago o plásmido invasor son integradas en la primera posición de la región CRISPR, de manera que se obtenga resistencia a invasiones futuras. **Fase 2 de inmunidad:** los espaciadores son empleados para detectar el DNA invasor con una secuencia similar. El pre-crRNA es procesado para obtener los crRNAs maduros que servirán como guía. En el caso del sistema tipo II, primero, el tracrRNA hibrida con las secuencias repetidas del pre-crRNA. Después, la RNasa III endógena corta liberando los crRNAs maduros que permanecen asociados con el tracrRNA y Cas9. El complejo corta en los protoespaciadores del DNA invasor con secuencias similares al crRNA. Imagen tomada de <sup>5</sup>

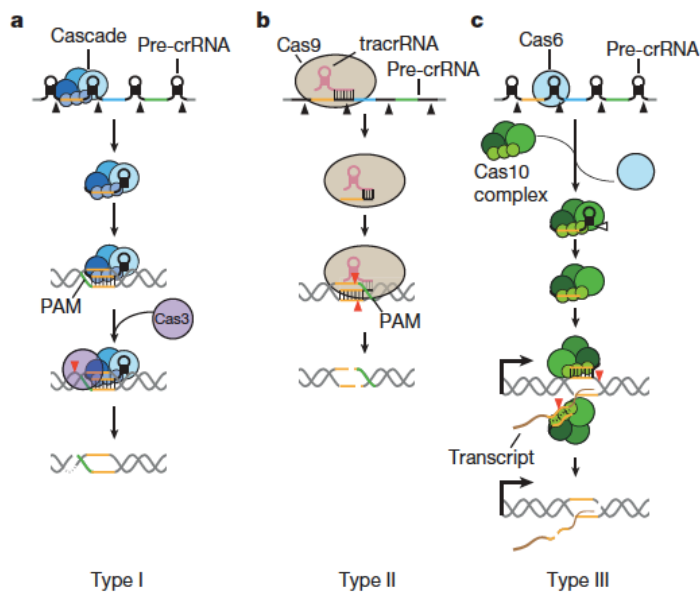
La inmunidad CRISPR-Cas tipo I (Ilus. 2a) está mediada por Cascade (complejo de proteínas cas) y la nucleasa Cas3. La endoribonucleasa encargada de cortar el precursor crRNA, generado tras la transcripción del locus CRISPR, en la base de la estructura secundaria en forma de lazo formada por las secuencias repetidas es la Cas6e. Tras el corte del pre-crRNA se obtiene los crRNAs guías y se forma un complejo Cascade-crRNA que recorre el DNA diana hasta encontrar la secuencia complementaria conocida como protoespaciador. Este protoespaciador



está flanqueado por un motivo protoespaciador adyacente o PAM. Estos PAM previenen las reacciones autoinmunitarias pues su ausencia en las secuencias repetidas previene la acción de CRISPR sobre los espaciadores<sup>2,4</sup>. La presencia de PAM permite la unión de Cascade a la diana y la hibridación del crRNA a la secuencia de DNA reconocida. Las primeras 8 bp del extremo 5' de esta unión crRNA-dsDNA es crítico en la inmunidad y cualquier mutación a este nivel confiere resistencia a la inmunidad CRISPR de tipo I en *E. coli*. La nucleasa Cas3 corta corriente abajo del PAM y también degrada la hebra opuesta<sup>2</sup>.

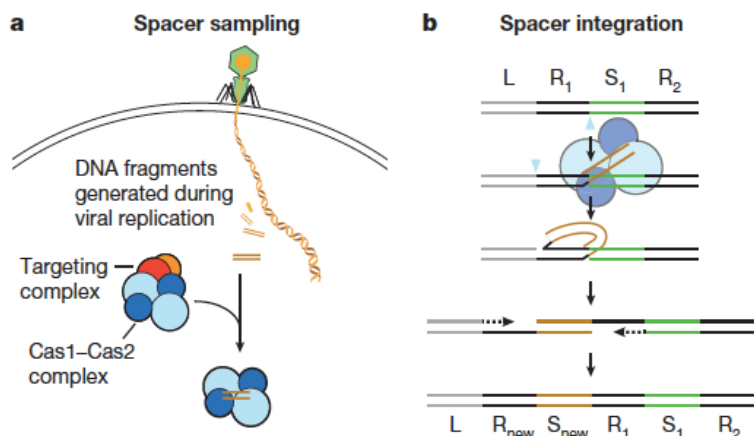
El sistema CRISPR-Cas tipo II (Ilus. 2b) requiere únicamente del gen cas9 para ejecutar la inmunidad en presencia de una secuencia de DNA extraño similar a un espaciador. En este caso son necesarios dos tipos de RNAs: el crRNA y el crRNA transcrito (tracrRNA). Este tracrRNA forma una estructura secundaria que permite su asociación con Cas9 y posee una región complementaria a las secuencias repetidas del locus CRISPR. Primero, tracrRNA hibrida con las repeticiones del pre-crRNAs y una RNasaIII endógena corta los crRNA-tracrRNAs híbridos. Después, se elimina el extremo 5' de los espaciadores para rindiendo los crRNAs maduros y se forma un complejo compuesto por Cas9, tracrRNA y el crRNA que recorre todo el DNA invasor hasta encontrar la secuencia diana complementaria al crRNA<sup>5</sup>. La inmunidad tipo II también requiere de PAM y las mutaciones a este nivel es la forma más común de evadirla. El PAM es reconocido por el dominio de unión a PAM de Cas9. Tras la unión transitoria de Cas9 a PAM se produce la separación de las hebras del DNA<sup>2</sup> y Cas9 corta en la doble hebra del DNA invasor mediante dos dominios de nucleasa: HNH, que corta en la hebra de diana complementaria al crRNA, y RuvC que corta en la hebra contraria<sup>6</sup>

En el sistema CRISPR-Cas tipo III (Ilus. 2c), el pre-crRNA es cortado por la endonucleasa Cas6, la cual no forma parte del complejo. Tras el procesamiento del pre-crRNA, 8 nucleótidos de la secuencia repetida permanecen en el extremo 5' del espaciador en la secuencia de crRNA y se denomina crRNA tag. El crRNA pasa a formar parte del complejo Cas10 donde es recortado en su extremo 3' dando lugar al crRNA maduro. En este caso el complejo Cas10 requiere de la transcripción de la diana y un crRNA complementario a la hebra del DNA invasor que no se había empleado como molde. En el sistema CRISPR-Cas tipo III, tanto el DNA como el RNA son diana, resultando en el corte de la hebra de DNA del protoespaciador no empleada como molde y de los transcritos procedentes de la secuencia guía complementaria al crRNA. Para evitar la autoinmunidad y que el propio locus CRISPR sea atacado, el sistema tipo III emplea la diferencia de apareamiento entre el crRNA tag y las secuencias que flanquean el protoespaciador. La completa complementariedad entre el crRNA tag y la secuencia repetida en el locus CRISPR previene del ataque al DNA propio<sup>2</sup>.



*Ilustración 2. Mecanismos de inmunidad de los diferentes tipos de CRISPR-Cas. . a, Sistema tipo I:* cabezas de flecha representan los puntos de corte de Cascade en el pre-crRNA, PAM en verde, cabeza de flecha roja representa el punto de corte de Cas3. *b, Sistema tipo II:* flechas negras, puntos de corte de Cas9 en el pre-crRNA; PAM, en verde; flechas rojas, punto de corte de Cas9 en el genoma invasor, protoespaciador y crRNA complementario en amarillo. *c. Sistema tipo III:* flechas negras son los puntos de corte de Cas6 en el pre-crRNA, crRNA maduro, cabeza de flecha blanca; puntos de corte de Cas10, cabezas de flecha rojas. Imagen tomada de <sup>2</sup>.

El mecanismo de incorporación (Ilus. 3) de nuevos espaciadores en el locus CRISPR consta de dos fases. Primero se realiza la selección de los protoespaciadores en el DNA invasor (Ilus. 3a). Este paso se realiza mediante la generación de cortes no específicos en el DNA durante la replicación del virus o el plásmido. Los fragmentos generados son capturados por el complejo Cas1-Cas2 con la participación de la maquinaria de reconocimiento de secuencias de DNA que incorporan un PAM funcional. En segundo lugar (Ilus. 3b), el complejo Cas1-Cas2 cataliza la incorporación del espaciador en la primera posición del locus CRISPR. Cas1 se encarga de delimitar dos lugares de ligación y corte donde el extremo 5' de cada hebra repetida es cortado y ligado al extremo 3' del espaciador. Este mecanismo da lugar a dos espacios ssDNA en las secuencias repetidas que flanquean el espaciador introducido las cuales son completadas por la DNA polimerasa. La primera adquisición facilita la adaptación a invasores con una secuencia relacionada a invasores previos y también a invasores que han conseguido escapar del sistema CRISPR. Además, la presencia de espaciadores pre-existentes parcialmente homólogos aumenta la tasa de incorporación de nuevos espaciadores<sup>2</sup>.



*Ilustración 3: Mecanismo de inmunización de CRISPR.* L: leader, R1: primera repetición, R2: segunda repetición, R<sub>new</sub>: nueva repetición, S1: primer espaciador, S2: segundo espaciador, S<sub>new</sub>: espaciador nuevo. Imagen tomada de <sup>2</sup>.

## APLICACIONES.

### *EDICIÓN GENÓMICA*

Desde 1980 se lleva investigando la forma de modificar el genoma de células eucariotas vivas. Para que esto fuese posible se observó que para la edición del genoma era necesario encontrar un método que produjese un corte en la doble hebra del DNA en los sitios de interés. Así pues, antes que el sistema CRISPR, se trabajó con dos estrategias diferentes:

- **Nucleasas de dedos de zinc (ZFNs).** Fue la primera estrategia utilizada y consistía en fusionar proteínas compuestas por un dedo de zinc unido con un dominio de unión a DNA y un dominio de corte del DNA, obtenido de una enzima de restricción de manera que se pudiese unir y cortar en un locus genómico. Sin embargo, resultó ser un método lento y complejo<sup>3,7</sup>.
- **Proteínas activadoras de la transcripción (TALENs).** Un sistema de edición genómica desarrollado en 2009. Este método emplea un código preciso de dominios modulados de manera que actuase sobre secuencias específicas de DNA. Seguía siendo una técnica compleja puesto que se necesitaba una nueva proteína para cada diana<sup>3,7</sup>.

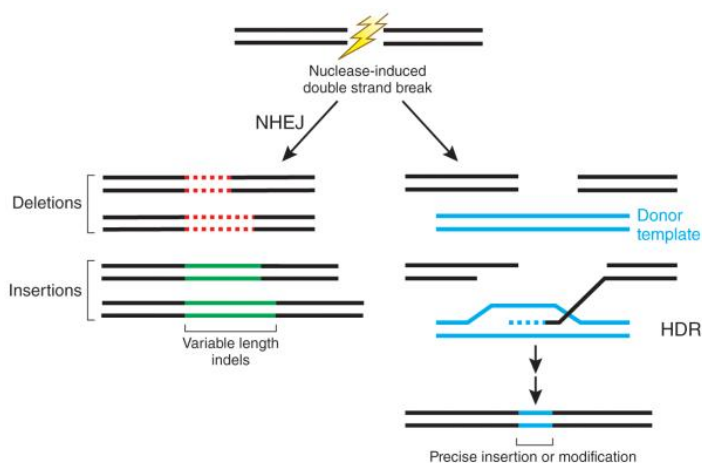
Actualmente, estas técnicas serán previsiblemente desplazadas por el sistema CRISPR ya que supone una herramienta más sencilla de diseñar, además de ser más eficaz y específica en el corte y la edición del genoma. Mientras ZFNs y TALENs consisten en nucleasas quiméricas compuestas por un dominio de unión a DNA programable y específico unido a un dominio de nucleasa que da lugar a un corte de la doble hebra, el sistema CRISPR trata de un loci con múltiples repeticiones cortas que emplea crRNAs y tracrRNAs para producir el silenciamiento de secuencias específicas de DNA extraño<sup>7</sup>.

De los tres sistemas CRISPR conocidos el más desarrollado de momento es el tipo II compuesto por el precursor de crRNA, la nucleasa Cas9 y el tracrRNA necesario para procesar el pre-crRNA y formar el complejo Cas9. Debido a la simplicidad y fácil modificación del sistema CRISPR tipo II este se ha empelado en edición genómica y en el control transcripcional de células procariotas y eucariotas. Cas9 es una endonucleasa de DNA guiada por RNA que puede ser programada para actuar en diferentes sitios del genoma cambiando la secuencia del crRNA guía<sup>6</sup>.

Cas9 está siendo empleada en múltiples estudios de edición genómica para determinar la función de determinados genes, la modulación de enfermedades y para desarrollar nuevos métodos terapéuticos en un amplio número de enfermedades genéticas e infecciosas<sup>6</sup>. Sin embargo, para trabajar en células humanas es necesario extraer una proteína Cas9 optimizada para el ser humano que posea una señal de localización nuclear. Además, el crRNA y el tracrRNA deben encontrarse en una misma quimera denominada sgRNA, junto con el promotor de la RNA polimerasa III. Las recientes aplicaciones del sistema tipo II se pueden agrupar en 5 grupos: Cas9 nativa empleada en modificación del genoma, modificación genómica mediada por Cas9 nickase, Cas9 activada o inactivada en el control transcripcional y Cas9 basada en el rastreo de amplio espectro genético<sup>8</sup>

### **Edición genómica mediada por Cas9 nativa**

La modificación del genoma empleando Cas9 requiere de dos pasos. En primer lugar, Cas9 reconoce una secuencia específica de 20 nucleótidos en el DNA complementaria al crRNA y realiza un corte en la doble hebra de DNA (DSB). En segundo lugar, el corte de doble hebra es reparado mediante recombinación homóloga (HDR) o mediante unión de los extremos no homóloga (NHEJ) (Ilus. 4). NHEJ produce mutaciones por inserción o selección en el DSB, y esto lleva normalmente a la pérdida de la función del gen, “knockout” (por ejemplo, causando el desplazamiento del cuadro de lectura del gen diana). Por el contrario, HDR permite generar el reemplazamiento del gen en el DSB mediante recombinación homóloga empleando una secuencia de DNA molde, produciendo selección del gen, mutagénesis, inserción o corrección del gen. Por tanto, el sistema CRISPR/Cas9 supone una herramienta prometedora en la edición genómica de secuencias específicas, incluyendo la pérdida de función o ganancia de función de un gen, así como mutaciones y correcciones en lugares específicos de la secuencia<sup>6,9</sup>.



*Ilustración 4. Edición genómica inducida por nucleasas. Imagen tomada de <sup>9</sup>*

El tracrRNA y el crRNA se han unido en una única molécula denominada sgRNA que posee la secuencia de reconocimiento del protoespaciador adyacente a PAM. En el caso de la nucleasa Cas9 más empleada, procedente de *Streptococcus pyogenes*, reconoce un PAM de dos bases únicamente (NGG) y puede cortar protoespaciadores cada 8 pares de bases<sup>8</sup>. Coexpresando Cas9 y sgRNAs individualizados se ha conseguido modificar el genoma humano. Sorprendentemente, elaborando un sistema de expresión mamífero con un codón Cas9 y sgRNA optimizado se ha conseguido una eficiencia en el reconocimiento de la diana de 2-25% en diferentes células humanas y de 36-48% en células embrionarias de ratón<sup>8</sup>.

El proceso de elección del gen diana envuelve diferentes pasos (Ilus. 5). En primer lugar, será necesario seleccionar un protoespaciador en el gen y diseñar los RNAs guías unidos a un espaciador que posea una secuencia idéntica a la diana. Después la proteína de la nucleasa Cas9 y el RNA guía son introducidos en la célula diana. El RNA guía dirige la nucleasa Cas9 al protoespaciador diana, el cual reconoce por el PAM adyacente, y se une al DNA mediante el espaciador del RNA guía. Si el protoespaciador y el espaciador son idénticos, o únicamente no coinciden en una pequeña parte del extremo 5', Cas9 corta en ambas hebras de DNA dando lugar a un corte de doble hebra con extremo romo que será reparado por HDR o NHEJ dando lugar a la disrupción del gen<sup>10</sup>. Además, Cas9 es útil para modificar ambos cromosomas y/o editar múltiples genes al mismo tiempo introduciendo múltiples sgRNAs simultáneamente. Por ejemplo, creando un doble DSBs en regiones próximas en el mismo cromosoma puede llevar a deleciones o inversiones en el segmento intermedio del DNA, y creando dos DSBs en cromosomas diferentes daría lugar a una translocación del cromosoma diana. Esta reorganización de las dianas puede ser útil para realizar modelos de enfermedades al imitar las reorganizaciones que tienen lugar en la enfermedad<sup>6,9</sup>.

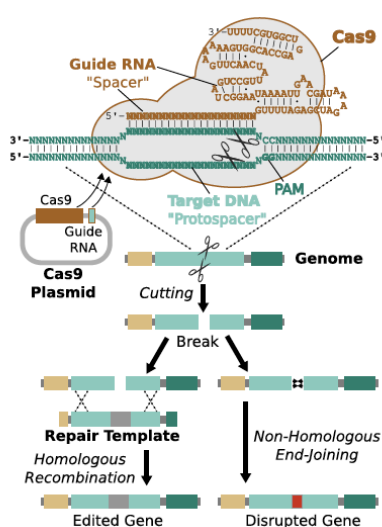


Ilustración 5. Edición genómica guiada por RNA vía Cas9. Imagen tomada de <sup>10</sup>.

El sistema Cas9 también posee el potencial para curar o tratar enfermedades como VIH, enfermedades genéticas y cáncer. Por ejemplo, cuando Cas9 es introducido en células infectadas junto con sgRNAs dirigidas a sitios del genoma viral imprescindibles, contribuye a inactivar o eliminar el genoma viral y, por tanto, se consigue defender a las células del organismo de la infección frente a VIH, virus de la hepatitis B, papiloma virus y Epstein-Barr virus. Además, se ha demostrado que usando CRISPR/Cas9 o ZNFs se puede editar los genes de los correceptores de VIH como el CCR5 en el genoma del huésped, el cual posee codificado el correceptor del VIH, dando lugar a una resistencia a la infección frente al virus VIH-1 siendo posible combatir la infección.

### **Edición genómica mediada por Cas9 nickasa (nCas9)**

Se pueden realizar mutaciones en uno de los dominios nucleasa (RuvC o HNH) de Cas9 de tal manera que en lugar de realizar un corte en la doble hebra de DNA se produce un corte en una de las dos hebras denominado nick. La producción de un doble nick requiere de dos Cas9 dirigidas por sgRNA diferentes a lugares adyacentes en hebras opuestas produciendo un corte de doble hebra<sup>6,8,9</sup>.

### **Factores influyentes en la aplicación de CRISPR/Cas en células humanas**

La manipulación de la nucleasa Cas9 supone un proceso delicado puesto que mutaciones en los sitios catalíticos puede dar lugar a localizaciones subcelulares inexactas, o una mala dosificación de la nucleasa puede afectar a la modificación del genoma. En células eucariotas, Cas9 procedente de procariotas suele fusionarse con una señal de localización celular (NLS) en el extremo N, C o ambos, para dirigir la translocación de la proteína al núcleo. Estudios recientes han demostrado que la inclusión de un adaptador o “linker” de 32 aminoácidos entre el NLS y Cas9 mejora la división del DNA genómico. Esto puede deberse a que se consigue una mejor localización subcelular, puesto que sin el adaptador el péptido NLS podría ser ocultado o protegido durante el plegamiento de Cas9. La optimización de codón también es esencial para que sea posible la expresión de una proteína Cas9 funcional en células humanas<sup>8</sup>.

### ***MODULACIÓN TRANSCRIPCIONAL BASADO EN Cas9 INACTIVADA***

El sistema CRISPR/Cas también se emplea en el control transcripcional a lo que se conoce como CRISPR “interference” (CRISPRi) (Ilu.6a) compuesto por una nucleasa Cas9 completamente inactivada (dCas9) y un sgRNA. En este caso los dos dominios nucleasa de

Cas9 están mutados; por tanto, dCas9 no posee actividad nucleasa, pero el reconocimiento y unión a la diana permanecen intactos. CRISPRi se controla por apareamiento de bases en genes localizados del genoma impidiendo la transcripción por impedimento estérico. Para mejorar la acción de dCas9 en células de mamífero, se ha observado que fusionándola con un dominio represor transcripcional se obtiene una supresión eficaz de genes endógenos<sup>6</sup>. El silenciamiento de genes puede ser potenciado o revertido empleando un promotor inducible, como por ejemplo un promotor inducible por tetraciclinas, dando lugar a la expresión de dCas9 y sgRNA. Por tanto, este método posee especial importancia en la modulación de la expresión de genes a nivel transcripcional<sup>8</sup>.

Además del sistema CRISPRi empleado en el silenciamiento de la expresión genética, también se ha desarrollado un sistema denominado CRISPR “activation” (CRISPRa) (Ilustración 6b) fusionando dCas9 con un activador transcripcional como por ejemplo VP64 o p65AD. Empleando la quimera dCas9-VP64 en células humanas desencadenar la expresión de genes.

El sgRNA puede modificarse fusionándose con aptámeros que recluyan proteínas de unión a RNA específicas (RBPs). Este sgRNA modificado se conoce como scRNA (scaffold RNA). Activadores y represores transcripcionales pueden ser fusionados a estas RBPs en lugar de al dCas9. De esta manera, se puede controlar la transcripción de varios genes al mismo tiempo empleando diferentes pares de aptámeros-RBPsg con diferentes sgRNAs; por ejemplo, un gen puede ser reprimido al emplear un scRNA con un aptámero que se una al represor transcripcional como KRAB, al mismo tiempo que otro es activado por VP64 la cual se ha unido al aptámero del scRNA a través del RBP (Ilustración 6c).

Por tanto, CRISPRi/a son sistemas simples y con alta especificidad que pueden controlar la transcripción de genes endógenos actuando en secuencias codificadas como no codificada. Esto permitiría regular la transcripción de secuencias específicas en células tumorales humanas.

### ***EDICIÓN EPIGENÉTICA MEDIADA POR Cas9 INACTIVADA***

El epigenoma está formado por compuestos químicos y proteínas, o etiquetas químicas, que pueden unirse al DNA y dirigir acciones tales como la activación o desactivación de genes, y el control de la producción de proteínas en células específicas. Fusionando dCas9 a enzimas moduladoras epigenéticas, como la LDS1, que actúen sobre marcadores epigenéticos próximos al sitio de unión de dCas9 se puede modificar el epigenoma. Mediante este método se consigue una regulación corriente abajo del gen de interés actuando sobre potenciadores o represores de la expresión génica. En el caso de la dCas9-LDS1 (Ilustración 6d) se produce la desmetilación de la

lisina 4 de la subunidad H3 de las histonas produciendo una represión de la expresión genética. Por el contrario, si se fusiona dCas9 con el dominio catalítico central de la histona acetiltransferasa p300 (Ilu.6d) se produce la metilación del resto de lisina 27 de la subunidad H3 de las histonas resultando en una activación de la expresión de genes de promotores distales<sup>11</sup>.

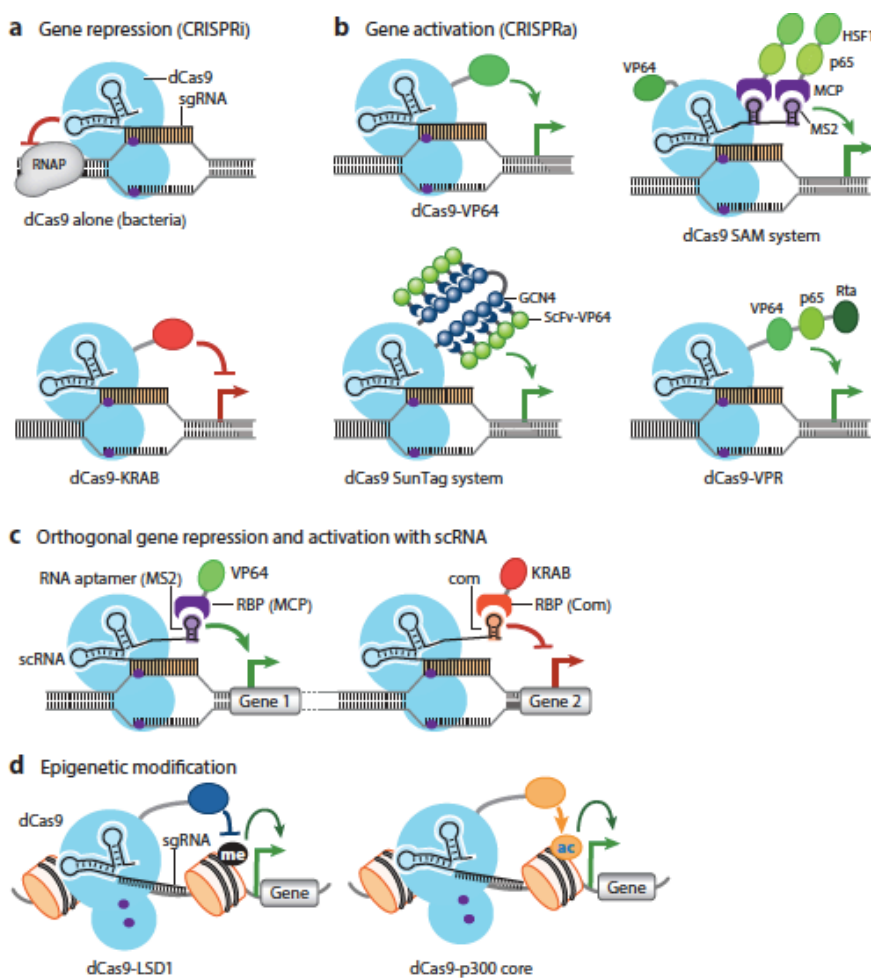


Ilustración 6. Aplicaciones de Cas9 inactivada (dCas9). Imagen tomada de <sup>6</sup>

### **RASTREO GENÉTICO DE AMPLIO ESPECTRO EMPLEANDO Cas9**

Se están empleando múltiples variantes de Cas9 para producir librerías de mutantes para el rastreo e identificación de elementos genéticos a nivel genómico o relacionados con el fenotipo. Se ha empleado un método de rastreo compatible con la selección de cepas tanto positivas como negativas en la expresión de un gen, empleando una librería de sgRNA en lentivirus como vector. La expresión de sgRNA fue incorporada al genoma obteniendo un pool de mutantes cuya secuencia fue monitorizada paralelamente<sup>8</sup>.



## ***ESCAÑO DEL GENOMA***

Muchos estudios han demostrado que la organización tridimensional del genoma en el núcleo posee un papel muy importante en la regulación de la expresión genética y en el control de la diferenciación celular. Por este motivo, los científicos estudian la correlación entre la información genética lineal en el DNA y su organización tridimensional. Cas9 se ha convertido en una herramienta prometedora en el estudio de la organización del genoma gracias a que es capaz de localizar secuencias específicas en el genoma y redireccionarlas a un loci genómico diferente<sup>6</sup>.

Cas9 combina dos técnicas convencionales de escaneo del genoma: apareamiento de bases e interacciones proteína-DNA. Una de las técnicas desarrolladas consiste en fusionar dCas9 con una proteína fluorescente para visualizar las secuencias codificantes y no codificantes en células humanas vivas, de tal manera que el sgRNA hibrida con la secuencia del gen diana y la proteína fluorescente unida a dCas9 emitirá una señal de luz que permitirá localizar el gen. En un estudio, secuencias repetidas del genoma fueron observadas durante el ciclo celular empleando un único sgRNA. Del mismo modo, un locus genómico no repetitivo también puede ser señalado empleando al mismo tiempo múltiples sgRNAs dirigidos a distintas secuencias no repetidas. Además, esta técnica permite estudiar diferentes genes al mismo tiempo empleando distintas dCas9 que tengan fusionadas diferentes proteínas fluorescentes y que posean diferentes PAM y un sgRNA correspondiente.

El uso de Cas9 en el escaneo del genoma permite estudiar cualquier parte del genoma, tanto secuencias repetidas como no repetidas; se puede emplear tanto en células vivas como modificadas por lo que permite obtener una imagen de las células vivas. Además, el diseño de los sgRNA es sencillo y se puede obtener una imagen de múltiples colores siendo posible estudiar diferentes secuencias al mismo tiempo. A pesar de las ventajas, esta técnica presenta como inconvenientes el uso de múltiples sgRNAs para obtener la imagen de secuencias no repetidas y, además, requiere de un PAM específico para cada Cas9<sup>6</sup>.

## ***CONTROL DE POBLACIONES***

Una de las posibles aplicaciones de CRISPR es el control de plagas mediante una técnica denominada “Gene Drive” con especial importancia en la agricultura. Sin embargo, a día de hoy esta técnica aún no se ha puesto en práctica. Lo que se pretende es erradicar enfermedades transmitidas por insectos, disminuir el uso de pesticidas y herbicidas y las resistencias a estos; conseguir la sostenibilidad y seguridad agrícola; y controlar el daño producido por especies

invasoras. Sin embargo, la posibilidad de efectos ecológicos no deseados y la expansión incontrolada de la especie modifica requieren una exhaustiva evaluación de cada aplicación.

El primero en proponer el “Gene Drive” fue Austin Burt basándose en genes de endonucleasas “homing” o buscadores de motivos de endonucleasas que poseen sitios de corte específicos dentro del genoma, de manera que al cortar en el cromosoma homólogo induce a la célula a copiar estos genes cuando se va a reparar el corte. La dificultad radica en dirigir el corte de estas endonucleasas a nuevas secuencias diana<sup>10</sup>. Este problema se puede solucionar mediante el desarrollo de genes drives guiados por RNA basados en la nucleasa Cas9.

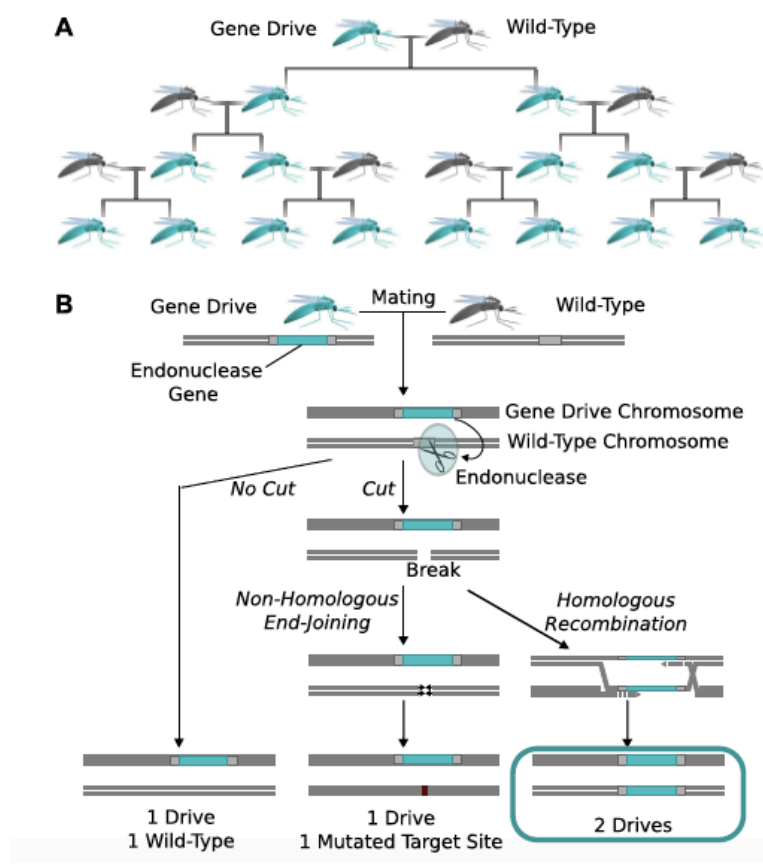
De forma natural, los genes de endonucleasas pueden dirigir su paso a la descendencia cortando en el locus correspondiente del cromosoma que carece de ellas. Esto lleva a la célula a reparar ese corte copiando el gen de la nucleasa en el cromosoma dañado mediante recombinación homóloga (Ilus. 7a). Este proceso de copia se denomina “homing” y la información que codifica la endonucleasa (endonuclease-containing cassette) se denomina “gene drive” o motor génico. De esta manera, se favorece que más de la mitad de la descendencia adquieran el gen de interés, pasando de generación en generación, aunque la tasa de reproducción de las poblaciones que lo contienen disminuya (Ilus. 7b). La autosostenibilidad del gen drive lleva a que, de un número pequeño de individuos, pase a todos los miembros de una población tras un largo periodo de tiempo, teniendo que pasar por múltiples generaciones<sup>10</sup>.

Para desarrollar una endonucleasa que pueda dirigir un gen de interés a las poblaciones, aumentando las probabilidades de que este sea heredado, es necesario insertar un transgen de la endonucleasa en el lugar de una secuencia natural que puede cortar. Originalmente, se diseñaron genes drive que llevasen codificados una endonucleasa y un transgen de manera que ese transgen fuese dirigido a las poblaciones silvestres. Otro método es diseñar genes drive que interrumpiesen o sustituyesen los genes existentes, de esta manera se conseguía una estabilidad evolutiva puesto que no se podía transmitir sin actuar sobre el gen diana. Por otro lado, si la copia del gen drive se produce en el cigoto o en el embrión temprano, todos los organismos que contengan el gen serán homocigotos en todos sus tejidos. Si tiene lugar, en las gónadas, en las células germinativas tardías que contribuyen al esperma o huevos, la descendencia seguirá siendo heterocigota en la mayoría de los tejidos, evadiendo las consecuencias inducidas por la interrupción dirigida del gen<sup>10</sup>.

Desarrollando genes dirigidos para interrumpir genes causando infertilidad o letalidad en caso de perderse ambas copias, se reduce el número de miembros de la población diana. Esto lleva a que se transmita rara vez a un pequeño número de heterocigotos y eventualmente ocasionaría la extinción de la especie por acumulo de mutaciones. Otra opción, sería dirigir la

selección natural a uno u otro sexo. En este modelo, el cromosoma Y llevaría codificado una endonucleasa que cortase y destruyese el cromosoma X durante la meiosis masculina, de esta manera se aseguraría que fuese el cromosoma Y el que se transmitiese en su mayoría consiguiendo una población mayormente masculina. Sin embargo, la progresiva disminución del género femenino llevaría a la extinción<sup>10</sup>.

Todas las técnicas de “gene drive” poseen especial interés hoy en día en la profilaxis de la malaria. Lo que se pretende es eliminar o sustituir los genes que llevan a la transmisión de la malaria, generando poblaciones homocigotas que no posean estos genes y por tanto no puedan transmitir la enfermedad; o generar poblaciones masculinas principalmente, de manera que no haya hembras responsables de transmitir *Plasmodium falciparum*. Aun así, habría que valorar las repercusiones tanto ambientales como biológicas que tiene el acabar con el género femenino del mosquito *Anopheles gambiae*, puesto que esto podría llevar a la extinción de la especie.



**Ilustración 7. La expansión de genes drives de endonucleasas.** A. Cuando un organismo con el gen drive de la endonucleasa (en azul) se aparea con un organismo silvestre (en gris), toda la descendencia heredará el gen drive; hasta llegar un momento que todos los miembros de la población poseen dicho gen. B. Al cortar la endonucleasa en el cromosoma homólogo del organismo silvestre, se emplea como molde el cromosoma del gen drive para reparar el corte, pasando de esta forma a la descendencia. En caso de que la endonucleasa no corte o se repare el corte por unión de los extremos, el gen drive no es copiado. Imagen tomada de <sup>10</sup>.

Lo que se pretende conseguir mediante edición genómica por vía nucleasa Cas9, es el mismo mecanismo realizado por el motor génico (“gene drive”) de la endonucleasa. La ventaja

de la nucleasa Cas9 es que se puede programar para cortar en una secuencia específica del DNA empleando una molécula de RNA como guía que provienen de la transcripción de un fragmento de DNA introducido en un elemento de memoria.

A pesar del potencial del sistema “Gene Drive”, este no deja de tener una serie de limitaciones que no deben pasar desapercibidas. La más importante, es que se necesitan muchas generaciones para conseguir que toda la población adquiera el gen de interés. El tiempo que tarda en extenderse a todos los miembros depende de diversos factores como son el número de individuos liberados con el gen drive, el tiempo de generación del organismo, la eficiencia de copia del gen o el impacto sobre los individuos de la población. Además, no es efectivo en especies que poseen una reproducción asexual exclusivamente como son todos los virus, bacterias y la mayoría de los organismos unicelulares. Por último, las alteraciones producidas en el genoma no son permanentes en la escala evolutiva ya que la presión selectiva favorecerá la transmisión de Cas9 y las RNAs guías hasta que el gen drive haya conseguido formar parte del genoma de forma natural<sup>10</sup>.

Genes dirigidos por RNA guía pueden revertir alteraciones del genoma que se han expandido a todas las poblaciones. Esto puede ser útil en caso de que los resultados obtenidos con el drive inicial no fuesen los esperados; por tanto, liberando un gen revertido puede eliminar uno o todos los cambios producidos en el genoma en un primer lugar.

Además, drives dirigidos por RNA puede bloquear la transmisión de otro drive recodificando secuencias del drive no deseado como dianas de manera que no pueda ser copiado. A este segundo drive liberado se le denomina drive de inmunización<sup>10</sup>.

## CONCLUSIONES

El sistema CRISPR/Cas constituye el sistema inmunitario adaptativo bacteriano y su descubrimiento ha supuesto un gran avance en ingeniería biológica al combinar en el mismo espacio-tiempo las tres moléculas más importantes: DNA, RNA y proteínas. Permite el estudio y la manipulación del genoma de gran variedad de organismos, incluyendo el genoma humano, obteniendo resultados más exactos de forma más rápida y sencilla en comparación con otras técnicas más complejas empleadas hasta el momento. Aunque la mayoría de los estudios con esta técnica continúan siendo teóricos, sus posibles aplicaciones en medicina, biología y biotecnología siguen aumentando. Por otro lado, todavía quedan por discutir las posibles repercusiones ambientales y éticas que podría tener la manipulación de ciertas secuencias génicas en diferentes especies.

## BIBLIOGRAFÍA

### Desarrollo histórico del CRISPR:

1. Sorek, R., Kunin, V. & Hugenholtz, P. CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* **6**, 181–186 (2008).
2. Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55–61 (2015).
3. Lander, E. S. The Heroes of CRISPR. *Cell* **164**, 18–28 (2016).

### Mecanismo molecular del sistema CRISPR:

2. Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55–61 (2015).
4. Barrangou, R. & Marraffini, L. A. CRISPR-cas systems: Prokaryotes upgrade to adaptive immunity. *Molecular Cell* **54**, 234–244 (2014).
5. Mali, P., Esvelt, K. M. & Church, G. M. Cas9 as a versatile tool for engineering biology. *Nat. Methods* **10**, 957–963 (2013).
6. Wang, H., La Russa, M. & Qi, L. S. CRISPR/Cas9 in Genome Editing and Beyond. *Annu. Rev. Biochem.* **85**, 227–264 (2016).

### Aplicaciones:

3. Lander, E. S. The Heroes of CRISPR. *Cell* **164**, 18–28 (2016).
6. Wang, H., La Russa, M. & Qi, L. S. CRISPR/Cas9 in Genome Editing and Beyond. *Annu. Rev. Biochem.* **85**, 227–264 (2016).
7. Gaj, T., Gersbach, C. A. & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology* **31**, 397–405 (2013).
8. Wang, X., Huang, X., Fang, X., Zhang, Y. & Wang, W. CRISPR-Cas9 System as a Versatile Tool for Genome Engineering in Human Cells. *Mol. Ther. Acids* **5**, e388 (2016).
9. Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355 (2014).
10. Esvelt, K. M., Smidler, A. L., Catteruccia, F. & Church, G. M. Concerning RNA-guided gene drives for the alteration of wild populations. *Elife* **3**, e03401 (2014).
11. Pratiksha I. Thakore, Joshua B. Black, Isaac B. Hilton, and C. A. G. Editing the Epigenome: Technologies for Programmable Transcriptional Modulation and Epigenetic Regulation. *Nat. Biotechnol.* **33**, 510–517 (2015).