



UNIVERSIDAD DE JAÉN
Facultad de Ciencias Experimentales

Trabajo Fin de Grado

Sistema CRISPR/Cas y su aplicación a la edición dirigida de genomas

Alumno: Míriam Expósito Martínez

Julio, 2018



Facultad de
Ciencias Experimentales

UNIVERSIDAD DE JAÉN

FACULTAD DE CIENCIAS EXPERIMENTALES

GRADO EN BIOLOGÍA

Trabajo Fin de Grado

***SISTEMA CRISPR/Cas y su aplicación a la edición
dirigida de genomas***

Míriam Expósito Martínez

Una firma manuscrita en tinta que parece decir "Míriam" con "E.M." debajo, sobre una línea punteada.

Universidad de Jaén, Julio de 2018

RESUMEN

CRISPR/Cas es un sistema inmune adaptativo bacteriano que actúa como defensa contra bacteriófagos u otros elementos genéticos móviles. Compuesto por una endonucleasa Cas y un ARN guía, permite degradar los ácidos nucleicos exógenos impidiendo su replicación en el interior de las bacterias. Debido a su simplicidad y su habilidad para cortar ADN, se ha convertido en una poderosa herramienta para edición genómica, poseyendo un gran número de aplicaciones de utilidad tanto para investigación como terapia génica. En esta revisión, se expone la estructura y funcionamiento de CRISPR/Cas como sistema inmune bacteriano, así como su adaptación para su uso en ingeniería genética, explicando algunas de sus aplicaciones y posibles inconvenientes que pueden darse en su uso.

Palabras clave: CRISPR/Cas, endonucleasa Cas, edición genómica, ARN guía, terapia génica.

ABSTRACT

CRISPR/Cas is an adaptive bacterial immune system that acts as a defense against bacteriophages or other mobile genetic elements. Composed by a Cas endonuclease and a guide RNA, degrades exogenous nucleic acids, preventing their replication inside the bacteria. Due to its simplicity and its ability to cleave DNA, it has turned into a very powerful tool for genomic editing, possessing many useful applications for both research and gene therapy. In this review, the structure and functioning of CRISPR/Cas as a bacterial immune system is exposed, as well as its adaptation for genetic engineering, explaining some of its applications and possible issues that can appear in its use.

Key words: CRISPR/Cas, Cas endonuclease, genomic edition, guide RNA, gene therapy

ÍNDICE

1. INTRODUCCIÓN	1
2. OBJETIVO.....	3
3. ANTECEDENTES.....	3
3.1 Nucleasas con dedos de Zinc (<i>Zinc Fingers Nuclease, ZFN</i>).....	3
3.2 TALEN (<i>Transcription activator-like effector nucleases</i>).....	4
4. El sistema CRISPR/Cas	5
4.1 Origen.....	5
4.2 Descubrimiento	5
4.3 Tipos de sistemas CRISPR/Cas.....	6
4.4 Estructura	7
4.5 Mecanismo de acción del sistema CRISPR/Cas	10
4.6 Familias de proteínas Cas	11
4.7 Mecanismos de defensa de virus: Inactivación del sistema CRISPR/Cas ..	14
5. EL SISTEMA CRISPR/CAS9 APLICADO A EDICIÓN GENÓMICA.....	15
5.1 Funcionamiento	16
5.2 Estrategias para editar genes con CRISPR/Cas9	18
5.3 Aplicaciones	20
5.3.1 Terapia génica	20
5.3.2 CRISPR/Cas9 – Silenciamiento de genes y CRISPRi	20
5.3.3 CRISPRa – Activación de genes.....	21
5.3.4 CRISPR/Cas9 – Edición de genomas	21
5.3.5 Reordenaciones cromosómicas y cáncer.....	22
5.3.6 Control de la replicación	23
5.3.7 Integración de rutas de señalización	24
5.3.8 Edición Múltiple	25
5.4 Inconvenientes	27
5.4.1 Efectos <i>Off-Target</i>	27
5.4.2 Preferencia de NHEJ sobre HDR	29
5.4.3 Mosaicismo	30
6. IMPLICACIONES ÉTICAS	31
7. DISCUSIÓN	32
8. CONCLUSIONES	33
9. REFERENCIAS	33

1. INTRODUCCIÓN

El descubrimiento de la estructura del ADN en 1953 marcó una nueva era para la biología. Se establecen las bases biológicas que permiten la transmisión de la información genética de una generación a otra y se propicia el origen de nuevas disciplinas, entre las que se encuentra la ingeniería genómica, que se refiere al proceso de realizar modificaciones específicas en el genoma, en él mismo o en sus productos (Hsu, Lander and Zhang, 2014).

Los avances en la manipulación genética han hecho posible el estudio de genes, favoreciendo el desarrollo de la ciencia básica, biotecnología, medicina y la secuenciación del genoma humano en el año 2003. Identificar los genes y conocer su estructura y su función, o analizar los efectos que puede tener una mutación en el genoma ha permitido identificar el origen de muchas enfermedades genéticas y desarrollar estrategias para su tratamiento.

Las tecnologías que pueden delecionar, insertar y modificar secuencias de ADN de células u organismos permiten diseccionar la función de genes específicos y elementos reguladores. De una manera similar, la manipulación de la regulación de la transcripción y los estados de la cromatina en un loci determinado puede revelar cómo se organiza y se utiliza ese material genético dentro de la célula, esclareciendo las posibles relaciones entre la arquitectura del genoma y sus funciones.

La capacidad para introducir cambios de forma dirigida en células y organismos es una herramienta muy poderosa. Durante muchos años, las estrategias para inducir alteraciones genómicas precisas han estado limitadas a solo algunos tipos de organismos (por ejemplo, la recombinación homóloga en levaduras o la sustitución dirigida en genes en ratones) y, a menudo, requerían marcadores seleccionables o dejaban “cicatrices” en las secuencias asociadas al método de modificación.

La edición dirigida del genoma utilizando nucleasas personalizadas supone un nuevo método general, aplicable a un amplio rango de organismos y tipos celulares. Estas enzimas, permiten inducir delecciones, inserciones y cambios precisos en una secuencia con una alta eficiencia, por lo que puede obviarse la necesidad de marcadores seleccionables, evitando la necesidad de manipulaciones adicionales para su eliminación (Sander and Joung, 2014).

El paso crucial para realizar una edición genómica específica es la creación de una rotura de doble hebra en el ADN en el locus que se quiere modificar (Carroll, 2011). La rotura de doble hebra es realizada por nucleasas y puede repararse por al menos dos métodos diferentes que funcionan en casi todos los tipos celulares y organismos: El método de reparación más frecuente es el ligamiento de extremos no homólogos, que puede introducir inserciones y deleciones de distintas longitudes en el momento en el que se produce la unión de los extremos. Por otro lado, la reparación puede estar dirigida por homología, si ésta se realiza utilizando una cadena intacta con la que los fragmentos generados tienen homología. De este modo pueden introducirse mutaciones puntuales o insertar secuencias deseadas mediante la recombinación del locus diana con ADN exógeno que incluya la mutación deseada.

La aparición de distintas endonucleasas de ADN específicas de sitio ha propiciado la edición genómica y ha dotado a los investigadores de la capacidad para manipular cualquier gen en un amplio espectro de tipos celulares y organismos. Las tecnologías más comúnmente utilizadas son ZFN, TALENs y CRISPR/Cas9. Estos sistemas generalmente se configuran para inducir una rotura de doble hebra (*"double strand break"*) en una secuencia específica del ADN. Esta rotura activa la maquinaria de reparación del ADN, dando lugar a la inactivación del gen (*knock-out*) por unión mutagénica de extremos no homólogos), o a la corrección o integración de un gen si tiene lugar una reparación dirigida por homología en presencia de un ADN molde exógeno (Liu *et al.*, 2015).

De la generación actual de tecnologías de edición de genoma, la que más rápidamente se ha desarrollado ha sido la basada en la endonucleasa guiada por ARN conocida como Cas9, basada en el sistema inmune bacteriano CRISPR (*Clustered regularly interspaced short palindromic repeats*), que puede dirigirse fácilmente frente a cualquier región genómica de elección utilizando un ARN guía único, fácil de diseñar y sintetizar. En contraposición a ZFN y TALEN que usan la interacción ADN-proteína, las nucleasas guiadas por ARN se basan en el apareamiento entre un ARN artificial y su diana específica en el ADN (Sander and Joung, 2014).

Estas tecnologías tienen el potencial de revolucionar la investigación en la biología básica y transformar la medicina, como se ha demostrado en el ensayo de una terapia

para VIH/SIDA basada en la modificación, mediante ZFN, del receptor de la quemoquina (motivo C-C) receptor 5 del VIH.

2. OBJETIVO

En esta revisión bibliográfica, se realizará un análisis del sistema CRISPR/Cas y su aplicación a la ingeniería genética, exponiendo las funciones que puede desempeñar y su utilidad en terapia génica. También, se valorarán los inconvenientes que presenta esta técnica en su aplicación, así como el debate ético que existe en torno a su uso.

3. ANTECEDENTES

3.1 Nucleasas con dedos de Zinc (*Zinc Fingers Nuclease, ZFN*)

Las nucleasas con dedos de zinc (ZFN) son proteínas híbridas formadas por la fusión del dominio de rotura del ADN no específico de la endonucleasa de restricción *FokI* con varios dedos de zinc. La formación de dímeros de ZFN da lugar a roturas de doble hebra en el ADN, estimulando los mecanismos de reparación. Mediante el diseño de los dominios en los dedos de zinc es posible dirigir la especificidad de unión de éstos al ADN, dirigiendo las ZFN a sitios específicos del genoma (Gaj, Gersbach and Barbas, 2013).

Un dominio en dedo de zinc individual consta de unos 30 aminoácidos con una configuración beta-beta-alfa muy conservada (Beerli and Barbas, 2002). Muchos aminoácidos en la superficie de la hélice alfa hacen contacto con 3 pares de bases en el surco mayor del ADN, con niveles variables de selectividad. La estructura modular de las proteínas con dedos de zinc permite el diseño de proteínas de unión al ADN personalizadas. La clave para la aplicación de los dedos de zinc al reconocimiento de regiones específicas del ADN fue el desarrollo de arrays artificiales que contienen más de 3 dominios de dedos de zinc. Esto fue posible gracias al descubrimiento de una secuencia *linker* muy conservada, que permitió la construcción de proteínas con dedos de zinc sintéticas que reconocían una secuencia de ADN de 9-18pb de longitud (Liu *et al.*, 1997), suficiente para conferir especificidad en genomas de 68 mil millones de pares de bases. Esta especificidad permitió seleccionar por primera vez secuencias específicas en el genoma humano (Kim and Pabo, 1998; Beerli, Dreier and Barbas 3rd, 2000).

Durante muchos años, la tecnología ZFN fue la única alternativa para crear nucleasas personalizadas con secuencias específicas de unión al ADN, que permiten la selección de prácticamente cualquier secuencia para la edición genómica.

3.2 TALEN (*Transcription activator-like effector nucleases*)

Las TALEN son una alternativa a las ZFN que también utiliza proteínas de unión al ADN. En este caso la nucleasa específica se obtiene fusionando el dominio de unión al ADN del efector TAL (*Transcription activator-like effector – TALE*) con el dominio de rotura del ADN de la nucleasa *FokI*. Con estas nucleasas se introducen roturas de doble hebra en secuencias específicas que activan la respuesta al daño del ADN y permiten alteraciones personalizadas en función del mecanismo de reparación utilizado por la célula.

En la naturaleza las proteínas TALE se han encontrado en bacterias del género *Xanthomonas*, patógenos de plantas, y contienen dominios de unión al ADN compuestos por 33-35 aminoácidos que reconocen un par de bases concreto. La especificidad de los TALE está determinada por dos aminoácidos muy variables, los di-residuos de repeticiones variables (RVDs) (Kaphingst, Persky and Lachance, 2010; Mak *et al.*, 2012).

El reconocimiento de pares de bases únicas por las repeticiones de unión al ADN de las proteínas TALE permiten una mayor flexibilidad en su diseño que las proteínas dedos de zinc (que reconocen tripletes). Sin embargo, la clonación de las repeticiones de los arrays de TALE presenta una mayor dificultad técnica debido al excesivo número de secuencias repetidas. Para superar este problema, se han desarrollado distintos métodos que permiten un ensamblaje rápido de arrays TALE personalizados. Estas estrategias incluyen el ensamblaje “Golden Gate” (Cermak *et al.*, 2011), el ensamblaje en base sólida (Briggs *et al.*, 2012; Reyon *et al.*, 2013), y técnicas de clonación independientes de ligamiento (Schmid-burgk *et al.*, 2014). La única limitación respecto a la unión a la secuencia diana es que ésta debe empezar con una timina (Gaj, Gersbach and Barbas, 2013).

La versatilidad de ZFN y TALEN viene dada por la posibilidad de personalizar el dominio de unión al ADN, lo que permite reconocer prácticamente cualquier secuencia. Además de con nucleasas, estos módulos de unión al ADN se pueden combinar con muchos otros dominios efectores que permiten afectar la estructura

genómica y su función, como por ejemplo activadores y represores de la transcripción, recombinasas, transposasas, metiltransferasas del ADN o acetiltransferasas de histonas. De este modo, la habilidad para generar alteraciones genéticas depende fundamentalmente de la especificidad de unión al ADN y la afinidad de los dedos de zinc o proteínas TALE diseñadas.

Aunque los sistemas de edición genómica ZFN y TALEN siguen siendo útiles, se piensa que en el futuro serán sustituidos plenamente por el sistema CRISPR/Cas, por su especificidad y por su facilidad de uso. Sin embargo, hay que tener en cuenta que, aun así, ninguna tecnología de edición genética es 100% específica, y todas tienen el riesgo de producir mutaciones fuera de la secuencia diana (Ivanov *et al.*, 2015; Kennedy and Cullen, 2016) .

4. El sistema CRISPR/Cas

4.1 Origen

El sistema CRISPR (*Clustered regularly interspaced short palindromic repeats*) tiene varias funciones en bacterias, como la regulación de comportamiento colectivo o patogenicidad (Makarova *et al.*, 2006), aunque es más conocido por formar parte del “sistema inmune” de procariotas (95% de las archeas y 48% de eubacterias, (Jore, Brouns and van der Oost, 2012)) que confiere protección contra bacteriófagos u otros elementos móviles.

Para las bacterias es importante la habilidad para distinguir sus ácidos nucleicos de ácidos nucleicos exógenos, que se pueden introducir en las células procariotas por infección, transducción, conjugación o transformación, pudiendo producir un daño en la célula. Entre los mecanismos de defensa innatos y no específicos que tienen las bacterias están los sistemas de restricción-modificación, la infección abortiva y los sistemas de exclusión superficiales (Samson *et al.*, 2013). Además, existen otros mecanismos que confieren inmunidad adaptativa y hereditaria contra ácidos nucleicos exógenos, como el sistema CRISPR/Cas, que imita en cierto modo el sistema inmune de eucariotas (Burmistrz and Pyrc, 2015; Ivanov *et al.*, 2015).

4.2 Descubrimiento

Los elementos del sistema CRISPR/Cas se conocían ya en 1987, aunque estas regiones repetidas no se asociaron con ninguna función específica de bacterias. (Ishino *et al.*, 1987). En 1993, fue un investigador español, Francis Mojica, el que

estudiando una bacteria extremófila encontró patrones de repetición de secuencias nucleotídicas. Concretamente observó 14 repeticiones de fragmentos de ADN de 30 bp de longitud y parcialmente palindrómicos que se sucedían a distancias regulares y que a su vez se encontraban separados entre sí por otras regiones nucleotídicas, encontrándose éstas en una región no codificante.

Tras el descubrimiento en el año 2002 del mismo sistema en otra bacteria (Jansen *et al.*, 2002), se postularon muchas funciones, incluyendo reordenaciones cromosómicas, modulación de la expresión genética y reparación del ADN (Babu *et al.*, 2012). Sin embargo, la similitud entre algunas secuencias espaciadoras de CRISPR y secuencias de virus y plásmidos llevaron a la conclusión de que los sistemas CRISPR están implicados en la inmunidad adaptativa frente a ácidos nucleicos exógenos (Makarova *et al.*, 2006). Finalmente, el descubrimiento de secuencias espaciadoras homólogas a genes cromosómicos del huésped, indica que los sistemas CRISPR/Cas, además de inmunidad frente a fagos, pueden tener un papel en autoinmunidad o como mecanismos reguladores.

En el año 2012, se describió la nucleasa Cas del sistema CRISPR/Cas. Se trata de una enzima que utilizando un ARN guía, es capaz de cortar secuencias de ADN con secuencia complementaria al ARN guía (Cruz and Freedman, 2018).

4.3 Tipos de sistemas CRISPR/Cas

Los sistemas CRISPR/Cas son muy diversos. Se clasifican en seis tipos distintos que se agrupan en dos amplias clases basadas en filogenia y mecanismos de actividad. Los sistemas de clase 1 (tipos I, III, IV) emplean complejos con actividad de supervisión y escisión (complejo crRNA-efector), mientras que los sistemas de clase 2 (tipos II, V, y VI) emplean una única proteína que cumple todas esas funciones (la nucleasa Cas9) (Makarova *et al.*, 2015).

Los sistemas de tipo II, son muy diferentes a los de tipo I y III, y son con diferencia los más simples respecto al número de genes. El gen que caracteriza a los sistemas de tipo II es el *cas9*, que codifica la nucleasa Cas9, una proteína con múltiples dominios que combina las funciones del complejo RNA-efector con la rotura en el ADN diana (Wang *et al.*, 2011), contribuyendo también en el proceso de adaptación (Heler *et al.*, 2015; Wei, Terns and Terns, 2015). Además de *cas9*, todos los loci de los sistemas de tipo II contienen *cas1* y *cas2* y la mayoría también codifican un *tracrRNA*, que es

parcialmente complementario a las repeticiones con el array CRISPR correspondiente (Chylinski, Le Rhun and Charpentier, 2013; Briner *et al.*, 2014; Chylinski *et al.*, 2014).

Además, debido a la similitud entre la secuencia de Cas9 y la de sus homólogos que no están relacionados con CRISPR, Cas9 no se puede usar como la única señal de identidad de los sistemas de tipo II. Sin embargo, la presencia de *cas9* en la vecindad de los genes *cas1* y *cas2* es característica de un loci de tipo II (Makarova *et al.*, 2015).

Los sistemas CRISPR-Cas de tipo II se pueden clasificar en tres subtipos. El subtipo II-A incluye un gen adicional, *csn2*, que se considera característico de este subtipo. Las variantes largas y cortas de *Csn2* forman clusters compactos cuando se superponen con la filogenia de Cas9 y parece que corresponden a dos variantes del subtipo II-A. Recientemente se ha demostrado que las cuatro proteínas Cas del subtipo II-A participan en la adquisición del espaciador. El subtipo II-B no tiene *csn2*, pero incluye *cas4*, que es típico de los sistemas de tipo I. El loci del subtipo II-C solo tiene tres genes codificantes para proteína (*cas1*, *cas2* y *cas9*) y es el sistema CRISPR-Cas de tipo II más común en bacterias. En la filogenia de Cas9, los subtipos II-A y II-B son monofiléticos mientras que el subtipo II-C es parafilético respecto a II-A (es decir, que el subtipo II-A se origina del II-C) (Chylinski, Le Rhun and Charpentier, 2013; Makarova *et al.*, 2015).

4.4 Estructura

Un locus típico del sistema CRISPR/Cas incluye un array CRISPR, formado por repeticiones directas cortas separadas por secuencias de ADN cortas y variables (espaciadores), que están flanqueados por diversos genes de proteínas Cas (Makarova *et al.*, 2015). El sistema CRISPR está comprendido por dos módulos: una endonucleasa Cas asociada a CRISPR que se encarga de inducir una rotura en el ADN de doble cadena y un ARN (crRNA) que especifica la secuencia de ADN diana donde va a cortar la nucleasa Cas (Luo, 2016).

Según un análisis funcional del sistema CRISPR/Cas (**Figura 1**) de *Streptococcus pyogenes* (Cruz and Freedman, 2018), para que el sistema se encuentre activo solo se necesitan 3 componentes: la endonucleasa Cas9, un crRNA específico complementario a la secuencia y un crRNA trans-activador estructural (tracrRNA). Los tres componentes pueden reducirse a dos uniendo tracrRNA y crRNA para formar un ARN guía único (sgRNA).

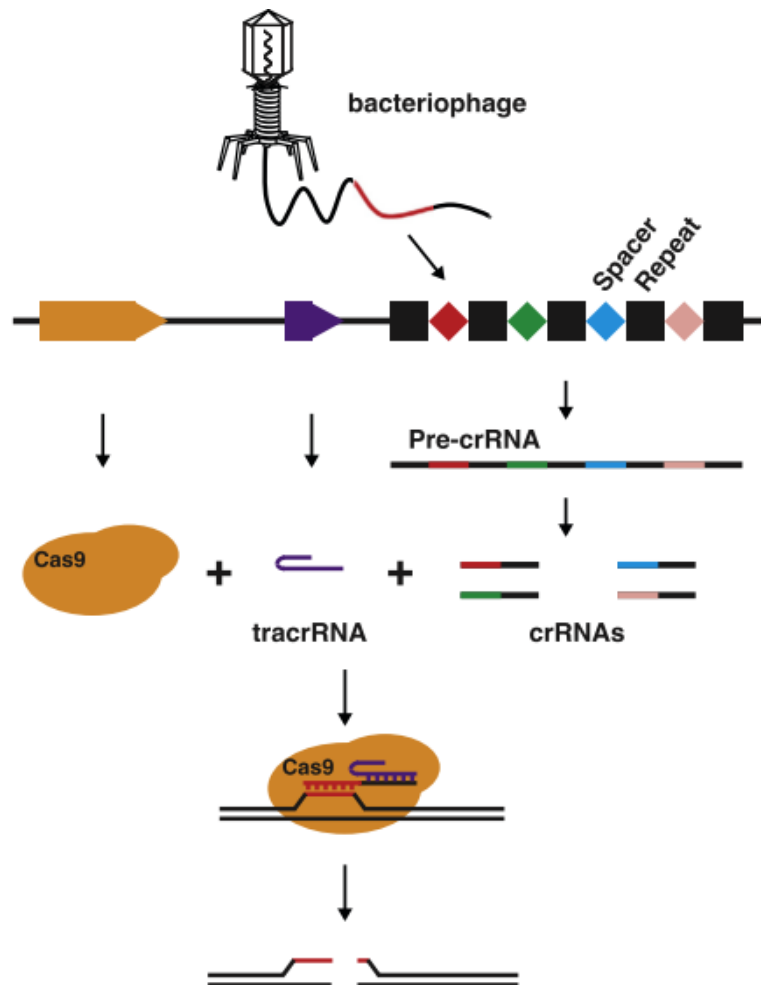


Figura 1. Esquema del sistema CRISPR/Cas (Cao, Xiao and Yan, 2018)

La endonucleasa Cas9 (**Figura 2**) incluye dos dominios de nucleasa (RuvC y HNH) y un cluster característico rico en arginina, y evolucionó a partir de genes de elementos transponibles que no están asociados con CRISPR (Chylinski *et al.*, 2014). El dominio RuvC corta cadenas de ADN no complementarias, mientras el dominio HNH corta en cadenas de ADN complementarias (Liu *et al.*, 2017).

El sgRNA (**Figura 3**) se compone de crRNA y tracrRNA. El crRNA contiene un protoespaciador de 20 nucleótidos y una secuencia adicional que es complementaria al tracrRNA. El tracrRNA hibrida con el crRNA y se une a la proteína Cas9, formando el complejo CRISPR-Cas9/sgRNA para cortar el ADN de doble cadena en las secuencias objetivas del genoma. El dúo tracrRNA-crRNA normalmente se prepara como una cadena única sgRNA que contiene dos segmentos imprescindibles: una estructura dúplex de RNA en el extremo 3' para unirse a Cas9 y una secuencia guía en el extremo 5' para unirse a la secuencia de ADN objetivo. Este sistema de dos

componentes es simple pero efectivo: sgRNA reconoce una secuencia específica en el genoma, y Cas9 actúa como un par de tijeras para cortar en esa secuencia de ADN.



Figura 2. Endonucleasa Cas9 (Liu et al., 2017).

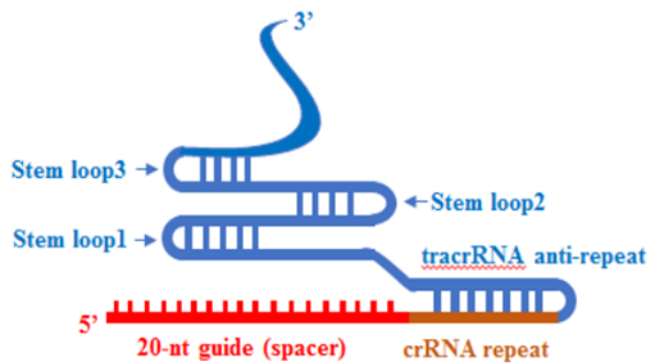


Figura 3. sgRNA (Liu et al., 2017).

El sgRNA combina una secuencia invariable 'scaffold' que se une a Cas y una secuencia espaciadora variable que aporta especificidad de secuencia. El espaciador tiene aproximadamente 20 nucleótidos de longitud y termina inmediatamente aguas arriba de una secuencia NGG característica llamada motivo adyacente de protoespaciador (PAM) (**Figura 4**). Como hay 4 posibles nucleótidos para cada posición en el sgRNA, el espaciador proporciona hasta 4^{20} combinaciones de secuencias posibles (>1 trillón). Esto excede en gran medida el número de pares de bases en cualquier genoma conocido, aportando suficiente especificidad para identificar secuencias específicas únicas (Mali et al., 2013; Cruz and Freedman, 2018).

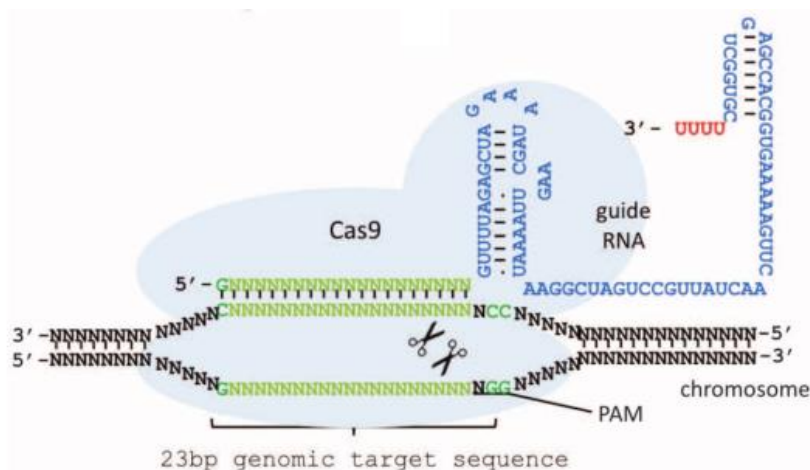


Figura 4. Cas9 desenrolla la doble hebra de ADN y corta ambas hebras tras el reconocimiento de la secuencia diana mediante sgRNA, solo si PAM está presente en el extremo 3' (Mali et al., 2013).

4.5 Mecanismo de acción del sistema CRISPR/Cas

La inmunidad mediante los sistemas CRISPR/Cas9 tiene lugar en tres etapas: adaptación, expresión e interferencia (**Figura 5**) (Burmistrz and Pyrc, 2015; Makarova *et al.*, 2015):

- La etapa de **adaptación** ocurre con la incorporación de fragmentos de ADN extraño (fragmentos de virus invasores o plásmidos, conocidos como protoespaciadores) en el array de CRISPR. Estos fragmentos constituyen nuevos espaciadores y funcionan como “secuencias de memorias” preparadas para luchar contra objetivos específicos en caso de que se produzcan invasiones subsiguientes de virus o plásmidos con secuencias similares. La adquisición de protoespaciadores en muchos sistemas CRISPR/Cas, requieren el reconocimiento de un motivo adyacente al protoespaciador (PAM) en el ADN diana
- Durante la etapa de **expresión**, el array CRISPR se transcribe en un precursor, pre-crRNA, que tras el proceso de maduración da lugar a CRISPR RNAs (crRNAs).

En la **interferencia**, los crRNAs junto a las proteínas Cas, funcionan como guías que seleccionan específicamente ácidos nucleicos de virus o plásmidos conocidos, que son posteriormente cortados por la nucleasa Cas. La unión a la secuencia diana depende del apareamiento con la región espaciadora del crRNA (Burmistrz and Pyrc, 2015; Makarova *et al.*, 2015).

En presencia de un ADN invasor, el sistema CRISPR/Cas adquiere pequeños fragmentos cortos del ácido nucleico extraño y los inserta como nuevos espaciadores en el inicio del array CRISPR, estando cada espaciador flanqueando a ambos lados por secuencias repetidas directas (K. I. Liu *et al.*, 2016; McMahon *et al.*, 2018).

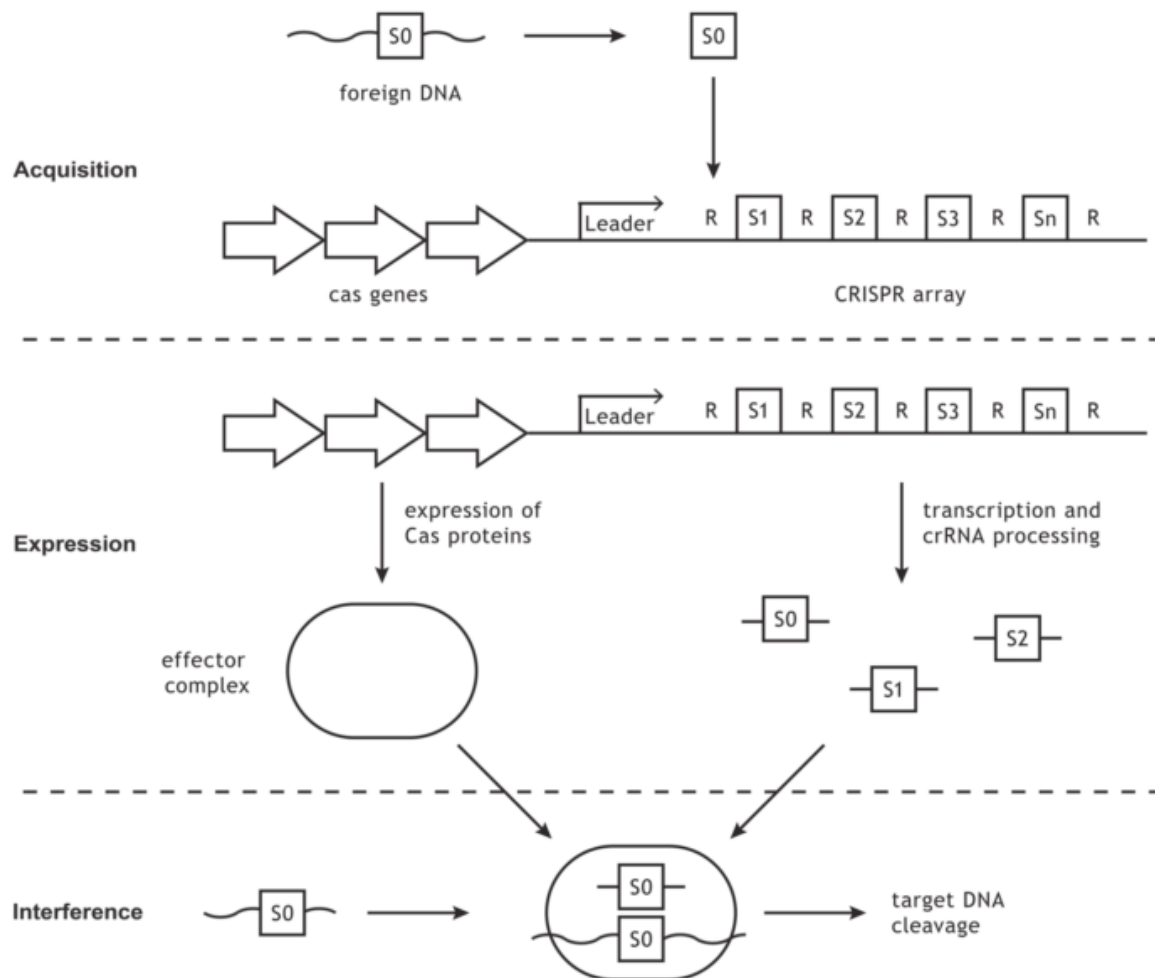


Figura 5. Etapas del funcionamiento del sistema inmune mediado por CRISPR/Cas (Burmistrz and Pyrc, 2015)

4.6 Familias de proteínas Cas

Hay una gran diversidad de proteínas Cas involucradas en las diferentes etapas de la actividad del sistema CRISPR. Se pueden dividir en cuatro grupos funcionales distintos (**Figura 6**): adaptación (adquisición del espaciador); expresión (procesamiento del crRNA y unión al objetivo); interferencia (rotura en la diana); y función auxiliar (otras funciones asociadas a CRISPR, como regulación de la expresión génica).

- El módulo de **adaptación** es muy uniforme en los sistemas CRISPR/Cas. Consiste en un complejo formado por las proteínas **Cas1** y **Cas2**, y a veces también involucra a **Cas4**. Cas1 es una integrasa que media en la inserción de nuevos espaciadores en los arrays de CRISPR cortando en sitios específicos en las repeticiones. La función de Cas2 no se conoce, pero es una proteína homóloga a las toxinas de ARNm

interferasa de sistemas toxina-antitoxina. Aunque Cas2 tiene actividad RNasa y DNasa, sus residuos catalíticos son dispensables para el proceso de adaptación, lo que indica que estas actividades no están involucradas en el proceso. En los sistemas CRISPR de tipo II, también aparece **Cas9** (Makarova *et al.*, 2015).

-Los módulos de **expresión e interferencia** están representados por los complejos multisubunidad crRNA-efector, o solo Cas9 en sistemas de tipo II. En los sistemas de tipo I y III, **Cas6** suele ser la ARN endonucleasa encargada de la maduración del crRNA, mientras que en los sistemas de tipo II una RNasa III y traRNA son los responsables de este proceso. También hay sistemas CRISPR/Cas de tipo II (*Neisseria meningitidis*) en los que el crRNA con el extremo 5' ya maduro se transcribe directamente mediante promotores internos, y no tiene lugar el procesamiento de crRNA.

-En el módulo de **interferencia**, el complejo crRNA-efector (en sistemas tipo I y tipo III) o Cas9 (en sistemas tipo II) combina actividad nucleasa con dominios de unión al ARN.

La **digestión en la diana** está catalizada por la nucleasa **Cas3** en sistemas tipo I, por la acción combinada de **Cas7** y **Cas10** en sistemas del tipo III, o por **Cas9** en los sistemas tipo II. En los sistemas de tipo I, el dominio de la nucleasa está fusionado con la helicasa Cas3' o está codificado por otro gen, *cas3''*, mientras que en los sistemas de tipo III, un dominio de HD nucleasa distinto se fusiona a Cas10 para cortar una cadena de ADN específica durante la interferencia. En los sistemas tipo II, el dominio de la nucleasa RuvC (pliegue RNasa H) y el dominio de la nucleasa de Cas9 cortan en cada una de las hebras del ADN diana. Hay que destacar que el multidominio de la proteína Cas9 es necesario en las tres etapas de la inmunidad basada en CRISPR (adaptación, expresión e interferencia) en los sistemas de tipo II y, por lo tanto, concentra varias funciones del sistema CRISPR-Cas en una sola proteína.

El módulo **auxiliar** es una combinación de varias proteínas y dominios que, con la excepción de Cas4, son mucho menos comunes que las proteínas principales Cas en los sistemas CRISPR/Cas. A parte de su papel en la adaptación, se piensa que Cas4 contribuye a la muerte celular programada mediada por CRISPR/cas. Otros componentes del módulo auxiliar incluyen: proteínas con pliegues del dominio de

Rossmann asociados a CRISPR, que pueden regular la actividad de CRISPR-Cas (en sistemas tipo I y tipo III), y el bucle inactivado P de la ATPasa Csn2, que forma un anillo homotetramérico que sitúa el ADN de doble cadena en el hueco central (sistemas tipo II). La proteína **Csn2** no es necesaria para la interferencia, pero tiene un papel en la integración del espaciador, probablemente para prevenir el daño de la rotura de doble cadena en el ADN cromosómico. Los genes del módulo auxiliar a menudo se encuentran fuera del loci de CRISPR-Cas, y sus funciones no han sido caracterizadas en profundidad (Makarova *et al.*, 2017).

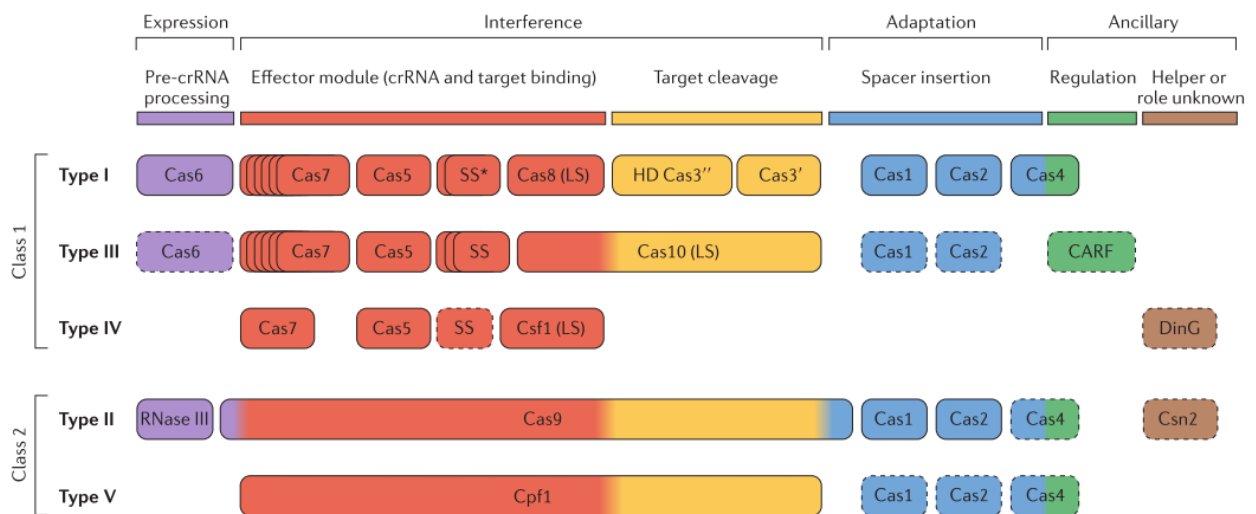


Figura 6. Clasificación funcional de las proteínas Cas. Componentes dispensables, con líneas discontinuas (Makarova *et al.*, 2015).

4.7 Mecanismos de defensa de virus: Inactivación del sistema CRISPR/Cas

En la batalla por la supervivencia, muchos virus han desarrollado mecanismos para evitar la defensa del sistema inmune bacteriano, entre éstos se incluyen mecanismos para evitar la acción del sistema CRISPR-Cas (van der Oost and Brouns, 2015).

Recientemente, se han identificado proteínas virales que suprimen la inmunidad proporcionada por CRISPR-Cas (Wiedenheft, 2013). En fagos específicos de *Pseudomonas aeruginosa* se ha identificado un amplio rango de proteínas anti-CRISPR (proteínas Acr) (Samson *et al.*, 2013; Bronkhorst and Van Rij, 2014; Bondydenomy *et al.*, 2016). La primera vez que se descubrieron las proteínas Acr fue en el análisis de cepas de *Pseudomonas* que contienen profagos en su cromosoma. La mayoría de ellas contienen un sistema CRISPR-Cas de tipo IF, que debería aportarles resistencia a fagos. Sin embargo, se observó que incluso teniendo espaciadores específicos para los fagos el sistema CRISPR-Cas no proporcionaba inmunidad a las bacterias. El análisis molecular de las cepas sensibles a los fagos identificó las proteínas Acr responsables (Bondydenomy *et al.*, 2016). En un estudio posterior, se demostró que hay tres proteínas Acr que inhiben el sistema CRISPR-Cas de tipo IF: AcrF3, que actúa uniéndose a la proteína Cas3 del sistema, evitando la actividad nucleasa-helicasa necesaria para la degradación del ADN del fago; AcrF1 y AcrF2, inhiben la unión del crRNA al ADN (**Figura 7**). Estas proteínas son específicas de *Pseudomonas* y del tipo de sistema CRISPR-Cas, por lo que no funcionan entre cepas diferentes u otros tipos de sistemas CRISPR-Cas (estas Acr son específicas para el sistema de *Pseudomonas* I-F, pero no sirven en el sistema I-F de *E. coli* o para el sistema I-E de *Pseudomonas* (van der Oost and Brouns, 2015)).

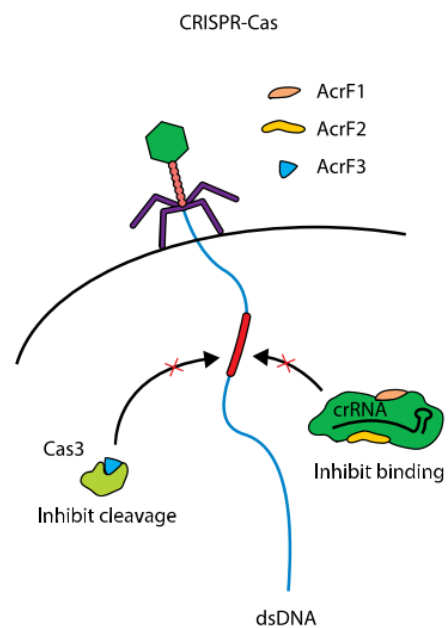


Figura 7. Inhibición mediada por proteínas AcrF (van der Oost and Brouns, 2015).

diferentes u otros tipos de sistemas CRISPR-Cas (estas Acr son específicas para el sistema de *Pseudomonas* I-F, pero no sirven en el sistema I-F de *E. coli* o para el sistema I-E de *Pseudomonas* (van der Oost and Brouns, 2015)).

Las proteínas Acr no son solo exclusivas del sistema CRISPR-Cas de tipo I-F. Así, aunque se han descubierto pocas hasta la fecha, también se han descubierto proteínas Acr para el sistema CRISPR-Cas de tipo II (Harrington *et al.*, 2017), de

utilidad para desactivar el sistema en su aplicación para edición genómica (Patterson *et al.*, 2016; Pawluk *et al.*, 2016).

Las proteínas Acr identificadas para el sistema CRISPR-Cas de tipo IIC (AcrIIC1, 2 y 3), son específicas de *Neisseria meningitidis* y actúan sobre Cas9. AcrIIC1 se une a Cas9, permitiendo que se una al ADN del fago, pero obstruyendo el sitio activo del dominio nucleasa HNH, por lo que evita que corte el ADN. AcrIIC3 inhibe la unión de Cas9 al ADN y causa su dimerización (**Figura 8**).

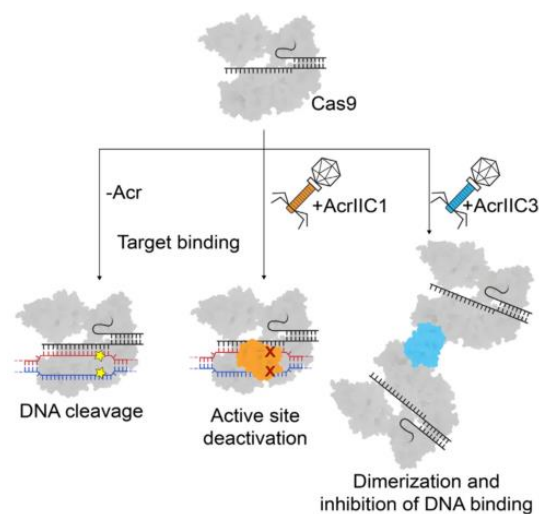


Figura 8. Inhibición mediada por proteínas AcrII (Harrington *et al.*, 2017)

Es interesante la posibilidad de que AcrIIC1 puede ser utilizada en aplicaciones de edición genómica, obviándose la necesidad de generar enzimas Cas9 defectuosas (dCas9) cuando se utiliza el sistema para regulación de la expresión génica (Harrington *et al.*, 2017).

5. EL SISTEMA CRISPR/CAS9 APLICADO A EDICIÓN GENÓMICA

CRISPR no es necesariamente más eficiente que otras tecnologías anteriores. Sin embargo, este sistema tiene muchas ventajas respecto a otros en términos de coste y facilidad de uso (Miyagi, Lu and Humphreys, 2016), sugiriendo que podría ser útil para edición genómica y aplicaciones para selección de genes (Gasiunas *et al.*, 2012; Jinek *et al.*, 2012). Rápidamente se demostró que se podía utilizar en células mamíferas (incluyendo humanas), aumentando la posibilidad de su uso para terapia génica (Freedman *et al.*, 2013, 2015). Esto ha provocado que se extienda el uso de CRISPR en laboratorios de investigación, y más recientemente, en laboratorios de ensayos clínicos (Cruz and Freedman, 2018).

Los sistemas CRISPR/Cas de tipo II son los que se utilizan preferentemente para la edición genética, ya que solo usan una única proteína Cas (Cas9) con múltiples dominios, por lo que es más simple y más fácil de manipular (Liu *et al.*, 2017).

5.1 Funcionamiento

Para editar un gen utilizando CRISPR/Cas9, lo primero que se necesita es diseñar un sgRNA para guiar a la endonucleasa Cas9 hacia la secuencia que se quiere modificar. Para ello existen varias herramientas online, como *CRISPy-Web*.

Una vez diseñado el sgRNA, el complejo Cas9-sgRNA se introduce en la célula (las estrategias de entrada se discuten en el apartado 5.2).

El complejo Cas9-sgRNA reconoce una secuencia de ADN de 20 nucleótidos que es complementaria al crRNA y aguas arriba de un motivo adyacente, el protoespaciador NGG (PAM). La unión entre Cas9 y PAM es esencial para que se produzca la iniciación del reconocimiento de la secuencia diana. Cas9 desenrolla la doble cadena de ADN aguas arriba de PAM para permitir que 20 nucleótidos de crRNA se introduzca en la cadena de ADN y busquen la secuencia diana. Cas9 forma un dúplex entre crRNA y el ADN diana y estimula la actividad nucleasa para generar un extremo romo, tras una rotura de doble hebra a 3 nucleótidos aguas arriba de PAM.

La posterior reparación del ADN puede tener lugar mediante la unión de extremos no homólogos (NHEJ) o la reparación dirigida por homología (HDR):

- La vía **NHEJ** ocurre durante el ciclo celular, siendo la primera por la que se reparan las lesiones en el ADN producidas por CRISPR (Cruz and Freedman, 2018). Normalmente da lugar a mutaciones del tipo indel (inserción o delección) en la secuencia original, que pueden alterar los marcos abiertos de lectura (*frameshifts*). Si estas mutaciones tienen lugar en exones, también pueden silenciar el gen diana.
- En la reparación mediante **HDR**, el ADN se corrige basándose en un ADN molde de secuencia similar. Este molde de ADN puede residir en algún lugar del genoma, (como en una cromátida hermana sin dañar después de la replicación de ADN (Freedman *et al.*, 2013, 2015; Ma *et al.*, 2017)) o bien puede ser suministrado por un investigador. En este caso, se prepara el molde de ADN para que contenga la mutación deseada (Luo, 2016), permitiendo utilizar el sistema CRISPR para

producir mutaciones específicas en el genoma y proporcionando un mecanismo de reparación más preciso que la reparación mediante NHEJ (**Figura 9**).

Normalmente, NHEJ es la vía más frecuente, aunque en ciertos tipos celulares puede mejorarse la proporción de HDR sincronizando las células en las etapas de replicación o división del ciclo celular (Lin *et al.*, 2014).

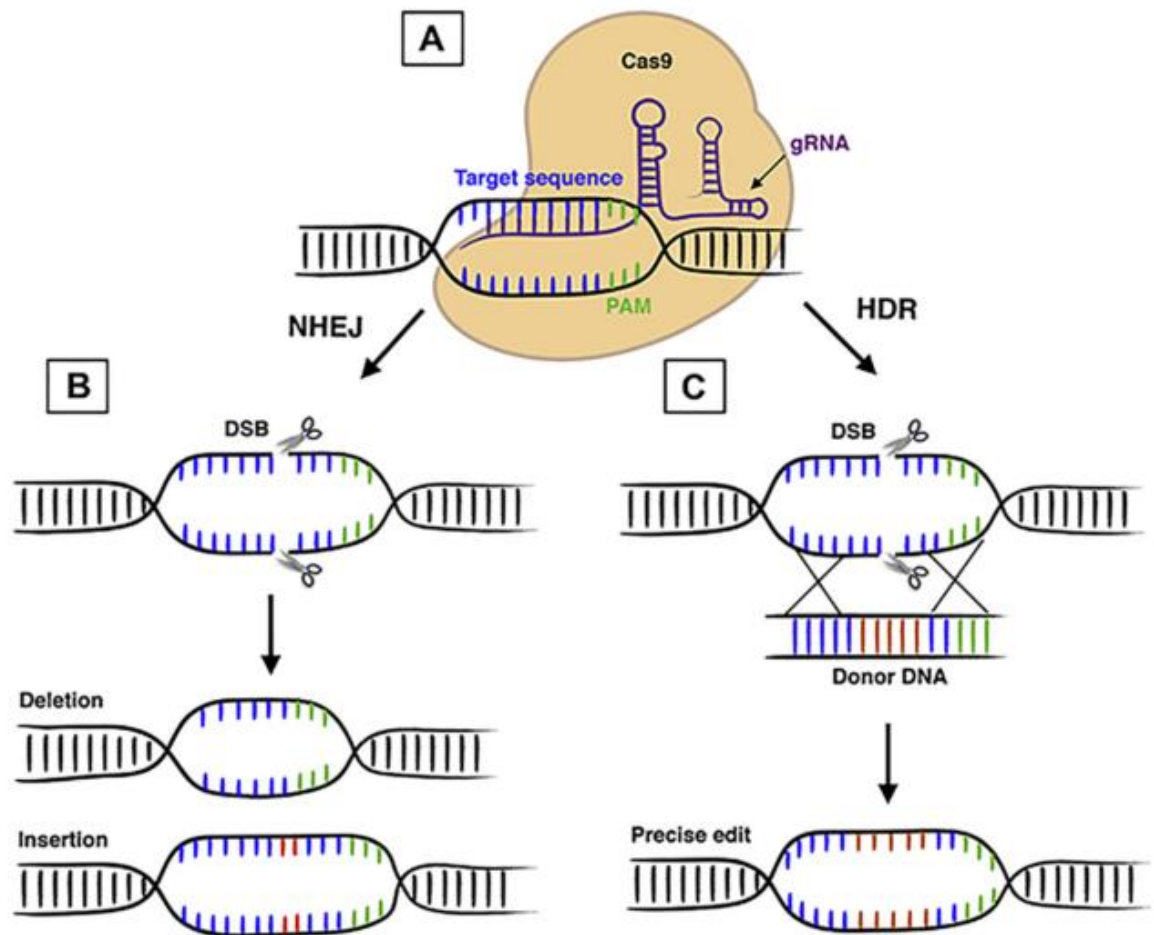


Figura 9. Esquema del sistema CRISPR para la edición genómica. (A) Un ARN guía único (sgRNA) se une a una secuencia de ADN complementaria aguas arriba de un motivo adyacente protoespaciador NGG (PAM). Esta unión dirige a la nucleasa Cas9 al ADN diana, donde produce una rotura de doble hebra en el ADN. (B) En ausencia de una secuencia homóloga para la reparación, la rotura de doble hebra se repara por apareamiento de extremos no homólogos (NHEJ), un proceso propenso a errores que a menudo genera mutaciones del tipo indel en el sitio de reparación. (C). En presencia de un ADN con homología que sirva de molde que esté presente un molde de ADN, se pueden crear ediciones precisas mediante reparación dirigida por recombinación homóloga (Cruz and Freedman, 2018).

Una de las características más destacadas de CRISPR/Cas9 es que los módulos de sgRNA y de Cas9 se codifican por separado. Esto, permite que cada uno se pueda

modificar, evolucionar y optimizar sin afectar a la función del otro. Además, el pequeño requerimiento de PAM hace que prácticamente todos los loci genómicos puedan ser diana para el sistema CRISPR/Cas9 (Luo, 2016).

5.2 Estrategias para editar genes con CRISPR/Cas9

Hay tres formas para editar un genoma utilizando CRISPR/Cas9:

- El sistema basado en plásmidos es la estrategia más sencilla y conveniente ya que evita la transfección de múltiples componentes a la misma célula. Además, el sistema CRISPR/Cas9 basado en plásmidos presenta mayor estabilidad que el sistema que combina el mRNA de Cas9 con sgRNA. En este grupo hay variantes que difieren en el vector utilizado. El proceso general de este método consiste en digerir con enzimas de restricción el plásmido y usarlos para ligarlos con oligonucleótidos complementarios a una secuencia diana concreta (Ran *et al.*, 2013). Este sistema presenta varios inconvenientes: el plásmido debe llegar al núcleo, que es complicado. Además, es necesaria la expresión de Cas9 en el interior de las células, lo que requiere más tiempo para que se produzca la modificación en la secuencia objetivo. Por otro lado, también produce mayor número de efectos no específicos por corte en secuencias distintas a la secuencia diana (efectos '*off-target*') (Ran *et al.*, 2013).
- La introducción directa del mRNA de Cas9 y el sgRNA en las células diana editan el genoma después de la expresión de la proteína Cas9. La ventaja de administrar el mRNA de Cas9 es la expresión de la proteína Cas9 por un corto periodo de tiempo, lo que limita la duración de la edición genética. Además, la introducción de mRNAs tiene menos efectos inespecíficos que el método anterior por roturas fuera de la secuencia diana. Finalmente, el mRNA solo tiene que entrar en el citoplasma para que ocurran sus efectos y el uso de mRNA codificante para Cas9 muestra poca citotoxicidad. Sin embargo, el mRNA es muy inestable, siendo ésta la principal limitación de este método.
- El último método, y el más estudiado en los últimos años, consiste en la introducción directa de la proteína Cas9 formando un complejo con sgRNA (Niu *et al.*, 2014). La proteína Cas9 purificada se carga positivamente y puede formar un complejo con sgRNA de manera eficiente, llamado complejo ribonucleoproteína Cas9/sgRNA (RNP). La entrada directa de RNPs tiene muchas ventajas,

incluyendo la acción rápida, la alta eficiencia de edición génica, la ausencia de requerimientos de optimización de uso codónico ni de selección del promotor, y los reducidos niveles de efectos '*off-target*', citotoxicidad y respuesta inmune (Liu *et al.*, 2017).

Para introducir los componentes de cada método en las células diana, existen diferentes procedimientos. Entre los mecanismos físicos de entrada se utilizan: electroporación, microinyección, transducción inducida por osmocitosis y propanobetano, deformación mecánica de la célula, inyección hidrodinámica, nanopartículas, péptido que penetra en la célula, nanoestructura del ADN, nanopartículas de oro... etc., Por otro lado, también pueden utilizarse mecanismos de entrada basados en vectores virales, como los derivados de Adenovirus (AAV) o de Lentivirus.

Si bien todos los métodos han conseguido utilizarse con éxito para la inserción de los sistemas CRISPR/Cas9 en las células, a pesar de la simplicidad y seguridad de los **métodos físicos**, éstos no se pueden utilizar en aplicaciones terapéuticas *in vivo* debido a su escasa eficacia (aunque siguen siendo factibles para generar *knock-out* en líneas celulares y modelos animales). Por otro lado, los **vectores virales** han sido desarrollados durante tres décadas para introducir en las células varios ácidos nucleicos basados en terapéuticos, habiendo sido aprobados para usos clínicos. Aunque existe preocupación por la seguridad y la posibilidad de introducir mutaciones no deseadas, los mecanismos de entrada basados en vectores virales son los mecanismos más eficientes para introducir ácidos nucleicos en células mamíferas *in vitro* e *in vivo*. Como resultado, es el método elegido y el más expandido para introducir el sistema CRISPR/Cas9 basado en plásmidos en células mamíferas.

Entre los vectores virales, AAV es el método preferido para terapia génica y ha mostrado ser una promesa en estudios preclínicos, ya que es seguro y puede infectar una gran variedad de células, incluyendo células que no se dividen (Cruz and Freedman, 2018). Estudios recientes están usando vectores de AAV para la entrada local y sistémica de la proteína Cas9 y sgRNAs para terapia génica mediada por CRISPR.

5.3 Aplicaciones

5.3.1 Terapia génica

CRISPR puede usarse para seleccionar y reparar mutaciones causantes de enfermedades específicas con mayor eficiencia y facilidad de implementación que otros métodos de terapia génica (Cruz and Freedman, 2018). Actualmente, se está utilizando para corregir enfermedades en modelos animales.

En un estudio reciente con ratones a los que se les trasplantó células inmunes humanas, se utilizó el sistema CRISPR para bloquear la replicación del VIH, permitiendo extirpar el virus una vez ya estaba integrado en el ADN. Esta aplicación también demostró ser de utilidad para ratones con infecciones más graves (De Silva Feelixge and Jerome, 2017). En ratones y ratas, CRISPR se ha usado para tratar un array de enfermedades genéticas *in vivo*, como la distrofia muscular de Duchenne, la enfermedad de Huntington, y enfermedades de degeneración de la retina. Sin embargo, las aplicaciones clínicas de la edición genómica mediante CRISPR están poco desarrolladas.

5.3.2 CRISPR/Cas9 – Silenciamiento de genes y CRISPRi

Una de las aplicaciones más importantes del sistema en mamíferos es el silenciamiento de genes mediante la producción de ‘*knock-out*’ en una célula cuando Cas9 actúa sobre un exón en un gen. En este caso, el corte de Cas9 y la posterior reparación por unión de extremos no homólogos resulta en inserciones/deleciones que, a menudo, causan un desplazamiento en el marco de lectura en el gen diana. Esto lleva a la producción de una proteína truncada o no funcional, o a la degradación del ARNm mutante mediante la degradación del ARNm sin sentido. De esta manera, CRISPR/Cas9 puede dar lugar a genes que pierden su función en su totalidad (Luo, 2016; Tian *et al.*, 2017).

Un enfoque alternativo al silenciamiento de genes mediante CRISPR, se conoce como CRISPR de interferencia (CRISPRi). Este sistema alternativo aprovecha la alta afinidad de unión de Cas9 a su secuencia diana. Insertando mutaciones en ambos dominios de la nucleasa Cas9, se consigue una nucleasa “muerta” (*nuclease-dead protein*, dCas9), que se une a la secuencia de ADN diana sin producir rotura de doble hebra. Cuando se fija como diana el sitio de inicio de la transcripción de un gen, dCas9 puede suprimir su expresión interfiriendo con la maquinaria de transcripción. El

silenciamiento de genes con dCas9 se puede potenciar fusionando esta proteína con un represor transcripcional, como el dominio Krüppel (KRAB) de la proteína Kox1. Esta fusión, Cas9-KRAB, puede reclutar factores modificadores de la cromatina al sitio de unión, consiguiendo un silenciamiento prolongado del gen. De esta manera, CRISPRi se asemeja más a un ARN de interferencia, provocando la supresión de un gen, pero no necesariamente completa. Sin embargo, a diferencia de la endonucleasa Cas9, que se encuentra en la naturaleza, CRISPRi basado en dCas9, no altera la secuencia del gen diana de manera permanente por lo que su efecto es reversible (Luo, 2016; Tian *et al.*, 2017).

5.3.3 CRISPRa – Activación de genes

Otro enfoque del sistema es el activador CRISPR (CRISPRa). En este caso se activa la expresión génica utilizando una proteína de fusión entre dCas9 y un activador de la transcripción (VP64), un sgRNA modificado en bucle capaz de reclutar factores de la transcripción, o una combinación de ambos. Cuando se une a la región promotora de un gen, CRISPRa promueve el ensamblaje de factores de transcripción para dar lugar a la sobreexpresión de genes. Curiosamente, tanto CRISPRa como CRISPRi se pueden introducir en la misma célula de manera que se consigue una regulación génica combinada.

De esta manera, CRISPR/Cas es de gran utilidad para *knock-out* o *knock-down* de rutas competitivas, gracias a su capacidad de suprimir o reducir la expresión de un gen específico mediante CRISPRi.

Cuando se quiere aumentar la producción de una sustancia concreta, la eliminación de la ruta competitiva puede ser de gran ayuda. La síntesis artificial de un sgRNA para seleccionar la secuencia específica correspondiente a la ruta metabólica de interés, supone un método muy sencillo de inhibición.

Respecto al *knock-down*, CRISPRi actúa en el ADN bloqueando el inicio de la transcripción o elongación dependiendo de si se selecciona un promotor o un marco abierto de lectura. Además, CRISPRi también es una herramienta importante para la identificación funcional de genes claves de rutas competitivas (Tian *et al.*, 2017)

5.3.4 CRISPR/Cas9 – Edición de genomas

Otra aplicación clave de CRISPR/Cas9 es la edición de genomas de manera precisa. Suministrando a la célula un molde de ADN para la reparación directa por homología

tras la rotura de doble hebra producida por la endonucleasa Cas9, prácticamente se puede conseguir cualquier alteración en la secuencia de ADN que se proponga: mutaciones puntuales, pequeñas inserciones y deleciones, deleciones a gran escala y reordenamientos cromosómicos. En el contexto de la biología del cáncer, este enfoque es muy útil en la construcción de líneas celulares y animales con alteraciones genéticas complejas que imiten las mutaciones que provocan tumores en humanos (Luo, 2016; Tian *et al.*, 2017).

5.3.5 Reordenaciones cromosómicas y cáncer

Las células cancerosas se caracterizan por presentar anomalías en el número de cromosomas y en su estructura, como reordenamientos, que pueden dar lugar a la expresión inapropiada de genes o a la activación de oncogenes (Choi and Meyerson, 2014).

Las reordenaciones cromosómicas son las responsables de muchos tipos de cáncer humanos muy graves (Blasco *et al.*, 2014; Gostissa *et al.*, 2014). La generación de ratones genéticamente modificados que recrean estas reordenaciones cromosómicas para su estudio es muy costosa y requiere mucho tiempo (Rabbitts *et al.*, 2001). Sin embargo, el sistema CRISPR/Cas9 ha podido usarse eficientemente para introducir reordenaciones cromosómicas de ratones adultos *in vivo* en hígado y músculo (Olson, 2015; Xue *et al.*, 2015).

Cuando tiene lugar la reparación de la rotura en la doble hebra de ADN mediante NHEJ, del mismo modo que se pueden originar mutaciones del tipo indel, también pueden producirse reordenaciones cromosómicas (Klein *et al.*, 2012; Alt *et al.*, 2013; Gostissa *et al.*, 2014). Se ha demostrado que CRISPR/Cas9 puede utilizarse para inducir reordenaciones cromosómicas asociadas a cáncer con una alta frecuencia en líneas celulares humanas, diseñando el sgRNA para el loci de interés (**Figura 10**) (Choi and Meyerson, 2014).

En principio, se podría utilizar CRISPR/Cas9 para recrear cualquier tipo de reordenación cromosómica (Choi and Meyerson, 2014). Sin embargo, la utilización de este sistema también se relaciona con cambios en el genoma no deseados, como deleciones y, posiblemente, duplicaciones (Blasco *et al.*, 2014). Otro problema de CRISPR/Cas9 es que no actúe en la secuencia diana, como se demostró en estudios *in vitro*, aunque probablemente esto es menos relevante *in vivo* (Duan *et al.*, 2014;

Kuscu *et al.*, 2014; Wu *et al.*, 2015) Por lo tanto, hay que tener en cuenta todas estas posibilidades durante la aplicación de la técnica.

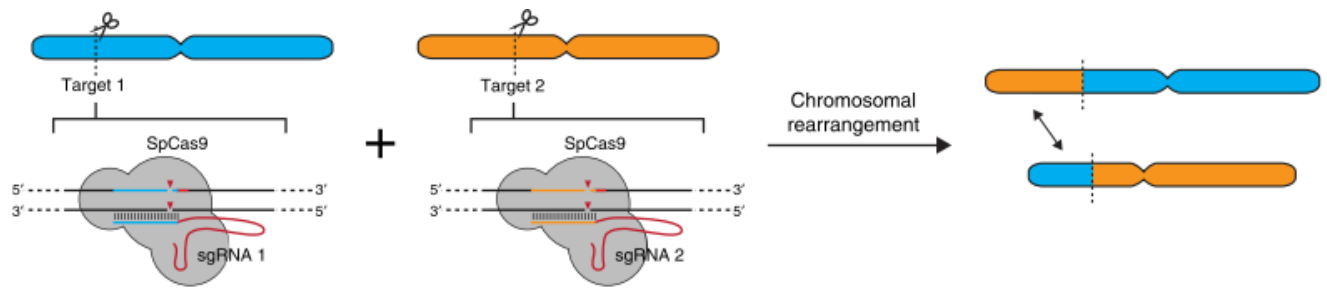


Figura 10. Esquema de la estrategia para la generación de ordenaciones cromosómicas mediante CRISPR/Cas9 (Choi and Meyerson, 2014).

5.3.6 Control de la replicación

Otra aplicación del sistema CRISPR/Cas9 conlleva la unión de dCas9 a cualquier posición del origen de la replicación, inhibiendo la replicación celular ya que evita que se unan otras proteínas necesarias para que se lleve a cabo el proceso. (Wiktor *et al.*, 2016).

Además, para minimizar los efectos no específicos que puedan tener lugar a causa de la alta expresión de dCas9, se le puede unir a ésta una secuencia señal que promueve su degradación, previniendo así la acumulación de cuerpos de inclusión y evitando el estrés celular que podría causar. Así, dCas9 seguirá activa, pero con una actividad más baja.

La proteína dCas9 con la secuencia señal unida (dCas9deg), se puede combinar con un sgRNA en el mismo plásmido para mejorar el rendimiento.

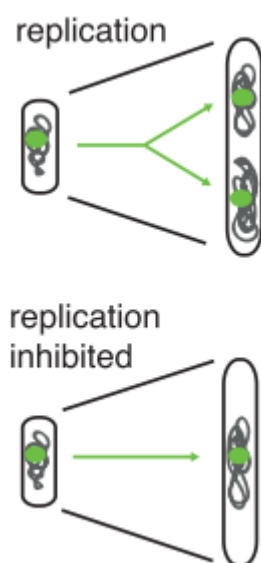


Figura 11. (Wiktor *et al.*, 2016)

Para comprobar cómo funciona el sistema con dCas9deg, se realizó un experimento en el que se pretendía inhibir la replicación de *E. coli* (Wiktor *et al.*, 2016). dCas9deg inhibe la replicación, pero no el crecimiento, por lo que las bacterias aumentarían de tamaño manteniendo un único cromosoma (**Figura 11**).

Los resultados del experimento fueron muy evidentes: cuando la actividad dCas9deg era inducida por anhidrotetraciclina (aTc), había crecimiento celular, pero sin

proceso de replicación. Por otro lado, en ausencia de aTc, dCas9deg no se encontraba activa y la replicación sí tenía lugar, observándose dos copias del cromosoma (**Figura 12**) (Wiktor *et al.*, 2016).

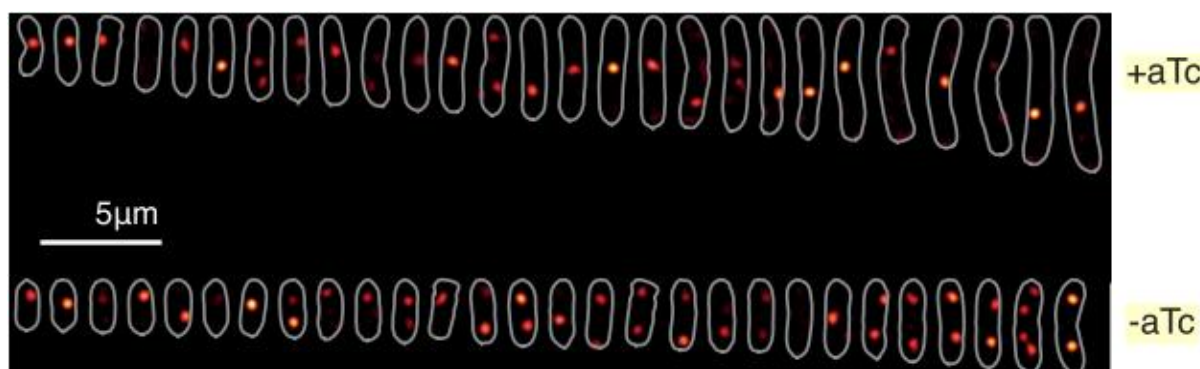


Figura 12. Representación del crecimiento celular con la actividad dCas9deg inducida por aTc (arriba) y con dCas9deg inactiva (abajo) (Wiktor *et al.*, 2016).

Además, mientras se hacían estos experimentos se descubrió que la inhibición mediada por CRISPR/dCas9 (y dCas9deg) no era funcional a 42°C. Las células cultivadas en un medio de agarosa con aTc eran capaces de formar colonias a 42°C, mientras que las células cultivadas a 37°C o 30°C no proliferaban en presencia de aTc. Esto indica que dCas9 y dCas9deg son termosensibles, propiedad que puede utilizarse para revertir la inhibición del proceso de replicación e incluso controlar el sistema CRISPR/dCas9 en otras aplicaciones (Wiktor *et al.*, 2016).

5.3.7 Integración de rutas de señalización

Aunque el sistema CRISPR/Cas9 demuestra ser muy útil en ingeniería genómica, su uso en redes de genes endógenos está limitado por la incapacidad de Cas9 para interactuar con señales biológicas. Para utilizar el sistema en el control y reasignación de señales, es posible construir plataformas que combinen riboswitches con respuesta a ligandos y sgRNAs para crear un nuevo tipo de dispositivos llamados conductores de señales (Y. Liu *et al.*, 2016).

Para llevar esto a cabo, hay que rediseñar el sgRNA integrando un aptámero de respuesta a señales, así se enlazan concentraciones de señales a eventos de expresión génica. CRISPR/Cas9 puede inactivar o activar la transcripción de cualquier gen de interés en respuesta a cualquier señal, y puede funcionar como conductor de señales (**Figura 13**). El sgRNA rediseñado usa el módulo tradicional para reclutar el represor transcripcional (dCas9) o activador (VP64) y el módulo del riboswitch para reconocer una señal específica. En ausencia de una señal específica, la región guía

del sgRNA original se aparea con la hebra antisentido diseñada y no puede unirse a la secuencia de ADN diana. Una señal específica podría unirse al aptámero e inducir un cambio conformacional que permite a la región guía del sgRNA interactuar con la secuencia de ADN diana para afectar a la transcripción de otra señal.

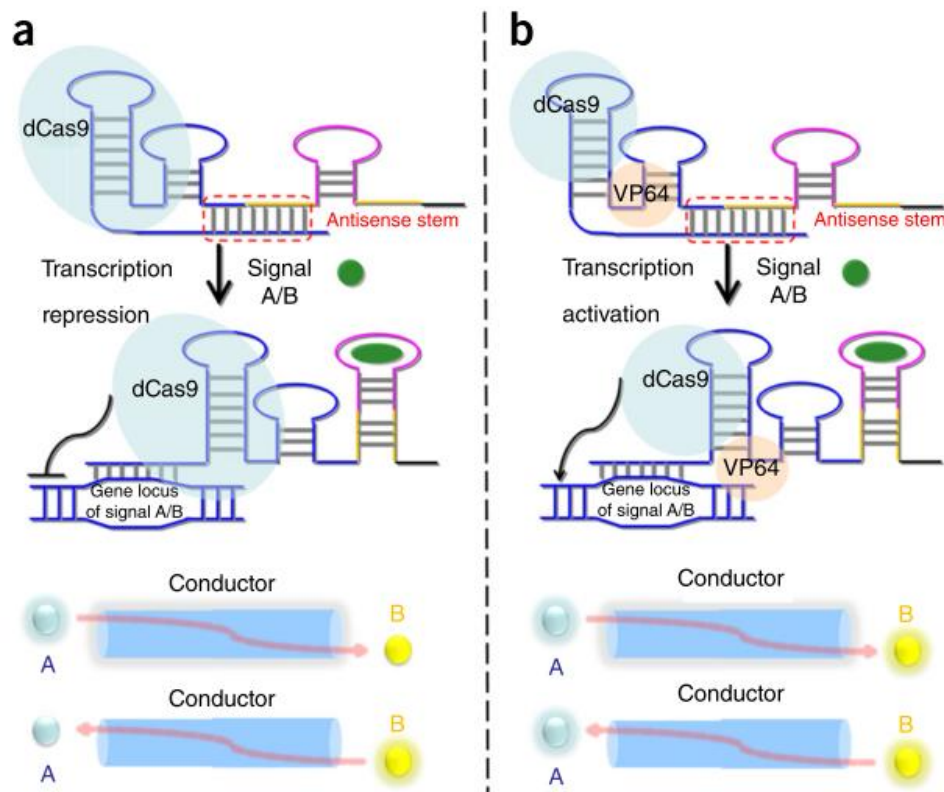


Figura 13. Diseño del conductor de señales basado en CRISPR/Cas9 que une una señal a otra. (a) Inactivación. (b) Activación. Secuencia antisentido del sgRNA en azul, aptámero en amarillo. En ausencia de señal, la región guía del sgRNA se aparea con la hebra antisentido y el sgRNA se encuentra inactivo. En presencia de señal, el sgRNA rediseñado cambia su conformación al estado activado. Entonces, la región guía del sgRNA se une a la secuencia diana del ADN y activa/inactiva la producción de la señal a través de dCas9/VP64 y la proteína dCas9 (Y. Liu et al., 2016).

5.3.8 Edición Múltiple

Los sistemas CRISPR pueden adaptarse fácilmente para editar varios genes, ya que se basan en el apareamiento entre el sgRNA y la secuencia complementaria en el ADN. Por lo tanto, si se introducen simultáneamente varios sgRNA, se podrán modificar varios genes al mismo tiempo (Cao, Xiao and Yan, 2018).

Cuando se quiere editar un único gen, se utilizan las estrategias vistas en el apartado 5.2 de este trabajo (RNPs, ARNm de Cas y sgRNA, o plásmidos). Para seleccionar varios genes al mismo tiempo bastaría con añadir sgRNAs adicionales. Sin embargo,

el modo de introducir los componentes necesarios y los múltiples sgRNAs en las limita la aplicabilidad de la edición múltiple. Teniendo esto en cuenta, se han desarrollado otras estrategias que generan sgRNAs múltiples a partir de un único plásmido y conseguir una edición multiplex:

- El método más utilizado para edición múltiple es la producción de sgRNAs a partir de cassettes de expresión separados. Cada cassette tiene su propio promotor, su secuencia guía, un sgRNA, y un terminador de la transcripción (**Figura 14**) (Cao, Xiao and Yan, 2018).

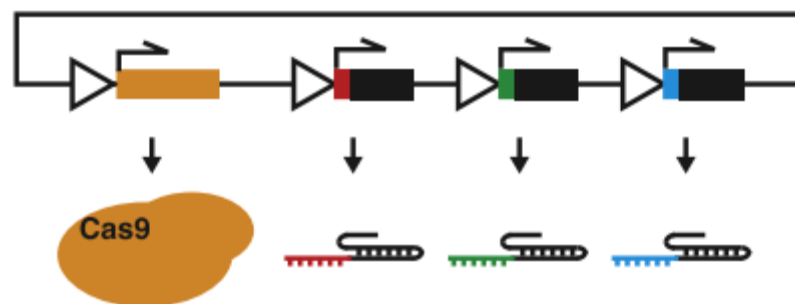


Figura 14. Cassettes sgRNA (Cao, Xiao and Yan, 2018)

- Otra estrategia para expresar múltiples sgRNAs es construir un precursor artificial único. Usar un único transcrito largo en vez de múltiples transcritos ofrece beneficios como evitar marcos de lectura abiertos en un plásmido que puedan afectar a la eficiencia de la transcripción; o la posibilidad de usar promotores de la polimerasa II, que permiten especificidad tisular o expresión inducible.

Una aplicación útil de la edición múltiple es la mejora de la tolerancia celular a estrés metabólico. Cuando se acumulan los metabolitos que necesita la célula, se acumulan subproductos, se produce una inhibición de metabolitos y disminuye el crecimiento celular. Existen reguladores que confieren tolerancia celular (Cheville *et al.*, 1996; Alper *et al.*, 2006; Abdullah-Al-Mahin *et al.*, 2010) y la manipulación genética de estos reguladores mejoran la tolerancia significativamente (Abdullah-Al-Mahin *et al.*, 2010; Chen, Stabryla and Wei, 2016).

La tolerancia celular se atribuye al loci QTL (rasgos cuantitativos), por lo que la edición multiplex CRISPR o la regulación génica de los QTL pueden mejorar la tolerancia

celular. En un estudio, se reemplazaron varios loci QTL de base múltiple utilizando un método basado en Cas9 y se diseñó una cepa celular con mayor tolerancia a la hidrolasa que la cepa de referencia (Maurer *et al.*, 2017).

Combinando estos sistemas con la diversidad de plataformas CRISPR, se conseguirá un control preciso del genoma, epigenoma, y transcriptoma, teniendo por lo tanto una gran implicación en investigación en biomedicina y en aplicaciones clínicas.

5.4 Inconvenientes

5.4.1 Efectos Off-Target

Hasta ahora, los efectos '*off-target*' (en sitios diferentes al sitio diana) parecen ser una de las mayores limitaciones de las aplicaciones de CRISPR (Lin *et al.*, 2014; Tian *et al.*, 2017). El sistema CRISPR/Cas9 puede desencadenar mutaciones y reordenaciones cromosómicas en loci no deseados que pueden dar lugar a alteraciones fenotípicas. Para su uso en ingeniería genómica, estos efectos '*off-target*' suponen un gran problema. Dado que están atribuidos principalmente al sgRNA, las secuencias diana deben ser diferentes de cualquier otra secuencia en, al menos, 3 nucleótidos de una secuencia de 20. En principio, la secuencia del sgRNA debe coincidir con el sitio de inicio de la transcripción, especialmente en la cadena antisentido.

Se está trabajando en estrategias para reducir los efectos *off-target*. Una estrategia es reducir la expresión de sgRNA y/o Cas9 en células humanas. Una de las estrategias se basa en la reducción de la expresión del sgRNA y/o la proteína Cas9 en células humanas. De esta manera se consigue un descenso alto en la proporción de efectos *off-target* con respecto a los efectos *on-target*.

El uso de sgRNAs con extremos 3' truncados o con dos nucleótidos de guanina extra en el extremo 5' también mejora la relación de efectos *on-target/off-target*, aunque también disminuye mucho la eficiencia de edición *on-target* (Cradick *et al.*, 2013; Hsu *et al.*, 2013).

Otro método que mejora la especificidad implica usar un par de "*nickasas*" con las que se generan mellas adyacentes a la secuencia diana. Para ello se usan dos sgRNAs y una nickasa Cas9 (nCas9) (**Figura 15**). Las nickasas Cas9 se unen a cadenas de ADN opuestas separadas por cuatro a cien pares de bases. De este modo se pueden inducir mutaciones del tipo inserción-delección, o eventos HDR utilizando un

oligonucleótido de ADN como molde en células humanas y de ratones. Se ha propuesto que los pares de nickasas crean una rotura en la doble hebra de ADN que se repara mediante NHEJ o HDR. La adición de un sgRNA secundario y la sustitución de la nickasa Cas9 por la nucleasa Cas9 podría dar lugar a un menor número de efectos *off-target* en sitios en los que previamente habían ocurrido efectos *off-target* con el sgRNA original. Sin embargo, una cuestión aún sin resolver es si el segundo sgRNA puede inducir por sí mismo mutaciones *off-target* mediadas por la nickasa Cas9.

Muchos estudios han mostrado que una única nickasa Cas9 puede funcionar por sí sola para inducir mutaciones indel en un cierto locus genómico, porque una mella puede convertirse en una rotura de doble hebra cuando la burbuja de replicación pasa por el locus.

Una forma muy importante de mejorar el sistema de pares de nickasas implicaría su modificación de manera que las actividades de las dos sgRNAs sean estrictamente interdependientes, es decir, que cada sgRNA solo esté activo para edición genómica cuando están ambos unidos al ADN muy cerca (Sander and Joung, 2014).

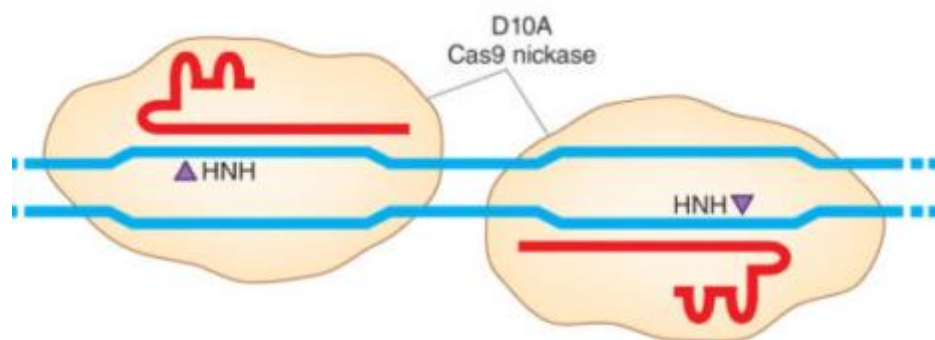


Figura 15. Esquematización del uso de nickasas pares para la reducción de los efectos *off-target*.
(Sander and Joung, 2014)

También se ha demostrado recientemente que los efectos *off-target* pueden reducirse usando sgRNAs que han sido acortados en el extremo 5' de su región complementaria. Estos sgRNAs truncados, (tru-sgRNAs), que tienen 17 o 18 nucleótidos de complementariedad con el ADN diana, generalmente funcionan eficientemente como un sgRNA de longitud máxima dirigiendo a la actividad Cas9, pero muestran una disminución de los efectos mutagénicos en sitios *off-target* mejorando la sensibilidad a apareamientos erróneos únicos o dobles en la unión

sgRNA-ADN. Los tru-sgRNAs son fáciles de implementar y esta estrategia evita las dificultades técnicas asociadas con la expresión de múltiples sgRNAs en una única célula por el método que utiliza pares de nickasas. Además, los tru-sgRNAs pueden ser usados junto a otras estrategias para mejorar la especificidad de Cas9 (se ha demostrado que los tru-sgRNAs mejoran la especificidad en pares de nickasas) (Sander and Joung, 2014).

5.4.2 Preferencia de NHEJ sobre HDR

De las dos vías de reparación de la rotura en la doble hebra de ADN, la HDR es la que emplea secuencias de ADN exógenas para producir inserciones, deleciones o sustituciones precisas, mientras que la NHEJ produce mutaciones variables no predecibles. En ingeniería genómica interesa, por tanto, que sea HDR la vía que prevalezca en la reparación. En la práctica, esto no es lo que ocurre, ya que la vía de reparación mediante NHEJ prevalece respecto a HDR, incluso en presencia de ADN molde (Freedman *et al.*, 2013, 2015; Lin *et al.*, 2014; Ma *et al.*, 2017).

Lo que determina que se emplee la vía NHEJ o HDR para la reparación de la rotura de doble hebra, es la fase del ciclo celular en la que se encuentre la célula. La NHEJ prevalece durante G1, S y G2, mientras que HDR no se encuentra activa en fase M y G1 temprana (para prevenir la fusión de telómeros (Orthwein *et al.*, 2014)), encontrándose activa solo en S y G2, cuando la replicación del ADN está completa y las cromátidas hermanas están disponibles como moldes para la reparación (Edalat, 2012), ya que son los moldes de preferencia para la reparación mediante HDR en levaduras y mamíferos (Kadyk and Hartwell, 1992; Dronkert *et al.*, 2000). Esta preferencia probablemente refleja un efecto próximo mediado por la cercana asociación de las cromátidas hermanas desde que se forman en la fase S del ciclo celular hasta que se segregan en anafase. La cohesión de cromátidas hermanas está mediada por cohesinas, e investigaciones recientes indican que las cohesinas migran a los sitios de reparación de roturas de doble hebra en el ADN, independientemente de la fase del ciclo celular en la que se encuentre la célula (Kadyk and Hartwell, 1992). Sin embargo, aunque HDR permanece activa solo en S y G2, NHEJ también permanece activa en estas fases del ciclo celular (Takata *et al.*, 1998), por lo que compite con HDR en la reparación, aunque haya cromátidas hermanas (es decir, moldes homólogos) próximas. Por tanto, que las nucleasas programables se

encuentren activas no implica que la edición genómica inducida por HDR se desarrolle de manera eficiente.

Se han estudiado varias formas de intentar combatir esta prevalencia de NHEJ sobre HDR. Una de las posibilidades es la interrupción química o genética de la vía NHEJ para favorecer HDR. El problema es que, aunque la interrupción química o genética de la vía NHEJ puede favorecer HDR (Shrivastav, De Haro and Nickoloff, 2008), tales manipulaciones pueden ser difíciles de llevar a cabo y dañinas para la célula.

Otra estrategia combina la sincronización del ciclo celular con nucleofección de la ribonucleoproteína Cas9 pre-ensamblada para conseguir un control de la acción de la nucleasa en la mejor fase del ciclo celular para HDR (**Figura 16**) (Lin *et al.*, 2014).

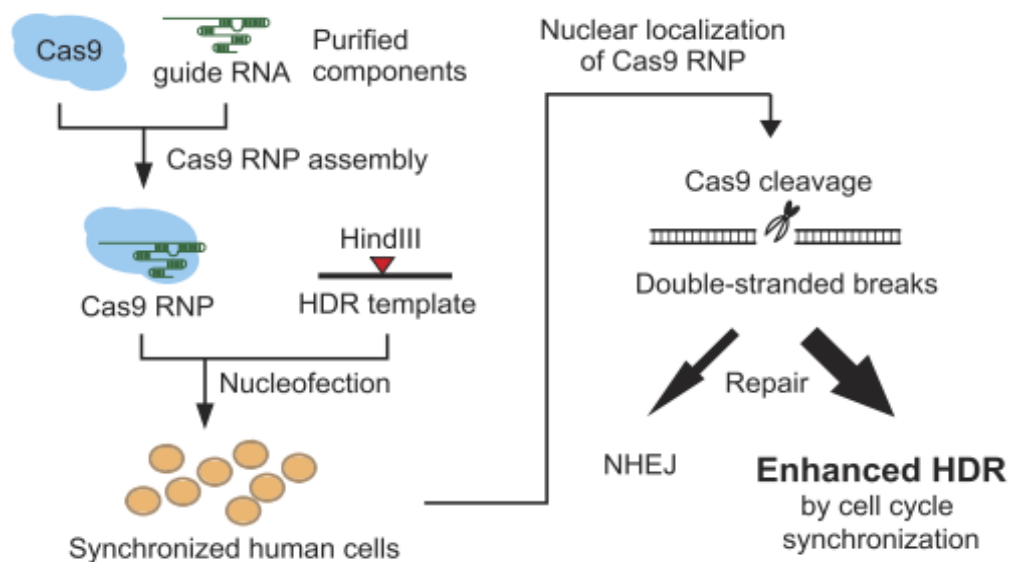


Figura 16. Aumento de los eventos HDR frente a NHEJ mediante sincronización del ciclo celular (Lin *et al.*, 2014)

Esta técnica incluye otras ventajas, como la presencia de actividad 'off-target' no detectable, la introducción de Cas9 RNP preensamblada en las células y la transfección simultánea de múltiples Cas9 RNPs y ADNs donantes. Con este sistema, se consigue maximizar la eficiencia de HDR hasta un 33%. Para mejorar la eficiencia más allá, se requeriría la manipulación de las proteínas involucradas en NHEJ y HDR (Humbert, Davis and Maizels, 2013).

5.4.3 Mosaicismo

La mayoría de modelos animales y embriones humanos modificados muestran mosaicismo, y eso incluye también a los manipulados con el sistema CRISPR/Cas9.

Muchos estudios previos han intentado producir animales genéticamente modificados sin mosaicismo en un único paso, incluyendo la inyección de proteína Cas9, o la inyección de un cóctel de sgRNAs. Sin embargo, estas aproximaciones dieron lugar a *Knock-out* en lugar de *Knock-in*. Recientemente, se ha conseguido corregir mutaciones heterocigóticas sin mosaicismo en embriones humanos usando CRISPR/Cas9, pero esto depende de la copia salvaje, por lo que no sirve para la corrección de mutaciones recesivas o la integración transgénica.

6. IMPLICACIONES ÉTICAS

La edición genética siempre ha sido un tema controvertido, ya que no hay un acuerdo claro sobre donde establecer los límites para su uso. El desarrollo de CRISPR/Cas9 ha causado tanto emoción como alarma en la comunidad científica internacional debido a su capacidad para modificar cualquier tipo de genoma con tal facilidad.

La aplicación más importante de CRISPR puede que sea su uso en terapia génica, prometiendo curar enfermedades genéticas de muy diversa índole (Brokowski and Adli, 2018). Existe un gran apoyo por parte de la comunidad científica a favor de su uso para ese fin. Sin embargo, no deja de haber preocupación por su utilización en ciertas áreas, como la modificación de la línea germinal.

En 2015, un grupo de investigadores y especialistas en ética se reunieron para discutir sobre los aspectos éticos, legales y biomédicos de los sistemas CRISPR (Baltimore *et al.*, 2015). El objetivo era estudiar cuándo, dónde y cómo esta tecnología podía ser utilizada en humanos. La discusión continuó en febrero de 2017 cuando un comité multidisciplinar de las NASEM (*National Academies of Science, Engineering and Medicine*) publicó un informe examinando numerosos aspectos de la edición genómica (Baltimore *et al.*, 2015), en el que, en resumen, permite la edición de líneas somáticas, pero no la modificación genética para cualquier tipo de mejora. Además, respecto a la modificación de la línea germinal, se estableció que solo podrá permitirse cuando se conozcan mejor los riesgos técnicos y de seguridad asociados.

Una de las principales preocupaciones son las limitaciones del sistema CRISPR/Cas mencionadas en este trabajo, como los efectos *off-target*. Aunque la tecnología está avanzando a pasos agigantados y es posible que en un futuro más o menos próximo no supongan ningún problema (Brokowski and Adli, 2018).

Otra aplicación de CRISPR con potencial para beneficiar y salvar muchas vidas es la “genética dirigida”, que consiste en utilizar CRISPR para acelerar la recombinación genética de tal manera que un gen de interés se distribuye rápidamente a toda la población más rápido que por herencia mendeliana. Por lo tanto, esta aplicación tiene el potencial de editar el genoma de poblaciones enteras o incluso especies. Usando genes dirigidos mediante CRISPR, los investigadores han demostrado que un alelo de un gen que aporta un fenotipo de defensa frente a parásitos en mosquitos podría transmitirse rápidamente en la población en una tendencia no mendeliana (Gantz *et al.*, 2015; Hammond *et al.*, 2016). De esta manera, se desconoce si los organismos genéticamente modificados se verán afectados indefinidamente y si los genes editados se transmitirán a generaciones futuras, afectándolas de manera inesperada. Por estos motivos, es complicado tomar una decisión moral acertada.

En definitiva, en la actualidad sigue sin haber un acuerdo claro sobre donde establecer los límites de la edición genética, por lo que habrá que esperar un tiempo a que se conozcan mejor los riesgos a largo plazo.

7. DISCUSIÓN

El sistema CRISPR/Cas, con origen en bacterias no solo supone un mecanismo de defensa de gran utilidad para los organismos procariotas, sino que tras muchos años de investigación se ha conseguido descubrir su importante aplicación a la ingeniería genética, siendo de gran importancia en muchas áreas en la actualidad, como la investigación o aplicaciones en terapia génica.

El gran impacto que supone este sistema en ciencia se debe principalmente a su simplicidad. Los sistemas basados en ZFN y TALEN también supusieron un gran avance, pero CRISPR/Cas (en concreto CRISPR/Cas9) es mucho más eficaz, de muy fácil diseño y una herramienta mucho más barata. Con CRISPR/Cas9, basta con modificar el sgRNA para especificar la secuencia diana, mientras que el empleo de ZFN y TALEN requiere el diseño de nuevos genes para cada región del genoma que quiera modificarse, complicando la técnica y volviéndola mucho más cara. Además, como se ha expuesto en la revisión, la proteína Cas9 también es fácil de modificar, dando lugar a nuevas aplicaciones gracias a la creación de proteínas alternativas como dCas9 y nCas9.

Gran parte del avance en el conocimiento de la técnica en los últimos años se debe principalmente a que la mayoría de los laboratorios enfocados a manipulación genética se centran en su estudio, ya que es la técnica de edición genética más barata y más sencilla conocida. Este es el motivo por el que la técnica avanza a pasos agigantados, suponiendo una gran promesa para la ingeniería genética. Por lo tanto, aunque aún presenta algunos problemas, poco a poco se están descubriendo estrategias para superarlos, existiendo grandes expectativas en la comunidad científica en torno al sistema CRISPR/Cas9.

8. CONCLUSIONES

Gracias al avance en la investigación en ciencia básica, se ha conseguido descubrir la estructura y función del sistema CRISPR/Cas, que ha permitido la creación de una herramienta muy poderosa para la ingeniería genética.

CRISPR/Cas ha demostrado ser un sistema muy simple y eficiente, por la facilidad de su manipulación y permitiendo su uso en distintas aplicaciones que pueden ser muy útiles tanto en investigación genética, para dilucidar cómo funcionan distintos genes y como se organizan, como en terapia génica, permitiendo el estudio de enfermedades y posibles tratamientos.

Aunque en la actualidad aún existe polémica en cuanto a su uso y posee algunos inconvenientes, en el futuro, a medida que avance la investigación, estos problemas podrían quedar atrás, haciendo de este sistema una herramienta muy precisa y segura para la manipulación genética.

9. REFERENCIAS

- Abdullah-Al-Mahin *et al.* (2010) 'Improvement of multiple-stress tolerance and lactic acid production in lactococcus lactis NZ9000 under conditions of thermal stress by heterologous expression of escherichia coli dnaK', *Applied and Environmental Microbiology*, 76(13), pp. 4277–85.
- Alper, H. *et al.* (2006) 'Engineering yeast transcription machinery for improved ethanol tolerance and production', *Science*, 314(5805), pp. 1565–68.
- Alt, F. W. *et al.* (2013) 'Mechanisms of programmed DNA lesions and genomic instability in the immune system', *Cell*, 152(3), pp. 417–29.
- Babu, M. *et al.* (2012) "A dual function of the CRISPR-Cas system in bacterial

- antivirus immunity and DNA repair', *Molecular Biology*, 79(2), pp. 484–502.
- Baltimore, D. *et al.* (2015) 'A prudent path forward for genomic engineering and germline gene modification', *Science*, 348(6230), pp. 36–38.
- Beerli, R. R., Dreier, B. and Barbas 3rd, C. F. (2000) 'Positive and negative regulation of endogenous genes by designed transcription factors', *Proceedings of the National Academy of Science of the USA*, 97(4), pp. 1495–500.
- Beerli, R. R. and Barbas, C. F. (2002) 'Engineering polydactyl zinc-finger transcription factors', *Nature Biotechnology*, 20(2), pp. 135–41.
- Blasco, R. B. *et al.* (2014) 'Simple and Rapid InVivo Generation of Chromosomal Rearrangements using CRISPR/Cas9 Technology', *Cell Reports*, 9(4), pp. 1219–27.
- Bondy-denomy, J. *et al.* (2016) 'Multiple mechanisms for CRISPR–Cas inhibition by anti–CRISPR proteins', *Nature*, 526(7571), pp. 136–39.
- Briggs, A. W. *et al.* (2012) 'Iterative capped assembly: Rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers', *Nucleic Acids Research*, 40(15), pp. 1–10.
- Briner, A. E. *et al.* (2014) 'Guide RNA functional modules direct Cas9 activity and orthogonality', *Molecular Cell*. Elsevier Inc., 56(2), pp. 333–39.
- Brokowski, C. and Adli, M. (2018) 'CRISPR ethics: Moral considerations for applications of a powerful tool', *Journal of Molecular Biology*. Elsevier Ltd.
- Bronkhorst, A. W. and Van Rij, R. P. (2014) 'The long and short of antiviral defense: Small RNA-based immunity in insects', *Current Opinion in Virology*. Elsevier B.V., 7(1), pp. 19–28.
- Burmistrz, M. and Pyrc, K. (2015) 'CRISPR-Cas systems in prokaryotes', *Polish Journal of Microbiology*, 64(3), pp. 193–202.
- Cao, J., Xiao, Q. and Yan, Q. (2018) 'The multiplexed CRISPR targeting platforms', *Drug Discovery Today: Technologies*. Elsevier Ltd, xxx(xx), pp. 1–9.
- Carroll, D. (2011) 'Genome engineering with zinc-finger nucleases', *Genetics*, 188(4), pp. 773–82.

- Cermak, T. *et al.* (2011) 'Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting (Nucleic Acids Research (2011) 39 (e82) DOI: 10.1093/nar/gkr218)', *Nucleic Acids Research*, 39(17), p. 7879.
- Chen, Y., Stabryla, L. and Wei, N. (2016) 'Improved acetic acid resistance in *Saccharomyces cerevisiae* by overexpression of the *WHI2* gene identified through inverse metabolic engineering', *Applied and Environmental Microbiology*, 82(7), pp. 3718-15.
- Cheville, A. M. *et al.* (1996) 'rpoS regulation of acid, heat, and salt tolerance in *Escherichia coli* O157:H7', *Applied and Environmental Microbiology*, 62(5), pp. 1822–24.
- Choi, P. S. and Meyerson, M. (2014) 'Targeted genomic rearrangements using CRISPR/Cas technology', *Nature Communications*. Nature Publishing Group, 5, pp. 1–6.
- Chylinski, K., Le Rhun, A. and Charpentier, E. (2013) 'The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems', *RNA Biology*, 10(5), pp. 726–37.
- Chylinski, K. *et al.* (2014) 'Classification and evolution of type II CRISPR-Cas systems', *Nucleic Acids Research*, 42(10), pp. 6091–105.
- Cradick, T. J. *et al.* (2013) 'CRISPR/Cas9 systems targeting β -globin and CCR5 genes have substantial off-target activity', *Nucleic Acids Research*, 41(20), pp. 9584–92.
- Cruz, N. M. and Freedman, B. S. (2018) 'CRISPR Gene Editing in the Kidney', *American Journal of Kidney Diseases*. Elsevier Inc, 20, pp. 1–10.
- De Silva Felixge, H. S. and Jerome, K. R. (2017) 'Excision of Latent HIV-1 from Infected Cells In Vivo: An Important Step Forward', *Molecular Therapy*. The American Society of Gene and Cell Therapy, 25(5), pp. 1062–64.
- Dronkert, M. L. G. *et al.* (2000) 'Mouse RAD54 Affects DNA Double-Strand Break Repair and Sister Chromatid Exchange', *Molecular and Cellular Biology*, 20(9), pp. 3147–56.
- Duan, J. *et al.* (2014) 'Genome-wide identification of CRISPR/Cas9 off-targets in

- human genome', *Cell Research*, 24(8), pp. 1009–12.
- Edalat, F. (2012) "Regulation of homologous recombination in eukaryotes", *Annual Review of Genetics*, 40(6), pp. 1301–15.
- Freedman, B. S. *et al.* (2013) 'Reduced Ciliary Polycystin-2 in Induced Pluripotent Stem Cells from Polycystic Kidney Disease Patients with PKD1 Mutations', *Journal of the American Society of Nephrology*, 24(10), pp. 1571–86.
- Freedman, B. S. *et al.* (2015) 'Modelling kidney disease with CRISPR-mutant kidney organoids derived from human pluripotent epiblast spheroids', *Nature Communications*. Nature Publishing Group, pp. 1–13.
- Gaj, T., Gersbach, C. A. and Barbas, C. F. (2013) 'ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering', *Trends in Biotechnology*. Elsevier Ltd, 31(7), pp. 397–405.
- Gantz, V. M. *et al.* (2015) 'Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*', *Proceedings of the National Academy of Sciences*, 112(49), pp. E6736–43.
- Gasiunas, G. *et al.* (2012) 'Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria', *Proceedings of the National Academy of Sciences*, 109(39), pp. 2579–86.
- Gostissa, M. *et al.* (2014) 'IgH class switching exploits a general property of two DNA breaks to be joined in cis over long chromosomal distances', *Proceedings of the National Academy of Sciences*, 111(7), pp. 2644–49.
- Hammond, A. *et al.* (2016) 'A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*', *Nature Biotechnology*. Nature Publishing Group, 34(1), pp. 78–83.
- Harrington, L. B. *et al.* (2017) 'A Broad-Spectrum Inhibitor of CRISPR-Cas9', *Cell*, 170(6), p. 1224–1233.e15.
- Heler, R. *et al.* (2015) 'Cas9 specifies functional viral targets during CRISPR-Cas adaptation', *Nature*, 519(7542), pp. 199–202.

- Hsu, P. D. *et al.* (2013) 'DNA targeting specificity of RNA-guided Cas9 nucleases', *Nature Biotechnology*, 31(9), pp. 827–32.
- Hsu, P. D., Lander, E. S. and Zhang, F. (2014) 'Development and applications of CRISPR-Cas9 for genome engineering', *Cell*. Elsevier, 157(6), pp. 1262–78.
- Humbert, O., Davis, L. and Maizels, N. (2013) 'Targeted Gene Therapies: Tools, Applications, Optimization', *Critical Reviews in Biochemistry and Molecular Biology*, 47(3), pp. 264–81.
- Ishino, Y. *et al.* (1987) 'Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product.', *Journal of Bacteriology*, 169(12), pp. 5429–33.
- Ivanov, Y. V. *et al.* (2015) 'A newly discovered *Bordetella* species carries a transcriptionally active CRISPR-Cas with a small Cas9 endonuclease', *BMC Genomics*. BMC Genomics, 16(1), pp. 1–10.
- Jansen, R. *et al.* (2002) 'Identification of genes that are associated with DNA repeats in prokaryotes', *Molecular Microbiology*, 43(6), pp. 1565–75.
- Jinek, M. *et al.* (2012) 'A Programmable Dual-RNA – Guided', *Science*, 337(August), pp. 816–22.
- Jore, M. M., Brouns, S. J. J. and van der Oost, J. (2012) 'RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements', *Cold Spring Harbor Perspectives in Biology*, 4(6), pp. 1–12.
- Kadyk, L. C. and Hartwell, L. H. (1992) 'Sister chromatids are preferred over homologs as substrates for recombinational repair in *Saccharomyces cerevisiae*', *Genetics*, 132(2), pp. 387–402.
- Kaphingst, K. A., Persky, S. and Lachance, C. (2010) 'Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors', 14(4), pp. 384–99.
- Kennedy, E. M. and Cullen, B. R. (2016) "Bacterial CRISPR/Cas DNA endonucleases: A revolutionary technology that could dramatically impact viral research and treatment', *Virology*, pp. 213–20.
- Kim, J.-S. and Pabo, C. O. (1998) 'Getting a handhold on DNA: Design of poly-zinc

- finger proteins with femtomolar dissociation constants', *Proceedings of the National Academy of Sciences*, 95(6), pp. 2812–17.
- Klein, I. A. *et al.* (2012) 'Translocation-Capture Sequencing Reveals the Extent and Nature of Chromosomal Rearrangements in B Lymphocytes', *Cell*, 147(1), pp. 95–106.
- Kuscu, C. *et al.* (2014) 'Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease', *Nature Biotechnology*. Nature Publishing Group, 32(7), pp. 677–83.
- Lin, S. *et al.* (2014) 'Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery', *eLife*, 3, p. e04766.
- Liu, C. *et al.* (2017) 'Delivery strategies of the CRISPR-Cas9 gene-editing system for therapeutic applications', *Journal of Controlled Release*. Elsevier B.V, 266, pp. 17–26.
- Liu, J. *et al.* (2015) 'Improved cell-penetrating zinc-finger nuclease proteins for precision genome engineering', *Molecular Therapy - Nucleic Acids*. IOP Publishing, 4(3).
- Liu, K. I. *et al.* (2016) 'A chemical-inducible CRISPR-Cas9 system for rapid control of genome editing', *Nature Chemical Biology*. Nature Publishing Group, 12(11), pp. 980–87.
- Liu, Q. *et al.* (1997) 'Design of polydactyl zinc-finger proteins for unique addressing within complex genomes', *Proceedings of the National Academy of Sciences*, 94(11), pp. 5525–30.
- Liu, Y. *et al.* (2016) 'Directing cellular information flow via CRISPR signal conductors', *Nature Methods*, 13(11), pp. 938–44.
- Luo, J. (2016) 'CRISPR/Cas9: From Genome Engineering to Cancer Drug Discovery', *Trends in Cancer*. Elsevier Inc., 2(6), pp. 313–24.
- Ma, H. *et al.* (2017) 'Correction of a pathogenic gene mutation in human embryos.', *Nature*. Nature Publishing Group.
- Mak, A. N. *et al.* (2012) 'The crystal structure of TAL effector PthXo1 bound to its DNA

- target', *Science*, 335(6069), pp. 716–19.
- Makarova, K. S. *et al.* (2006) 'A putative RNA-interference-based immune system in prokaryotes: Computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action', *Biology Direct*, 1, pp. 1–26.
- Makarova, K. S. *et al.* (2015) 'An updated evolutionary classification of CRISPR-Cas systems', *Nature Reviews Microbiology*, 13(11), pp. 722–36.
- Makarova, K. S. *et al.* (2017) 'HHS Public Access', 13(11), pp. 722–36.
- Mali, P. *et al.* (2013) 'RNA-Guide Human Genome Engineering via Cas9', *Science*, 339(2013), pp. 823–26.
- Maurer, M. J. *et al.* (2015) 'An updated evolutionary classification of CRISPR-Cas systems', *Nature Reviews Microbiology*, 13(11), pp. 722–36.
- McMahon, M. A. *et al.* (2018) 'Chemically Modified Cpf1-CRISPR RNAs Mediate Efficient Genome Editing in Mammalian Cells', *Molecular Therapy*. Elsevier Inc.
- Miyagi, A., Lu, A. and Humphreys, B. D. (2016) 'Gene Editing: Powerful New Tools for Nephrology Research and Therapy', *Journal of the American Society of Nephrology*, 27(10), pp. 2940–47.
- Niu, Y. *et al.* (2014) 'Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos', *Cell*, 156(4), pp. 836–43.
- Olson, E. N. (2015) 'Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA', *Science*, 345(6201), pp. 1184–88.
- Orthwein, A. *et al.* (2014) 'Mitosis inhibits DNA double-strand break repair to guard against telomere fusions', *Science*, 344(6180), pp. 189–93.
- Patterson, A. G. *et al.* (2016) 'Quorum Sensing Controls Adaptive Immunity through the Regulation of Multiple CRISPR-Cas Systems', *Molecular Cell*. Elsevier Inc., 64(6), pp. 1102–08.
- Pawluk, A. *et al.* (2016) 'Naturally occurring off-switches for CRISPR-Cas9 HHS Public Access', *Cell*, 16711(7), pp. 1829–38.

- Rabbitts, T. H. *et al.* (2001) 'Mouse models of human chromosomal translocations and approaches to cancer therapy', *Blood Cells, Molecules, and Diseases*, 27(1), pp. 249–59.
- Ran, F. A. *et al.* (2013) 'Genome engineering using the CRISPR-Cas9 system', *Nature protocols*, 8(11), pp. 2281–308.
- Reyon, D. *et al.* (2013) 'FLASH Assembly of TALENs Enables High-Throughput Genome Editing', *Nature Biotechnology*, 30(5), pp. 460–65.
- Samson, J. E. *et al.* (2013) 'Revenge of the phages: Defeating bacterial defences', *Nature Reviews Microbiology*. Nature Publishing Group, 11(10), pp. 675–87.
- Sander, J. D. and Joung, J. K. (2014) 'CRISPR-Cas systems for genome editing, regulation and targeting', *Nature Biotechnology*, 32(4), pp. 347–55.
- Schmid-burgk, J. L. *et al.* (2014) 'A ligation-independent cloning technique for high-throughput assembly of transcription activator – like effector genes', *Nature Biotechnology*, 31(1), pp. 76–81.
- Shrivastav, M., De Haro, L. P. and Nickoloff, J. A. (2008) 'Regulation of DNA double-strand break repair pathway choice', *Cell Research*, 18(1), pp. 134–47.
- Takata, M. *et al.* (1998) 'Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells', *EMBO Journal*, 17(18), pp. 5497–508.
- Tian, P. *et al.* (2017) 'Fundamental CRISPR-Cas9 tools and current applications in microbial systems', *Synthetic and Systems Biotechnology*. Elsevier Ltd, pp. 6–12.
- van der Oost, J. and Brouns, S. J. J. (2015) 'CRISPR sabotage', *Genome Biology*. Genome Biology, 16(1), pp. 15–17.
- Wang, R. *et al.* (2011) 'Interaction of the Cas6 ribonuclease with CRISPR RNAs: Recognition and cleavage', *Structure*. Elsevier Ltd, 19(2), pp. 257–64.
- Wei, Y., Terns, R. M. and Terns, M. P. (2015) 'Cas9 function and host genome sampling in type II-A CRISPR–cas adaptation', *Genes and Development*, 29(4),

pp. 356–61.

Wiedenheft, B. (2013) 'In defense of phage: Viral suppressors of CRISPR-mediated adaptive immunity in bacteria', *RNA Biology*, 10(5), pp. 886–90.

Wiktor, J. *et al.* (2016) 'CRISPR-mediated control of the bacterial initiation of replication', *Nucleic Acids Research*, 44(8), pp. 3801–10.

Wu, X. *et al.* (2015) 'Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells', *Nature Biotechnology*, 32(7), pp. 670–76.

Xue, W. *et al.* (2015) 'CRISPR-mediated direct mutation of cancer genes in the mouse liver', *Nature*, 514(7522), pp. 380–84.