

IDENTIFICACION

Titulo

Español: Sistema CRISPR/Cas: De la investigación básica al premio Nobel 2020

Ingles: CRISPR/Cas system: From basic research to Nobel Prize 2020

Nombre y apellido del autor o los autores

1. Steve Dany Jesus Palomino
2. Solange Rosa Paredes Moscosso
3. Claudio Nicolas Villegas Llerena

Código ORCID de cada autor

1. 0000-0002-9624-2973
2. 0000-0001-8461-2546
3. 0000-0002-2634-5504

Correo electrónico de cada autor

1. steve.jesus.p@upch.pe
2. sparedesm@usmp.pe
3. cvillegas1@usmp.pe

Afiliación institucional de los autores

1. Universidad de San Martín de Porres, Facultad de Medicina Humana, Instituto de Investigación, Centro de Investigación de Genética y Biología Molecular, Lima, Perú.
2. Universidad de San Martín de Porres, Facultad de Medicina Humana, Instituto de Investigación, Centro de Investigación de Genética y Biología Molecular, Lima, Perú.
3. Universidad de San Martín de Porres, Facultad de Medicina Humana, Instituto de Investigación, Centro de Investigación de Genética y Biología Molecular, Lima, Perú.

Dirección postal

Universidad de San Martin de Porres, Facultad de medicina Humana
Av. Alameda del Corregidor 1517-1531, Urb. Sirius – III Etapa – La Molina, Lima 15024
Lima , Perú

Resumen (español)

El sistema CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) ha revolucionado el campo de la investigación biológica. Inicialmente identificado como clústeres de repeticiones interespaciadas por elementos genéticos sin homología aparente

en el cromosoma de bacterias y arqueas, fue eventualmente identificado como un sofisticado sistema de defensa inmune en procariontes dirigido a evitar la infección por fagos y elementos genéticos extra cromosómicos. Su descubrimiento e investigación por parte de numerosos grupos de investigadores alrededor del mundo son un ejemplo claro de como la investigación básica puede dar paso al descubrimiento de herramientas moleculares con utilidad en todos los campos de las ciencias biológicas. Sus aplicaciones van más allá de la edición genética, incluyendo la regulación de la expresión génica, regulación epigenética, la toma de imágenes y manipulación de la topología de la cromatina, el diagnóstico de enfermedades infecciosas y en terapia génica. Todas estas aplicaciones sumadas a la versatilidad y facilidad de uso del sistema CRISPR, condujeron a dos de las investigadoras pioneras en el uso del sistema a la obtención del premio Nobel en química en el año 2022. Aquí, presentaremos una breve reseña de los principales descubrimientos y hallazgos que ayudaron a develar el mecanismo y función del sistema CRISPR. Así mismo, describiremos el desarrollo de una amplia gama de herramientas moleculares basadas en CRISPR. Concluiremos con direcciones futuras y el impacto más amplio de las tecnologías CRISPR.

Resumen (inglés)

The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) system has revolutionized the field of biological research. Initially identified as clusters of repeats interspaced by genetic elements with no apparent homology and located within the chromosomes of bacteria and archaea, it was eventually identified as an advanced immune defense system in prokaryotes aimed at avoiding infection by phages and extrachromosomal genetic elements. Its discovery and investigation by numerous groups of researchers around the world are a clear example of how basic research can lead to the discovery of molecular tools that are useful in all fields of biological sciences. Its applications go beyond gene editing, including regulation of gene expression, epigenetic regulation, imaging and manipulation of chromatin topology, diagnosis of infectious diseases, and gene therapy. All these applications, added to the versatility and ease of use of the CRISPR system, led two of the pioneering researchers in the use of this system to obtain the Nobel Prize in Chemistry in 2022. Here, we will present a brief review of the main discoveries that helped to unveil the mechanism and function of the CRISPR system. Likewise, we will describe the development of a wide range of molecular tools based on CRISPR. We will conclude with future directions and the broader impact of CRISPR technologies.

Palabras clave

Español: CRISPR, proteína cas, edición génica, herramienta molecular

Inglés: CRISPR, cas protein, gene editing, molecular tool

Información sobre la contribución

SDJP: Escritura - Preparación del borrador original,

SRPM: Redacción-revisión y edición

CNVLL: Conceptualización, Redacción-revisión y edición

Información sobre conflicto de intereses

Los autores no declaran ningún conflicto de interés.

Fuente de financiamiento

El estudio fue financiado por PROCENCIA (a través del contrato N°153-2018-FONDECYT-BM-IADT-MU) y la Universidad de San Martín de Porres.

Agradecimientos

Ninguno

Sobre ética en investigación y normas legales

Los autores declaran que no violaron ni omitieron normas éticas o legales en esta investigación.

EL CUERPO DEL TRABAJO (COMENTARIO)

Hasta antes del año 2005, nadie tenía conocimiento que bacterias y arqueas, poseían un sistema de defensa adaptativo o específico contra fagos y plásmidos extra-cromosómicos. Este sistema inmune de memoria en procariotas se conoce hoy con el nombre de sistema CRISPR/Cas por sus siglas en inglés; Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) - CRISPR associated proteins (Cas) que en su forma traducida sería Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas - Proteína Asociada a CRISPR, respectivamente.

En el año 2020, Emmanuelle Charpentier y Jennifer A. Doudna fueron premiadas con el premio Nobel de Química por el desarrollo de un nuevo método de alta precisión para la edición de genomas (2022), el mismo que ha tenido un alto impacto en las ciencias biomédicas. No cabe duda del gran aporte científico por parte de las laureadas con el Nobel 2020, al utilizar el conocimiento del sistema CRISPR de procariotas como una nueva herramienta muy precisa de corte a nivel de ADN, sin embargo, no se debe dejar de lado la importancia de muchos otros investigadores que han sido pieza clave para la comprensión del sistema CRISPR en procariotas e incluso de aquellos investigadores que lograron adaptar y aplicar esta nueva forma de edición génica en eucariotas.

En consecuencia, el propósito de esta revisión es hacer un recuento de los investigadores que impulsaron la “revolución CRISPR”, para así conocer “quién es quién” y sus respectivas contribuciones en el camino a este premio Nobel del 2020.

La primera evidencia del sistema CRISPR/Cas se dio en una publicación de Ishino et al. (1987). Estos investigadores tenían como objetivo identificar la proteasa putativa codificada por el gen *iap*, así como identificar la estructura primaria de la proteasa *iap*, la cual sería responsable de mediar la conversión de la isoenzima 1 de la fosfatasa alcalina de *Escherichia coli* en la isoenzima 2 y 3. Dentro de sus resultados hallaron una estructura inusual en la región flanqueante del extremo 3' del gen *iap* (ver Figura 1). Se halló cinco secuencias altamente homólogas de 29 nucleótidos, cada una en arreglos de tipo “repeticiones directas”, las cuales estaban interespaciadas por secuencias de 32 nucleótidos. Estas últimas poseían secuencias diferentes entre sí. Al final de la publicación, los investigadores mencionaban que no se encontraron secuencias homólogas a las secuencias interespaciadas de 29 nucleótidos en procariotas; y que su función biológica era desconocida.

TGAAAATGGGAC	GGAGTTCTACCGCAGAGGCGGGGAACTC	CAAGTGATATCCATCATCGCATCCAGTGCGCC	(1,451)
(1,452)	CGGTTTATCCCGCTGATGCGGGGAACAC	CAGCGTCAGGCGTAAATCTCACCGTCGTTGC	(1,512)
(1,513)	CGGTTTATCCCTGCTGGCGGGGAACTC	TCGGTTCAGGCGTTGCAAACCTGGCTACCGGG	(1,573)
(1,574)	CGGTTTATCCCGCTAACGCGGGGAACTC	GTAGTCCATCATTCACCTATGTCTGAACTCC	(1,634)
(1,635)	CGGTTTATCCCGCTGGCGGGGAACTCG		(1,664)

consensus: CGGTTTATCCCGCT^{GG}_{AA}CGCGGGGAACTC

Figura 1. Alineamiento de secuencias tipo repeticiones directas de la región flanqueante terminal 3' del gen *iap*. En el recuadro rojo se observa una de las secuencias homólogas de 29 nucleótidos y en color verde se encuentra una de las secuencias no homólogas de 32 nucleótidos. En el recuadro naranja, se observa subrayado, 14 pares de bases (bp) con simetría de diada. Adaptado de Ishino et al. (1987).

En 1989, el mismo grupo de investigación de la universidad de Osaka publicó un artículo donde se describe la presencia de un arreglo nucleotídico inusual con secuencias repetidas en el cromosoma de la *Escherichia coli* K12 (Nakata et al. 1989). En dicho artículo, se ubicaron 14 secuencias repetitivas altamente conservadas de 29bp cada una e interespaciadas por secuencias no homólogas de 32-33bp. Por otro lado, 24 kilopares de bases (Kbp) downstream de las 14 secuencias descritas, se halló otras 7 secuencias repetitivas altamente conservadas de 29bp, también interespaciadas por secuencias 32bp.

Este grupo también buscó determinar si las secuencias repetitivas altamente conservadas de 29bp estaban presentes en otras especies de *E. coli* así u otras bacterias gram negativas, todo esto a través de análisis por sondas nucleotídicas (Southern blot). Se determinó que tanto *E. coli* C600 como *E. coli* Ymel, *Salmonella typhimurium* TA1535 y *Shigella dysenteriae* poseen las secuencias repetitivas altamente conservadas de 29bp, y no así *Klebsiella pneumoniae* o *Pseudomonas aeruginosa*.

En un estudio similar, Hermans et al. (1991) reportaron clústeres de secuencias repetitivas altamente homólogas interespaciadas por secuencias diferentes entre sí por primera vez en bacterias Gram positivas (las reportadas por Nakata et al. (1989) fueron Gram negativas). En dicho estudio, se encontró que *Mycobacterium bovis* presentaba 49 secuencias homólogas entre sí (llamadas por aquel entonces repeticiones directas) de las cuales 20 eran idénticas en sus secuencias y el resto casi idénticas. Las secuencias homólogas se encuentran separadas por secuencias no homólogas de 35-41bp. Las secuencias de repeticiones directas de *M. bovis* también fueron halladas en *Mycobacterium tuberculosis*, más no en otras especies de *Mycobacterium* como *M. avium*, *M. asiaticum*, *M. chitae*, *M. gordonae*, *M. Kansasii* o *M. terrae*. De esta manera los investigadores concluyeron que las secuencias repetitivas homólogas entre sí servían como región de integración de las cepas de *M. tuberculosis*.

En el año 1993, el investigador español Francisco Mojica presenta el primer reporte de secuencias repetitivas altamente homólogas en arqueas (Mojica et al. 1993). El objetivo del proyecto era estudiar el comportamiento de algunos genes y sus niveles de expresión a diferentes concentraciones de salinidad para determinar si existían diferentes patrones de

expresión en la arquea halofílica *Haloferax mediterranei* clon m122 y m61. En el clon m122 se observó la presencia de secuencias repetitivas altamente homólogas, cada una de ellas conformadas por 30–34bp idénticos e interespaciadas por secuencias únicas de 34-39bp. Los autores de la publicación sugieren que las repeticiones en tandem podrían ser una posible secuencia señal de ADN.

En el año 1995, en un intento por identificar por primera vez la función a las secuencias repetidas en tandem, Mojica et al. (1995) reportaron la presencia de largos segmentos de repeticiones en tándem en el genoma de *Haloferax mediterranei* y *Haloferax volcanii*. Tenían como objetivo determinar la extensión y recurrencia de las repeticiones en el genoma de estas Haloarqueas. Llamó la atención como estas repeticiones largas en tándem tienen una secuencia y número de nucleótidos conservados y simetría de diada, además de encontrarse separados por distancias nucleotídicas constantes. El primer nombre que se le da a estas secuencias repetitivas altamente homólogas, antes de ser denominadas CRISPR, es el de TREPs (del inglés tándem repeats) o repeticiones en tándem. El análisis detallado de estas secuencias repetitivas en tándem del clon m122 de *H. Mediterranei* mostró que poseía secuencias de 30bp, repetidas hasta 21 veces cada una de ellas y separadas por secuencias únicas de 33-39bp. Similarmente, *H. volcanii* poseía secuencias repetidas altamente conservadas de 30bp y separadas por secuencias equivalentes a las observadas en *H. mediterranei*. La diferencia entre las secuencias repetidas de ambas especies de Haloferax era de tan solo de tres nucleótidos. Debido a estos hallazgos, se asoció a estas repeticiones en tandem a una posible función con los replicones que forman parte de la división celular de procariontes. El autor del estudio sugiere que estas estructuras repetitivas en tándem parcialmente palindrómicas podrían ser de importancia biológica debido a su amplia distribución en dos dominios como son Arquea y Bacteria, haciendo alusión a esta última por la *E. coli* donde se encontraron también estructuras repetitivas en tándem (Ishino et al. 1987; Nakata et al. 1989).

En el año 2000, en una microcorrespondencia de Mojica et al. (2000), publicada en Molecular Microbiology, se daba cuenta de secuencias repetitivas en tandem en diferentes procariontes. A estas secuencias repetitivas se les denominó Short Regularly Spaced Repeats (SRSRs) que en castellano se traduce como: repeticiones cortas regularmente espaciadas. En dicha correspondencia se mostró que tanto arquea como bacterias poseen estas secuencias cortas parcialmente palindrómicas y conservadas, los cuales pueden hallarse desde arreglos repetitivos que van desde 1 hasta más de 10 clústeres. Al ver que las secuencias repetitivas en tandem están presentes en una gran variedad de procariontes, los autores plantean la pregunta de si las SRSRs tienen una función en común en los procariontes o si son secuencias remanentes de secuencias antepasadas (ver Figura 2).

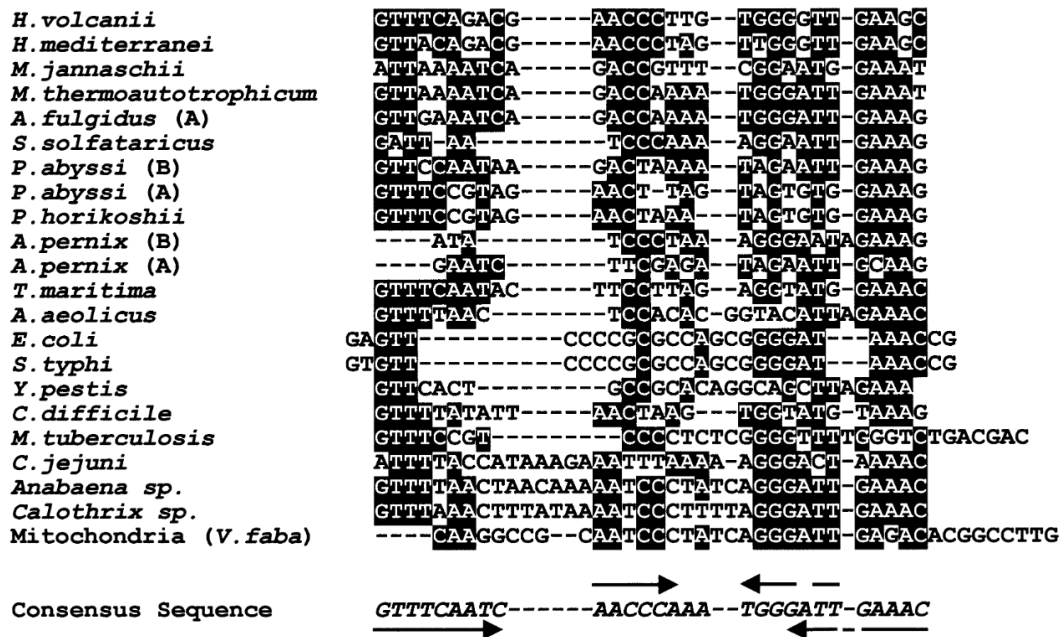


Figura 2. Alineamiento de SRSR provenientes de diferentes organismos procariotas. Los bloques resaltados indican las posiciones ocupadas por las bases más frecuente en cada posición de las secuencias alineadas. El autor solo ha considerado el tipo más abundante de elemento SRSR para *M. jannaschii* y *Clostridium difficile*. Así mismo, se alinearon dos tipos de SRSR (A y B) presentes en *P. abyssi*, *A. pernix* y *A. fulgidus*. Se muestra la secuencia consenso con la base más frecuente en cada posición del alineamiento en la parte inferior de la figura. Las flechas indican el carácter palindrómico de los SRSR. Reproducido de Mojica et al. (2000).

En un estudio subsecuente, Jansen et al. (2002) reportaron la identificación de cuatro genes asociados a las secuencias repetitivas de ADN en varios procariotas. Además, reportan una secuencia leader contigua por un sólo lado a estas secuencias repetitivas. Junto a Francisco Mojica (fue él quien propuso el nombre) acordaron dar un nombre un nuevo nombre a esta familia de secuencias repetitivas como **Clustered Regularly Interspaced Short Palindromic Repeats** o **CRISPR**. En dicho estudio, el objetivo fue determinar la presencia de los loci CRISPR entre diferentes microorganismos procariotas; y es a través de esta comparación que se identifica la secuencia leader y los cuatro genes asociados a los loci CRISPR. Estas secuencias leader, se encuentran flanqueando por un solo lado a las secuencias CRISPR y se caracterizan por ser un 80% idénticas en su secuencia entre procariotas de la misma especie y por no presentar homología con secuencias de especies no relacionadas taxonómicamente. Con respecto a los genes asociados a CRISPR (cas por sus siglas en ingles), se identificaron cuatro genes con los nombres de cas1, cas2, cas3 y cas4. Estos genes cas pueden o no encontrarse contiguos a los loci CRISPR y es el gen cas1 el que siempre se encuentra presente acompañado de uno o más genes cas. Los genomas de procariotas estudiados carentes de secuencias CRISPR tienen ausencia de genes cas. Esto último nuevamente planteo preguntas sobre la función de las secuencias CRISPR, los genes cas y el origen de las secuencias espaciadoras.

En un nuevo estudio, Tang et al. (2002) tenían como objetivo identificar RNAm no codantes, denominados small non-mRNAs (snmRNAs), en *Archaeoglobus fulgidus*, una arquea hipertermófila que metaboliza sulfuro. Dicho estudio reportó 86 candidatos a snmRNAs que aún no habían sido caracterizados. Los 86 snmRNAs candidatos fueron caracterizados en dos clases, basado en la presencia (clase I) o ausencia (clase II) de motivos estructurales conocidos. A su vez, cada snmRNAs fue asignado a un grupo respectivo según su estructura y/o locación genómica. Es así como en el grupo 3 de la clase 1 se agruparon 22 snmRNAs los cuales eran transcritos de repeticiones cortas regularmente espaciadas. El primer locus SRSR, poseía 48 secuencias idénticas de 30 nucleótidos cada una interespaciadas por secuencias únicas de 38bp. El segundo locus SRSR poseía 60 secuencias idénticas de 30bp, cada una interespaciadas por secuencias únicas de 38 nucleótidos. El tercer locus SRSR, poseía 42 secuencias casi idénticas de 36 nucleótidos cada una interespaciadas por secuencias únicas de 41bp. Esta fue la primera evidencia de que los loci CRISPR dan lugar a elementos RNA no codantes, transcritos conocidos hoy como CRISPR RNAs (crRNA).

Mojica et al. (2005) en el 2003 (no se publicó hasta el 2005) demostraron por primera vez que los espaciadores CRISPR derivan de secuencias preexistentes provenientes de elementos genéticos foráneos, y se propone por primera vez el concepto de inmunidad en procariontas que poseen CRISPR loci. El objetivo de este proyecto fue determinar si las secuencias espaciadoras CRISPR tenían alguna similitud significativa con secuencias nucleotídicas almacenadas en el GeneBank. De las 4500 secuencias espaciadoras CRISPR comparadas, correspondientes a 67 cepas de 36 géneros de arquea y bacterias, 88 espaciadores CRISPR eran homólogos con secuencias de 47 fagos, 10 plásmidos extracromosómicos y 31 secuencias cromosómicas no directamente relacionado a ADN foráneo. Los investigadores, basados en estudios previos de infectividad de fagos y plásmidos foráneos en procariontas, revelaron cómo dichos elementos extra-cromosómicos son incapaces de infectar a cepas de arquea o bacterias que poseen los espaciadores CRISPR homólogos a dichos elementos. Estos hallazgos sugieren una relación entre los loci CRISPR e inmunidad en arquea y bacterias. El autor y su equipo hacen una importante declaración proponiendo que los transcritos (snmRNA) de los loci CRISPR descubiertos por Tang et al. (2002), es decir los CRISPR-RNA, actúan como ARN reguladores que reconocen específicamente su secuencia target debido a la homología del ARN de la secuencia espaciadora CRISPR y las secuencias de fagos o plásmidos foráneos, con la finalidad de evitar una infección.

En 2005, Pourcel et al. (2005) se trazan como objetivo analizar los elementos CRISPR de cepas de *Yersinia pestis* de tres biovars con orígenes geográficos diferentes, además de otras nueve cepas de *Yersinia pseudotuberculosis*. En el mencionado estudio se hacen 2 observaciones importantes. Primeramente, los investigadores concluyen que la adición de secuencias no repetitivas o nuevos espaciadores en un locus CRISPR es polarizada, es decir, que la inserción siempre se da por el extremo de la secuencia leader. En segundo lugar, los investigadores proponen que los elementos CRISPR son estructuras capaces de tomar partes de ADN foráneo como parte de un mecanismo de defensa y que los elementos CRISPR representan una memoria de pasadas agresiones genéticas.

Bolotin et al. (2005) analizaron la secuencia del genoma completo de *Streptococcus thermophilus* en donde hallaron un locus CRISPR formado por 42 repeticiones de 36bp separadas por secuencias únicas de 30bp. Un análisis *In silico* reveló la presencia de tres genes asociados a CRISPR en *S. thermophilus*: cas1B, cas5 y cas6. También encontraron homología en las secuencias de los espaciadores de los diferentes loci CRISPR con varios fagos y plásmidos. Además, evaluaron si existía una correlación entre la resistencia a fagos por parte de *S. thermophilus* y el número de espaciadores en un determinado locus CRISPR. Los autores observaron una correlación negativa entre la proporción de fagos propagados y el número de espaciadores presentes. En consecuencia, los autores concluyen que la resistencia a infección por fagos observada esta mediada por los espaciadores CRISPR de *S. thermophilus*, y proponen que esta se encuentra mediada por un mecanismo de inhibición vía ARN anti-sentido contra la expresión de genes de fagos. Uno de los aportes más importantes de este artículo fue la identificación de una secuencia consenso adyacente a las secuencias complementarias a los espaciadores que se encuentran en diferentes fagos y plásmidos. Esta secuencia sería luego conocida como PAM (protospacer adjacent motif) o motivo adyacente al protoespaciador.

En 2007, Barrangou et al. (2007) demostraron por primera vez de manera experimental (Mojica lo demostró *In Silico*), que el sistema CRISPR de procariontas provee resistencia adquirida contra bacteriófagos. Para ello decidieron evaluar si el locus CRISPR de *S. thermophilus* se altera durante el proceso de resistencia a fagos. Los resultados mostraron que en el locus CRISPR se insertaron de 1 a 4 espaciadores adicionales como respuesta a una infección inducida (**ver Figura 3**). Además, la inserción siempre se dio a un determinado extremo del locus. El análisis de secuencia demostró que los espaciadores insertados en el locus CRISPR eran homólogos a secuencias del genoma de los fagos utilizados en estos experimentos. Como una comprobación de su hipótesis, los autores eliminaron los nuevos espaciadores del locus CRISPR, observándose que el *S. thermophilus* perdía su resistencia a los fagos. Adicionalmente, se inactivó el gen cas5, observando nuevamente pérdida de la resistencia de *S. thermophilus* a los fagos experimentales. Con este último experimento, se demostró que los genes cas tienen implicancia en la resistencia a bacteriófagos mediada por elementos CRISPR.

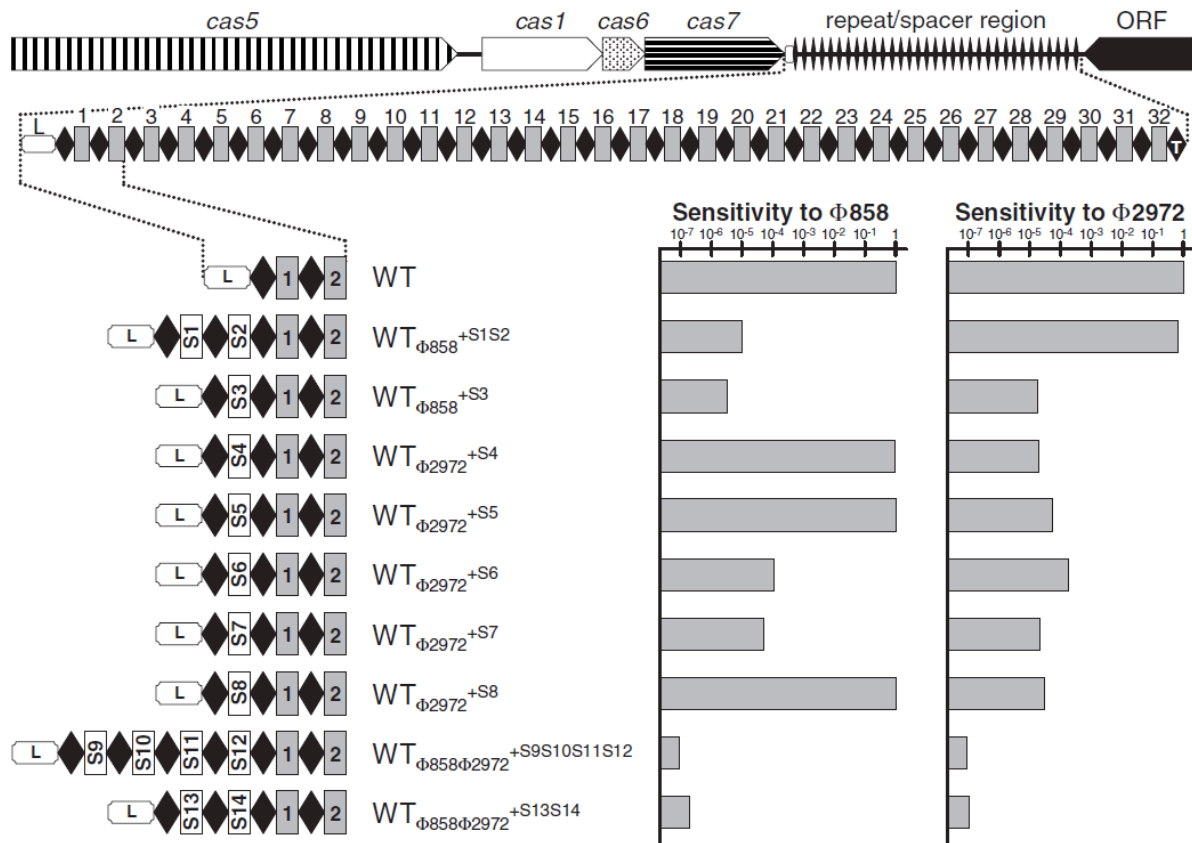


Figura 3. Esquema general del locus CRISPR de *Streptococcus thermophilus*, espaciadores recién adquiridos en mutantes resistentes a fagos y sensibilidad a los fagos correspondiente. La figura muestra el locus CRISPR1 de la cepa DGCC7710 (WT) en parte superior. Así mismo, se muestra la región de repeticiones/espaciadores de *S. thermophilus* en el medio: repeticiones (rombos negros), espaciadores (cuadros grises numerados), secuencia líder (L, recuadro blanco) y repetición terminal (T, rombo negro). (Arriba a la derecha). Se detalla el contenido de espaciadores en el lado de la secuencia líder del locus CRISPR1 en mutantes resistentes a fagos, con espaciadores recién adquiridos (recuadros blancos, S1 a S14). (Abajo a la izquierda). Finalmente, se muestra la sensibilidad de cada cepa de *S. thermophilus* a los fagos $\Phi 858$ y 2972. La sensibilidad se representa como un histograma de la eficiencia de formación de placas (EOP), que es la proporción de recuento de placas de una cepa mutante con respecto a la de tipo salvaje. Reproducido de

En 2008, el grupo de investigación liderado por John van der Oost publica los resultados experimentales que demuestran la participación de transcritos CRISPR (llamados crARNs) como parte del sistema de defensa de procariontes contra bacteriófagos (Brouns et al. 2008). Entre los resultados obtenidos, se describe como los transcritos CRISPR se sintetizan de manera inmadura (pre-crARN) y que necesitan de unas proteínas llamadas cascada (complejo asociado a CRISPR para la defensa antiviral) para su maduración. Estas proteínas cascada, encargadas del procesamiento y maduración de nuevos transcritos CRISPR más cortos (crRNAs), forman parte de los genes cas. Así mismo, se plantea la hipótesis de que estos crRNAs se unen directamente a la secuencia de ADN o ARN viral, ya que se comprobó

experimentalmente que aquellas bacterias expuestas al bacteriófago Lambda y que poseían secuencias CRISPR anti- lambda no eran afectadas por dicho bacteriófago.

En una investigación publicada en la revista Science en 2008, el argentino Luciano Marraffini demostró que el sistema CRISPR tiene como diana molecular de interferencia secuencias de ADN y no de ARN como se había pensado (Marraffini and Sontheimer 2008). Además, se demostró que el sistema CRISPR, interfiere con secuencias de fagos para evitar la infección, así como también lo hace con plásmidos, que pueden ser transmitidos de manera horizontal ya sea en bacterias o arqueas.

Es así que hasta el 2008, ya se conocían bastantes principios y elementos que eran parte del sistema CRISPR. Ya se conocía que los arreglos CRISPR poseen una secuencia líder por donde se da la incorporación de nuevos espaciadores. Además, contiguos a los arreglos CRISPR encontramos a los genes cas, de los cuales cas1 y cas2 se encuentran presentes en los diferentes tipos de arreglos CRISPR descritos hasta la fecha. Así también, se conocía que la interferencia contra elementos móviles genéticos foráneos en bacterias o arqueas es mediada por transcritos CRISPR (crRNA) que contienen la secuencia espaciadora.

Es así como en 2009, el grupo de investigación liderado por Mojica describe la presencia de un motivo de tres nucleótidos adyacente a cada protoespaciador (Mojica et al. 2009). Dicho trinucleotido de secuencia NGG, es nombrado PAM (proto-spacer adjacent motifs). El motivo PAM permite que las secuencias protoespaciadoras no sean elegidas al azar para luego ser incorporadas en los arreglos CRISPR como un espaciador. Además, en la adquisición de nuevos espaciadores, el motivo PAM determinará la orientación del espaciador, es decir, si tenemos dos futuros protoespaciadores, de un bacteriófago cualquiera, que tienen orientación opuesta, una respecto a la otra, veremos que cuando forman parte de los arreglos CRISPR como espaciadores, éstos poseen la misma orientación con respecto al motivo PAM, el cual a su vez está orientada hacia la secuencia líder. Este trabajo es una confirmación experimental del trabajo de Bolotin et al. (2005).

En 2010, Garneau et al. (2010) publican en la revista Nature datos experimentales que demuestran que el sistema inmune bacteriano CRISPR/Cas puede producir un corte molecular preciso en el ADN de bacteriófagos, así como de plásmidos. Este corte se caracteriza por ser un corte de doble hebra en el protoespaciador (ADN invasor) y además ocurren en sitios específicos. Este trabajo corrobora el trabajo realizado por Marraffini y Sontheimer (2008) en *Staphylococcus*.

Un año después, el grupo de investigación liderado por Emmanuelle Charpentier describe y caracteriza la presencia de un nuevo transcritto semi-complementario a los transcritos crARN, denominado, trans-activating CRISPR RNA o tracrRNA (Deltcheva et al. 2011). Como se sabe, el transcritto crARN está formado por una secuencia espaciadora y un segmento de una de las secuencias repetitivas, siendo esta última secuencia complementaria a la secuencia del recién caracterizado tracrRNA. Dicha investigación demostró que la función del tracrARN (ARN no codante) es la de dirigir la maduración del otro ARN no codante, es decir, el pre-crARN quien madurará a su forma activa como varios crRNAs. Así mismo, se evidencio la participación de una ARNasa III en el procesamiento de pre-crARN a crARN, así

como la participación de una proteína Cas inicialmente denominada Csn1, posteriormente nombrada Cas9.

Para agosto del 2012, una investigación dirigida por Emmanuelle Charpentier y Jennifer Doudna (Jinek et al. 2012), y publicada en la revista Science, propone una nueva metodología para edición de genomas. Esta nueva metodología se propone como alternativa a las ya conocidas ZFN (Zinc-finger nucleases) o TALEN (Transcription Activator-Like Effector Nucleases).

En este trabajo se determinó el mecanismo por el cual el sistema CRISPR lleva a cabo la interferencia del ADN invasor (protoespaciador). Dicha interferencia depende de la formación de una estructura formada por dos ARNs (crRNA y tracrRNA); que junto a una endonucleasa llevan a cabo el corte de doble hebra en un sitio específico del protoespaciador. Los dos ARNs involucrados (crARN y tracrARN) se unen de forma complementaria hebra a hebra y unidos a la proteína Cas9 permiten que se lleve a cabo el corte de la doble hebra del protoespaciador de manera específica. Para que la endonucleasa Cas9 reconozca el protoespaciador se requiere de dos cosas; 1) que la secuencia del crARN sea complementaria al protoespaciador y 2) que adyacente al protoespaciador se encuentre el motivo PAM (NGG) (ver Figura 4).

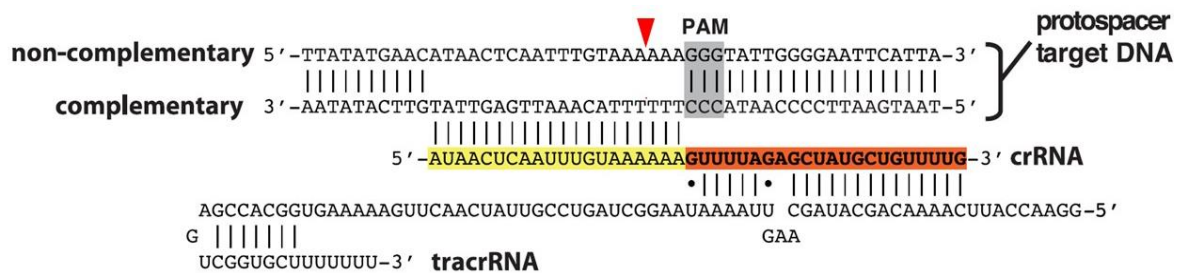


Figura 4. La Cas9 es una endonucleasa de ADN guiada por dos moléculas de ARN. Representación esquemática de secuencias de tracrRNA, crRNA y su respectivo protoespaciador. Se representan regiones de complementariedad de crARN con tracrARN (naranja) y el protoespaciador (amarillo). La secuencia PAM se muestra en gris. El sitio de corte de doble hebra está representado por una punta de flecha roja.

En la misma publicación, los investigadores demostraron que se podía programar a la endonucleasa Cas9 a través de un único transcripto de ARN, es decir un híbrido que fusionaba al tracrARN y crARN; conocido como single guide RNA (sgRNA) que serviría como guía a la Cas9 para encontrar y llevar a cabo el corte de doble hebra en cualquier secuencia de ADN de interés. Esto último, es lo que llevaría a las autoras a ser galardonadas con el premio Nobel en química en el año 2020 (ver Figura 5).

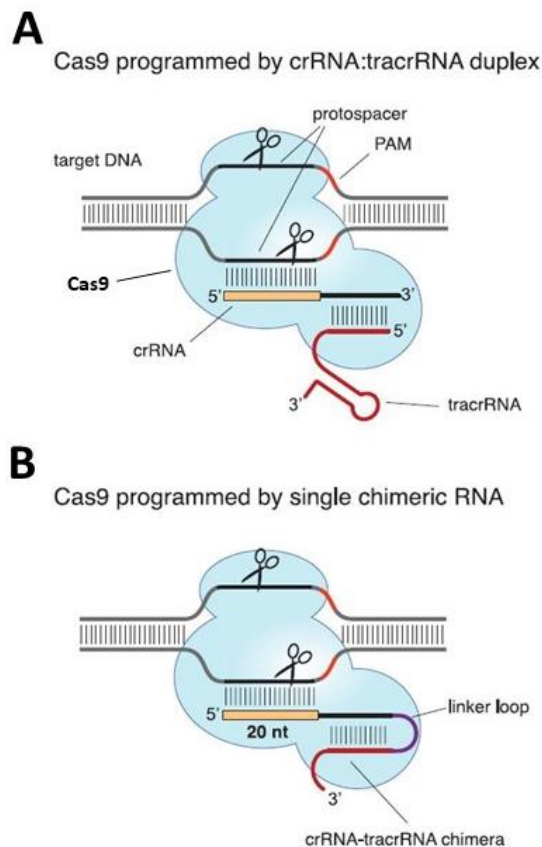


Figura 5. Componentes del sgRNA. El panel **A** muestra como la unión de crRNA y tracrRNA se alcanza únicamente por complementariedad de bases nucleotídicas, como ocurre en procariontas. El panel **B** muestra que el crRNA y tracrRNA pueden ser fusionados para formar una sola molécula -sgRNA- que servirá de guía para la Cas9. Adaptado de Jinek et al. (2012).

En enero del 2013, el laboratorio liderado por Feng Zhang en el Instituto BROAD del MIT, reportó por primera vez el uso del CRISPR en células de mamífero en cultivo, demostrándose por primera vez que el sistema CRISPR, sistema inmune de procariontas, era capaz de llevar a cabo la edición génica en células eucariotas (Cong et al. 2013).

Desde su utilización por el grupo de Feng Zhang, la proteína Cas9, derivada del sistema inmunitario bacteriano CRISPR tipo II (para una revisión completa del sistema de clasificación de los diferentes sistemas CRISPR ver Makarova et al. (2018)), se ha convertido en una poderosa herramienta para modificar el genoma en diversos organismos. Siendo que la Cas9 es una endonucleasa de ADN guiada por una molécula de ARN, es posible programarla fácilmente para cortar nuevos sitios diana alterando su secuencia de ARN guía (crRNA) (Wang et al. 2016), lo que ha facilitado y democratizado el uso de la edición genética en diversos laboratorios de investigación alrededor del mundo.

Por otro lado, se han venido buscando nuevas formas de utilizar la Cas9 en otras aplicaciones, dando lugar a una versión de Cas9 donde su función de nucleasa ha sido desactivada. Esta Cas9 desactivada (dCas9) proporciona una plataforma versátil para el “targeting” de secuencias específicas de ADN guiada por una molécula de ARN, lo que puede ser usado para regular la expresión o estado epigenético de genes y/o porciones

específicas del genoma, además de permitir obtener imágenes del genoma (ver Figura 6). Para una revisión extensa de las nuevas herramientas generadas a partir del sistema CRISPR revisar Adli (2018).

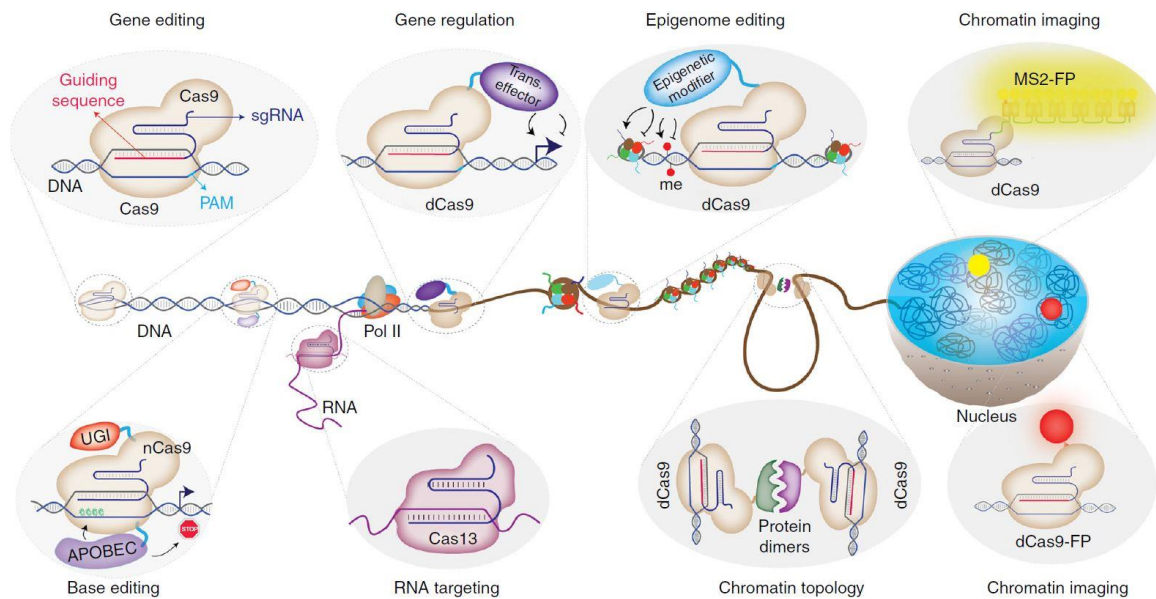


Figura 6. Tecnología CRISPR: más allá de la edición genética. La Cas9 “wild-type” (WT) permite la edición del genoma a través de su actividad endonucleasa guiada por una molecular de ARN. Sin embargo, las enzimas Cas9 catalíticamente desactivadas (dCas9) se han aprovechado para permitir la regulación de la expresión génica, la alteración del estado epigenético, la toma de imágenes y manipulación de la topología de la cromatina. Además, la enzima nickase Cas9 (nCas9) catalíticamente alterada se ha utilizado como plataforma para la edición de bases sin ruptura de doble cadena. Finalmente, también se han descrito nuevos sistemas CRISPR/Cas dirigidos a moléculas de ARN (Cas13). Reproducido de Adli (2018).

Además de la nueva “caja de herramientas” moleculares que ofrece el descubrimiento y desarrollo de la tecnología CRISPR, su potencial más esperanzador es su uso en terapia génica. La terapia génica es una técnica que busca otorgar un beneficio terapéutico a través de la modificación (disrupción, corrección o reemplazo) de genes en una persona para tratar o curar una enfermedad. El campo de la terapia génica ha visto tanto éxitos tempranos (Blaese et al. 1995; Bordignon et al. 1995; Rosenberg et al. 1990) como fracasos trágicos en ensayos clínicos (Cavazzana-Calvo et al. 2000; Raper et al. 2003; Sibbald 2001). Sin embargo, la llegada del sistema CRISPR/Cas9 a la clínica ha brindado una segunda oportunidad para que la terapia génica se recupere del estigma de sus fracasos tempranos y la posición como una valiosa estrategia terapéutica. El advenimiento de la tecnología CRISPR en los ensayos clínicos está allanando el camino para que surja la nueva era de la terapia génica basada en el sistema CRISPR (Uddin et al. 2020).

Más recientemente, el sistema CRISPR ha sido utilizado como herramienta diagnóstica en respuesta a la pandemia mundial desatada por el SARS-CoV-2 (Uddin et al. 2020). Para este fin se ha utilizado un ensayo basado en el uso de la proteína Cas12. Dicho ensayo, conocido como SARS-CoV.2 DETECTR, tan solo requiere 40min para arrojar resultados con un 95% de precisión reportada (Broughton et al. 2020) y se presenta como alternativa rápida y precisa

a los 2 métodos convencionales de detección del SARS-CoV-2 (molecular y serológico). Ya en 2018, el grupo dirigido por Feng Zhang había desarrollado SHERLOCK, una herramienta de diagnóstico que se basa en los mismos principios que DETECTR, pero depende de la actividad de la nucleasa Cas13, que reconoce y corta específicamente solo ARN, en lugar de ADN (como Cas12). Tanto SHERLOCK como su versión mejorada SHERLOCKv2 han demostrado su capacidad para el diagnóstico de virus como dengue y zika (Gootenberg et al. 2018) y SARS-CoV-2 (Joung et al. 2020). Para una revisión (y comparación) detallada sobre el funcionamiento de los ensayos DETECTR y SHERLOCK ver Mustafa and Makhawi (2021).

Pero el sistema CRISPR, no solo se presenta como una alternativa para diagnóstico en esta pandemia, si no también podría ser utilizada como una alternativa terapéutica para enfrentar el COVID-19. Recordemos que la proteína Cas13 posee el potencial de unirse y romper únicamente cadenas de ARN. Utilizando esta novedosa Cas13 se ha desarrollado una herramienta terapéutica conocida como PAC-MAN (Prophylactic Antiviral CRISPR in huMAN cells), la cual es capaz de tener varias regiones diana simultáneamente, lo que permite una degradación más efectiva del ARN viral (Abbott et al. 2020). Estos nuevos avances, producto del descubrimiento de nuevas variantes del sistema CRISPR, podrían ser nuevamente utilizadas para cumplir el propósito original del sistema CRISPR, como mecanismo de defensa contra virus (esta vez contra un virus específico como SARS-CoV-2) para brindar ayuda durante esta pandemia (Uddin et al. 2020).

Conclusiones y avances futuros

El descubrimiento del sistema CRISPR en procariontes ha sido un proceso largo de varios años (e incluso décadas) donde las contribuciones de diferentes investigadores alrededor del mundo ayudaron a develar el funcionamiento del sistema. Este camino hacia el premio Nobel 2020, inicio con estudios de investigación básica donde se buscaba identificar la estructura primaria de la proteasa Cas9 en *E. coli* (Ishino et al. 1987) y ha desencadenado una revolución de herramientas moleculares novedosas (Pickar-Oliver and Gersbach 2019), que sigue expandiéndose en la actualidad. Esto debe alentar a las instituciones encargadas de evaluar y financiar ciencia en nuestro país, a seguir financiando investigación básica como fuente inagotable de nuevas ideas y recursos que puedan convertirse luego en investigaciones aplicadas. Esto es especialmente cierto en un país megadiverso y con escasos recursos financieros como el nuestro.

La utilización del sistema CRISPR para su uso en células eucariotas revolucionó y continúa alborotando el campo de la ingeniería genómica. La facilidad de uso, especificidad y reproducibilidad de esta tecnología han revolucionado la edición del genoma para investigaciones que van desde la ciencia fundamental hasta la medicina traslacional. Más aún, los éxitos iniciales han permitido descubrir nuevos sistemas para identificar y manipular ácidos nucleicos, incluidos los de los ortólogos Cas9, Cas12 y Cas13. Además, las proteínas Cas inactivas han sido fusionadas con una variedad de proteínas efectoras para regular la expresión génica, las modificaciones epigenéticas y las interacciones de la cromatina. En conjunto, los nuevos avances están mejorando considerablemente nuestra comprensión de los procesos biológicos, al mismo tiempo que se viene probando su potencial para aplicaciones clínicas en terapias génicas y celulares. Aunque los resultados preclínicos son prometedores, aun se debe evaluar la seguridad y la eficacia de estos métodos para asegurar su avance clínico.

Finalmente, durante la pandemia varios grupos de investigación en nuestro país apostaron por la utilización del sistema CRISPR como herramienta de diagnóstico para SARS-CoV-2. Dos de los proyectos más conocidos dieron resultados positivos (Alcantara et al. 2021; del Prado et al. 2021), sin embargo, debido a la inmadurez tecnológica de nuestro medio, los prototipos quedaron sin desarrollo. Nuestras autoridades y entidades encargadas de promover el avance científico no deben dejar que este tipo de esfuerzos se ahoguen por falta de apoyo. Actualmente, existen al menos una docena de grupos nacionales que vienen llevando a cabo proyectos de investigación basados en la utilización del sistema CRISPR para edición génica y modelamiento de mutaciones genéticas, terapia génica y diagnóstico de enfermedades, estos esfuerzos deben ser canalizados apropiadamente por los mismos investigadores y las autoridades competentes.

LITERATURA CITADA

2022. NobelPrize.org. Nobel Prize Outreach AB; [accessed 27 Agosto de 2022].
<https://www.nobelprize.org/prizes/chemistry/2020/summary/>.
- Abbott TR, Dhamdhare G, Liu Y, Lin X, Goudy L, Zeng L, Chemparathy A, Chmura S, Heaton NS, Debs R et al. 2020. Development of crispr as an antiviral strategy to combat sars-cov-2 and influenza. *Cell*. 181(4):865-876 e812.
<https://doi.org/10.1016/j.cell.2020.04.020>
- Adli M. 2018. The crispr tool kit for genome editing and beyond. *Nature communications*. 9(1):1911. <https://doi.org/10.1038/s41467-018-04252-2>
- Alcantara R, Penaranda K, Mendoza-Rojas G, Nakamoto JA, Martins-Luna J, Del Valle-Mendoza J, Adauí V, Milon P. 2021. Unlocking sars-cov-2 detection in low- and middle-income countries. *Cell reports methods*. 1(7):100093.
<https://doi.org/10.1016/j.crmeth.2021.100093>
- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007. Crispr provides acquired resistance against viruses in prokaryotes. *Science*. 315(5819):1709-1712. <https://doi.org/10.1126/science.1138140>
- Blaese RM, Culver KW, Miller AD, Carter CS, Fleisher T, Clerici M, Shearer G, Chang L, Chiang Y, Tolstoshev P et al. 1995. T lymphocyte-directed gene therapy for ada- scid: Initial trial results after 4 years. *Science*. 270(5235):475-480.
<https://doi.org/10.1126/science.270.5235.475>
- Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. 2005. Clustered regularly interspaced short palindrome repeats (crisprs) have spacers of extrachromosomal origin. *Microbiology*. 151(Pt 8):2551-2561. <https://doi.org/10.1099/mic.0.28048-0>
- Bordignon C, Notarangelo LD, Nobili N, Ferrari G, Casorati G, Panina P, Mazzolari E, Maggioni D, Rossi C, Servida P et al. 1995. Gene therapy in peripheral blood lymphocytes and bone marrow for ada- immunodeficient patients. *Science*. 270(5235):470-475.
<https://doi.org/10.1126/science.270.5235.470>
- Broughton JP, Deng X, Yu G, Fasching CL, Servellita V, Singh J, Miao X, Streithorst JA, Granados A, Sotomayor-Gonzalez A et al. 2020. Crispr-cas12-based detection of sars-cov-2. *Nature biotechnology*. 38(7):870-874. <https://doi.org/10.1038/s41587-020-0513-4>
- Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, van der Oost J. 2008. Small crispr rnas guide antiviral defense in prokaryotes. *Science*. 321(5891):960-964.
<https://doi.org/10.1126/science.1159689>
- Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G, Gross F, Yvon E, Nusbaum P, Selz F, Hue C, Certain S, Casanova JL et al. 2000. Gene therapy of human severe combined immunodeficiency (scid)-x1 disease. *Science*. 288(5466):669-672.
<https://doi.org/10.1126/science.288.5466.669>
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA et al. 2013. Multiplex genome engineering using crispr/cas systems. *Science*. 339(6121):819-823. <https://doi.org/10.1126/science.1231143>
- del Prado JAN, Reyes AQ, La Torre JB, Gutiérrez Loli R, Pinzón Olejua A, Chamorro Chirinos ER, Loza Mauricio FA, Maguiña JL, Leon J, Rodríguez Aliaga P et al. 2021. Clinical validation of rcsms: A rapid and sensitive crispr-cas12a test for the molecular detection of sars-cov-2 from saliva. *medRxiv.2021.2004.2026.21256081*.
<https://doi.org/10.1101/2021.04.26.21256081>

- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. 2011. Crispr rna maturation by trans-encoded small rna and host factor rnaase iii. *Nature*. 471(7340):602-607. <https://doi.org/10.1038/nature09886>
- Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadan AH, Moineau S. 2010. The crispr/cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*. 468(7320):67-71. <https://doi.org/10.1038/nature09523>
- Gootenberg JS, Abudayyeh OO, Kellner MJ, Joung J, Collins JJ, Zhang F. 2018. Multiplexed and portable nucleic acid detection platform with cas13, cas12a, and csm6. *Science*. 360(6387):439-444. <https://doi.org/10.1126/science.aaq0179>
- Hermans PW, van Soolingen D, Bik EM, de Haas PE, Dale JW, van Embden JD. 1991. Insertion element is987 from mycobacterium bovis bcg is located in a hot-spot integration region for insertion elements in mycobacterium tuberculosis complex strains. *Infection and immunity*. 59(8):2695-2705. <https://doi.org/10.1128/iai.59.8.2695-2705.1991>
- Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. 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):5429-5433. <https://doi.org/10.1128/jb.169.12.5429-5433.1987>
- Jansen R, Embden JD, Gastra W, Schouls LM. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular microbiology*. 43(6):1565-1575. <https://doi.org/10.1046/j.1365-2958.2002.02839.x>
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-rna-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 337(6096):816-821. <https://doi.org/10.1126/science.1225829>
- Joung J, Ladha A, Saito M, Kim NG, Woolley AE, Segel M, Barretto RPJ, Ranu A, Macrae RK, Faure G et al. 2020. Detection of sars-cov-2 with sherlock one-pot testing. *The New England journal of medicine*. 383(15):1492-1494. <https://doi.org/10.1056/NEJMc2026172>
- Makarova KS, Wolf YI, Koonin EV. 2018. Classification and nomenclature of crispr-cas systems: Where from here? *The CRISPR journal*. 1(5):325-336. <https://doi.org/10.1089/crispr.2018.0033>
- Marraffini LA, Sontheimer EJ. 2008. Crispr interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*. 322(5909):1843-1845. <https://doi.org/10.1126/science.1165771>
- Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of molecular evolution*. 60(2):174-182. <https://doi.org/10.1007/s00239-004-0046-3>
- Mojica FJ, Diez-Villasenor C, Soria E, Juez G. 2000. Biological significance of a family of regularly spaced repeats in the genomes of archaea, bacteria and mitochondria. *Molecular microbiology*. 36(1):244-246. <https://doi.org/10.1046/j.1365-2958.2000.01838.x>
- Mojica FJ, Ferrer C, Juez G, Rodriguez-Valera F. 1995. Long stretches of short tandem repeats are present in the largest replicons of the archaea haloferax mediterranei and haloferax volcanii and could be involved in replicon partitioning. *Molecular microbiology*. 17(1):85-93. https://doi.org/10.1111/j.1365-2958.1995.mmi_17010085.x

- Mojica FJ, Juez G, Rodriguez-Valera F. 1993. Transcription at different salinities of haloferax mediterranei sequences adjacent to partially modified pstI sites. *Molecular microbiology*. 9(3):613-621. <https://doi.org/10.1111/j.1365-2958.1993.tb01721.x>
- Mojica FJM, Diez-Villasenor C, Garcia-Martinez J, Almendros C. 2009. Short motif sequences determine the targets of the prokaryotic crISPR defence system. *Microbiology*. 155(Pt 3):733-740. <https://doi.org/10.1099/mic.0.023960-0>
- Mustafa MI, Makhawi AM. 2021. Sherlock and detectr: CrISPR-cas systems as potential rapid diagnostic tools for emerging infectious diseases. *Journal of clinical microbiology*. 59(3). <https://doi.org/10.1128/JCM.00745-20>
- Nakata A, Amemura M, Makino K. 1989. Unusual nucleotide arrangement with repeated sequences in the escherichia coli k-12 chromosome. *Journal of bacteriology*. 171(6):3553-3556. <https://doi.org/10.1128/jb.171.6.3553-3556.1989>
- Pickar-Oliver A, Gersbach CA. 2019. The next generation of crISPR-cas technologies and applications. *Nature reviews Molecular cell biology*. 20(8):490-507. <https://doi.org/10.1038/s41580-019-0131-5>
- Pourcel C, Salvignol G, Vergnaud G. 2005. CrISPR elements in yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*. 151(Pt 3):653-663. <https://doi.org/10.1099/mic.0.27437-0>
- Raper SE, Chirmule N, Lee FS, Wivel NA, Bagg A, Gao GP, Wilson JM, Batshaw ML. 2003. Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Molecular genetics and metabolism*. 80(1-2):148-158. <https://doi.org/10.1016/j.ymgme.2003.08.016>
- Rosenberg SA, Aebersold P, Cornetta K, Kasid A, Morgan RA, Moen R, Karson EM, Lotze MT, Yang JC, Topalian SL et al. 1990. Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. *The New England journal of medicine*. 323(9):570-578. <https://doi.org/10.1056/NEJM199008303230904>
- Sibbald B. 2001. Death but one unintended consequence of gene-therapy trial. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne*. 164(11):1612.
- Tang TH, Bachelier JP, Rozhdestvensky T, Bortolin ML, Huber H, Drungowski M, Elge T, Brosius J, Huttenhofer A. 2002. Identification of 86 candidates for small non-messenger rnas from the archaeon archaeoglobus fulgidus. *Proceedings of the National Academy of Sciences of the United States of America*. 99(11):7536-7541. <https://doi.org/10.1073/pnas.112047299>
- Uddin F, Rudin CM, Sen T. 2020. CrISPR gene therapy: Applications, limitations, and implications for the future. *Frontiers in oncology*. 10:1387. <https://doi.org/10.3389/fonc.2020.01387>
- Wang H, La Russa M, Qi LS. 2016. CrISPR/cas9 in genome editing and beyond. *Annual review of biochemistry*. 85:227-264. <https://doi.org/10.1146/annurev-biochem-060815-014607>