

El sistema CRISPR-CAS.

Vargas Hernández Joel Alberto,² Hinojosa Juárez Araceli Consuelo,^{1,2} Mendieta Zerón Hugo,²

Centro Estatal de Vigilancia Epidemiológica y Control de Enfermedades, Regulación Sanitaria.¹
Universidad Autónoma del Estado de México.²

Resumen

La metodología conocida como CRISPR-Cas, se ha convertido en un instrumento que hace posible cambios dirigidos a secuencias genéticas específicas. La relevancia del sistema CRISPR-Cas como una herramienta para hacer edición genómica, recae en su capacidad para guiar y marcar dianas en el ADN, no mediante interacciones ADN-proteína, sino mediante un apareamiento ADN-ARN que resulta más específico. A través del desarrollo de la técnica de uso y la optimización y depuración de la misma, se prevé que en un futuro, sea posible la aplicación rutinaria de la tecnología CRISPR-Cas tanto en investigación básica, como en biotecnología y medicina, entre otros campos.

Palabras clave: *Modificación genómica; CRISPR-Cas; Cas9.*

Antecedentes

Los humanos hemos sido ingenieros genéticos desde hace cientos o quizá miles de años, sin siquiera habernos dado cuenta de lo que hacíamos al respecto. Hemos llevado a cabo modificaciones en plantas y animales al dirigir ciertas hibridaciones selectivas.

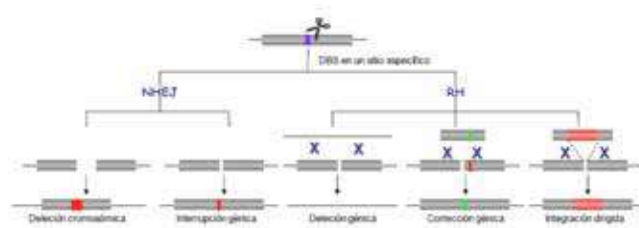
Sin embargo, los cambios provocados por el avance en el estudio de la biología molecular y, en específico, en el campo de la genómica y la genética; no tienen precedentes: por su magnitud, relevancia e impacto, tanto en el ámbito científico como en el tecnológico. Desde mediados del siglo XIX, ya se estaba gestando un cambio drástico en la biología con las contribuciones notables del inglés Charles Darwin. Empero, un gran salto ocurrió cuando Watson y Crick postularon su modelo de doble hélice del ADN en la década de los cincuenta. Además, recientemente, se ha podido conocer toda la secuenciación del ADN humano gracias al proyecto del genoma humano.

De igual forma, en los años setenta se dio inicio a una nueva era en la biología con la implementación de la tecnología del ADN recombinante,³ se propulsó que la disciplina biológica hacia la vanguardia en investigación, dando pie al nacimiento de la ingeniería genética. A partir de este momento, fue posible manipular el material genético, así como poder estudiar de forma más detallada la regulación y expresión de los genes.

Conforme fueron desarrollándose las investigaciones en el campo de la ingeniería genética, aparecieron nuevas técnicas mucho más eficaces. La mayoría de estas metodologías se basaron en enzimas nucleasas (meganucleasas, nucleasas con dedos de zinc, entre otras). Estas herramientas permitieron solucionar el problema de la inserción génica aleatoria pudiendo hacer modificaciones específicas del genoma.

Las técnicas que utilizan las enzimas nucleasas se basan en la introducción de una rotura en la doble cadena (DSB, Double Strand Break) en un lugar específico del ADN, que después de ser reparadas por una "Recombinación homóloga" (RH) o por una "Reparación por unión de extremos no homólogos" (NHEJ), se logran hacer cambios dirigidos en la secuenciación del código genético.⁵

Figura 1. Resultados generados en la secuencia genómica tras la rotura de la doble hebra.





La necesidad de mejorar los procesos de ingeniería genética ha impulsado la búsqueda de nuevas metodologías con la capacidad de modificar o controlar los genes. Avances recientes, han permitido la incorporación de herramientas basadas en el sistema CRISPR-Cas (Clustered regularly interspaced short palindromic repeats) como una nueva herramienta para introducir alteraciones en el genoma.

El Sistema CRISPR-CAS.

El sistema CRISPR-Cas es un método de ingeniería genética que hace posible la modificación de prácticamente cualquier gen. Genes específicos pueden ser activados, desactivados o simplemente ser editados. A pesar de que esta metodología ha sido utilizada principalmente en investigación básica, las aplicaciones potenciales son enormes: tales como medicina genómica en humanos; así como modificaciones en ciertos organismos para que funcionen como productores de medicamentos, combustibles biológicos, etc. Es una tecnología que está resultando más simple y barata que cualquier otra disponible hasta el momento.

A pesar de lo innovadora que resulta esta técnica, el mecanismo en el cual se basa proviene de hace muchos millones de años, siendo parte del sistema inmunológico de los microorganismos, a través de una especie de auto inoculación. Estos van incorporando a su propia secuencia genética fragmentos de código provenientes de virus invasores, siendo estos como una especie de memoria de las agresiones sufridas. Así los microorganismos utilizan estas secuenciaciones con la finalidad de identificar al virus que pretende atacarlos. Posteriormente despliegan alguna enzima, que corta la secuencia genómica en sitios específicos del virus invasor reduciendo el ataque antes de que pueda matar a la célula.

A finales de la década de los ochentas, Ishino y colaboradores descubrieron en una cepa del microorganismo *Escherichia coli*, unas repeticiones regularmente espaciadas: 29 nucleótidos espaciados por 32 nucleótidos dentro de un fragmento de ADN; diferentes a cualquier otro tipo recurrente conocido.²

Figura 2. Secuencia de una agrupación CRISPR del genoma de *E. coli*.

```

aacgggttat atggtggttt atccccctg gcccgggga ctcgacagaa cggcctcagt
agtctegtca ggctccggtt tateccccct gcccgggga acacctgttt tcgcaaatct
atggactatt gctattcggg tttatccccg tggccgggg aacacggggc cacggaatac
aaagccgtgt atctgctcgg tttatccccg ctggccgggg gaacactggc tctgcaacag
cagcaccat gaccacgtcg gtttatcccc cctggccggg ggaacacgaa atgctgggta
gcgttaatgc cgcaaacaca gtttatcccc cctggccggg ggaacacat tacgcctttt
tgcgattgcc cggtttttgc cgtttatccc cctggccgc ggaacacat ctaaacaata

```

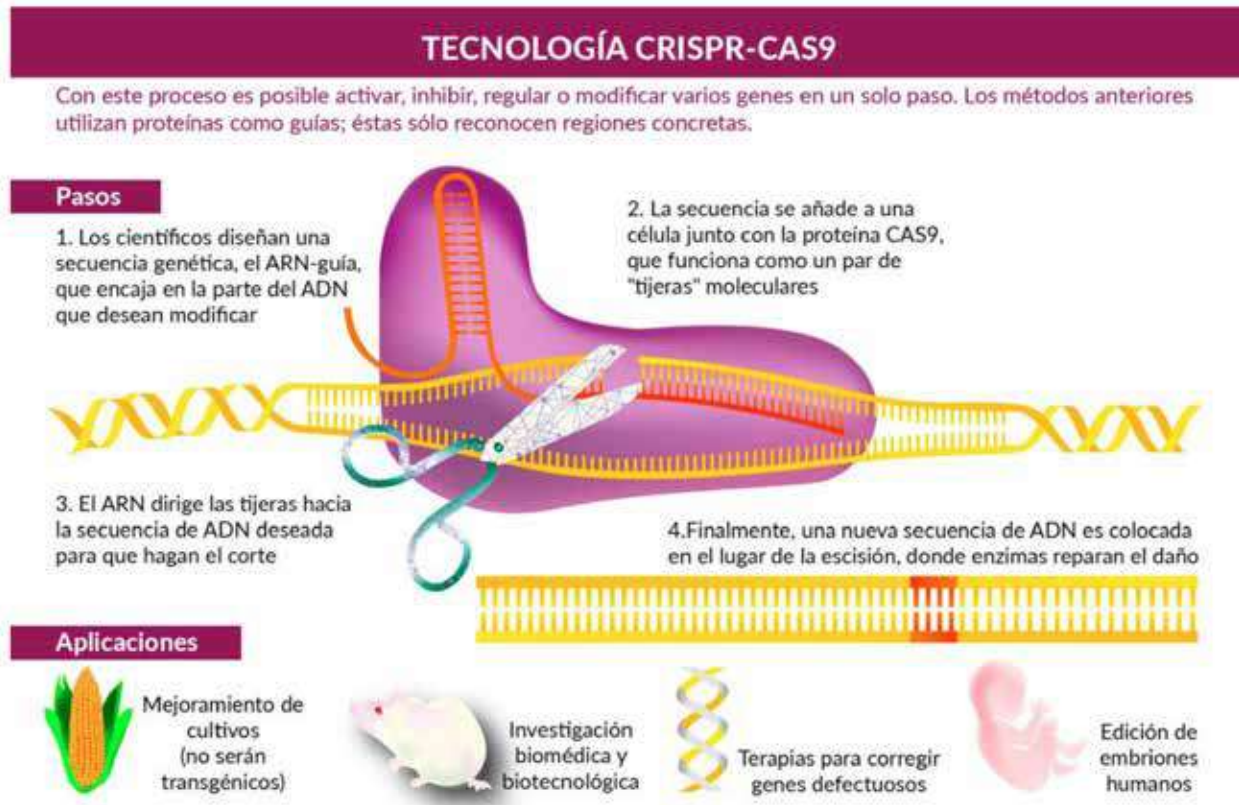
La conclusión a la que llegaron los investigadores, a partir de estos fragmentos, era que constituían la secuencia de terminación de la transcripción. También confirmaron la presencia de estas secuencias en otros microorganismos relacionados (otras cepas de *E. coli*, *Shigella dysenteriae* y *Salmonella entérica*).¹⁰

Durante los siguientes años se desarrollaron varios estudios donde se describían otras secuencias interespaciadas repetidas con una organización similar (*Mycobacterium tuberculosis*, *Haloferax mediterranei*, *Anabaena sp.*, *Methanocaldococcus jannaschii*, y *Streptococcus pyogenes*).^{7,8}

El análisis de estas secuencias, a finales de los años 90, en conjunto con las nuevas tecnologías de ingeniería genética disponibles, pudo describir una nueva familia de repeticiones denominadas como SRSR (short regularly spaced repeats) (identificando estas secuencias en un rango amplio de procariontes); para después ser renombradas como CRISPR (clustered regularly interspaced short palindromic repeats).^{4,9}



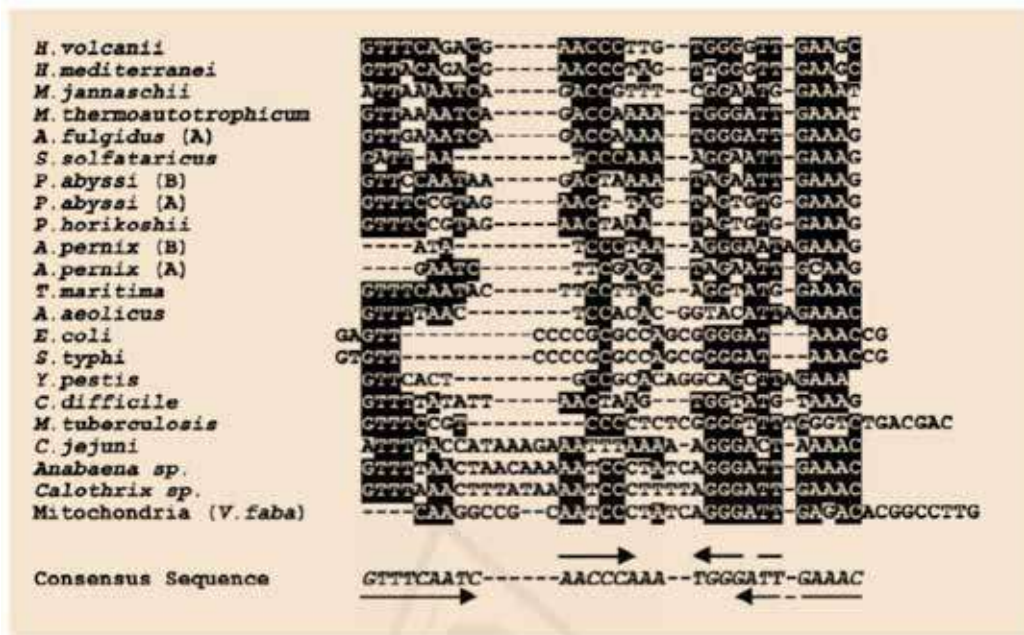
Figura 3. Tecnología CRISPR-Cas9.⁶



En esta primera aproximación, a pesar de que se habían hecho grandes progresos en la revelación de la secuencia CRISPR en múltiples genomas de microorganismos, todavía no se tenía la certeza de la funcionalidad de sus componentes. Lo que sí se pudo establecer para esta agrupación de unidades repetidas, era un carácter parcialmente palindrómico.⁹

Se fue descubriendo que las SRSR presentes en distintos loci de un mismo cromosoma podían ser prácticamente las mismas. Dentro de cada grupo, las CRISPR estaban espaciadas y los espaciadores poseían secuencias únicas en el genoma. Resultaba evidente que este espaciamiento debía tener algún significado biológico.⁹

Figura 4. Alineamiento de secuencias CRISPR de distintas especies.⁹



Además se pudo verificar que de un solo locus de CRISPR se podían transcribir múltiples ARNs. De igual forma, fueron identificados cuatro genes Cas (CRISPR-associated) localizados en regiones adyacentes a los locus de CRISPR que no estaban presentes en cepas de microorganismos carentes de los mismo; pudiendo concluir que de alguna forma los Cas estaban relacionados con alguna funcionalidad a la de CRISPR.⁴

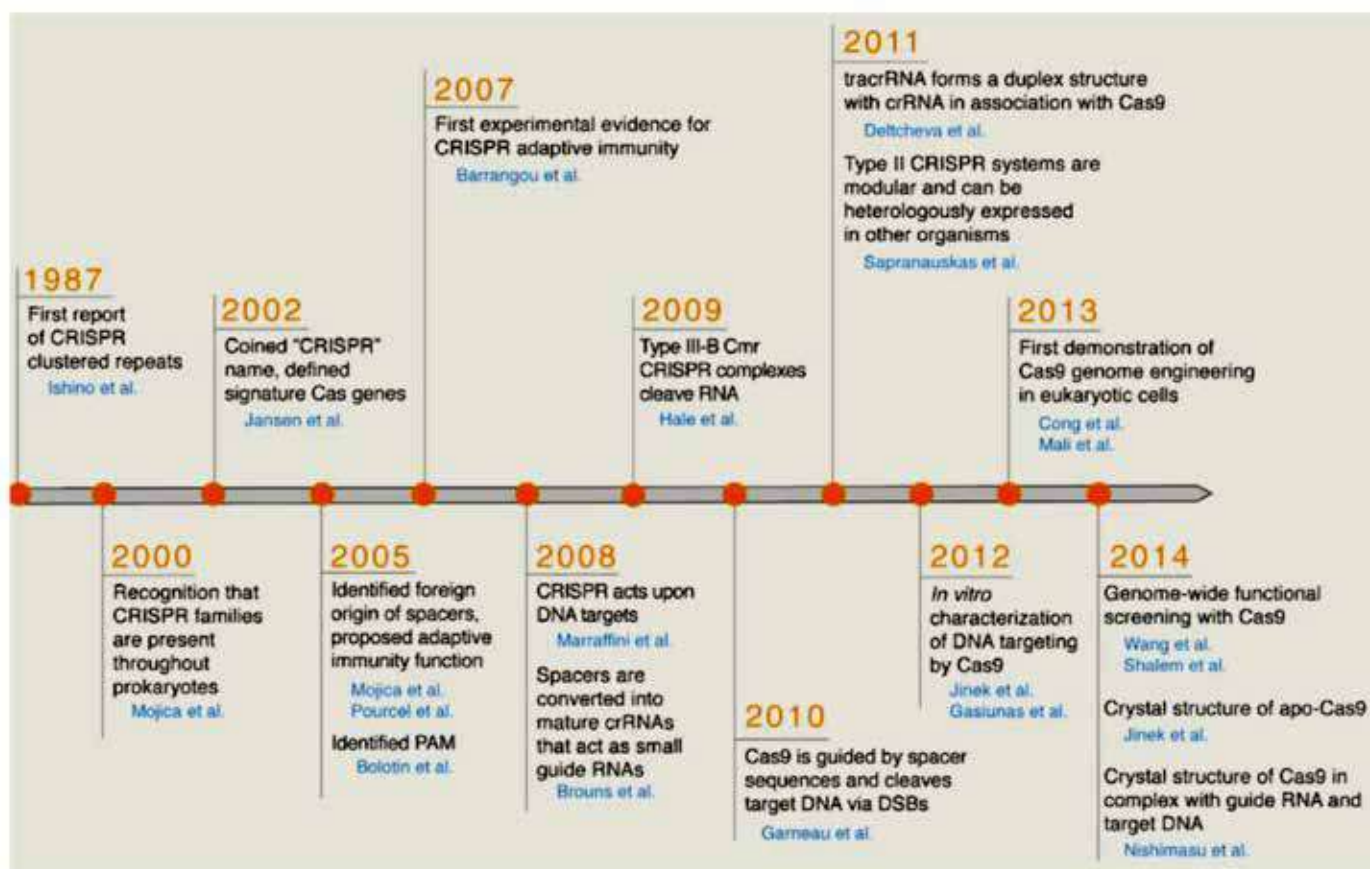
En el año 2005 se publicaron investigaciones en donde se describía lo que es la función del sistema CRISPR-Cas. Se descubrió que las secuencias espaciadoras eran muy parecidas a segmentos genéticos relacionados con el ADN de ciertos virus (que de ninguna otra forma pudieron haber sido transferidos), estableciendo la teoría de que la funcionalidad del sistema CRISPR forma parte del sistema inmune primitivo de ciertos microorganismos. Es decir, estos al poder "recordar" infecciones virales pasadas almacenadas en su ADN; podían utilizar una enzima, como si fuera unas tijeras moleculares, para cortar al virus en lugares precisos de su genoma antes de que pudiese provocar algún daño al

microorganismo. Existen una gran variedad de enzimas Cas, siendo la más utilizada en estos momentos la Cas9.

El sistema CRISPR-Cas, por lo simple de su metodología y lo barato que resulta ser, atrajo la atención de forma inmediata. A partir del 2011, es que empezaron a proponerse diferentes soluciones basadas en este sistema, aplicaciones relacionadas principalmente con la ingeniería genética. Se ha utilizado para modificar los genomas de muchas especies de plantas y animales, así como de células humanas in vitro. De hecho, el sistema CRISPR-Cas puede ser una alternativa viable para las terapias génicas, es decir, encontrar una cura real a las enfermedades genéticas.

Entre otras aplicaciones, se podrán activar o desactivar ciertos genes, para poder estudiar la expresión o no de estos en sus respectivas proteínas. También se podrán explorar las alternativas que pueden proporcionar aquellas partes del código que tiene secuencias no codificantes. La amplitud de posibilidades para este método se vuelven prácticamente infinitas.

Figura 5. Esquema de la cronología de los avances realizados en referencia al sistema CRISPR-Cas, desde su descubrimiento en 1987 hasta su aplicación como herramienta de edición y manipulación genómica.¹





Referencias bibliográficas

1. Hsu, P.D., Lander, E.S. y Zhang, F., 2014. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*, 157, pp.1262-1278.
2. 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.5429-5433.
3. Jackson, D. a, Symons, R.H. & Berg, P., 1972. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 69(10), pp.2904-2909.
4. Jansen, R. et al., 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 43(6), pp.1565-1575.
5. Jensen NM, Dalsgaard T, Jakobsen M, Nielsen RR, Sorensen CB, Bolund L, and Jensen TG: An update on targeted gene repair in mammalian cells: methods and mechanisms. *J Biomed Sci*, 18:10.
6. <http://www.invdes.com.mx/>
7. Mojica, F.J., Juez, G. y Rodríguez-Valera, F., 1993. Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites. *Molecular microbiology*, 9(3), pp.613-621.
8. Mojica, F.J.M. et al., 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), pp.85-93.
9. Mojica, F.J. et al., 2000. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology*, 36(1), pp.244-246.
10. Nakata, A., Amemura, M. y Makino, K., 1989. Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. *Journal of Bacteriology*, 171(6), pp.3553-3556.
11. Tafurt Y, Marín MA. Principales mecanismos de reparación de daños en la molécula de ADN. *Revista Biosalud* 2014; 13(2): 95-110.