

## ¿CÓMO FUNCIONA?

## CRISPR/Cas9: La enésima revolución

José Joaquín Serrano

Alumno del Máster en Biología Celular y Molecular de la Universidad de Málaga  
[conguino@hotmail.com](mailto:conguino@hotmail.com)

207

Al comenzar a escribir este artículo me vienen a la mente conversaciones recientes con algunos buenos amigos; conversaciones que giraban alrededor de si la ciencia avanzaba a saltos o, por el contrario, lo hacía de manera gradual. La respuesta podría ser que siguió un proceso gradual hasta hace algo más de un siglo, momento en el que experimenta un salto espectacular, sobre todo en el campo de la física, con la revolución relativista y cuántica. Además de en el ámbito de la Física, en Biología ya se venía fraguando un cambio con olor a revolución desde mediados del siglo XIX (con la figura notable de Charles Darwin), pero no es hasta los años 70 del siglo pasado cuando tiene lugar el “boom” de la Biología, unos años después de que Watson y Crick propusieran el modelo de doble hélice del DNA. El advenimiento de la tecnología del DNA recombinante propulsó la disciplina biológica hacia la vanguardia en investigación. Desde entonces, una gran cantidad de metodologías, que mejoraban las condiciones para realizar estudios de investigación básica en Biología, han ido apareciendo, permitiendo a los investigadores trabajar, cada vez de manera más frecuente, dentro del contexto “organismo”. Lo que aquí quiero presentar hoy es una de las últimas metodologías en aparecer y revolucionar, con su irrupción, la forma de realizar investigación en Biología. Dividiré en varios puntos el artículo, comenzando con una breve introducción sobre sistemas parecidos que se han venido usando hasta ahora para modificar la expresión génica (el proceso de activación o inactivación de genes y/o sus productos, sean proteínas o RNA de diversos tipos), explicando a continuación en que consiste este sistema y acabando con las aplicaciones del mismo en la investigación más actual, barajando ventajas e inconvenientes.

La investigación biológica ha estado siempre lastrada por las limitaciones de la traslación de los resultados obtenidos *in vitro* (en el laboratorio, bajo condiciones experimentales definidas por el investigador, de forma que se parezcan lo más posible al organismo) hacia el modelo *in vivo* (el propio organismo). En el caso de la expresión génica, la problemática aumenta debido a la gran cantidad de factores que intervienen en la misma. El método para abordar el estudio *in vivo* de la expresión génica de un determinado gen (o genes) se basaba en la realización de los llamados “knockout”, modelos ani-

males (por lo general, ratones) en los cuales se había eliminado o modificado un gen concreto mediante un proceso de recombinación entre la secuencia del gen endógena y una secuencia exógena, complementaria al gen, normalmente presente en un vector de expresión (plásmidos, por ejemplo) (Capecchi, 1989). Sin embargo, este sistema tenía el poderoso inconveniente de que, en muchos casos, mutaba genes que eran indispensables para el desarrollo del animal (también la frecuencia de recombinación es baja en estas tecnologías). La necesidad de mejorar los procesos de modificación génica empujó la búsqueda de alternativas con las que poder controlar los genes que queremos eliminar, o simplemente mutar. El siguiente paso fue la generación de mutantes mediante el sistema Cre-LoxP, una metodología experimental que hace uso de una enzima con actividad recombinasa capaz de reconocer una secuencia de unos 34 pb, los denominados sitios LoxP, catalizando una recombinación intra o intermolecular entre dos sitios de este tipo, escindiendo la secuencia que se encuentra entre ambos, generando una molécula de DNA circular covalentemente cerrada que será degradada. Mediante ingeniería genética se pueden ubicar secuencias LoxP en el sitio del genoma que se quiera eliminar y generar el mutante adecuado. El siguiente paso en la modificación génica vino desde el mundo vegetal, curiosamente, a finales de los 90 del siglo XX y principios del siglo XXI. Estoy hablando del sistema de interferencia por RNA (RNAi, *RNA interference*, del inglés). Este es un sistema basado en el reconocimiento de secuencias de RNA bicatenario que sufren un procesamiento enzimático, generando unos RNA de pequeño tamaño que permiten a un complejo macromolecular, que los utiliza de guía, reconocer secuencias complementarias a ellos y degradarlas, provocando lo que se conoce como silenciamiento de la expresión génica (inactivar el proceso que lleva a que un gen de lugar a un elemento funcional). Este método ha resultado de mucha utilidad para conseguir silenciar familias multigénicas, de forma que se evitan los efectos compensatorios que los distintos miembros de dicha familia tienen sobre la función de un gen de la misma (John et al., 2003).

En la década de los 90 del siglo pasado varios investigadores habían mostrado que el uso de nucleasas para generar roturas dobles en la cadena de DNA podría estimular un pro-

ceso de edición del genoma (modificación de las secuencias de nucleótidos del DNA endógeno) a través de procesos de recombinación mediada por "recombinación homóloga directa", usando un molde que presenta homología al sitio que se quiere modificar (explicado en el párrafo anterior) o, en ausencia de dicho molde, inducir deleciones o inserciones (mutaciones que elimina o introducen, respectivamente, nucleótidos de la secuencia original del gen endógeno) a través de otro mecanismo de reparación de roturas dobles en la hebra de DNA, la "unión de extremos no homólogos" (NHEJ), generándose codones de parada prematuros o cambios en el marco abierto de lectura del gen (Hsu et al., 2014). En los últimos 8 años (aproximadamente) se han generado hasta cuatro metodologías de éste tipo, siendo las que más repercusión han tenido los sistemas ZFN (*zinc finger nucleases*, del inglés, nucleasas dedo de zinc), los TALEN (*transcription activator-like effector nucleases*, del inglés, nucleasas efectoras con actividad similar a los activadores de la transcripción) y aquella a la que vamos a dedicar más contenido en este artículo, CRISPR/Cas9 (Gaj et al., 2013). Si estas tres metodologías sirven para lo mismo ¿dónde está la diferencia entre ellas? Las dos primeras tecnologías se basan en quimeras proteicas que presentan dominios de unión al DNA y un dominio catalítico de una endonucleasa, mientras que CRISPR/Cas9 es una nucleasa por sí misma, que es guiada por un pequeño RNA (sgRNA, *small guide RNA*) hacia el DNA diana, interaccionando a través de apareamientos de Watson-Crick, apareamientos clásicos entre bases que tienen lugar en los ácidos nucleicos (Ran et al., 2013). Esto ya supone una ventaja, sobre todo a la hora de modificar nuestra diana, ya que con el sistema CRISPR bastaría con introducir un nuevo sgRNA específico para la secuencia de un gen que queramos modificar, mientras que en el caso de los otros dos métodos sería necesario reconstruir el sistema completo de nuevo (Hsu et al., 2014), algo que incrementaría los costes del proceso, probablemente. Así que, sin más dilación, vamos a centrar el resto del artículo en la presentación de la tecnología de edición genómica basada en el sistema CRISPR/Cas9.

CRISPR-Cas es un locus génico formando por un operón, que codifica las proteínas Cas, y una serie de secuencias repetidas idénticas separadas por otras secuencias conocidas como espaciadoras, que son reconocidas por moléculas de DNA intruso (Chylinski et al., 2013). El sistema, presente en bacterias y arqueas, actúa de manera análoga a una inmunidad adaptativa frente a elementos genéticos móviles extraños, ya sean plásmidos (moléculas de DNA circulares con capacidad autorreplicante) o genoma de fagos (virus con capacidad para infectar bacterias). Su descubrimiento tuvo lugar en 1987, pero no fue hasta la segunda mitad de la primera década del siglo XXI cuando se comenzó a dilucidar su funcionamiento (Hsu et al., 2014). Dicho funcionamiento puede dividirse en 3 fases. En una primera fase, conocida como fase de adaptación, una parte del ácido nucleico extraño se incorpora a la zona de

espaciadores del locus, de forma que la infección queda algo así como memorizada para posibles futuros ataques. Esta incorporación, un proceso de transferencia horizontal (transmisión de información genética entre organismos pertenecientes a distintas especies o dominios), está mediado por las propias proteínas Cas, que degradan el DNA extraño, en un primer momento de forma inespecífica, para guiar su incorporación, a posteriori, como nuevos espaciadores en el locus CRISPR-Cas (Gasiunas et al., 2013). En una segunda fase se produce la transcripción de CRISPR-Cas, generándose un precursor denominado CRISPR-RNA o pre-crRNA, que sufre un procesamiento que va a dar lugar a crRNAs de pequeño tamaño que son complementarios a la secuencia de DNA foráneo, de forma que en la tercera fase (conocida como de interferencia) las proteínas Cas, usando como guía a los crRNAs, detectan las secuencias intrusas y las degradan (Chylinski et al., 2013). Se conocen hasta 3 tipos de sistemas CRISPR-Cas, siendo el que tiene más interés, por sus aplicaciones biotecnológicas, el CRISPR-Cas de tipo II. Este sistema tiene una serie de peculiaridades con respecto a los otros dos. La primera diferencia la encontramos en los elementos que llevan a cabo la degradación de los ácidos nucleicos, ya que en los sistemas I y III se trata de un complejo formado por varias proteínas, mientras que en el caso del sistema tipo II solo es necesaria una proteína, Cas. Además de esto, el proceso de maduración del pre-crRNA es diferente entre los tipos I y III y el tipo II. En el caso del tipo II se requiere de la participación de otra molécula de RNA pequeña, no codificante, que presenta complementariedad de bases con el pre-crRNA y que se expresa desde un lugar cercano al locus CRISPR-Cas. Este RNA, denominado tracrRNA, forma dúplex con el pre-crRNA que son reconocidos y cortados por la endoribonucleasa III en presencia de Cas9, que parece ser la única enzima implicada en el corte del DNA diana (Garneau et al., 2010). Los crRNA que se generan sufren otro procesamiento que genera los crRNAs pequeños, en un mecanismo que aún no se comprende bien (Chylinski et al., 2013). Cada uno de estos crRNA va a permanecer unido al correspondiente traCRNA, y el dúplex formado por ambos se asociará con Cas9 (presente solo en bacterias, no en arqueas), formando un complejo ternario que será el encargado de reconocer y eliminar las secuencias de DNA foráneo que sean complementarias a la del crRNA (Sampson & Weiss 2013). El principal requerimiento para que este sistema funcione en bacterias y arqueas es la presencia de una pequeña secuencia (en el caso de *Streptococcus pyogenes* es 5'-NGG-3') conocida como PAM (del inglés, *protospacer adjacent motif*) que se encuentra adyacente al locus CRISPR en dirección 3' y que parece esencial para que las proteínas Cas presenten especificidad por la secuencia de los crRNA (Ran et al., 2013a). Se ha observado que la carencia de dichos PAM dentro de la secuencia de repeticiones directas de CRISPR parece impedir que el propio sistema CRISPR-Cas se ataque a sí mismo (Deveau et al., 2008). Además, los PAM muestran especificidad para cada ortólogo de Cas9 que se conoce (Hsu et al., 2014).

Todo lo comentado en el párrafo anterior ha llevado a varios grupos a plantearse la posibilidad de usar este sistema como método para conseguir una eficiente edición del genoma, permitiendo al investigador adquirir una mayor fiabilidad a la hora de modificar determinados genes. Usando como base el sistema de tipo II, los investigadores han modificado el dúplex crRNA:tracrRNA sintetizando un RNA diana de doble cadena que combina los rasgos de ambos RNAs en un solo RNA, lo que más arriba habíamos denominado sgRNA (Sampson & Weiss, 2013). De esta forma no hace falta ni el procesamiento por la RNAasa III ni la presencia del tracrRNA. Sin embargo, Cas9, al igual que otras nucleasas, posee el inconveniente de cortar en secuencias del genoma que tengan una cierta similitud con la secuencia diana que porta el sgRNA, lo que podría llevar a resultados no deseados. Además, el doble corte en la hebra de DNA que produce Cas9, debido a que posee dos centros con actividad catalítica, va a promover la reparación por NEHJ, con lo que son frecuentes las deleciones e inserciones. Con objeto de minimizar este hándicap, en un primer momento lo que se hizo fue mutar Cas9 para que solo poseyera un sitio catalítico. Esta nueva versión de la enzima se denomina Cas9 D10A e introduce un solo corte en la hebra de DNA. El proceso va a requerir de dos sgRNA, uno para cada cadena del DNA diana, lo que incrementa la especificidad del corte gracias a que existe ahora un mayor número de bases que necesitan ser reconocidas de manera específica; además, los nuevos cortes serán reparados por otro sistema molecular que presenta una alta fiabilidad, la reparación por escisión de bases (Ran et al., 2013a, 2013b). Para poder facilitar un doble corte eficiente con este sistema, los sgRNAs deben de sintetizarse de forma que al producirse se originen extremos salientes en 5' (que sobresalen de la cadena de DNA). Los sitios de corte de la enzimas Cas9 deben de estar separados al menos 20 pares de bases para que se generen dichos extremos salientes (Ran et al., 2013a). Además de poder editar el genoma, este sistema permite también ejercer un control sobre la propia expresión génica, sin llevar a cabo una alteración permanente del genoma celular. Esto se puede conseguir inactivando los dos dominios proteicos con actividad nucleasa que posee Cas9 (en esta forma la enzima se denomina dCas9), de manera que cuando el sgRNA guía a Cas9 hacia la secuencia diana específica, en lugar de procesar el segmento génico correspondiente lo que hace es unirse al mismo y bloquear la unión de la RNA polimerasa, de forma que no se puede producir la elongación de la transcripción (Sampson & Weiss, 2013). Esta forma sin actividad nucleasa puede aprovecharse de igual modo para justamente el proceso contrario, activar la transcripción mediante la fusión de la dCas9 a un activador transcripcional (Sampson & Weiss, 2013).

Las aplicaciones del sistema CRISPR/Cas9 que se están desarrollando son muchas en muy poco tiempo y su número no deja de crecer. El desarrollo de esta tecnología está permitiendo, por ejemplo, que se puedan manipular genéticamente

especies que hasta ahora habían sido esquivas a las técnicas de manipulación genética más al uso (esto podría aumentar el rango de organismos modelo). Supone un avance, además, en tanto en cuanto es viable pensar que podría permitir la modificación o generación de líneas celulares determinadas, además de la generación de nuevos mutantes in vivo. Entre las aplicaciones más prometedoras se encuentra la de poder recapitular las mutaciones genéticas encontradas en grupos de pacientes afectados por una enfermedad, para de esta forma poder determinar con mayor seguridad la causa de las mismas (Hsu et al., 2014). La posibilidad que ofrece esta tecnología de inactivar ambos alelos en un mismo gen abre una nueva vía en la investigación de enfermedades genéticas: usando Cas9 se puede estudiar, por ejemplo, el efecto de una variante génica en una enfermedad que muestra una fuerte asociación con haplotipos identificados por estudios de asociación genómica a gran escala en humanos (Wang et al., 2014). También puede utilizarse para modificar varios genes de manera individual (Wang et al., 2013) o para estudiar el efecto que tiene la edición de dichos genes sobre un determinado tipo celular (Hsu et al., 2014). Entre las enfermedades donde esta nueva metodología permitiría obtener mejores resultados se encuentra el cáncer. Ya se están realizando ensayos donde se observa cómo utilizando este sistema se puede fenocopiar (copiar el fenotipo) el desarrollo de tumores en ratones, así como comprobar el efecto sinérgico de dos mutaciones claves en la formación de un cáncer hepático, también en ratones (Xue et al., 2014). Dado que Cas9 solo requiere del PAM para que su funcionamiento sea óptimo, se pueden realizar construcciones in vitro de plásmidos con casetes que incluyan varios sgRNA que reconozcan diferentes secuencias en el genoma, con lo que se puede modificar, al mismo tiempo, varios genes que previamente habían mostrado una fuerte correlación con algún fenotipo propio de alguna enfermedad (Sakuma et al., 2014; Hsu et al., 2014). Del mismo modo que se pueden dirigir sgRNAs para editar secuencias génicas, se puede usar esta nueva tecnología para la edición de secuencias reguladoras o no-codificantes (Hsu et al., 2014), con vistas a poder descubrir nuevas secuencias de este tipo o aclarar la función de alguna de ellas. Aprovechando la capacidad que describimos antes para unir Cas9 a reguladores de la transcripción, podría fusionarse a modificadores del estado epigenético y de esta forma avanzar en el conocimiento de cómo influye la metilación o los estados de la cromatina en el desarrollo de determinadas enfermedades (Hsu et al., 2014).

A pesar de las numerosas ventajas que este sistema ofrece, aún se encuentra en período de prueba. Ya hemos visto que Cas9 presenta una cierta variabilidad en el corte, lo que provoca que, en sus primeros intentos por hacer uso de esta tecnología, los investigadores observaran la existencia de una relativa diversidad de secuencias modificadas -secuencias con cierta similitud con la secuencia del sgRNA o del crRNA usado para guiar a Cas9. También queda por establecer cómo afecta

la accesibilidad a la cromatina a este sistema, aunque ya se ha observado que dCas9 se une con relativa eficacia a la cromatina inaccesible (Perez-Pinera et al., 2013). Para el final he dejado el que, a priori, supone el principal problema: la eficacia del propio sistema CRISPR-Cas9. Como con otros métodos de transgénesis, lo más utilizado hasta ahora era el uso de vectores virales para la transfección (introducir genes de interés, que no tienen que ser del organismo con el que se trabaja, en dicho organismo). Sin embargo, la eficacia en la transformación es relativamente baja (3% señalan Xue et al. en su artículo), con lo cual es necesaria una mejora en dicho aspecto. Recientemente se ha llevado a cabo una investigación en la que han conseguido introducir (sin modificar) Cas9 unido al sgRNA, aumentando la eficacia de modificación hasta en un 80% y aumentando la especificidad de dicha modificación hasta 10 veces si se compara con la transfección por plásmidos (Zuris et al., 2014). Esto lo llevan a cabo aprovechando la

constelación de cargas que poseen ambas moléculas, tanto de Cas9 como del sgRNA. Los mismos autores habían probado con anterioridad tecnologías para la introducción de ácidos nucleicos basada en la conjugación o fusión con moléculas catiónicas (como péptidos sin estructura o proteínas con muchas cargas positivas) que facilitan la endocitosis (Cronican et al., 2010; Thompson et al., 2012). Pero eso ya es otra historia...

Cabe decir, por último, que durante la elaboración de esta introducción al sistema CRISPR/Cas9, la cantidad de bibliografía ha ido aumentando, con lo que puede que haya más información al respecto de la que aquí se cubre. Sin embargo, la idea principal de cómo funciona el sistema creo que queda suficientemente cubierta.

#### Bibliografía citada:

- \* Capecchi, M.R. (1989). Altering the genome by homologous recombination. *Science* 244, 1288–1292.
- \* Chylinski, K., Le Rhun, A., & Charpentier, E. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biology*, 10(5), 726-737.
- \* Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339(6121), 819-823
- \* Cronican, J. J., Thompson, D. B., Beier, K. T., McNaughton, B. R., Cepko, C. L., & Liu, D. R. (2010). Potent delivery of functional proteins into Mammalian cells in vitro and in vivo using a supercharged protein. *ACS Chemical Biology*, 5(8), 747-752.
- \* Deveau, H., Barrangou, R., Garneau, J.E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology* 190, 1390–1400
- \* Gaj, T., Gersbach, C. A., & Barbas III, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*, 31(7), 397-405.
- \* Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadañ, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71.
- \* Gasiunas, G., Sinkunas, T., & Siksnys, V. (2014). Molecular mechanisms of CRISPR-mediated microbial immunity. *Cellular and Molecular Life Sciences*, 71(3), 449-465.
- \* Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*, 157(6), 1262-1278.
- \* John, M., Geick, A., Hadwiger, P., Vornlocher, H. P., & Heidenreich, O. (2003). Gene silencing by RNAi in mammalian cells. *Current Protocols in Molecular Biology*, 26-2.
- \* Li, W., Teng, F., Li, T., & Zhou, Q. (2013). Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nature Biotechnology*, 31(8), 684-686
- \* Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nature Methods* 10, 973–976.
- \* Ran, F., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E. & Zhang, F. (2013a). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, 154(6), 1380-1389.
- \* Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013b). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), 2281-2308.
- \* Sakuma, T., Nishikawa, A., Kume, S., Chayama, K., & Yamamoto, T. (2014). Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Scientific Reports*, 4, 5400.
- \* Sampson, T. R., & Weiss, D. S. (2014). Exploiting CRISPR/Cas systems for biotechnology. *Bioessays*, 36(1), 34-38.
- \* Thompson, D. B., Villaseñor, R., Dorr, B. M., Zerial, M., & Liu, D. R. (2012). Cellular uptake mechanisms and endosomal trafficking of supercharged proteins. *Chemistry & Biology*, 19(7), 831-843.
- \* Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F., & Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, 153(4), 910-918.
- \* Wang, T., Wei, J. J., Sabatini, D. M., & Lander, E. S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, 343(6166), 80-84.
- \* Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N. S., & Jacks, T. (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature*, 514, 380-396.
- \* Zuris, J. A., Thompson, D. B., Shu, Y., Guilinger, J. P., Bessen, J. L., Hu, J. H., & Liu, D. R. (2014). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nature Biotechnology*, 33, 73-80