

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

**ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA
AGRONÒMICA I DEL MEDI NATURAL**



**Expresión en plantas de un sistema de edición genómica
CRISPR mediante vectores virales**

**Grado en Biotecnología
Curso 2016-2017**

Autora:

Beatriz Gayubas Balaguer

Director:

José Antonio Daròs Arnau

Tutor:

Carmelo López del Rincón



Valencia, junio de 2017

Expresión en plantas de un sistema de edición genómica CRISPR mediante vectores virales

Autora: Dña. Beatriz Gayubas Balaguer.

Localidad y fecha: Valencia, junio de 2017

Director: José Antonio Daròs Arnau

Tutor: Carmelo López del Rincón

Tipo de Licencia: Licencia Creative Commons. “Reconocimiento no Comercial Sin Obra Derivada”

Resumen:

La edición dirigida de genomas en plantas va a revolucionar en los próximos años el campo de la mejora de las variedades vegetales y la biotecnología vegetal. De entre las diferentes técnicas de edición de genomas, el sistema CRISPR-Cas ha sido la que ha destacado notablemente. En origen, CRISPR-Cas forma parte del sistema inmune adaptativo de bacterias y arqueas y les permite atacar a cualquier material genético invasor. Existen dos clases de sistemas CRISPR-Cas y, dentro de cada una de ellas, distintos tipos. De entre ellos, el sistema de tipo II CRISPR-Cas9 ha sido el que se ha utilizado para su adaptación a la edición de genomas en células de mamífero y en plantas. Sin embargo, últimamente ha surgido el sistema de tipo V CRISPR-Cpf1 para la edición de genomas, que tiene algunas características que lo hacen más favorable, como la utilización de un único RNA para la edición. De entre las diferentes proteínas de la familia de la Cpf1, las procedentes de *Acidaminococcus sp.* (AsCpf1) y *Lachnospiraceae bacterium* (LbCpf1) han demostrado poseer una actividad editora de genes comparable a la de la Cas9 en células de mamífero y en plantas. En este trabajo se ha explorado la posibilidad de expresar un sistema de edición CRISPR-Cpf1 en plantas mediante vectores virales. Para ello se construyó un vector viral derivado del virus del grabado del tabaco (*Tobacco etch virus*, TEV) al cual se le insertó el gen de la nucleasa LbCpf1. La construcción del vector se realizó a partir del plásmido pGTEVΔEE, diseñado con el objetivo de crear un vector viral derivado del TEV estable en *Escherichia coli* y apto para agroinocular plantas. Además, se sabe que a este vector viral se le puede eliminar el gen de la RNA polimerasa viral (N1b) y su función puede ser complementada *in trans* en células que expresen dicha proteína. Se comprobó que el vector derivado del TEV en el que N1b se reemplazó por LbCpf1 es infeccioso tras agroinocularlo en plantas de *Nicotiana tabacum* que expresan N1b. Sin embargo, los análisis de la progenie viral en las plantas infectadas indicaron que el virus pierde el inserto de la LbCpf1. No sabemos si el virus expresará la nucleasa un tiempo suficiente para desarrollar su labor de edición genómica. También se construyeron cuatro vectores virales derivados del virus X de la patata (*Potato virus X*, PVX) a los cuales se les insertaron diferentes construcciones con RNAs CRISPR (crRNAs) para modificar el gen de la fitoeno desaturasa (PDS). Los insertos se componían de las diferentes combinaciones de las repeticiones directas específicas para las nucleasas AsCpf1 o LbCpf1 con las secuencias guía para modificar la PDS de *N. tabacum* y *Nicotiana benthamiana*. Mediante agroinoculación de plantas de *N. benthamiana*, se comprobó que el vector viral que expresa el crRNA correspondiente a LbCpf1 en esta especie es infeccioso. Cuando se dispongan de plantas de *N. benthamiana* que expresen N1b, cuya construcción está en curso, estos vectores virales se utilizarán para ensayar si es posible editar el genoma de la planta.

Palabras clave:

CRISPR, virus de plantas, vectores virales, edición genoma

Abstract:

Genome editing in plants will revolutionize the field of plant breeding and plant biotechnology in the coming years. Among the different genome editing techniques, the CRISPR-Cas system has stood out so far. In origin, CRISPR-Cas is part of the adaptive immune system of bacteria and archaea and allows them to attack invading genetic material. There are two types of CRISPR-Cas systems and, within each of them, different types. Among them, the CRISPR-Cas9 type II system has been adapted to genome editing in mammalian cells and plants. However, more recently the CRISPR-Cpf1 type V system has emerged, with some advantageous features, such as the need of a single RNA. Among the different proteins of the Cpf1 family, those from *Acidaminococcus sp.* (AsCpf1) and *Lachnospiraceae bacterium* (LbCpf1) have comparable gene editing activity to that of Cas9 in mammalian cells and plants. In this work, we have explored the possibility of expressing a CRISPR-Cpf1 editing system in plants using viral vectors. We constructed a viral vector derived from the *Tobacco etch virus* (TEV) in which the nuclease gene LbCpf1 was inserted. The vector was made from the plasmid pGTEV Δ EE, designed with the objective of creating a TEV-derived viral vector stable in *Escherichia coli* and suitable for plant agroinoculation. In addition, it is known that from this viral vector, the viral RNA polymerase (NIB) gene can be deleted while the function can be complemented *in trans* in cells expressing this protein. The TEV-derived vector in which NIB was replaced by LbCpf1 was infectious after agroinoculation of NIB-expressing *Nicotianatabacum* plants. However, the analysis of the viral progeny in the infected plants indicated that the virus loses the LbCpf1 insert. We do not know if the virus will express the nuclease long enough to develop its genomic editing work. Four *Potato virus X* (PVX)-derived virus vectors were also built and different constructs were inserted inside with CRISPR-RNAs (crRNAs) to modify the phytoene desaturase (PDS) gene. The inserts were composed of the different combinations of the direct repeats specific for the nucleases AsCpf1 or LbCpf1 with the guide sequences to modify the PDS of *N. tabacum* and *Nicotiana benthamiana*. By agroinoculation of *N. benthamiana* plants, we confirmed that the viral vector that expresses the crRNA corresponding to LbCpf1 is infectious in this species. When *N. benthamiana* plants expressing NIB, which are under construction, are available, these viral vectors will be used to test whether it is possible to edit the genome of the plant.

Keywords:

CRISPR, plant viruses, viral vectors, genome edition

AGRADECIMIENTOS

A José Antonio, por invitarme a formar parte de su laboratorio, su entrega hacia este proyecto y sus consejos.

A mis compañeros de laboratorio, en especial a Danielle y a Miguel, por su ayuda incondicional y sus ánimos.

A Carmelo por ofrecerme su ayuda.

A mis compañeras de piso, por su apoyo y por los desahogos después de cenar.

A mi familia y amigos, por hacer siempre todo más llevadero.

ÍNDICE

1. INTRODUCCIÓN.....	1
1.1. Edición de genomas en plantas.....	1
1.2. El sistema CRISPR-Cas.....	3
1.3. El sistema CRISPR-Cpf1.....	4
1.4. Vectores virales.....	6
1.5. Virus del grabado del tabaco.....	7
1.6. Virus X de la patata.....	8
2. OBJETIVOS.....	10
3. MATERIALES Y MÉTODOS.....	11
3.1. Transformación de bacterias por electroporación.....	11
3.2. Purificación de plásmidos.....	11
3.3. Amplificación DNA mediante PCR.....	11
3.4. Electroforesis en gel de agarosa.....	12
3.5. Manipulación plásmidos mediante enzimas de restricción y DNA ligasa.....	12
3.6. Elución de ácidos nucleicos de un gel de agarosa.....	13
3.7. Purificación de ácidos nucleicos de una solución.....	13
3.8. Construcción de plásmidos mediante ensamblaje Gibson.....	13
3.9. Hibridación de dos cebadores.....	13
3.10. Agroinoculación de plantas.....	14
3.11. Purificación RNA de plantas infectadas.....	14
3.12. Amplificación de RNA por RT-PCR.....	14
3.13. Secuenciación DNA y análisis computacional de las secuencias.....	15
4. RESULTADOS Y DISCUSIÓN.....	16
4.1. Construcción de vectores virales derivados del TEV que expresan Cpf1.....	16
4.2. Construcción de vectores virales derivados del PVX que expresan el RNA guía.....	21
4.3. Agroinoculación de plantas.....	25
4.4. Análisis de tejido foliar infectado.....	28
5. CONCLUSIONES.....	31
6. BIBLIOGRAFÍA.....	32
7. ANEXO: SECUENCIAS.....	35

1. INTRODUCCIÓN

1.1. Edición de genomas en plantas

Los genomas de las diferentes especies de plantas han ido evolucionando a lo largo de su historia y, en el caso de las cultivadas, los cambios se han ido acelerando sobre todo a partir de su domesticación. Esto último ocurre desde que los hombres dejan de ser nómadas para ser sedentarios. Otra forma de que se produzcan cambios en el genoma son los cruzamientos espontáneos o las mutaciones. De esta forma han surgido, por ejemplo, las diferentes variantes del género *Brassica*, como el repollo, la coliflor, la col de Bruselas o el brócoli. Junto con el descubrimiento de que las plantas tienen sexo se inició la era de la mejora científica, en la que el hombre ya empezó a manipular las plantas para producir mejores variedades y comenzaron a realizarse los cruzamientos artificiales (“cruza lo mejor con lo mejor y espera lo mejor”). Pero no fue hasta el redescubrimiento de las leyes de la herencia de Mendel y la difusión de las ideas de Darwin sobre la evolución de las especies a principios del siglo XX cuando comenzó la mejora genética y con ella los primeros pasos en manipulación de genomas. Así, empezaron a obtenerse nuevas variedades por hibridación sexual seguida de cuidadosa selección, bien entre especies del mismo género o bien entre especies de distintos géneros; por ejemplo, el triticale, híbrido sexual del trigo y el centeno. También es posible obtener plantas genéticamente modificadas por medio de mutaciones inducidas provocadas por un agente exógeno como radiaciones ionizantes, radiaciones no ionizantes o agentes alquilantes. Además, estas variedades modificadas por mutaciones inducidas no son consideradas organismos genéticamente modificados (OGMs).

Tras la puesta a punto de las técnicas de cultivo *in vitro* de plantas surgieron otras estrategias de modificación genética como el mejoramiento por haploides mediante cultivo de anteras o cultivo de microsporas, o el mejoramiento por cultivo de tejidos, a partir del cual pueden surgir con una probabilidad relativamente alta mutaciones de DNA, metilación del DNA, inducción de transposones, aberraciones cromosómicas, cambios en el nivel de ploidía, cambios epigenéticos, etc.

Actualmente, ya nos encontramos con la tecnología transgénica, la cual ha permitido superar la barrera del sexo o de especie. El inicio de la ingeniería genética viene con la obtención del primer OGM, que fue una bacteria *Escherichia coli* transformada. Diez años después se originó la primera planta genéticamente modificada (GM), una planta de tabaco con el gen de resistencia al antibiótico kanamicina, lograda por un equipo de científicos de la conocida multinacional Monsanto. En 1986 se emplea el primer OGM en el campo, Frostban, un spray que contenía bacterias genéticamente modificadas, a las cuales se les había introducido un gen de resistencia a las heladas. Se pulverizó sobre cultivos de fresa en California para protegerlas del daño provocado por las heladas. El primer OGM en el mercado fue el tomate FLAVR SAVR producido por Calgene. Salió al mercado en EEUU en el año 1994 y la modificación genética le permitía tener una maduración retardada.

De esta forma, la edición de genomas ha permitido introducir mejoras en variedades de interés como la de poseer una mayor productividad, mantener la producción en condiciones de alta densidad, tolerar distintos tipos de estrés tanto biótico como abiótico (estrés hídrico, estrés salino, etc.), mejorar la resistencia a enfermedades y plagas, adaptarse a suelos específicos, etc. Sin embargo, la modificación genética en plantas ha dado un importante giro al descubrirse técnicas que han permitido poner a punto la edición dirigida de genomas. Es decir, ser capaces de introducir una determinada mutación en una determinada base en un determinado gen, o incluso de eliminar o añadir un nuevo gen en una posición concreta del genoma. Muchas de estas técnicas fueron recogidas por la Comisión Europea, la cual creó en 2007 un grupo de trabajo para evaluar cuáles de estas técnicas podrían ser aceptadas por la legislación y que desde entonces se conocen como *New Plant Breeding Techniques* (NPBT). Entre las NPBT destacan las nuevas técnicas de mutagénesis, y entre ellas, las técnicas de edición de genomas basadas en nucleasas específicas (*site directed nucleases*, SDNs) (EUROPEAN COMMISSION, 2017). Las

SDNs son nucleasas que cortan DNA de doble cadena de manera específica en un sitio seleccionado del genoma gracias a que son capaces de reconocer secuencias específicas. La reparación del sitio reconocido por las nucleasas por parte de los mecanismos propios de la célula suele resultar en la inserción de mutaciones en dicho sitio diana. Si además de la SDN se introduce en la célula un DNA molde que lleve ya incorporados cambios específicos, estos serán introducidos en el sitio de corte y será posible introducir las mutaciones deseadas en sitios concretos del genoma, o introducir secuencias de DNA en el sitio diana de la nucleasa si se añaden a ambos lados de esta secuencia, secuencias homólogas a las del sitio de corte. Estas distintas funciones que se le pueden dar a las SDNs se denominan SDN1, o mutación aleatoria en un sitio predefinido; SDN2, o mutación específica en un sitio predefinido; y SDN3, o incorporación de una secuencia en un sitio predefinido (Casacuberta y Puigdomènech, SEBBM).

Los tipos más destacados de SDNs que han sido adaptados para su uso en edición de genomas en plantas son las nucleasas de dedos de zinc (ZFNs), las nucleasas efectoras activadoras de la transcripción (TALENs) y el sistema CRISPR-Cas (Endo *et al.*, 2016). Las ZFNs son proteínas quiméricas compuestas por un dominio de unión al DNA llamado dominio de dedos de zinc, que reconocen de manera específica un triplete de bases del DNA, y una nucleasa, que normalmente suele ser una nucleasa FokI (Casacuberta y Puigdomènech, SEBBM). De esta forma, las ZFNs reconocen un sitio concreto en el genoma y provocan una rotura en el DNA de doble cadena. Esta rotura puede ser reparada por mecanismos propios de la célula que introducirán mutaciones en la secuencia diana, o se puede introducir en la célula una secuencia de interés que queramos introducir en el sitio diana por recombinación homóloga (Fig. 1).

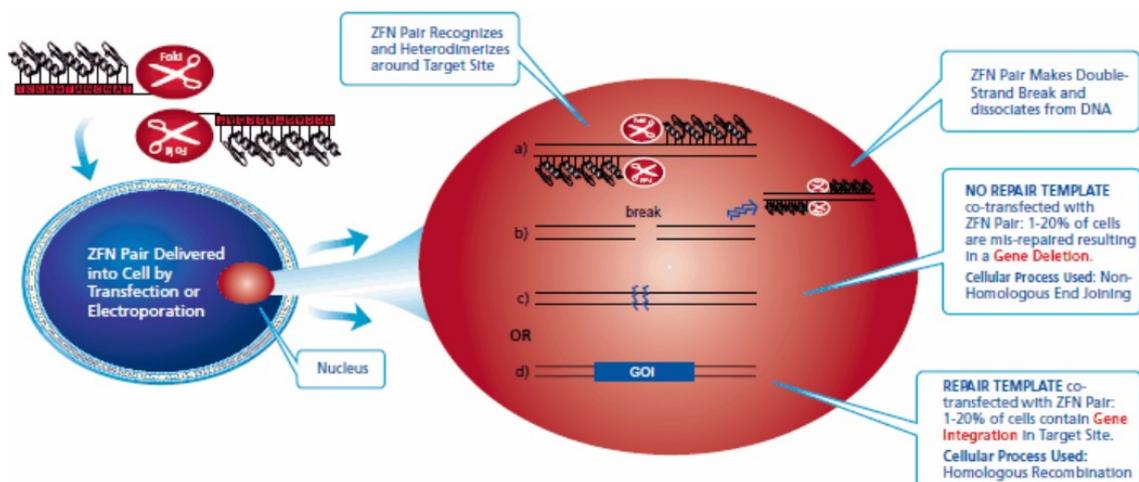


Figura 1. Esquema del funcionamiento de las ZFNs para edición de genomas.

No obstante, la flexibilidad de las ZFN es limitada y resulta complicado obtener ZFNs específicas, por lo que se vieron superadas por las TALENs. Las TALENs son también proteínas quiméricas compuestas por un dominio nucleasa que, del mismo modo que en las ZFNs, suele ser FokI, y un dominio de unión al DNA derivado de los factores TAL de bacterias del género *Xanthomonas* (Casacuberta y Puigdomènech, SEBBM). En la práctica, se pueden generar secuencias de TALENs para reconocer secuencias específicas del genoma y fusionarles la nucleasa FokI para generar el dímero funcional TALEN. De igual manera que en el caso de las ZFNs, se pueden introducir mutaciones aleatorias debido a los mecanismos de reparación de la célula o una secuencia específica de interés en el sitio de corte.

El sistema CRISPR-Cas, descubierto más recientemente, entre otros, por el microbiólogo Francisco Juan Martínez Mojica (Mojica *et al.*, 2009), se ha convertido, gracias a los hallazgos de la bioquímica francesa Emmanuelle Charpentier y la química estadounidense Jennifer

Doudna (Doudna y Charpentier, 2014), en la tecnología de edición de genomas que ha destacado notablemente frente al resto durante los últimos años.

1.2. El sistema CRISPR-Cas

El CRISPR-Cas es un mecanismo de inmunidad adaptativa propio de bacterias y arqueas que les proporciona una defensa frente a material genético invasor. El mecanismo funciona incorporando secuencias cortas del genoma de virus y otros elementos genéticos móviles en el locus CRISPR del huésped, que son transcritas y procesadas en pequeños RNAs que guían a las nucleasas para destruir el material genético invasor (Hryhorowicz *et al.*, 2016). CRISPR corresponde a, en inglés, *clustered regularly interspaced short palindromic repeats*, o a, en español, repeticiones palindrómicas cortas y regularmente interespaciadas, y contiene las secuencias adquiridas de virus, plásmidos o cualquier otro material genético detectado como perjudicial para el microorganismo, separadas por repeticiones directas. Y Cas corresponde a un locus asociado al CRISPR que codifica para proteínas nucleasas (López Mancheño, 2015).

El funcionamiento del CRISPR-Cas en estado natural como mecanismo de defensa se divide en tres pasos: adquisición de CRISPRs, biosíntesis del RNA guía e interferencia con el DNA invasor. En la etapa de adquisición, el material genético invasor es procesado por una nucleasa Cas y convertido en pequeños fragmentos de DNA, llamados secuencias protoespaciadoras, que posteriormente son incorporadas al locus CRISPR del genoma bacteriano para formar un nuevo espaciador. La selección de los protoespaciadores depende en parte del reconocimiento específico por parte de las nucleasas Cas de motivos adyacentes de protoespaciador (PAMs), que se encuentran presentes en el genoma viral, pero que no forman parte de los protoespaciadores incorporados al locus CRISPR. A continuación, en la etapa de biosíntesis, el locus CRISPR se transcribe y se sintetiza un RNA CRISPR precursor (pre-crRNA). El llamado crRNA trans-activador (tracrRNA) hibrida con las repeticiones directas del pre-crRNA y una RNasa III endógena rompe la unión entre ambos RNAs y se producen los crRNAs maduros, cada uno de los cuales contiene una secuencia espaciadora y una repetición directa parcial. Por último, en la etapa de interferencia, el crRNA maduro guía a la nucleasa hacia la secuencia complementaria al espaciador del DNA invasor, desencadenando la degradación de dicha secuencia extraña (Hryhorowicz *et al.*, 2016).

En cuanto a su clasificación, existen dos clases de sistemas CRISPR-Cas dependiendo de la configuración de sus módulos efectoros, y dentro de ellos diferentes tipos dependiendo del tipo de nucleasa asociada. Los sistemas CRISPR-Cas de clase I se caracterizan por la presencia de un módulo efector con múltiples subunidades de crRNA (utilizan complejos multiproteicos) e incluye los sistemas CRISPR-Cas de tipo I y de tipo III, así como el nuevo tipo IV. Todos los sistemas CRISPR-Cas de tipo I contienen el gen *cas3* o su variante *cas3'*, el cual codifica una helicasa con capacidad para desenrollar DNA de doble cadena y complejos de DNA-RNA. Frecuentemente, el dominio helicasa se fusiona con un dominio endonucleasa situado en la proteína Cas3 que es responsable de la ruptura del DNA diana. También son características de este tipo de sistema CRISPR-Cas las proteínas Cas8 y Cas4 (Fig. 2). Los sistemas CRISPR-Cas de tipo III poseen el gen *cas10*, que codifica para proteínas que muestran una gran variación de secuencia entre los diversos sistemas CRISPR-Cas de tipo III. Todos los loci de tipo III, al igual que los de tipo I, también codifican para una proteína Cas5 y varias proteínas Cas7 (Fig. 2). De igual manera que en los sistemas CRISPR-Cas de tipo I, Cas10 también suele estar fusionada con un dominio nucleasa, pero este es distinto de los de tipo I. Varios genomas bacterianos también poseen los llamados sistemas de tipo IV, que no están funcionalmente caracterizados y suelen presentarse en plásmidos. Este sistema no posee los genes *cas1* y *cas2*, y, por tanto, posee una estructura minimalista que lo distingue de los sistemas de tipo I y de tipo III. Además, posee una pequeña subunidad *csf1* única en este tipo de sistema CRISPR-Cas que también justifica que haya sido separado de los otros tipos (Makarova *et al.*, 2015) (Fig. 2).

Los sistemas CRISPR-Cas de clase 2 están definidos por la presencia de un módulo efector con una sola subunidad de crRNA y por funcionar con proteínas Cas de un solo componente para mediar la interferencia. Esta clase incluye los sistemas de tipo II y de tipo V. Los sistemas CRISPR-Cas de tipo II son muy diferentes de los de tipo I y III, y son los más simples en cuanto al número de genes. El gen característico de los sistemas de tipo II es *cas9*, el cual codifica una proteína multidominio que combina las funciones de complejo efector de crRNA con la de escisión del DNA diana. También contienen los genes *cas1* y *cas2* (Fig. 2), y muchos de ellos codifican para un tracrRNA (Makarova *et al.*, 2015). Cabe destacar que el sistema CRISPR-Cas de tipo II es el que normalmente se utiliza en edición dirigida de genes debido a su sencillez (Cong *et al.*, 2013), ya que a diferencia de los sistemas de clase 1, depende de Cas9, una sola proteína. Cas9 es una endonucleasa que funciona con el crRNA y el tracrRNA, formando un complejo que va examinando el DNA invasor para encontrar una secuencia diana de unos 20 nucleótidos, al igual que el correspondiente PAM, y entonces corta dicha secuencia. La secuencia PAM, en este caso, se encuentra localizada inmediatamente aguas abajo de la secuencia diana, y varía para cada tipo de proteína Cas9 (Endo *et al.*, 2016).

El nuevo sistema CRISPR-Cas de tipo V presenta como distintivo el gen *cpf1*, presente en varios genomas bacterianos y un solo genoma arqueal (por ejemplo, en *Francisella novicida*), y se encuentra junto a los genes *cas1* y *cas2* (Fig. 2) (Makarova *et al.*, 2015). El aumento del número de publicaciones sobre edición de genomas en plantas mediante el sistema CRISPR-Cas9 es muy notable pero no es una novedad. Sin embargo, cada vez más investigadores están apostando por el nuevo sistema CRISPR-Cpf1, debido a sus ventajas frente al tradicional sistema CRISPR-Cas9.

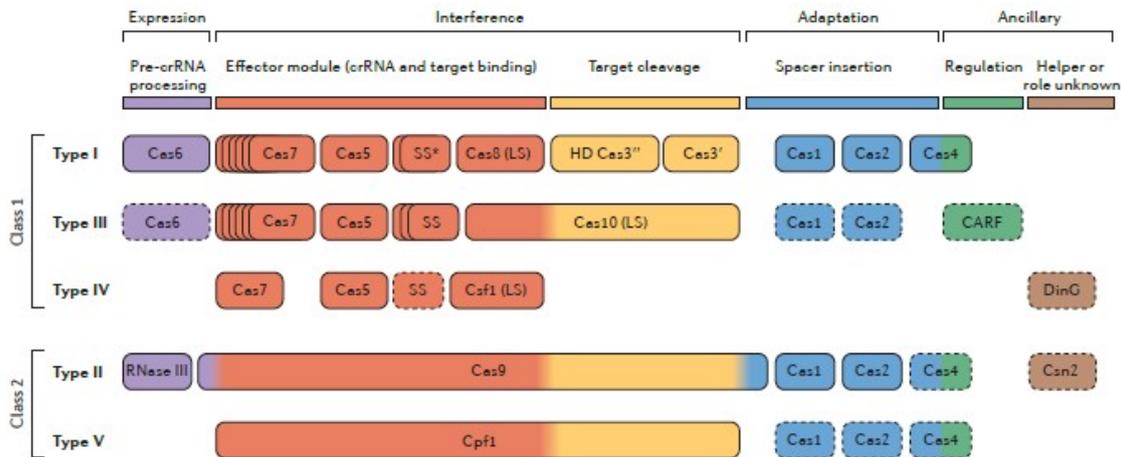


Figura 2. Clasificación funcional de las proteínas Cas y estructura general de las diferentes clases y tipos de sistemas CRISPR (Makarova *et al.*, 2015).

1.3. El sistema CRISPR-Cpf1

El sistema CRISPR de *Prevotella* y *Francisella 1* (Cpf1) se basa en una endonucleasa guiada por RNA que ha sido recientemente caracterizada y que difiere de Cas9 en numerosas de sus características. Estas diferencias en términos de estructura y función proveen a Cpf1 de importantes ventajas para su aplicación en edición de genomas. En primer lugar, los sistemas CRISPR-Cpf1 no requieren del tracrRNA para procesar el crRNA o para mediar la interferencia, si no que la Cpf1 procesa el crRNA independientemente de un tracrRNA y los complejos crRNA-Cpf1 no requieren de ningún otro RNA adicional para destruir el DNA diana. Además, el RNA guía de Cpf1 es más simple y corto (alrededor de 42 nucleótidos frente a los 100 nucleótidos de los RNA de sistemas basados en Cas9) y consiste en un solo bucle en la repetición directa. Esta puede que sea la característica más notable de Cpf1, ya que puede

simplificar en gran medida el diseño y distribución de este tipo de tecnologías de edición genómica y, además, los oligonucleótidos más cortos de RNA son más fáciles y baratos de sintetizar (Zetsche *et al.*, 2015). Zetsche *et al.* (2017) probaron que resultaba mucho más sencillo expresar el RNA guía de Cpf1 que el de Cas9 para la edición de múltiples genes, ya que esto requiere construcciones relativamente grandes, estrategias de clonaje complejas, o la distribución simultánea de varias construcciones con múltiples plásmidos. En cambio, para la expresión del RNA guía de Cpf1, basta con un simple transcrito precedido por un promotor adecuado (Zetsche *et al.*, 2017).

Es más, Zetche *et al.* (2015) también destacaron las diferencias entre los crRNAs maduros de Cpf1 y Cas9. De modo que los crRNAs maduros de unos 42-44 nucleótidos propios de sistemas de tipo V, empiezan con 19 nucleótidos de repetición directa seguidos de 23-25 nucleótidos de secuencia espaciadora, mientras que en los sistemas CRISPR-Cas9 el crRNA maduro empieza con 20-24 nucleótidos de secuencia espaciadora seguidos por unos 22 nucleótidos de repetición directa (Zetsche *et al.*, 2015). Asimismo, en el caso de sistemas CRISPR-Cas9, el procesamiento del pre-crRNA es catalizado por la RNasa III, que reconoce una estructura que se forma con el tracrRNA y que no está presente en el pre-crRNA de los sistemas CRISPR-Cpf1 (Fig. 3). Esto indica que la RNasa III no es responsable del procesamiento del pre-crRNA en este tipo de sistemas.

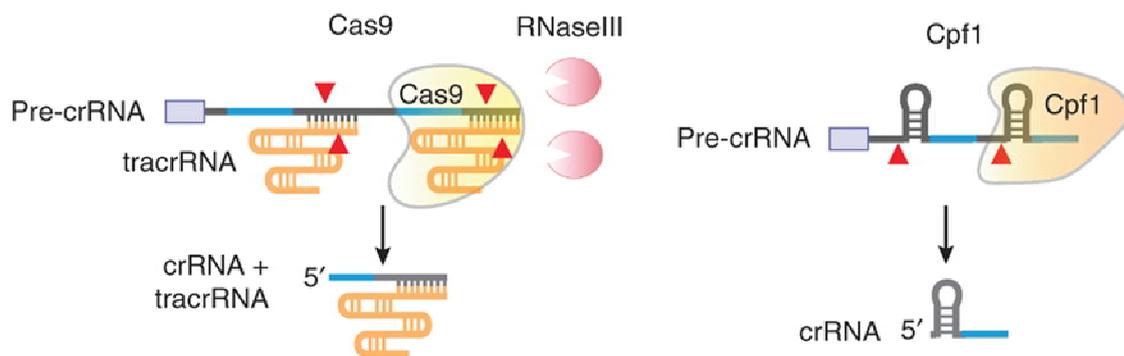


Figura 3. Esquema del procesamiento del pre-crRNA para Cas9 y Cpf1 (Zetsche *et al.*, 2017).

Segundo, el PAM que reconocen los sistemas Cas9 es rico en guaninas (G) y se encuentra inmediatamente aguas debajo del sitio diana, en cambio, el PAM que reconocen los complejos crRNA-Cpf1 es rico en timinas (T), solo variando en el número de Ts, y se localiza inmediatamente aguas arriba de la secuencia diana de 24 nucleótidos de Cpf1. Hasta el descubrimiento de este nuevo sistema de tipo V, todas las proteínas de edición de genomas requerían la presencia de al menos una G, por lo que los PAMs ricos en T de las proteínas de la familia Cpf1 expanden el rango de secuencias diana de las nucleasas de edición de genomas guiadas por RNA, por ejemplo, permitiendo su aplicación en organismos con genomas ricos en adenina y timina, como *Plasmodium falciparum*. Las secuencias de reconocimiento PAM de Cpf1 son diferentes en las distintas especies bacterianas que la poseen. Tercero, la Cpf1 provoca una rotura en el DNA que produce extremos cohesivos con una porción sobresaliente de 4 o 5 nucleótidos en el extremo 5', a diferencia de la Cas9 que produce extremos romos. Esta estructura formada tras la escisión con Cpf1 puede ser ventajosa para facilitar la inserción de genes en el genoma. Si se programara la secuencia exacta de un extremo cohesivo se podría diseñar el inserto de DNA de forma que se integre en el genoma en la orientación deseada.

Otra característica potencialmente útil de Cpf1 que podría ayudar a la introducción de fragmentos de DNA es que esta enzima rompe el DNA lejos de la secuencia diana. Por tanto, los indels que hayan podido formarse estarán lejos del sitio diana y este estará disponible para subsiguientes rondas de modificación con Cpf1. Con Cas9, cualquier indel formado afectaría al

sitio diana, eliminando la posibilidad de insertar nuevo DNA en este sitio en particular. Además, Cpf1 contiene un único dominio nucleasa, el dominio RuvC, que contiene todos los residuos catalíticos de esta familia de endonucleasas, al contrario de Cas9 que presenta dos dominios nucleasa (Zetsche *et al.*, 2015). De entre las diferentes proteínas de la familia de Cpf1, la Cpf1 de *Francisella novicida* (FnCpf1) ha sido utilizada en algunas publicaciones para edición de genomas tanto de células de mamíferos (Zetsche *et al.*, 2015), como de plantas (Endo *et al.*, 2016). Otras proteínas Cpf1 destacadas en ciertas publicaciones son la proveniente de *Acidaminococcus sp.* (AsCpf1) y la de *Lachnospiraceae bacterium* (LbCpf1), que se vio que, de entre un grupo de Cpf1s, eran las que mostraron niveles detectables (y comparables con los obtenidos con Cas9) de indels en células humanas (Zetsche *et al.*, 2015).

FnCpf1 reconoce un PAM más corto (TTN) que el resto de proteínas Cpf1 conocidas. Esto es una característica favorable en lo que se refiere a mutagénesis dirigida, ya que la frecuencia de secuencias diana para FnCpf1 en el genoma será mayor. No obstante, al reconocer un PAM más corto, aumenta la frecuencia de que aparezcan secuencias diana en el genoma, pero también la probabilidad de que se produzcan modificaciones fuera del sitio diana. De manera que conforme la secuencia del PAM se alarga, se reduce el número de sitios diana, pero la fidelidad aumenta. AsCpf1 tiene más fidelidad que FnCpf1, pues la secuencia de su PAM es más larga (TTTN), y de esta manera induce mutaciones fuera de diana solo en casos muy extraños (Endo *et al.*, 2016).

Por otra parte, tanto AsCpf1 como LbCpf1, que también reconoce TTTN como secuencia PAM, se ha visto, tras realizar ensayos en células humanas, que poseen una actividad más alta para inducir mutaciones que otras enzimas Cpf1, incluyendo FnCpf1 (Zetsche *et al.*, 2015). Entre estas dos enzimas, Tang *et al.* (2017) comprobaron con un ensayo que LbCpf1 mostraba una actividad superior a la de AsCpf1 (Tang *et al.*, 2017). Además de su actividad editora de genes, también se ha comprobado la actividad de procesamiento del pre-crRNA in vitro de AsCpf1 y LbCpf1 con éxito (Zetsche *et al.*, 2017).

1.4. Vectores virales

Los virus son elementos genéticos infectivos que tienen la capacidad de replicarse independientemente del cromosoma de la célula huésped a la que infectan. El genoma del virus puede ser de RNA o de DNA y está rodeado de la denominada cápside o cubierta proteica. El material genético de los virus puede ser monocatenario o bicatenario, entre los que se encuentran moléculas lineales de DNA, moléculas de DNA circular, RNA segmentado, o RNA monocatenario de polaridad positiva o negativa. El conjunto de los ácidos nucleicos y la cápside se denomina nucleocápside. Cuando se encuentran en estado extracelular, se denominan viriones (Legorreta Herrera *et al.*, 2012).

Obtener clones infecciosos de virus en plásmidos es una base importante en la investigación de su biología molecular e interacción con el huésped y vectores de transmisión. Además, en muchas ocasiones estas construcciones son el punto de partida para conseguir diseñar vectores virales de expresión con múltiples aplicaciones biotecnológicas (Bedoya y Daròs, 2010). La biotecnología vegetal depende de dos estrategias para la infección y la expresión de genes heterólogos en plantas: la transformación genética estable y la expresión transitoria usando vectores virales (Marillonnet *et al.*, 2005). Actualmente, se investigan diferentes virus para producir vectores virales que proporcionen expresión permanente o transitoria de transgenes. Al introducir un vector viral con un gen de interés en una planta huésped, éste se replicará y la proteína de interés podrá producirse. El uso de vectores virales presenta una serie de ventajas: (1) el producto proteico se produce más rápido y con mayor rendimiento y, (2) el gen no tiene por qué incorporarse en el genoma de la planta, no se transmite a la descendencia y está contenido (Gleba *et al.*, 2007). Los vectores virales se pueden clasificar en de primera y de segunda generación. Los vectores virales de primera generación son virus funcionales que expresan, además de todos los genes naturales del virus (también llamado estrategia del virus completo), la secuencia codificante de una proteína de interés, la cual se expresa bajo la

influencia de un promotor fuerte. El gen de interés es transfectado a las células vegetales usando copias infecciosas de ácidos nucleicos del vector, por ejemplo, plásmidos. El proceso suele llevarse a cabo por medio de infección mediada por *Agrobacterium tumefaciens*, y dependiendo del tipo de vector y su habilidad para moverse por la planta, se requieren de entre una y tres semanas para que la mayoría de los tejidos de la planta se infecten. Este método proporciona alta velocidad de infección con un protocolo muy simple, pero el rendimiento de proteína obtenido es bastante bajo. Otras limitaciones de la estrategia del virus completo se basan en que los insertos más grandes de 1 kb normalmente no se pueden expresar, los vectores nunca infectan todas las partes de la planta, la infección avanza a diferentes velocidades en diferentes partes de la planta, y el vector suele ser inestable, por consiguiente, muchos tejidos infectados no expresan la proteína de interés. Los vectores virales de segunda generación, en lugar de utilizar genomas virales completos, son virus deconstruidos en los que solo se mantienen los elementos virales requeridos para la expresión de la secuencia de interés, y las funciones perdidas se proporcionan usando componentes no virales. Por ejemplo, la proteína de la cápside es un elemento del virus que le permite moverse sistémicamente por la planta, sin embargo, gran parte de los recursos metabólicos de la célula van destinados a la síntesis de grandes cantidades de la proteína de la cápside, y esto reduce el rendimiento de producción de la proteína de interés. Por tanto, una solución simple en cuanto a esto es eliminar la proteína de la cápside del vector y reemplazar la capacidad de moverse sistémicamente por la planta por la infección de toda la planta mediante *A. tumefaciens*. Además, como el vector viral no contiene el gen de la proteína de la cápside, puede expresar genes más largos (Gleba *et al.*, 2007). Los tipos de virus más utilizados históricamente para construir vectores virales en plantas han sido virus de RNA monocatenario de polaridad positiva, como, por ejemplo, el virus del mosaico del tabaco (TMV) y el virus X de la patata (PVX).

1.5. Virus del grabado del tabaco

Tradicionalmente, los potyvirus (género *Potyvirus*, familia *Potyviridae*) han resultado ser de gran interés en biotecnología vegetal. Son virus de RNA monocatenario de polaridad positiva de unas 10 kb, con una proteína viral unida en el extremo 5' de su genoma y una cola poli-A en el extremo 3'. Varias características de los potyvirus les hacen buenos candidatos para ser vectores de expresión. Por ejemplo, permiten la producción de proteínas heterólogas en cantidades equimolares al resto de proteínas virales debido a que expresan sus proteínas a partir de una poliproteína que es procesada por una proteasa viral para dar los productos génicos maduros (Bedoya y Daròs, 2010).

El virus del grabado del tabaco (*Tobacco etch virus*, TEV) pertenece al género *Potyvirus* dentro de la familia *Potyviridae*. Su sintomatología incluye el mosaico con enanismo o clorosis severa con enanismo (Velasquez *et al.*, 2015). Este virus se ha usado en numerosas ocasiones como modelo para la investigación de la biología molecular de virus vegetales de RNA, en concreto de los potyvirus. Por ejemplo, ha servido para la identificación y caracterización de proteasas virales (Carrington *et al.*, 1989), la localización subcelular de proteínas virales (Restrepo *et al.*, 1990), algunos importantes progresos en replicación viral (Daròs *et al.*, 1999), o para la descripción del primer supresor viral del silenciamiento por RNA (Kasschau y Carrington, 1998).

Al principio, solo estaba disponible para la investigación un plásmido con un clon completo del TEV (pTEV7DA) que permitía la producción *in vitro* de transcritos infecciosos y resultaba ser muy inestable al propagarlo en *E. coli*. Bedoya y Daròs (2010) usaron el clon infeccioso del TEV contenido en el pTEV7DA para diseñar y construir tres nuevos plásmidos que permitían infectar plantas y cuyo diseño estaba pensado para mejorar su estabilidad en *E. coli*. De los tres, el pGTEV permitía infectar plantas por agroinoculación. Este vector lleva un casete de expresión con el cDNA del TEV flanqueado por el promotor y terminador 35S del virus del mosaico de la coliflor (*Cauliflower mosaic virus*, CaMV), insertado entre los dos bordes del T-DNA (Fig. 4A) (Bedoya y Daròs, 2010).

Un importante descubrimiento sobre la biología molecular del TEV fue realizado por Li y Carrington (1995), quienes hallaron que un mutante del TEV con una delección de todo el cistrón de la RNA polimerasa viral (*nuclear inclusión 'b' protein*, Nib) (TEV Δ Nib) puede ser complementado en *trans* en una planta transgénica que exprese esta proteína (Li y Carrington, 1995). Es decir, el vector solo se replica y se mueve sistémicamente en plantas cuando la actividad de Nib se complementa en *trans* por un transgén u otro vector viral. Basándose en este hallazgo, Bedoya *et al.* (2010) construyeron un vector viral basado en el TEV y derivado del pGTEV (Bedoya y Daròs, 2010) en el cual se reemplazó la RNA polimerasa viral (Nib) por distintas proteínas heterólogas, y demostraron el potencial uso del pGTEV como vector de expresión en plantas (Fig. 4B) (Bedoya *et al.*, 2010).

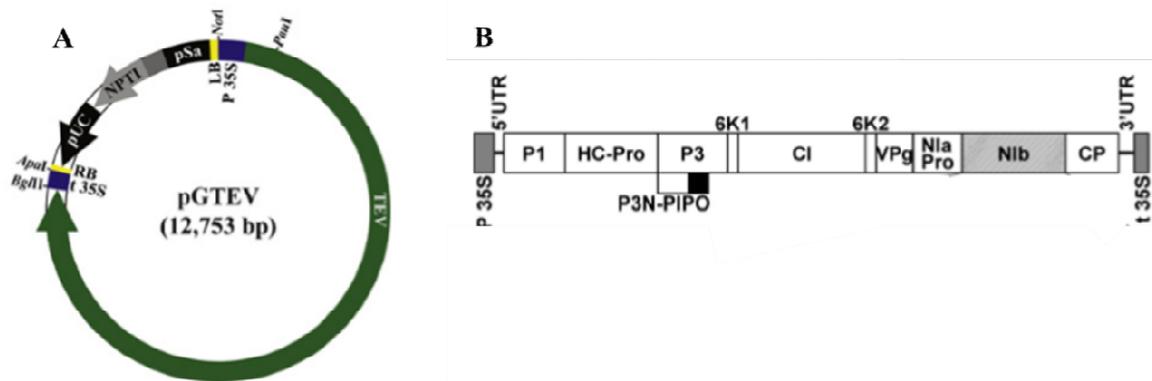


Figura 4. A) Estructura del plásmido binario pGTEV, que contiene un cDNA infeccioso del TEV (Bedoya y Daròs, 2010). B) Diagrama del vector de expresión TEV, con el cistrón de Nib en gris, el cual se puede sustituir por un gen de interés para formar el vector de expresión TEV Δ Nib (Bedoya *et al.*, 2010).

1.6. Virus X de la patata

El virus X de la patata (*Potato virus X*, PVX) pertenece al género *Potexvirus* de la familia *Flexiviridae*, históricamente investigado debido a sus perjudiciales efectos en patata, el cuarto producto alimentario más importante del mundo. Se trata de un virus de estructura filamentosa formado por un genoma de RNA monocatenario de polaridad positiva de unas 6.4 kb, el cual se encuentra envuelto en aproximadamente 1300 unidades de la proteína de la cápside (Lico *et al.*, 2015). Su genoma se compone de 5 grandes marcos de lectura abierta (*open reading frames*, ORFs), el primero de los cuales es seguido de tres ORFs parcialmente superpuestos, conocidos como el triple bloque de genes (TGB), los cuales codifican tres proteínas de 25, 12 y 8 kDa (Krishnamurthy *et al.*, 2002). El PVX es fácilmente transmisible por la savia y en la naturaleza por contacto mecánico. Infecta a más de 240 especies de 16 familias, pero la mayoría de sus huéspedes pertenecen a la familia *Solanaceae*. Su sintomatología en patata consiste en un suave mosaico o se encuentra latente y en tabaco causa moteado o aros de necrosis. Los efectos que causa pueden verse empeorados por la co-infección con otros virus, en especial con el virus Y de la patata (*Potato virus Y*, PVY; es un fenómeno denominado sinergismo). En particular *N. tabacum* y otras especies de *Nicotiana* son especies diagnósticas y de propagación del PVX. En estas plantas, las hojas que primero se infectan suelen desarrollar anillos necróticos, y las hojas que se infectan más tarde muestran moteado clorótico o necrótico, mosaico o clorosis de las venas (Fig. 5). Además, gracias a los síntomas que el virus causa en tabaco, se pueden distinguir las variantes del mismo (Koenig y Lesemann, 1989).

El PVX ha sido utilizado como modelo para el estudio de la interacción planta-patógeno, para elucidar los mecanismos de defensa de la planta y el contraataque del virus, como por ejemplo la supresión del silenciamiento. Para llevar a cabo estos estudios, se tuvieron que desarrollar vectores virales de expresión (Chapman *et al.*, 1992). Muchos de estos vectores han encontrado

su aplicación en diferentes campos, como el de la genómica funcional, en el cual se ha desarrollado la técnica denominada silenciamiento génico inducido por virus (Ruiz *et al.*, 1998); en '*molecular farming*', se están usando vectores basados en PVX para inducir la expresión de genes exógenos en plantas, los cuales codifican biomoléculas de alto valor añadido (Komarova *et al.*, 2010); y también se están empezando a utilizar en nanotecnología (Lico *et al.*, 2015).

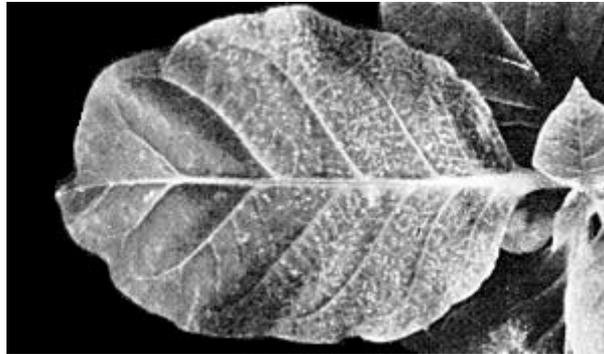


Figura 5. Clorosis de las venas provocada por el PVX en una hoja de *N. tabacum* (Koenig y Lesemann, 1989).

2. OBJETIVOS

El sistema CRISPR-Cas ha sido un descubrimiento relativamente reciente que ha estado protagonizando los proyectos de investigación de muchos laboratorios en los últimos años. A pesar de que el primero que se adaptó a la edición de genomas fue el sistema CRISPR-Cas9, el nuevo sistema CRISPR-Cpf1 promete facilitar la modificación genética debido a sus características. La inserción de este sistema en un vector viral podría permitir una rápida expresión del mismo para modificar exitosamente genomas de plantas. En este trabajo de fin de grado se ha pretendido:

- a) Construir vectores virales derivados del TEV y del PVX para expresar el sistema CRISPR-Cpf1 en plantas de *Nicotiana tabacum* y *Nicotiana benthamiana*.
- b) Comprobar la infectividad de los diferentes vectores virales construidos.
- c) Comprobar si los vectores virales que expresan la nucleasa Cpf1 son capaces de conservar el inserto tras infectar plantas.

3. MATERIALES Y MÉTODOS

3.1. Transformación de bacterias por electroporación

Se construyeron vectores virales a partir de dos plásmidos derivados del TEV, el plásmido intermedio p15TV3-BB3 y el plásmido final pGTEV Δ EE. Estos dos plásmidos se modificaron para insertar los cDNAs de las nucleasas AsCpf1 (*Acidaminococcus sp.*) y LbCpf1 (*Lachnospiraceabacterium* ND2006). También se construyeron vectores virales a partir de dos plásmidos derivados del PVX, el plásmido intermedio pMPVXBa-PdamiBZB y el plásmido final pPVX. Estos dos plásmidos se modificaron para insertar las repeticiones directas y las secuencias guía del crRNA para AsCpf1 y LbCpf1.

Los diferentes plásmidos con los clones virales se electroporaron en la cepa DH5 α *E. coli* competentes. Se utilizaron entre 0.2 y 8 μ l de cada plásmido en 40 μ l de células competentes sobre un tubo Eppendorf de 1.5 μ l. Se traspasó todo a una cubeta de 0.1 cm enfriada previamente en hielo y se electroporó a 1500 V durante 5 ms en un electroporador ECM399 (BTX). Rápidamente, se añadió 1 ml de SOC (20 g/l triptona, 5 g/l extracto de levadura, 0.5 g/l NaCl, 2.5 mM KCl, 10 mM MgCl₂, glucosa a 20 mM, pH 7.0), se recogieron las células con una pipeta y se devolvieron al tubo Eppendorf. Las células se incubaron a 37 °C en agitación (200 rpm) durante 1 h. Posteriormente se plaquearon, 10 μ l y el resto, en dos placas de Petri distintas con medio de selección LB suplementado con antibióticos, y se dejaron incubando 24 h a 37 °C (48 h a 37 °C los plásmidos más grandes).

Una vez construidos y purificados los diferentes plásmidos finales modificados, se electroporaron en *A. tumefaciens* competentes. Se utilizó 1 μ l de plásmido en 40 μ l de células competentes sobre un tubo Eppendorf de 1.5 μ l. Se electroporó siguiendo el mismo protocolo descrito previamente para *E. coli*, pero las células se incubaron a 28 °C durante 2-3 h tras la electroporación. Después, se plaquearon 10 μ l y el resto en dos placas de Petri distintas con medio de selección LB suplementado con tetraciclina 7.5 μ g/ml, rifampicina 50 μ g/ml y kanamicina 50 μ g/ml, y se dejaron incubando 5 días a 28 °C.

3.2. Purificación de plásmidos

La purificación de plásmidos se llevó a cabo utilizando el kit GeneJet Plasmid Miniprep (ThermoScientific). En primer lugar, se traspasaron 2 ml de cultivo saturado de células en un tubo Eppendorf, se centrifugaron durante 2 min a 13000 rpm y se descartó el sobrenadante. Se resuspendió el sedimento utilizando vortex en 250 μ l de solución de resuspensión con RNasa A conservada a 4 °C. Seguidamente, se añadieron 250 μ l de solución de lisis y se mezcló por inversión al menos 5 veces. Después, se añadieron 350 μ l de solución de neutralización, se mezcló exhaustivamente por inversión al menos 5 veces y se centrifugó durante 5 min a 13000 rpm. A continuación, se transfirió el sobrenadante a una columna sobre un tubo Eppendorf de 2 ml, se centrifugó durante 1 min a 13000 rpm y se descartó el filtrado. Posteriormente, se lavó con 500 μ l de solución de lavado centrifugando durante 30 s a 13000 rpm y se descartó el filtrado. Se realizó un segundo lavado igual al descrito anteriormente y después se centrifugó durante 1 min para eliminar restos de la solución de lavado. Finalmente, se traspasó la columna a un tubo Eppendorf de 1.5 ml, se añadieron 50 μ l de tampón de elución, se incubó durante 2 min a temperatura ambiente y se centrifugó 2 min a 13000 rpm antes de eliminar la columna.

3.3. Amplificación DNA mediante PCR

El DNA se amplificó mediante PCR utilizando la DNA polimerasa Phusion (ThermoFisherScientific), la cual tiene 50 veces más fidelidad de copia (tasa de error $4.4 \cdot 10^{-7}$) que la *Taq* DNA polimerasa y produce productos con extremos romos. Para ello, se mezclaron

sobre hielo, en cabina de PCR y utilizando puntas de filtro, 4 μ l de tampón HF (ThermoFisherScientific) (x5), 0.6 μ l de DMSO, 0.4 μ l de dNTPs (10 mM), 0.4 μ l del cebador 1 (25 μ M), 0.4 μ l del cebador 2 (25 μ M), 0.2 μ l de DNA polimerasa Phusion (2 U/ μ l), DNA molde y H₂O hasta 20 μ l, para cada reacción. Cada PCR se incubó en un termociclador con el siguiente programa: 30 s a 98 °C para la desnaturalización, seguida de 30 ciclos de 10 s a 98 °C para desnaturalización, 30 s a 55 °C para el anillamiento, 15 s/1 kb a 72 °C para la extensión y 10 min a 72 °C para asegurar la correcta extensión de los productos finales.

3.4. Electroforesis en gel de agarosa

Para analizar los plásmidos y los fragmentos de DNA se realizó electroforesis en gel de agarosa 1%. Dependiendo de la cantidad de muestras a analizar se utilizaron geles pequeños (8 x 65 x 0.5 cm) o geles más grandes (15 x 11 x 0.5 cm). Para realizar la electroforesis en los geles pequeños se vertieron 25 ml de agarosa fundida, y se dejó enfriar en presencia de un peine con un número de pocillos variable durante 15-20 min. Para la electroforesis en los geles grandes se vertieron 75 ml de agarosa fundida, y se dejó enfriar en presencia del peine durante el mismo tiempo que los geles pequeños. Una vez solidificado el gel, se retiró el peine y el gel se colocó en su correspondiente cubeta de electroforesis sumergido en tampón TAE (x1) (40 mM Tris, 20 mM acetato de sodio, 1 mM EDTA). A las muestras a analizar se les añadió una cantidad proporcional de tampón de carga LB (x10) (12.5 mg azul bromofenol, 12.5 mg xilencianol, 28.75 g 87% glicerol, 1 ml 0,5 M EDTA pH 8.0, H₂O hasta 50 ml). Las muestras se cargaron en los pocillos. Se corrió la electroforesis a 60 V durante 75 min para los geles pequeños, y a 90V durante 75 min para los geles grandes.

Para detectar el DNA separado por electroforesis, se transfirió el gel a una cubeta de plástico con 200 ml de agua destilada y 10 μ l de bromuro de etidio (10 mg/ml) y se dejó en agitación 15 min. Después, se realizaron tres lavados con agua destilada para eliminar el bromuro de etidio sobrante y se visualizó el DNA bajo luz ultravioleta.

3.5. Manipulación plásmidos mediante enzimas de restricción y DNA ligasa

Se digirieron diferentes plásmidos con distintas enzimas de restricción. Estas digestiones se realizaron en tubos Eppendorf de 0.5 μ l añadiendo una cantidad adecuada de DNA junto con 1 μ l de la enzima o enzimas de restricción correspondientes, un volumen adecuado del tampón correspondiente a las enzimas y H₂O hasta alcanzar un volumen de 20 o 40 μ l. Finalmente, las digestiones se incubaron a 37 °C durante 1 h.

También se realizó una reacción de digestión parcial a 37 °C durante 1 h, recogiendo alícuotas de la reacción a diferentes tiempos. Estas alícuotas se depositaron en tubos Eppendorf de 0.2 μ l que contenían EDTA, que detiene la reacción de digestión al secuestrar el magnesio que la enzima necesita para poder funcionar.

Para construir los plásmidos también se realizaron diferentes reacciones de ligación. Estas reacciones se realizaron en tubos Eppendorf de 0.2 μ l añadiendo las cantidades adecuadas de DNA (con relación molar de 1 molécula de plásmido por 3 moléculas de inserto), 1 μ l de la enzima DNA ligasa T4 (ThermoFisherScientific), tampón DNA ligasa T4 (x10), y H₂O, si es necesario, hasta alcanzar un volumen de 10 μ l. Finalmente, las ligaciones se incubaron a 22 °C durante 1 h.

3.6. Elución de ácidos nucleicos de un gel de agarosa

Se utilizó el kit de columnas ZymoResearch durante todo el proceso de elución. Primero, se cortó y se pesó el trozo de gel correspondiente que contenía el fragmento de DNA a eluir y se depositó en un tubo Eppendorf de 1.5 ml. A continuación, se añadieron 3 volúmenes de Agarose Dissolving Buffer (ADB) y se incubó la muestra durante 5-10 min a 55 °C. Una vez disuelta toda la agarosa, se traspasó la mezcla a una columna sobre un tubo Eppendorf de 2 ml, se centrifugó durante 10 s a 13000 rpm y se retiró el filtrado. Seguidamente se realizaron dos lavados con 200 µl de Washing Buffer (WB), el primero se centrifugó durante 10 s a 13000 rpm y el segundo durante 30 s a 13000 rpm, y se retiró el filtrado en ambos. Posteriormente se traspasó la columna a un nuevo tubo Eppendorf de 1.5 ml, se añadieron entre 8 y 11 µl de Elution Buffer (EB) (x 1). Se incubó durante 1 min a temperatura ambiente, se centrifugó 1 min a 13000 rpm y se recogió el eluido.

3.7. Purificación de ácidos nucleicos de una solución

Utilizando el kit de columnas ZymoResearch, primero, se añadieron a la solución 100 µl de DNA Binding Buffer (DBB), se traspasó la mezcla a una columna sobre un tubo Eppendorf de 2 µl, se centrifugó durante 10 s a 13000 rpm y se retiró el filtrado. Después, se hicieron dos lavados con 200 µl de WB, el primero se centrifugó durante 10 s a 13000 rpm y el segundo durante 30 s a 13000 rpm, y se retiró el filtrado en ambos. En el caso de necesitar la solución purificada para electroporar, se lavó adicionalmente con 200 µl de etanol al 80%. A continuación, se traspasó la columna a un nuevo tubo Eppendorf de 1.5 ml, se añadieron entre 8 y 11 µl de EB (x 1) (en el caso de que la solución purificada se utilice para electroporar, se utiliza EB (x 0.1), se incubó durante 1 min a temperatura ambiente y se centrifugó 1 min a 13000 rpm. Por último, se recogió el eluido.

3.8. Construcción de plásmidos mediante ensamblaje Gibson

Un método alternativo a la construcción de plásmidos mediante una reacción de ligación es el ensamblaje Gibson. Este método consiste en el ensamblaje de fragmentos de DNA, que pueden ser productos de reacciones de restricción o productos de PCR, con extremos solapantes. Para que el ensamblaje tenga lugar, el DNA se mezcla con un combinado de tres enzimas (exonucleasa, DNA polimerasa y DNA ligasa) y tampón.

Se realizaron los ensamblajes Gibson con el kit NEBuilderHiFi DNA Assembly Master Mix (NEB). Se añadieron los distintos fragmentos de DNA de manera que la cantidad total de DNA estuviera entre 0.02-0.5 pmol (ensamblaje de 2-3 fragmentos) y un volumen de Gibson Assembly Master Mix (2X) en un tubo Eppendorf de 0.2 µl. Las reacciones se incubaron a 50 °C durante 1 h. Después, se purificó el producto de la reacción mediante columnas de gel de sílice (Zymo) y se transformó *E. coli* con el mismo mediante electroporación tal y como se describe en el apartado 3.1.

3.9. Hibridación de dos cebadores

Para obtener los productos con la secuencia del RNA guía homóloga a la secuencia diana a modificar del genoma de *N. tabacum* y *N. benthamiana*, se realizó una hibridación de dos cebadores que contenían las secuencias deseadas. Para ello se mezclaron los componentes de una reacción de PCR con DNA polimerasa Phusion, pero sin dNTPs, polimerasa y DNA molde. A cada reacción se le añadieron 4 µl de tampón HF (x5), 0.6 µl de DMSO, 0.4 µl del cebador 1, 0.4 µl del cebador 2, y H₂O hasta alcanzar un volumen final de 20 µl, es decir, 14.6 µl de H₂O. Las reacciones se incubaron en un termociclador con un programa para desnaturalizar DNA y

renaturalizar lentamente: 2 min a 95 °C, rampa hasta 25 °C a razón de 2 °/min (35 min), y mantener a 25 °C por tiempo indefinido.

3.10. Agroinoculación de plantas

Se seleccionaron colonias aisladas de *A. tumefaciens* y se inocularon en 10 ml de medio LB líquido con kanamicina (50 µg/ml). Los tubos se incubaron a 28 °C en agitación hasta que se observó el crecimiento de la bacteria. Se midió la DO de los cultivos líquidos con un espectrofotómetro y se seleccionaron aquellos que presentaron una DO a 600 nm de entre 0.5 y 1. Se transfirió un volumen de 600 µl de los cultivos seleccionados a tubos Eppendorf de 1.5 ml que contenían 600 µl de glicerol 20%, para conservarlos a -80 °C. A continuación, se sedimentaron las células traspasando la cantidad adecuada de cada uno de los cultivos seleccionados a tubos Eppendorf de 2 ml, centrifugando durante 2 min a 13000 rpm. Antes de sedimentar las células, se calculó la cantidad necesaria de cultivo que había que centrifugar para obtener una DO de 0.5 tras resuspender las células. Después, se retiró el sobrenadante y se resuspendieron las células en 1.8 ml de una disolución compuesta por 10 mM Mes-NaOH pH 5.6 y 10 mM MgCl₂. A cada una de estas preparaciones se le añadieron 2.7 µl de acetosiringona a 150 µM. Posteriormente, los cultivos se incubaron durante 2 h a 28 °C y se infiltraron las plantas.

Se infiltraron las hojas por el envés apoyando una jeringa sin aguja cargada con la disolución con el cultivo de *A. tumefaciens* y presionando ligeramente con el dedo en el haz, comprobando la entrada del cultivo en el apoplasto, lo que provoca un cambio de color en la hoja. En *N. tabacum*, se inocularon 4 plantas, 2 con cada clon seleccionado del vector pGTEVΔEE-LbCpf1. En cada planta se inocularon 3 hojas. También se seleccionó una planta control sin inocular.

En *N. benthamiana*, se inocularon 4 plantas también, 2 con cada clon seleccionado del vector pPVX-crRNA para la LbCpf1. En cada planta se inocularon 3 o 4 hojas. Se seleccionó de igual manera que en el caso de *N. tabacum*, una planta control sin inocular.

3.11. Purificación RNA de plantas infectadas

Para extraer y purificar RNA a partir de tejido foliar infectado, en primer lugar, se trituró hasta 0.1 g de tejido en un tubo Eppendorf de 2 ml con 1 bolita de acero. Para ello, se congelaron los tubos con el tejido en N₂ líquido y se molió con Tissuelyser durante 1 min a 30 s⁻¹. Una vez el tejido estuvo triturado, se añadió a cada tubo Eppendorf 1 ml de TEX (4 M tiocianato de guanidinio, 0.1 M acetato sódico pH 5.5, 10 mM EDTA, 0.1 M 2-mercaptoetanol), se mezcló con ayuda de vórtex y se centrifugó durante 5 min a 13000 rpm. Después, se recogieron 600 µl del sobrenadante, a los cuales se añadieron 390 µl de 96% etanol. Se mezcló con vórtex y se centrifugó 1 min a 13000 rpm. A continuación, se transfirieron 700 µl de sobrenadante a una minicolumna Zymo de gel de sílice (Zymo-Spin I columns, ZymoResearch) posicionada sobre un tubo Eppendorf de 2 ml, se centrifugó durante 1 min a 13000 rpm y se descartó el filtrado. Posteriormente, se lavó con 500 µl de TLA (70% etanol, 10 mM acetato sódico pH 5.5) centrifugando 30 s y descartando el filtrado. Se realizó un segundo lavado con 500 µl de TLA centrifugando 1 min y transfiriendo después la columna cuidadosamente a un tubo Eppendorf de 1.5 ml. Por último, se eluyó el RNA con 10 µl de TEL (10 mM Tris-HCl pH 8.5), se incubó 1 min a temperatura ambiente, se centrifugó durante 2 min a 13000 rpm y se descartó la columna.

3.12. Amplificación de RNA por RT-PCR

El primer paso para la amplificación de RNA por RT-PCR es la transcripción inversa (RT) del RNA a cDNA y posteriormente se lleva a cabo la amplificación de este cDNA mediante PCR.

Tanto la preparación de las RT como de las PCR es recomendable realizarla en cabina de PCR y con puntas de filtro.

Para el diagnóstico del TEV, se realizaron 15 reacciones de RT y un control, utilizando para cada una de ellas 3 μl de RNA previamente purificado. Se añadió el RNA, 1 μl del cebador D179 (5 pmol/ μl) y H_2O hasta 6.5 μl en tubos Eppendorf de 0.2 μl . Para la reacción control se añadieron 3 μl más de agua en lugar de RNA. Después, las reacciones se incubaron en un termociclador durante 1.5 min a 98 °C. En el momento en que la temperatura del termociclador bajó hasta los 42 °C, se pausó el programa y se añadió a cada reacción 3.5 μl de una mezcla que contenía: 2 μl de tampón RT (x5) [M-MuLV] (250 mM Tris-HCl, pH 8.3, 250 mM KCl, 20 mM MgCl_2 , 50 mM DTT), 1 μl de dNTPs 10 mM, 0.25 μl de inhibidor de RNasas (~40 U/ μl) y 0.25 μl de M-MuLV RT (200 U/ μl). A continuación, se reanudó el programa y se incubaron las reacciones 45 min a 42 °C, 10 min a 50 °C, 5 min a 60 °C y, por último, 15 min a 70 °C.

Seguidamente, se realizaron las 15 reacciones de PCR en tubos Eppendorf de 0.2 ml mezclando sobre hielo 0.4 μl de dNTPs 10 mM, 0.8 μl del cebador 1 (25 μM), 0.8 μl del cebador 2 (25 μM), 0.2 μl de *Tth* DNA polimerasa (5 U/ μl ; Biotools), 2 μl del correspondiente tampón (x10), 1 μl de reacción RT y H_2O hasta alcanzar un volumen final de 20 μl . Para la reacción control se añadieron las mismas cantidades de reactivos. A continuación, se incubaron las reacciones en un termociclador siguiendo el siguiente programa: 2 min a 94 °C, 30 ciclos de 40 s a 94 °C, 30 s a 55 °C y 4 min a 72 °C, y 10 min a 72 °C.

3.13. Secuenciación DNA y análisis computacional de las secuencias

Se secuenciaron algunas secuencias de los plásmidos construidos como comprobación. Para ello se recurrió al servicio de secuenciación del IBMCP. Las muestras a secuenciar se proporcionaron en alícuotas de 6 μl , junto con el correspondiente cebador. Para el análisis de las secuencias se utilizaron los recursos del portal de internet TheSequenceManipulation Suite (<http://bioinformatics.org/sms/>).

4. RESULTADOS Y DISCUSIÓN

El objetivo de este trabajo fue la construcción de dos vectores virales derivados del TEV y del PVX para la expresión de un sistema CRISPR-Cpf1 y comprobar la infectividad de los mismos. Ya que los resultados de expresión de la nucleasa Cas9 mediante vectores derivados de TEV están resultando prometedores, se intentó mejorar estos resultados expresando Cpf1, nucleasa que forma parte de un sistema CRISPR que utiliza un solo RNA (Zetsche *et al*, 2015), con lo cual es más sencillo de manejar y expresar correctamente, y además se trata de una proteína ligeramente más pequeña que Cas9, lo cual facilita la construcción de plásmidos y su manejo. Se utilizó un vector derivado del TEV para la expresión de la Cpf1, en concreto las Cpf1 de *Acidaminococcus sp.* BV3L6 (AsCpf1) y de *Lachnospiraceae* bacterium ND2006 (LbCpf1). La elección de estas dos nucleasas se debe a que han demostrado la mejor actividad de edición en células humanas de entre las diferentes proteínas de la familia Cpf1. Los genes de ambas nucleasas se clonaron en TEV Δ N1b, sustituyendo en el vector pGTEV Δ EE el cistron N1b por el cDNA de las Cpf1. El cistron N1b codifica la RNA polimerasa viral y eliminándola del genoma del virus se consigue imposibilitar su replicación y movimiento a través de la planta. Sin embargo, este gen expresado en *trans* en las células de la planta complementa al virus que no expresa la polimerasa. Para inducir la edición del genoma en plantas, se necesita construir un segundo vector viral derivado del PVX para expresar el crRNA. Se decidió incluir en la construcción del RNA guía la secuencia homóloga al gen de la fitoeno desaturasa (PDS) de *N. tabacum* y *N. benthamiana*, ya que las publicaciones de CRISPR en plantas sugieren que se trata del marcador visual más versátil, cuyo alelo nulo produce un blanqueamiento fácilmente detectable a simple vista. Se buscó un RNA guía de 23 nucleótidos para editar el primer exón del gen de la PDS de *N. benthamiana* y *N. tabacum* a través de la página web Benchling (<https://benchling.com/academic>). Una vez identificadas las secuencias guía deseadas, se construyó un derivado de PVX en el que en un casete de expresión con un promotor duplicado se insertaron las dos repeticiones directas de AsCpf1 y LbCpf1 con un polylinker con dos sitios BsaI separados por LacZ2. Posteriormente este gen LacZ2 se eliminó y se introdujeron unos fragmentos producto de hibridación de cebadores, los cuales se correspondían con las secuencias guía previamente identificadas. Estos plásmidos se construyeron para agroinfiltrar plantas de *N. tabacum* y *N. benthamiana* y así comprobar, en primer lugar, que los clones eran infecciosos y, en segundo lugar, si los clones derivados del TEV perdían o no el fragmento correspondiente al gen de la nucleasa.

4.1. Construcción de vectores virales derivados de TEV que expresan Cpf1

Para la expresión de la proteína Cpf1 en la planta, se utilizó un plásmido que contenía un vector viral derivado del TEV, el pGTEV Δ EE, con gen de resistencia a kanamicina, y un plásmido intermedio, el p15TV3-BB3, que contiene un gen de resistencia a cloranfenicol, para facilitar la clonación del gen de la proteína en el plásmido final. Se transformaron células de *E. coli* por medio de electroporación con cada uno de estos plásmidos, que estaban conservados a -20 °C. Una vez transformadas las bacterias, se plaquearon en placas Petri con medio LB con kanamicina en el caso de las células transformadas con el pGTEV Δ EE, o medio LB con cloranfenicol en el caso de las células transformadas con el p15TV3-BB3, para poder seleccionar en cada caso las bacterias que habían incorporado el plásmido. A partir de las colonias resultantes de estas transformaciones se purificaron los plásmidos por duplicado y se midió su concentración con un espectrofotómetro NanoDrop ND-1000 (Tabla 1). Las preparaciones de plásmidos con una concentración más alta se seleccionaron para llevar a cabo futuros experimentos.

Tabla 1. Concentraciones de plásmidos pGTEVΔEE y p15TV3-BB3 purificados a partir de diferentes cultivos. Los cultivos seleccionados para continuar con el experimento se muestran en negrita y con un asterisco.

Construcción	Cultivo	Concentración (ng/μl)
pGTEVΔEE	1*	110.2
	2	17.2
p15TV3-BB3	1	43
	2*	49

Se pidió un plásmido a Addgene (pY010, pcDNA3.1-hAsCpf1) (Anexo) que contiene la secuencia de AsCpf1 optimizada a codones humanos con una etiqueta de localización nuclear (NLS) y tres epítopos HA (3xHA). Este plásmido se utilizó como molde para amplificar la AsCpf1 por PCR (1 minuto de extensión) con los cebadores D2507 y D2508 (Tabla 2). También se realizó una reacción de digestión del plásmido p15TV3-BB3 seleccionado (49 ng/μl) y la enzima BsaI-HF (HF corresponde a 'High Fidelity'), para después realizar un ensamblaje Gibson con el plásmido digerido y el producto de PCR de AsCpf1. Se transformaron células de *E. coli* cepa DH α por electroporación con el producto del ensamblaje Gibson y se plaquearon en medio LB con cloranfenicol. Posteriormente se cultivaron 6 colonias en medio líquido y se purificaron los plásmidos de cada una de ellas. Se realizó una separación en gel de agarosa para comprobar si las construcciones purificadas eran correctas (Fig. 6A). Se siguió el mismo procedimiento a partir del plásmido de Addgene que contenía la secuencia de la LbCpf1 (pY016, pcDNA3.1-hLbCpf1) (Anexo). Los cebadores para amplificar la LbCpf1 por PCR fueron el D2521 y D2508 (Tabla 2). También se realizó una electroforesis en gel de agarosa para comprobar si las construcciones eran correctas (Fig. 6B). De entre los plásmidos intermedios purificados, se seleccionaron uno con cada nucleasa Cpf1 (en ambos casos se seleccionó la construcción 1) (Fig. 6), con el fin de utilizarlos para traspasar la secuencia de la proteína al plásmido final, el pGTEVΔEE.

Tabla 2. Secuencias de los cebadores utilizados para realizar las correspondientes PCRs para amplificar las nucleasas AsCpf1 y LbCpf1 a partir de los plásmidos pY010 y pY016 (Anexo), respectivamente.

Fragmento amplificado	Cebador 1	Cebador 2
AsCpf1	D2507: 5' -ATTGGTGTACTCGCAAGGGGA GAAGATGACACAGTTCGAGGGC-3'	D2508: 5' -TCTGAAAATAAAGATTCTCA GTCGTGGCATAGTCGGGGACATC-3'
LbCpf1	D2521: 5' -ATTGGTGTACTCGCAAGGGG AGAAGATGAGCAAGCTGGAGAAG-3'	D2508: 5' -TCTGAAAATAAAGATTCTCA GTCGTGGCATAGTCGGGGACATC-3'

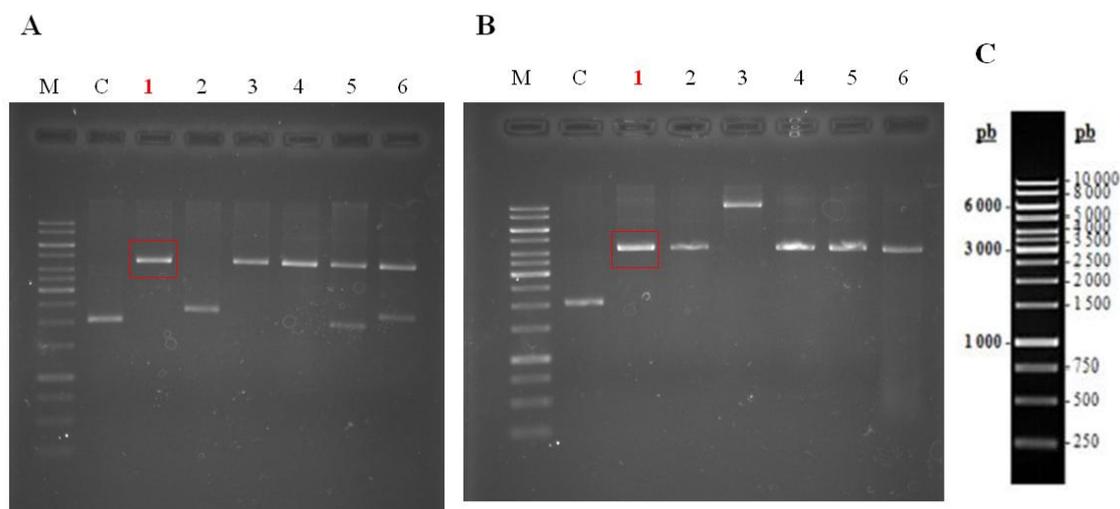


Figura 6. **A)** Gel de electroforesis de las construcciones p15TV3-AsCpf1 y el plásmido p15TV3-BB3 como control. **B)** Gel de electroforesis de las construcciones p15TV3-LbCpf1 y el plásmido p15TV3-BB3 como control. En ambos casos, la construcción marcada en rojo (1) fue seleccionada para realizar los pasos siguientes del experimento. **C)** Marcador DNA 1 kb Thermo (SM0311).

Para la inserción de AsCpf1, el pGTEV Δ EE se digirió con las enzimas Eco81I y BlnI, y el plásmido p15TV3-AsCpf1 se digirió también con las mismas enzimas de restricción. Ambas enzimas tienen un único sitio de restricción en p15TV3-AsCpf1. Los productos de ambas digestiones se utilizaron para llevar a cabo una reacción de ligación con la DNA ligasa T4. El producto de esta ligación se utilizó para transformar células de *E. coli* por electroporación que posteriormente se plaquearon en medio LB con kanamicina. Las colonias tardaron más de 24 h en crecer a 37 °C debido al gran tamaño de la construcción pGTEV Δ EE-AsCpf1, que provoca que a la bacteria le cueste mucho trabajo replicarlo y, por lo tanto, crecer. Sin embargo, los resultados no fueron los esperados y se procedió a realizar las construcciones mediante ensamblaje Gibson.

Para digerir el plásmido p15TV3-LbCpf1 se necesitó realizar una digestión parcial con Eco81I y ApaI, ya que la enzima Eco81I tiene dos sitios de restricción, uno de ellos interno en LbCpf1 al principio del cDNA. ApaI tiene un único sitio de restricción. En primer lugar, se digirió p15TV3-LbCpf1 con ApaI. Después, se realizó la digestión parcial con Eco81I a partir del producto de la digestión anterior. Se recogió una alícuota a tiempo 0 y, posteriormente, se añadió 1 Ude Eco81I y se inició la reacción de digestión a 37 °C. Se recogieron alícuotas a 5, 10, 20 y 60 min y se depositaron inmediatamente en tubos Eppendorf con tampón de carga (x10) y EDTA a 29 mM. El EDTA se utilizó para parar la reacción conforme se iban cogiendo las alícuotas, ya que tiene la capacidad de secuestrar el Mg²⁺ que necesita la enzima para funcionar. Esta digestión debería dar como productos dos bandas en un gel de agarosa, una de 5549 pb y otra de 5028 pb. El producto deseado se corresponde con la banda de 5549 pb. Una vez obtenido el producto de la digestión parcial, también se digirió el pGTEV Δ EE con ApaI y Eco81I. Ambos productos se utilizaron para realizar una ligación con la DNA ligasa T4. Tras llevar a cabo la reacción, se electroporaron células de *E. coli* y se plaquearon en medio LB con kanamicina. En este caso tampoco se obtuvieron buenos resultados mediante ligación, por tanto, también se probó el nuevo abordaje experimental mediante ensamblaje Gibson.

En este caso, primero se digirió el pGTEV Δ EE con dos enzimas de restricción que realizan cortes únicos en este plásmido, Eco81I y BlnI. Por otra parte, se diseñaron dos cebadores para amplificar por PCR el fragmento del p15TV3-Cpf1 que contiene el cDNA de las nucleasas Cpf1. Estos cebadores contenían secuencias que permitieron insertar por ensamblaje Gibson el gen de la Cpf1 en el pGTEV Δ EE (Fig. 7).

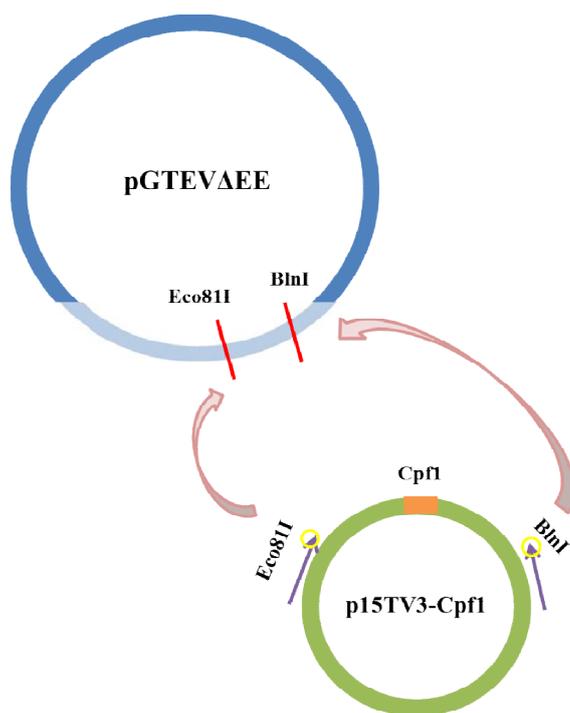


Figura 7. Abordaje experimental para insertar el cDNA de la Cpfl en el pGTEVΔEE a partir de la construcción intermedia p15TV3-Cpfl mediante ensamblaje Gibson.

Siguiendo el nuevo abordaje experimental, se digirió por un lado el pGTEVΔEE con Eco81I y BlnI, y por otra parte se realizaron dos PCRs, una para cada uno de los plásmidos intermedios p15TV3-AsCpfl y p15TV3-LbCpfl, con el mismo par de cebadores (Tabla 3) y 1:15 minutos de extensión, ya que se esperaba obtener productos de PCR de unas 5000 pb. A continuación, y una vez se tuvieron el pGTEVΔEE digerido y los fragmentos de PCR con las Cpfl, se realizó el ensamblaje Gibson.

Tabla 3. Secuencias de los cebadores utilizados para amplificar las nucleasas AsCpfl y LbCpfl mediante PCR a partir de los plásmidos intermedios p15TV3-AsCpfl y p15TV3-LbCpfl.

Fragmento amplificado	Cebador 1	Cebador 2
AsCpfl y LbCpfl	D2564: 5' -GGGACATGATAATTA TTCGCATGCCTAAGGATT TCCCACCATTTC-3'	D2567: 5' -TGTAATGTTTCT CTGTAGACCATAC-3'

Tras varios intentos, el producto de un ensamblaje Gibson de pGTEVΔEE con LbCpfl permitió obtener varias construcciones aparentemente correctas al realizar un análisis electroforético (Fig. 8A). Las bandas del gel de agarosa de este análisis mostraron demasiada intensidad, con lo cual no se podía determinar con exactitud si se había introducido el cDNA de la LbCpfl en el vector pGTEVΔEE. Para posteriormente inocular plantas se han de seleccionar uno o dos plásmidos, y el trabajo que conlleva es mucho, con lo cual se ha de estar muy seguro de que los plásmidos seleccionados son buenos. Así pues, se realizaron una serie de digestiones prueba para comprobarlo. Para ello, se seleccionaron 4 de las construcciones y se hicieron 5 digestiones, una con el plásmido pGTEVΔEE como control y 4 con las construcciones

seleccionadas, con las mismas enzimas de restricción con las que se introdujo la Cpf1 (Eco81I y BlnI). El análisis electroforético de las digestiones mostró resultados positivos para todos los pGTEV Δ EE-LbCpf1 seleccionados, mostrando una banda correspondiente a la nucleasa y otra correspondiente al resto del plásmido (Fig. 8B). De entre estos 4 vectores virales, se seleccionaron dos para proceder a la agroinoculación de plantas.

Ninguno de los ensamblajes Gibson realizados con el pGTEV Δ EE junto a AsCpf1 dieron resultados positivos. No se consiguió insertar esta proteína en el plásmido. Esto puede ser debido a algún error al originar el fragmento con el cDNA de la proteína, en alguno de los pasos anteriores, incluso en los pasos iniciales, o simplemente debido a la mala suerte y el limitado tiempo. Debido a esto, solo se llevó a cabo la inoculación de plantas con la construcción pGTEV Δ EE-LbCpf1. Aun así, se reservaron las muestras de AsCpf1 para futuros proyectos.

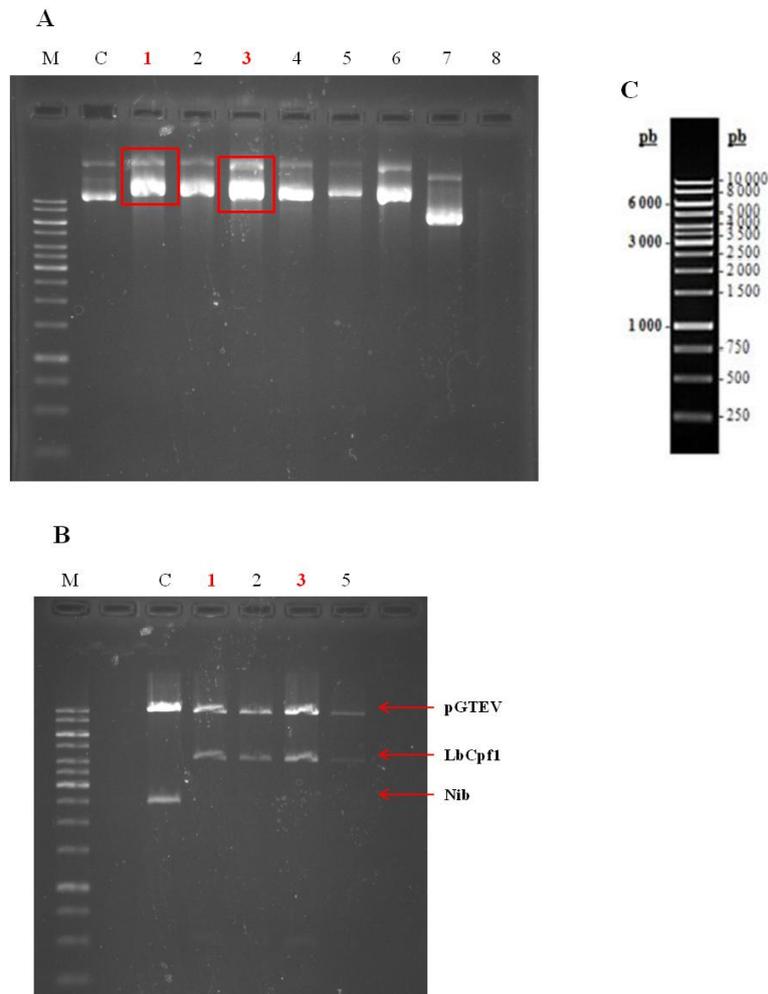


Figura 8. A) Gel de electroforesis de las construcciones pGTEV Δ EE-LbCpf1 y el plásmido pGTEV Δ EE como control. Las construcciones marcadas en rojo (1 y 3) fueron seleccionadas para realizar los pasos siguientes del experimento. B) Digestiones prueba de las construcciones de pGTEV Δ EE-LbCpf1 seleccionadas (1, 2, 3 y 5) para comprobar que realmente han incorporado el inserto de LbCpf1 deseado. La preparación control corresponde con el plásmido pGTEV Δ EE, y en ella se observa una banda de unas 2500 pb correspondiente a la RNA polimerasa viral (Nlb) y otro DNA de mayor tamaño que corresponde al resto del plásmido. En las preparaciones de pGTEV Δ EE-LbCpf1 se observa una banda de unas 4500 pb que corresponde al inserto de LbCpf1, y la banda correspondiente al resto del plásmido que también se observa en la preparación del plásmido control. Las construcciones marcadas en rojo (1 y 3) fueron seleccionadas para realizar los pasos siguientes del experimento. C) Marcador DNA 1 kb Thermo (SM0311).

4.2. Construcción de vectores virales derivados de PVX que expresan el RNA guía

El crRNA se expresó en un PVX. Como este virus se replica en el citoplasma, el crRNA se fusionó a una secuencia correspondiente a un fragmento del viroide del tubérculo fusiforme de la patata (PSTVd) para conseguir que viaje al núcleo. Para ello, se utilizó un vector que contenía dicho virus, el pPVX, el cual contiene un gen de resistencia a kanamicina, y un plásmido intermedio, el pMPVXBa-PdamiBZB, con un gen de resistencia a ampicilina, para facilitar el proceso. Se siguió el mismo procedimiento que con los plásmidos derivados del TEV para la transformación de células de *E. coli* y la purificación, aunque esta vez se transformaron las bacterias con 0.5 µl de preparación de plásmido (-20 °C) y se plaquearon las bacterias en medio LB con kanamicina para bacterias transformadas con el pPVX, o con ampicilina y X-gal para bacterias transformadas con el pMPVXBa-PdamiBZB. Los plásmidos se purificaron por triplicado y se midió la concentración de los mismos con un NanoDrop ND-1000 para así seleccionar uno de cada y continuar con ellos los experimentos (Tabla 4). Se seleccionaron las preparaciones de plásmidos de acuerdo a su concentración, las bandas del gel y la curva resultante al analizar con el NanoDrop.

Tabla 4. Concentraciones de plásmidos pPVX y pMPVXBa-PdamiBZB purificados a partir de diferentes cultivos. Los cultivos seleccionados para continuar con el experimento se muestran en negrita y con un asterisco.

Construcción	Cultivo	Concentración (ng/µl)
pPVX	1*	163
	2	181.6
	3	401.3
pMPVXBa-PdamiBZB	1	66.3
	2*	65.1
	3	62.5

En primer lugar, se insertaron las repeticiones directas para cada una de las Cpf1 en el plásmido intermedio pMPVXBa-PdamiBZB. Para ello, se realizaron dos PCR (1 min de extensión), cada una con el par de cebadores correspondiente a cada una de las nucleasas (Tabla 5). Los cebadores ya llevaban incorporadas las repeticiones directas que se deseaba introducir en el vector, que son las secuencias del RNA guía que reconocen las Cpf1 y entre las cuales se debe situar la secuencia homóloga a la secuencia del genoma que se desea editar; es decir, la secuencia diana del sistema CRISPR.

Tabla 5. Secuencias de los cebadores utilizados para insertar las repeticiones directas (RD) para las nucleasas *AsCpf1* y *LbCpf1* en el plásmido pMPVXBa-PdamiBZB mediante dos reacciones de PCR.

Fragmento amplificado	Cebador 1	Cebador 2
pMPVXBa-PdamiBZB-RDAsCpf1	D2528: 5' -GCGGTCTCCATCTAC AAGAGTAGAAATTTAAAACC CTGTTTTCGGCGGGAAAT-3'	D2529: 5' -CGGGTCTCGTAATTTTC TACTCTTGTAGATTAGGGT TTGTTAAGTTTCCCT-3'
pMPVXBa-PdamiBZB-RDLbCpf1	D2530: 5' -GCGGTCTCCATCTACA CTTAGTAGAAATTTAAAACC CTGTTTTCGGCGGGAA-3'	D2531: 5' -CGGGTCTCGTAATTTTC TACTAAGTGTAGATTAGGG TTTGTAAAGTTTCCC-3'

Una vez analizados los productos de PCR mediante electroforesis en gel de agarosa y viendo que eran correctos, se insertó el gen *LacZ2* entre las dos repeticiones directas introducidas en el plásmido pMPVXBa-PdamiBZB. En realidad, lo que se desea insertar entre las dos repeticiones directas es la secuencia homóloga a la secuencia diana a editar en el genoma del organismo huésped, pero se inserta el gen *LacZ2* para así poder seleccionar las colonias en un medio con X-gal. Esto se llevó a cabo mediante una reacción de ligación para cada uno de los plásmidos con la DNA ligasa T4. Esta ligación se plaqueó en medio LB con ampicilina y X-gal. Se seleccionó una colonia azul para continuar construyendo el crRNA, ya que eran las que habían incorporado un vector con el gen *LacZ2* y por tanto forman el característico precipitado azul.

La secuencia guía para editar el primer exón de la PDS se originó a partir de la hibridación de dos pares de cebadores, un par para formar la secuencia homóloga a la secuencia de la PDS de *N. tabacum*, y otro par para la de *N. benthamiana* (Tabla 6). Mediante la unión de estos dos pares de cebadores se formaron los fragmentos componentes de las secuencias guía deseadas para introducirlas en el vector pMPVXBa-PdamiBZB, entre las dos repeticiones directas, para formar los crRNAs. Por otra parte, se digirió el plásmido pMPVXBa-LacZ2 seleccionado anteriormente con *BsaI*-HF, ya que se debía retirar el gen *LacZ2* e introducir en su lugar la secuencia guía para modificar la PDS en *N. tabacum* y *N. benthamiana*.

Tabla 6. Secuencias de los cebadores utilizados en la hibridación para obtener las guías para modificar mediante el sistema CRISPR-*Cpf1* el gen de la PDS de *N. benthamiana* y *N. tabacum*.

Secuencia guía PDS	Cebador 1	Cebador 2
<i>N. benthamiana</i>	D2542: 5' -AGATTCAATCTTC GGGTCGTGGCACTT-3'	D2543: 5' -ATTAAAGTGCCAC GACCCGAAGATTGA-3'
<i>N. tabacum</i>	D2544: 5' -AGATGGAAGTCAA AGTCAAGATGGTCA-3'	D2545: 5' -ATTATGACCATCT TGACTTTCAGTTCC-3'

Una vez se dispuso de los cebadores hibridados y el pMPVXBa-LacZ2 digerido, se realizaron dos ligaciones, una con cada par de cebadores hibridados, con alícuotas de ambos productos y la DNA ligasa T4, y una ligación control. Las ligaciones se purificaron y se electroporaron células

de *E. coli* con ellas para hacerlas crecer en medio LB con ampicilina y X-gal. En este caso interesa seleccionar colonias blancas para continuar el proceso, ya que éstas habrán perdido el gen LacZ2. De entre los plásmidos de las colonias blancas seleccionadas, se escogió uno con los cebadores hibridados para cada planta, y se continuó con ellos el experimento (Fig. 9).

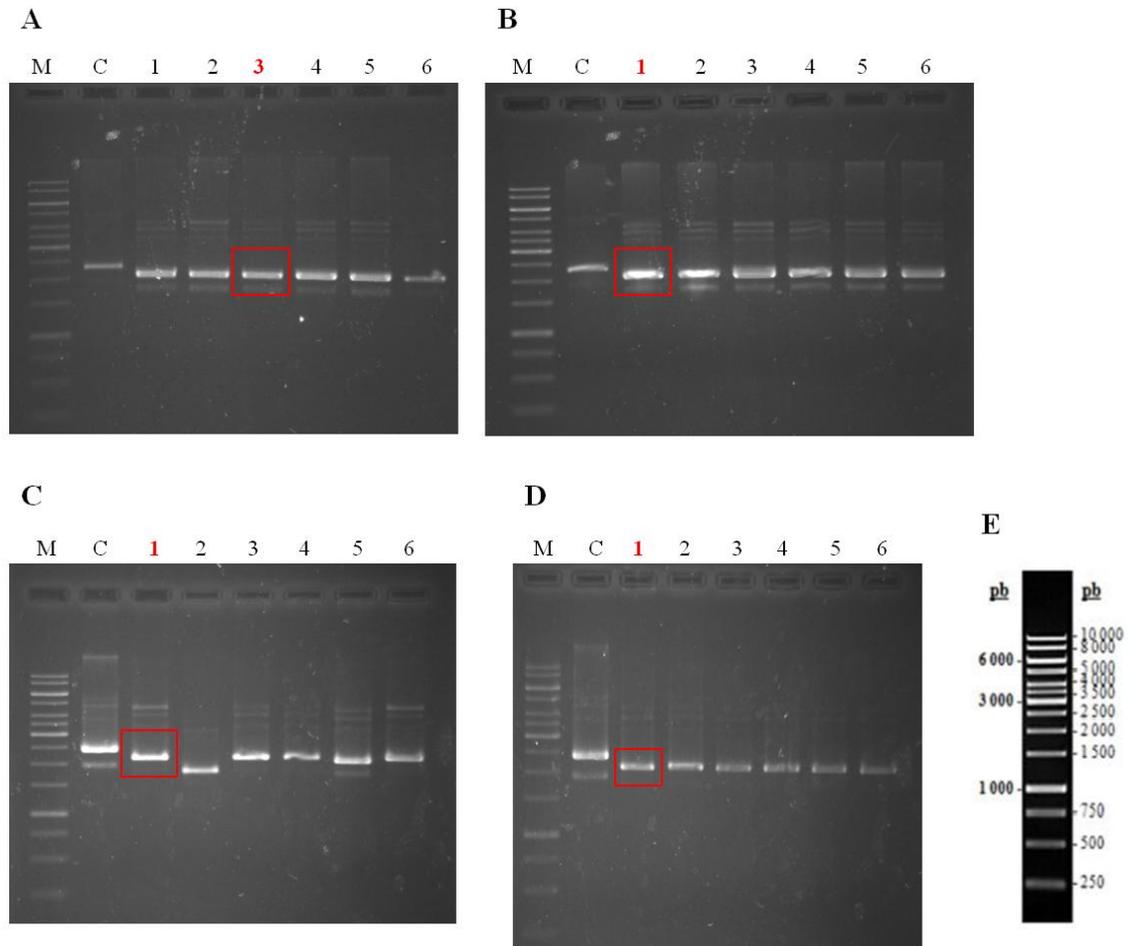


Figura 9. **A)** Gel de electroforesis de las construcciones *pMPVXBa-crRNA-NbPDS-AsCpf1* y el plásmido *pMPVXBa-LacZ2* como control (*NbPDS* corresponde a la *PDS* de *N. benthamiana*). **B)** Gel de electroforesis de las construcciones *pMPVXBa-crRNA-NtPDS-AsCpf1* y el plásmido *pMPVXBa-LacZ2* como control (*NtPDS* corresponde a la *PDS* de *N. tabacum*). **C)** Gel de electroforesis de las construcciones *pMPVXBa-crRNA-NbPDS-LbCpf1* y el plásmido *pMPVXBa-LacZ2* como control. **D)** Gel de electroforesis de las construcciones *pMPVXBa-crRNA-NtPDS-LbCpf1* y el plásmido *pMPVXBa-LacZ2* como control. En todos los casos, los plásmidos marcados en rojo fueron seleccionados para realizar los pasos siguientes del experimento. **E)** Marcador DNA 1 kb Thermo (SM0311).

Para finalmente insertar cada crRNA en el pPVX, en primer lugar, se digirió el pPVX con *ApaI* y *XhoI*. Por otro lado, se hizo una digestión de cada uno de los 4 plásmidos seleccionados anteriormente (Fig. 9) con las mismas enzimas de restricción. Después, se llevaron a cabo las 4 ligaciones con la DNA ligasa T4, y una ligación control a la que únicamente se le añadió el pPVX digerido. Los productos de las ligaciones se plaquearon en medio LB con kanamicina. Un análisis electroforético reveló que los productos de las ligaciones aparentemente eran los esperados, ya que el patrón de bandas de los clones era muy similar al del plásmido control pPVX y el fragmento que se insertó era muy pequeño (Fig. 10 y 11). Sin embargo, el marcador y las muestras cargadas en los geles correspondientes a los clones de pPVX-crRNA-PDS-AsCpf1 migraron de forma extraña, más de lo normal (Fig. 10A y 10B). Por este motivo, se

llevaron a cabo una serie de digestiones con *ApaI* de dos clones de pPVX-crRNA-NbPDS-AsCpf1, dos clones de pPVX-crRNA-NtPDS-AsCpf1, y del pPVX como control. Los productos de las digestiones se cargaron en un gel de agarosa que reveló que, efectivamente, la ligación había salido bien y en las construcciones se había incorporado el RNA guía al pPVX, ya que los productos de las digestiones de los clones eran de tamaño casi igual al pPVX control (Fig. 10C).

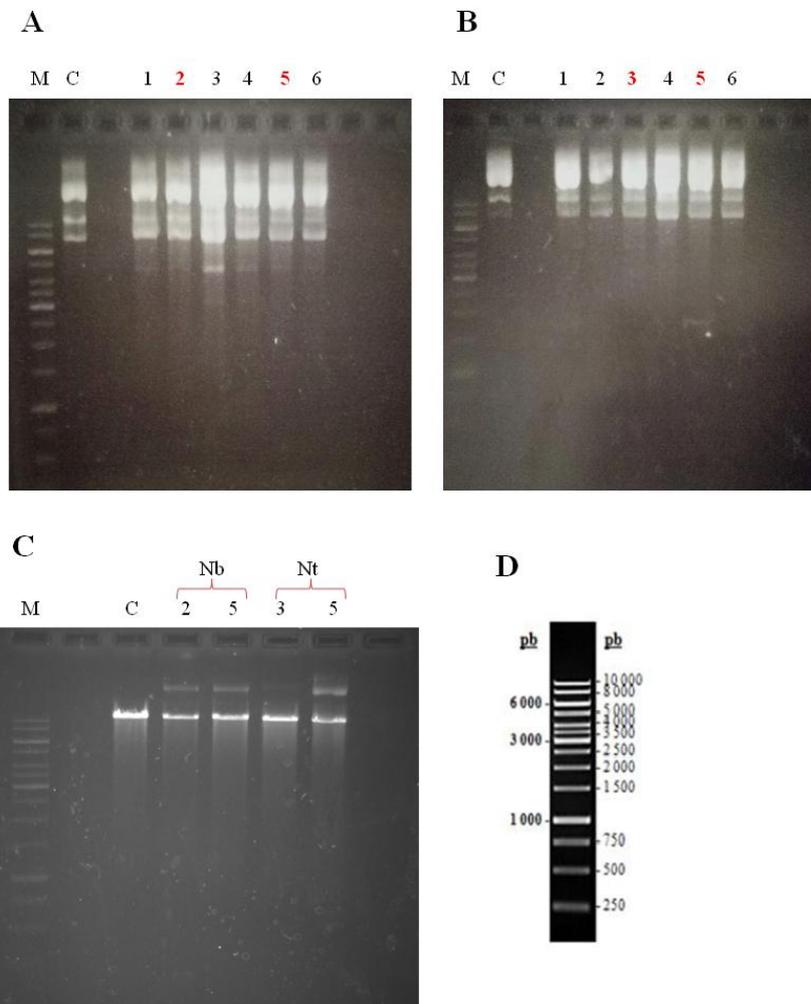


Figura 10. **A)** Gel de electroforesis de las construcciones pPVX-crRNA-NbPDS-AsCpf1 y el plásmido pPVX como control (NbPDS corresponde a la PDS de *N. benthamiana*). Las construcciones marcadas en rojo (2 y 5) fueron seleccionadas para realizar las digestiones prueba y para realizar los siguientes pasos del experimento. **B)** Gel de electroforesis de las construcciones pPVX-crRNA-NtPDS-AsCpf1 y el plásmido pPVX como control (NtPDS corresponde a la PDS de *N. tabacum*). Las construcciones marcadas en rojo (3 y 5) fueron seleccionadas para realizar las digestiones prueba y para realizar los siguientes pasos del experimento. **C)** Digestiones prueba de las construcciones de pPVX-crRNA-NbPDS-AsCpf1 y pPVX-crRNA-NtPDS-AsCpf1 seleccionadas (2 y 5, y 3 y 5, respectivamente) para comprobar que realmente han incorporado el inserto deseado. La preparación control corresponde con el plásmido pPVX linealizado. **D)** Marcador DNA 1 kb Thermo (SM0311).

Debido a que no se pudo construir el vector pGTEVΔEE-AsCpf1, solo se seleccionaron dos clones del pPVX-crRNA-NbPDS-LbCpf1 para agroinocular *N. benthamiana* (Fig. 11), que es una planta más susceptible a las infecciones virales que *N. tabacum*, ya que es mutante en el gen de la RDR (RNA polimerasa RNA dependiente) que hace que la amplificación del silenciamiento de RNA se vea comprometida, y por eso suele usarse como planta modelo en

virología vegetal (Bally et al, 2015; Carbonell, 2015). Por este motivo y debido al límite de tiempo se inoculó solo *N. benthamiana*.

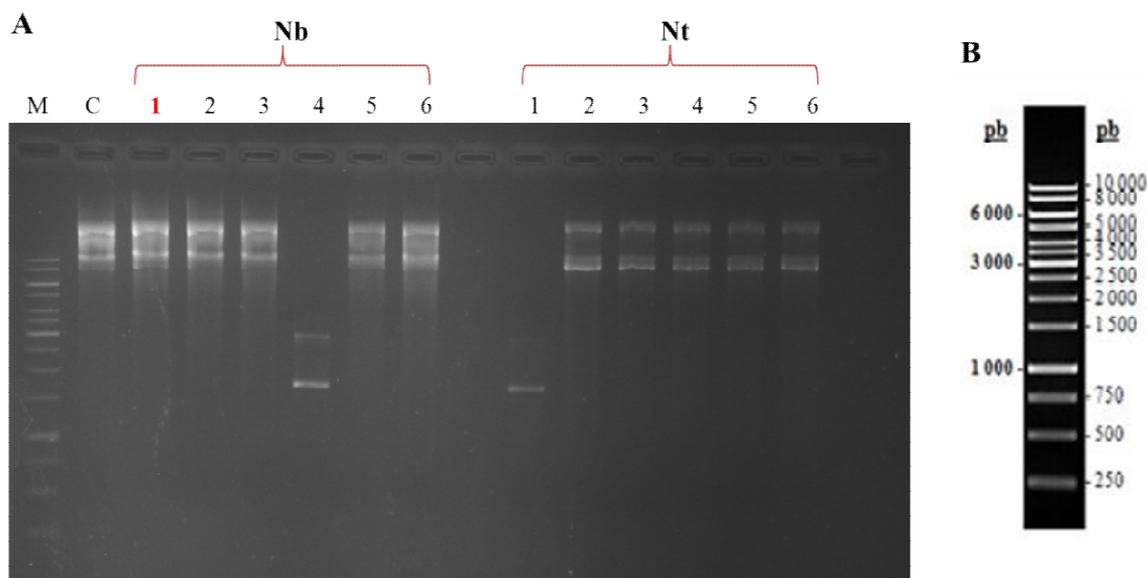


Figura 11. A) Gel de electroforesis de las construcciones *pPVX-crRNA-NbPDS-LbCpf1*, *pPVX-crRNA-NtPDS-LbCpf1* y el plásmido *pPVX* como control (*NbPDS* y *NtPDS* corresponden a la *PDS* de *N. benthamiana* y de *N. tabacum*, respectivamente). El patrón de bandas de todas las construcciones excepto la del *pPVX-crRNA-NbPDS-LbCpf1 4* y el *pPVX-crRNA-NtPDS-LbCpf1 1*, es muy similar al del plásmido control, indicando que probablemente correspondan con la construcción deseada. La construcción marcada en rojo (1) fue seleccionada para inocular plantas. **B)** Marcador DNA 1 kb Thermo (SM0311).

4.3. Agroinoculación de plantas

Se transformó *A. tumefaciens* con los plásmidos *pGTEVΔEE-LbCpf1 1* y *3* por electroporación de la misma forma que *E. coli*. Tras transformar las células, se dejaron crecer los cultivos durante un mínimo de 5 días a 28 °C en medio LB con tetraciclina, rifampicina y kanamicina. Tras el crecimiento de las bacterias, se traspasaron 3 clones de cada plásmido a medio LB líquido con kanamicina. De este total de 6 cultivos, se midió su D.O. y se seleccionaron dos de cada, preferiblemente aquellos que tuvieran una D.O. próxima a 1 (Tabla 7). Los cultivos seleccionados por una parte se glicerizaron para conservarlos, y por otra parte se prepararon para la agroinoculación.

Para los plásmidos de *pPVX-crRNA-NbPDS-LbCpf1 1* y *2* seleccionados anteriormente se siguió el mismo procedimiento, con la diferencia de que estos cultivos tardaron menos tiempo en crecer debido al menor tamaño del plásmido. Cabe destacar también que las D.O. de los cultivos líquidos fueron muy altas, así que se tuvieron que seleccionar aquellos que presentaran una D.O. más baja, y no aquellos con una D.O. próxima a 1 (Tabla 7).

Con los clones derivados de *pGTEVΔEE-LbCpf1* se inocularon 3 plantas de *N. tabacum* de 3 semanas con cada cultivo seleccionado, es decir, un total de 12 plantas, y también se reservaron 3 plantas sin inocular como control.

Tabla 7. Densidades ópticas de los cultivos de *A. tumefaciens* transformadas con los plásmidos pGTEVΔEE-LbCpf1 1 y 3 (Fig. 8), y pPVX-crRNA-NbPDS-LbCpf1 1 (Fig. 11).

Construcción	Cultivo	D.O.
pGTEVΔEE-LbCpf1-1	1*	1
	2*	0.71
	3	-
pGTEVΔEE-LbCpf1-3	1*	1.05
	2	0.77
	3*	1.05
pPVX-crRNA-NbPDS-LbCpf1-1	1*	1.42
	2	<4.00
	3*	3.48

Con los clones derivados de pPVX-crRNA-NbPDS-LbCpf1 se inocularon plantas de *N. benthamiana* de 2 semanas para comprobar su capacidad infectiva. Se agroinocularon 2 plantas con cada cultivo seleccionado, es decir, un total de 8 plantas, y también se reservaron otras 2 plantas sin inocular como control.

Los resultados de la agroinoculación mostraron que ambos vectores virales eran infecciosos. En el caso de las plantas infectadas con el vector viral derivado del TEV, estas presentaron como síntoma característico el empaldecimiento de las venas (Fig. 12). En cuanto a las plantas de *N. benthamiana* inoculadas con el vector viral derivado del PVX, estas mostraron síntomas de encrespamiento y reducción del tamaño de las hojas (Fig. 13).

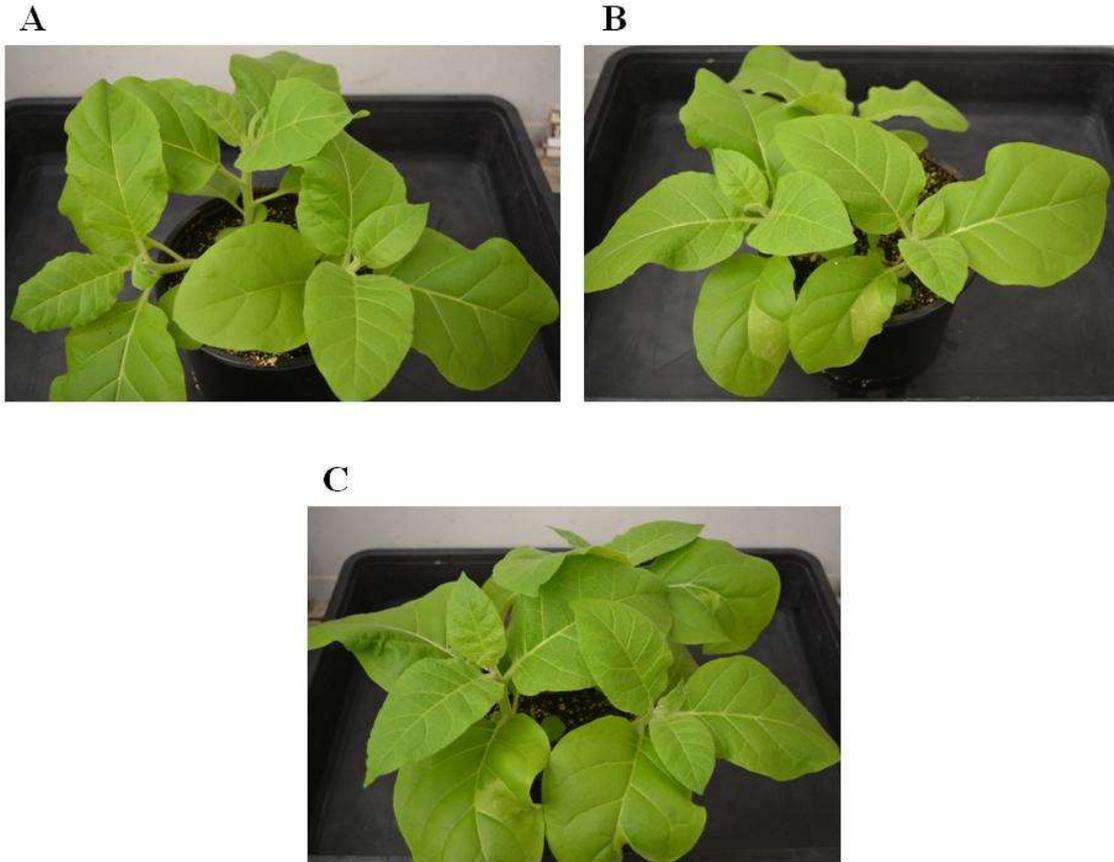


Figura 12. *A)* Plantas de *N. tabacum* control. *B)* Plantas de *N. tabacum* infectadas con el clon viral derivado de pGTEVΔEE-LbCpfI-1. *C)* Plantas de *N. tabacum* infectada con el clon viral derivado de pGTEVΔEE-LbCpfI-3.



Figura63. *A)* Plantas de *N. benthamiana* control. *B)* Plantas de *N. benthamiana* infectadas con el clon viral derivado de pPVX-crRNA-NbPDS-LbCpfI-1.

4.4. Análisis de tejido foliar infectado

Se recogió como muestra una hoja de cada una de las plantas infectadas y de las plantas control. Las hojas se recogieron de la zona baja de la planta, donde antes aparece la infección (Fig. 14). Se pesó un trozo de ~0.5 g de cada hoja y se conservó a -80 °C para futuros experimentos. Además, se depositó una pequeña porción de tejido de la zona sintomática de cada hoja en tubos Eppendorf de 2 ml para la realización de las RT-PCR.

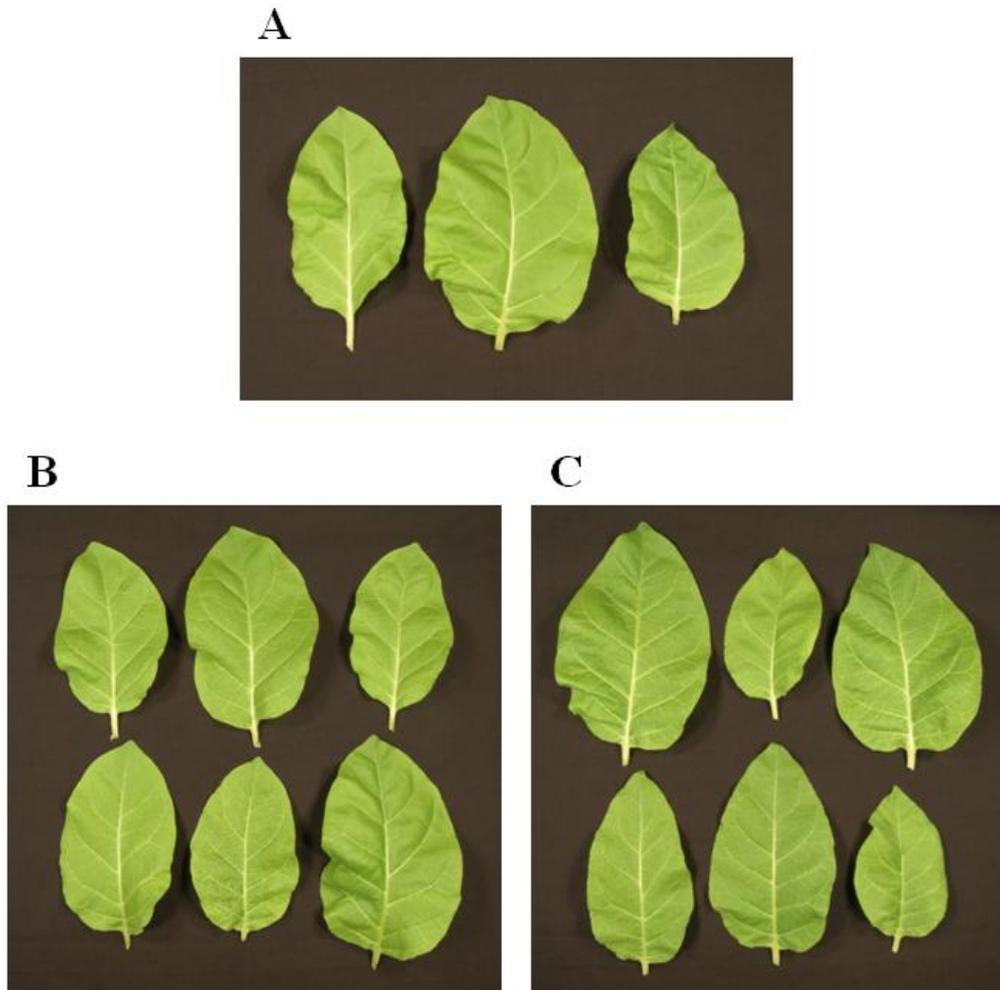


Figura 14: A) Hojas de las plantas de *N. tabacum* control. B y C) Hojas de las plantas de *N. tabacum* infectadas con los clones virales derivados de (B) pGTEVΔEE-LbCpf1-1 y (C) pGTEVΔEE-LbCpf1-3, 7 días después de la inoculación.

Posteriormente, se extrajo el RNA de las muestras de tejido foliar guardadas en los tubos Eppendorf de las plantas de *N. tabacum* control y las infectadas con el pGTEVΔEE-LbCpf1 y se realizó una RT-PCR con los cebadores D179, D1835 y D1949 (Tabla 8) para comprobar si el virus había conservado el inserto correspondiente a la Cpf1 o no. Las muestras de tejido infectado con pPVX-crRNA-NbPDS-LbCpf1 se conservaron a -80 °C para realizar el análisis de la RT-PCR en otra ocasión. Un análisis electroforético reveló que el vector viral había perdido el inserto de la proteína durante la infección (Fig. 15). El gel muestra bandas de distintos tamaños y ninguna se corresponde con el producto esperado en el caso de conservar el cDNA de Cpf1.

Tabla 8. Secuencias de los cebadores utilizados para la RT-PCR del vector pGTEV Δ EE-LbCpf1.

Cebador RT	Cebador 1 PCR	Cebador 2 PCR
D179: 5'-CTCGCACTACA TAGGAGAATTAGAC-3'	D1835: 5'-ATTCTGGAAGC ATTGGATTCAA-3'	D1949: 5'-TTGATCTTTCT TCTTACCAGCG-3'

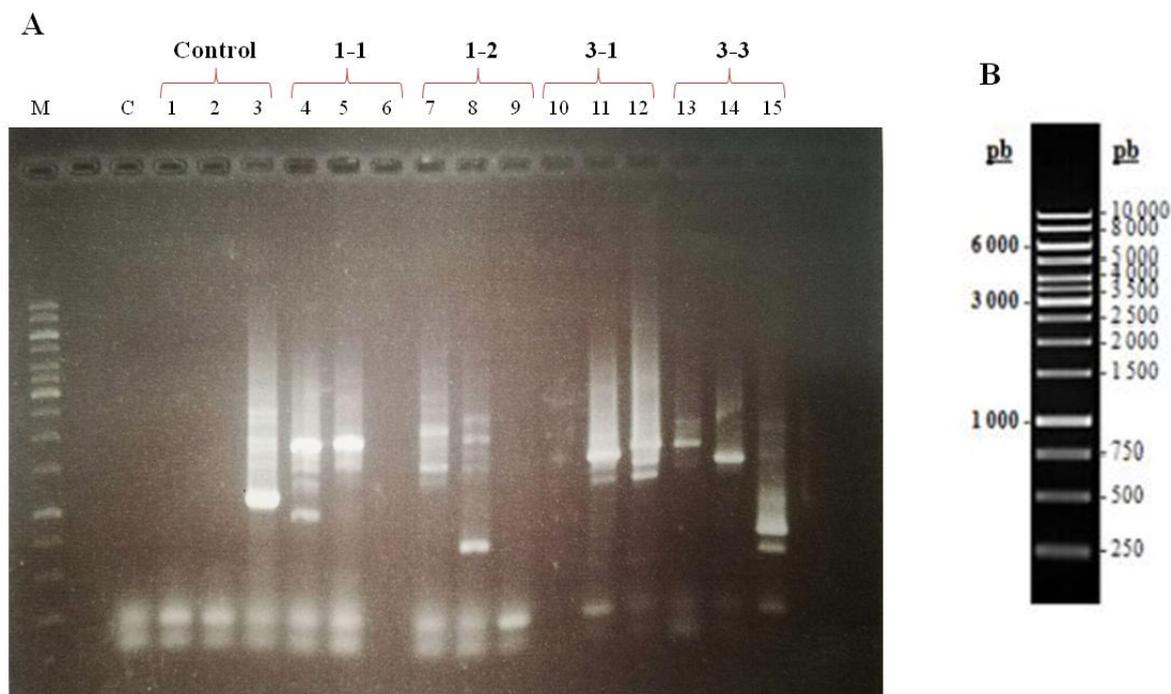


Figura 15. A) Gel de electroforesis de las RT-PCR realizadas a partir de las muestras de RNA de las hojas de *N. tabacum* recogidas tras la inoculación con pGTEV Δ EE-LbCpf1 1 y 3, y las muestras de las plantas control. Las preparaciones 6 y 10 se perdieron durante la realización de la PCR, y la 1 no presenta ningún cDNA. B) Marcador DNA 1 kb Thermo (SM0311).

Después de todo este trabajo, faltaría co-infiltrar plantas con los dos vectores virales construidos para así comprobar si el sistema CRISPR construido es funcional. El virus puede haber perdido el inserto con el gen de la LbCpf1 debido a la presión de selección opuesta por la planta de *N. tabacum*, por ejemplo, mediante los mecanismos de silenciamiento por RNA (Kościańska *et al.*, 2005). Debido a esto, se debería de probar la infiltración de plantas de *N. benthamiana* con los mismos vectores virales, que, a pesar de que también poseen los mismos mecanismos de silenciamiento (Kościańska *et al.*, 2005), suelen ser plantas más débiles y ejercen una menor presión de selección. Esto se debe a que la línea de *N. benthamiana* que se utiliza normalmente en los laboratorios de investigación es un mutante natural para el gen de la RDR (RNA polimerasa RNA dependiente), que se recogió casualmente de un ambiente extremo de Australia central. Esto afecta al mecanismo de silenciamiento de RNA y hace a la planta susceptible a un gran número de especies de virus vegetales de muchos grupos taxonómicos diferentes. En consecuencia, esta especie se utiliza frecuentemente como planta modelo en muchos proyectos de investigación de virología vegetal (Bally *et al.*, 2015; Carbonell, 2015). Además de eso, en nuestro grupo de investigación se está trabajando en la obtención de plantas de *N. benthamiana* transgénicas que expresan la NIB viral y así podrán complementar en *trans* a los vectores TEV Δ NIB a los que se les ha insertado el gen de la Cpf1. Si el virus siguiera perdiendo el inserto, se podría plantear como solución utilizar plantas transgénicas en las cuales se vean afectados todavía más los mecanismos de silenciamiento (Voinnet *et al.*, 2003).

A pesar de las dificultades, parece prometedor seguir apostando por el desarrollo de este sistema CRISPR-Cpf1 basado en vectores virales, ya que el sistema CRISPR-Cpf1 ya ha sido utilizado con éxito para modificar genomas de plantas (Endo *et al.*, 2016), incluso con la nucleasa LbCpf1 específicamente ya se ha modificado el gen de la PDS en arroz (Xu *et al.*, 2016). Los vectores virales se caracterizan por la velocidad a la que son capaces de alcanzar altos niveles de expresión de proteínas exógenas, y últimamente ya se están utilizando para expresar proteínas del sistema CRISPR-Cas (Zetsche *et al.*, 2017; Ortinski *et al.*, 2017).

La verdadera utilidad práctica de este sistema CRISPR para modificar una secuencia específica de un determinado gen de una planta, se basa en generar mutantes estables con características que pueden resultar de interés en sectores como el de la agricultura, el ornamental o el agroalimentario. De este modo podrían modificarse plantas, por ejemplo, silenciando genes, aumentando el nivel de expresión de otros genes de interés como los de resistencia a enfermedades o aquellos que disminuyen el estrés hídrico o salino, o incluso insertar genes exógenos para añadir alguna característica de interés a una planta.

5. CONCLUSIONES

1. Se ha construido con éxito un vector viral derivado del TEV que lleva insertado el gen de la nucleasa LbCpf1 en el lugar del gen de la RNA polimerasa viral (NIb).
2. Este vector viral es capaz de infectar plantas de *Nicotianatabacum* produciendo síntomas fácilmente detectables.
3. El análisis de las plantas de *N. tabacum* inoculadas con este vector viral mostró que el virus pierde el inserto de la LbCpf1 7 días después de la inoculación debido a la presión de selección ejercida por la planta. Inoculando *Nicotiana benthamiana* con este vector viral puede que se obtengan mejores resultados.
4. Se han construido con éxito cuatro vectores virales derivados del PVX que llevan insertado un crRNA. Los insertos han resultado de la combinación de las repeticiones directas específicas para las nucleasas AsCpf1 y LbCpf1, con las secuencias guía para modificar el gen de la fitoeno desaturasa de *N. tabacum* y *N. benthamiana*.

6. BIBLIOGRAFÍA

Bally, J., Nakasugi, K., Jia, F., Jung, H., Ho, S. Y., Wong, M., Paul, C. M., Naim, F., Wood, C. C., Crowhurst, R. N., Hellens, R. P., Dale, J. L., y Waterhouse, P. M. (2015). The extremophile *Nicotiana benthamiana* has traded viral defence for early vigour. *Nature plants*, *1*, 15165.

Bedoya, L. C., y Daròs, J. A. (2010). Stability of Tobacco etch virus infectious clones in plasmid vectors. *Virus research*, *149*(2), 234-240.

Bedoya, L., Martínez, F., Rubio, L., y Daròs, J. A. (2010). Simultaneous equimolar expression of multiple proteins in plants from a disarmed potyvirus vector. *Journal of biotechnology*, *150*(2), 268-275.

Carbonell, A. (2015). Molecular Ecology: Trading defence for vigour. *Nature plants*, *1*, 15174.

Carrington, J. C., Cary, S. M., Parks, T. D., y Dougherty, W. G. (1989). A second proteinase encoded by a plant potyvirus genome. *The EMBO journal*, *8*(2), 365.

Casacuberta, J. y Puigdomènech, P (2017). Las nuevas herramientas de edición genómica y la mejora genética de plantas. *SEBBM*. <<http://www.sebbm.es/revista/articulo.php?id=270&url=las-nuevas-herramientas-de-edicion-genomica-y-la-mejora-genetica-de-plantas>>[Consulta: 28 de mayo de 2017].

Chapman, S., Kavanagh, T., y Baulcombe, D. (1992). Potato virus X as a vector for gene expression in plants. *The Plant Journal*, *2*(4), 549-557.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., y Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*(6121), 819-823.

Daròs, J. A., Schaad, M. C., y Carrington, J. C. (1999). Functional analysis of the interaction between VPg-proteinase (NIa) and RNA polymerase (NIb) of tobacco etch potyvirus, using conditional and suppressor mutants. *Journal of virology*, *73*(10), 8732-8740.

Doudna, J. A., y Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213), 1258096.

Endo, A., Masafumi, M., Kaya, H., y Toki, S. (2016). Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from *Francisella novicida*. *Scientific reports*, *6*.

EUROPEAN COMMISSION. New plant breeding techniques - Food Safety. <https://ec.europa.eu/food/plant/gmo/legislation/plant_breeding_en> [Consulta: 28 de mayo de 2017].

Gleba, Y., Klimyuk, V., y Marillonnet, S. (2007). Viral vectors for the expression of proteins in plants. *Current Opinion in Biotechnology*, *18*(2), 134-141.

Hryhorowicz, M., Lipiński, D., Zeyland, J., y Słomski, R. (2016). CRISPR/Cas9 Immune System as a Tool for Genome Engineering. *Archivum Immunologiae et Therapiae Experimentalis*, 1-8.

Jia, H., y Wang, N. (2014). Targeted genome editing of sweet orange using Cas9/sgRNA. *PLoS one*, *9*(4), e93806.

- Kasschau, K. D., y Carrington, J. C. (1998). A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. *Cell*, 95(4), 461-470.
- Krishnamurthy, K., Mitra, R., Payton, M. E., y Verchot-Lubicz, J. (2002). Cell-to-cell movement of the PVX 12K, 8K, or coat proteins may depend on the host, leaf developmental stage, and the PVX 25K protein. *Virology*, 300(2), 269-281.
- Koenig, R., y Lesemann, D. E. (1989). Potato virus X. *Descriptions of plant viruses*, (385).
- Komarova, T. V., Baschieri, S., Donini, M., Marusic, C., Benvenuto, E., y Dorokhov, Y. L. (2010). Transient expression systems for plant-derived biopharmaceuticals. *Expert Review of Vaccines*, 9(8), 859-876.
- Kościańska, E., Kalantidis, K., Wypijewski, K., Sadowski, J., y Tabler, M. (2005). Analysis of RNA silencing in agroinfiltrated leaves of *Nicotiana benthamiana* and *Nicotiana tabacum*. *Plant molecular biology*, 59(4), 647-661.
- Legorreta Herrera, M., Martínez Flores, F., Hernández Sánchez, F., y Zentella Dehesa, A. (2012). Los vectores virales y la transgénesis. *Revista Especializada en Ciencias de la Salud*, 15(1), 5-14.
- Li, X. H., y Carrington, J. C. (1995). Complementation of tobacco etch potyvirus mutants by active RNA polymerase expressed in transgenic cells. *Proceedings of the National Academy of Sciences*, 92(2), 457-461.
- Lico, C., Benvenuto, E., y Baschieri, S. (2015). The two-faced Potato Virus X: from plant pathogen to smart nanoparticle. *Frontiers in plant science*, 6.
- López Mancheño, Y. (2015). *Ingeniería genómica mediante sistemas CRISPR-Cas*. Trabajo fin de Grado. Universidad de Alicante.
- Majer, E., Salvador, Z., Zwart, M. P., Willemsen, A., Elena, S. F., y Daròs, J. A. (2014). Relocation of the *nib* gene in the tobacco etch potyvirus genome. *Journal of virology*, 88(8), 4586-4590.
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., Backofen, R., y Koonin, E. V. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews. Microbiology*, 13(11), 722.
- Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V., y Gleba, Y. (2005). Systemic *Agrobacterium tumefaciens*–mediated transfection of viral replicons for efficient transient expression in plants. *Nature biotechnology*, 23(6), 718-723.
- Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J., y Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, 155(3), 733-740.
- Ortinski, P. I., O'Donovan, B., Dong, X., y Kantor, B. (2017). Integrase-Deficient Lentiviral Vector as an All-in-One Platform for Highly Efficient CRISPR/Cas9-Mediated Gene Editing. *Molecular Therapy-Methods y Clinical Development*, 5, 153-164.
- Restrepo, M. A., Freed, D. D., y Carrington, J. C. (1990). Nuclear transport of plant potyviral proteins. *The Plant Cell*, 2(10), 987-998.

Ruiz, M. T., Voinnet, O., y Baulcombe, D. C. (1998). Initiation and maintenance of virus-induced gene silencing. *The Plant Cell*, 10(6), 937-946.

Tang, X., Lowder, L. G., Zhang, T., Malzahn, A. A., Zheng, X., Voytas, D. F., Zhong, Z., Chen, Y., Ren, Q., Li, Q., Kirkland, E. R., Zhang, Y., y Qi, Y. (2017). A CRISPR–Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nature Plants*, 3, 17018.

Velasquez, N., Hossain, M. J., y Murphy, J. F. (2015). Differential disease symptoms and full-length genome sequence analysis for three strains of Tobacco etch virus. *Virus genes*, 50(3), 442-449.

Voinnet, O., Rivas, S., Mestre, P., y Baulcombe, D. (2003). Retracted: an enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant Journal*, 33(5), 949-956.

Xu, R., Qin, R., Li, H., Li, D., Li, L., Wei, P., y Yang, J. (2016). Generation of targeted mutant rice using a CRISPR–Cpf1 system. *Plant Biotechnology Journal*.

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., van der Oost, J., Regev, A., Koonin, E. V., yZhang, F. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 163(3), 759-771.

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., Winblad, N., Choudhury, S. R., Abudayyed, O. O., Gootenberg, J. S., Wu, W. Y., Scott, D. A., Severinov, K., van der Oost, J., y Zhang, F. (2017). Multiplex gene editing by CRISPR–Cpf1 using a single crRNA array. *Nat. Biotechnol.*, 35, 31-34.

7. ANEXO: SECUENCIAS

>pY010 (pcDNA3.1-hAsCpf1), secuencia de hAsCpf1 en mayúscula **negrita**

gacggatcgggagatctcccgatcccctatggtgcaactctcagtacaatctgctctgatgccg
atagttaagccagtatctgctccctgcttggtgtgtggtggaggtcgtgagtagtgcgcgagcaaa
atthaagctacaacaaggcaaggcttgaccgacaattgcatgaagaatctgcttagggtaggc
gttttgcgctgcttcgcgatgtacggggccagatatacgcgcttgacattgattattgactagtta
ttaatagtaatcaattacggggtcattagttcatagccatataatggagttccgcggttacataa
cttacggtaaatggcccgcctggctgaccgcccacgacccccgccattgacgtcaataatga
cgtatggtcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacg
gtaactgcccacttggcagtacatcaagtgtatcatatgccagtagcccccctattgacgtc
aatgacggtaaatggcccgcctggcattatgccagtagacacattatgggactttcctactt
ggcagtagacatctacgtatttagtcatcgctattaccatggtgatgcggttttggcagtagacaa
tgggctggtgatagcgggttgactcacggggatttccaagtctccaccccattgacgtcaatggg
agtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtacaactccgccccattga
cgcaaatgggctgtaggctgtacgggtgggaggtctatataagcagagctctctggctaactag
agaaccactgcttactggcttatcgaaattaatacgaactcactatagggagacccaagctggc
tagcgtttaaacttaagcttggtagccacc**ATGACACAGTTCGAGGGCTTTACCAACCTGTA**
TCAGGTGAGCAAGACTGCGTTTTGAGCTGATCCCACAGGGCAAGACCCTGAAGCACATCCAG
GAGCAGGGCTTCATCGAGGAGGACAAGGCCCGCAATGATCACTACAAGGAGCTGAAGCCCATCA
TCGATCGGATCTACAAGACCTATGCCGACCAGTGCCTGCAGCTGGTGCAGCTGGATTGGGAGAA
CCTGAGCGCCGCCATCGACTCCTATAGAAAAGGAGAAAACCGAGGAGACAAGGAACGCCCTGATC
GAGGAGCAGGCCACATATCGCAATGCCATCCACGACTACTTCATCGGCCGGACAGACAACCTGA
CCGATGCCATCAATAAGAGACACGCCGAGATCTACAAGGGCCTGTTCAAGGCCGAGCTGTTTAA
TGGCAAGGTGCTGAAGCAGCTGGGCACCGTGACCACAACCGAGCACGAGAACGCCCTGCTGCCG
AGCTTCGACAAGTTTTACAACCTACTTCTCCGGCTTTTATGAGAACAGGAAGAACGTGTTTCAGCG
CCGAGGATATCAGCACAGCCATCCCACACCGCATCGTGCAGGACAACCTCCCAAGTTTAAGGA
GAATTGTACATCTTCACACGCCTGATCACCGCCGTGCCAGCCTGCGGGAGCACTTTGAGAAC
GTGAAGAAGGCCATCGGCATCTTCGTGAGCACCTCCATCGAGGAGGTGTTTTCTTCCCTTTTT
ATAACCAGCTGCTGACACAGACCCAGATCGACCTGTATAACCAGCTGCTGGGAGGAATCTCTCG
GGAGGCAGGCACCGAGAAGATCAAGGGCTGAACGAGGTGCTGAATCTGGCCATCCAGAAGAAT
GATGAGACAGCCACATCATCGCCTCCCTGCCACACAGATTTCATCCCCCTGTTTAAAGCAGATCC
TGTCCGATAGGAACACCCTGTCCTTTCATCTGGAGGAGTTTAAAGAGCGACGAGGAAGTGATCCA
GTCCTTCTGCAAGTACAAGACTGCTGAGAAACGAGAACGTGCTGGAGACAGCCGAGGCCCTG
TTTAAACGAGCTGAACAGCATCGACCTGACACACATCTTCATCAGCCACAAGAAGCTGGAGACAA
TCAGCAGCGCCCTGTGCGACCCTGGGATACACTGAGGAATGCCCTGTATGAGCGGAGAATCTC
CGAGCTGACAGGCAAGATCACCAAGTCTGCCAAGGAGAAGGTGCAGCGCAGCCTGAAGCACGAG
GATATCAACCTGCAGGAGATCATCTCTGCCGACGCAAGGAGCTGAGCGAGGCCCTTCAAGCAGA
AAACCAGCGAGATCCTGTCCACGCACACGCCGCCCTGGATCAGCCACTGCCTACAACCCTGAA
GAAGCAGGAGGAGAAGGAGATCCTGAAGTCTCAGCTGGACAGCCTGCTGGGCCTGTACCACCTG
CTGGACTGGTTTGCCGTGGATGAGTCCAACGAGGTGGACCCCGAGTTCTCTGCCCGCTGACCG
GCATCAAGCTGGAGATGGAGCCTTCTCTGAGCTTCTACAACAAGGCCAGAAATTATGCCACCAA
GAAGCCCTACTCCGTGGAGAAGTTCAAGCTGAACCTTTCAGATGCCTACACTGGCCTCTGGCTGG
GACGTGAATAAGGAGAAGAACAATGGCGCCATCCTGTTGTGAAGAACGGCCTGTACTATCTGG
GCATCATGCCAAAGCAGAAGGGCAGGTATAAGGCCCTGAGCTTCGAGCCCACAGAGAAAACCAG
CGAGGGCTTTGATAAGATGTACTATGACTACTTCCCTGATGCCGCCAAGATGATCCCAAAGTGC
AGCACCCAGCTGAAGGCCGTGACAGCCACTTTCAGACCCACACAACCCCATCCTGCTGTCCA
ACAATTTTCATCGAGCCTCTGGAGATCACAAGGAGATCTACGACCTGAACAATCCTGAGAAGGA
GCCAAAGAAGTTTCAGACAGCCTACGCCAAGAAAACCGGCCGACCAGAAGGGCTACAGAGAGGCC
CTGTGCAAGTGGATCGACTTCACAAGGGATTTTCTGTCCAAGTATAACCAAGACAACCTCTATCG
ATCTGTCTAGCCTGCCGCCATCCTCTCAGTATAAGGACCTGGGGCAGTACTATGCCGAGCTGAA
TCCCCTGCTGTACCACATCAGCTTCCAGAGAATCGCCGAGAAGGAGATCATGGATGCCGTGGAG
ACAGGCAAGCTGTACCTGTTCCAGATCTATAACAAGGACTTTGCCAAGGGCCACCACGGCAAGC
CTAATCTGCACACACTGTATTGGACCGCCCTGTTTTCTCCAGAGAACCTGGCCAAGACAAGCAT

CAAGCTGAATGGCCAGGCCGAGCTGTTCTACCGCCCTAAGTCCAGGATGAAGAGGATGGCACAC
CGGCTGGGAGAGAAGATGCTGAACAAGAAGCTGAAGGATCAGAAAACCCCAATCCCCGACACCC
TGTACCAGGAGCTGTACGACTATGTGAATCACAGACTGTCCACGACCTGTCTGATGAGGCCAG
GGCCCTGCTGCCAACGTGATCACCAAGGAGGTGTCTCACGAGATCATCAAGGATAGGCGCTTT
ACCAGCGACAAGTTCTTTTTCCACGTGCCTATCACACTGAACTATCAGGCCGCCAATTTCCCAT
CTAAGTTCAACCAGAGGGTGAATGCCTACCTGAAGGAGCACCCGAGACACCTATCATCGGCAT
CGATCGGGGCGAGAGAAACCTGATCTATATCACAGTGATCGACTCCACCGGAAGATCCTGGAG
CAGCGGAGCCTGAACACCATCCAGCAGTTTGATTACCAGAAGAAGCTGGACAACAGGGAGAAGG
AGAGGGTGGCAGCAAGGCAGGCCTGGTCTGTGGTGGGCACAATCAAGGATCTGAAGCAGGGCTA
TCTGAGCCAGGTCATCCACGAGATCGTGGACCTGATGATCCACTACCAGGCCGTGGTGGTGTCTG
GAGAACCTGAAATTCGGCTTTAAGAGCAAGAGGACCGGCATCGCCGAGAAGGCCGTGTACCAGC
AGTTCGAGAAGATGCTGATCGATAAGCTGAATTGCCTGGTGTCTGAAGGACTATCCAGCAGAGAA
AGTGGGAGGCGTGTGAACCATAACCAGCTGACAGACCAGTTCACCTCCTTTGCCAAGATGGGC
ACCCAGTCTGGCTTCCTGTTTTACGTGCCTGCCCCATATACATCTAAGATCGATCCCCTGACCG
GCTTCGTGGACCCCTTCGTGTGGAAAACCATCAAGAATCACGAGAGCCGCAAGCACTTCCTGGA
GGGCTTCGACTTTCTGCACCTACGACGTGAAAACCGGCGACTTCATCCTGCACTTTAAGATGAAC
AGAAATCTGTCTTCAGAGGGGCCTGCCCGCTTTATGCCTGCATGGGATATCGTGTTCGAGA
AGAACGAGACACAGTTTGACGCCAAGGGCACCCCTTTCATCGCCGGAAGAGAATCGTGCCAGT
GATCGAGAATCACAGATTCACCGGCAGATACCGGGACCTGTATCCTGCCAACGAGCTGATCGCC
CTGCTGGAGGAGAAGGGCATCGTGTTCAGGGATGGCTCCAACATCCTGCCAAGCTGCTGGAGA
ATGACGATTCTCACGCCATCGACACCATGGTGGCCCTGATCCGCAGCGTGTGCAGATGCGGAA
CTCCAATGCCGCCACAGGCGAGGACTATATCAACAGCCCCGTGCGCGATCTGAATGGCGTGTGC
TTCGACTCCCGTTTTAGAACCCAGAGTGGCCATGGACGCCGATGCCAATGGCGCCTACCACA
TCGCCCTGAAGGGCCAGCTGCTGCTGAATCACCTGAAGGAGAGCAAGGATCTGAAGCTGCAGAA
CGGCATCTCCAATCAGGACTGGCTGGCCTACATCCAGGAGCTGCGCAACAAAAGGCCGGCGGCC
ACGAAAAGGCCGGCCAGGCCAAAAGAAAAGGGATCCTACCCATACGATGTTCCAGATTACG
CTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCC**TAA**Gaatt
ctgcagatatccagcacagtgggcgccgctcgagtctagagggcccgtttaaaccgctgatca
gcctcgactgtgccttctagttgccagccatctgttgttgcccctccccgctgccttcttga
ccctggaaggtgccactcccactgtcctttcctaataaaaatgaggaaatgcatcgcatgtct
gagtaggtgtcattctattctggggggtgggggtggggcaggacagcaagggggaggattgggaa
gacaatagcaggcatgctggggatgctgggtgggctctatggcttctgaggcggaaagaaccagct
ggggctctaggggggatccccacgcgccctgtageggcgcatlaagcgcggcgggtgtgggtgt
tacgcgcagcgtgaccgctacacttgccagcgccttagcgcgccgctcctttcgtttcttccct
tcctttctcgccaggttcgcccgtttccccgtcaagctctaaatcgggggctccctttaggt
tccgatttagtgctttacggcacctcgacccccaaaaaacttgattaggtgatggttcacgtag
tggccatcgccctgatagacggtttttcgccccttgacggttgagtcacgcttctttaaag
ggactcctgttccaaactggaacaacactcaaccctatctcggctctattcttttgattataag
ggatthtggcagatthcggcctattgggttaaaaaatgagctgatttaacaaaaatthaacgcgaa
ttaattctgtggaatgtgtgtcagttaggtgtggaaagtccccaggctccccagcaggcagaa
gtatgcaaagcatgcatctcaattagtcagcaaccagggtgtggaaagtccccaggctccccagc
aggcagaagatgcaaagcatgcatctcaattagtcagcaaccatagtcggccccctaactccg
ccatccccgcccctaactccgcccagttccgcccattctccgccccatggctgactaatttttt
ttatthtatgcagaggccgaggccgctctgctctgagctattccagaagtagtgaggaggctt
ttttggaggcctaggcttttgcaaaaagctcccgaggcttgatataatccatthtccgatctgat
caagagacaggatgaggatcgthtgcagatgattgaacaagatggattgcacgcagggttctccgg
ccgcttgggtggagaggctattcggctatgactgggcacaacagacaatcggctgctctgatgc
cgccgtgttccggctgtcagcgcaggggcgcccggttctthtgtcaagaccgacctgtccggt
gcctgaatgaactgcaggacgaggcagcgcggctatcgtggctggccacgacggggcgttccct
gcgagctgtgctcgacggtgtcactgaagcgggaagggactggctgctattgggcgaagtgc
ggggcaggatctcctgtcatctcaccttgcctcctgcccagaaagtatccatcatggctgatgca
atgcggcggctgcatacgttgatccggctacctgcccattcgaccaccaagcgaacatcgca
tcgagcagcagcactcggatggaagccggtcttctgctcagatcaggatgatctggacgaagagca
tcaggggctcgcgccagccgaactgttcgcccaggctcaaggcgcgcatgcccagcggcaggat
ctcgtcgtgacctatggcgatgctgcttgcgcaatatcatgggtggaaaatggccgctthtctg
gattcatcgactgtggccggctgggtgtggcggaccgctatcaggacatagcgttggctaccg

tgatattgctgaagagcttggcggcgaatgggctgaccgcttcctcgtgctttacgggatcgcc
gctcccgattcgcagcgcacgccttctatcgccttcttgacgagttcttctgagcgggactct
ggggttcgaaatgaccgaccaagcgcacgccaacctgccatcacgagatttcgattccaccgcc
gccttctatgaaaggttgggcttcggaatcgttttccgggacgcgggctggatgatcctccagc
gcggggatctcatgctggagttcttcgccaccccaactgtttattgcagcttataatggtta
caaataaagcaatagcatcacaatttcacaataaagcatttttttactgcattctagttgt
ggtttgtccaaactcatcaatgtatcttatcatgtctgtataaccgctcgacctctagctagagct
tggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaa
catacgagccggaagcataaagtgtaaagcctggggcctaatgagtgagctaactcacatta
attgcggttgcgctcactgcccgccttccagtcgggaaacctgtcgtgccagctgcattaatgaa
tcggccaacgcgcggggagaggcgggttgcgtattgggcgctcttccgcttcctcgcctactga
ctcgcctgcgctcggctcgttcggctgcggcgcgagcggatcagctcactcaaaggcggtaatacgg
ttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggcca
ggaaccgtaaaaaggccgcgcttgcctggcgttttccataggtccgccccctgacgagcatca
caaaaatcgacgctcaagtcaaggtggcgaaacccgcagaggactataaagataaccaggcgtt
ccccctggaagctccctcgtgcgctctcctgttccgacctgccgcttaccggatacctgtccg
cctttctccctcgggaagcgtggcgccttctcatagctcacgctgtaggtatctcagttcggg
gtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgctcagcccagccgctgcgcc
ttatccggtaactatcgtcttgagtcacaacccggtaagacacgacttatcgccactggcagcag
ccactggtaacaggattagcagagcgcaggtatgtaggcgggtgtacagagttcttgaagtggg
gcctaactacggctacactagaagaacagtatgttggatctgcgctctgctgaagccagttacc
ttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcgggttttttg
tttgcaagcagcagattacgcgcagaaaaaaggatctcaagaagatcctttgatcttttctac
ggggtctgacgctcagtggaacgaaaactcacgtaagggattttggatcatgagattatcaaaa
aggatcttcacctagatccttttaaatataaaatgaagtttaaatcaatctaaagtatatatg
agtaaacttggctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtct
atctcgttcatccatagttgctgactcccgcgctgtagataactacgatacgggagggctta
ccatctggccccagtgctgcaatgataccgcgagaccacgctcacccgctccagatttatcag
caataaaccagccagccggaagggccgagcgcagaagtggctcctgcaactttatccgcctccat
ccagcttattaattgttggcgggaagctagagtaagtagttcgccagttaatagtttgcgcaac
gttgttgccattgctacaggcatcgtgggtgtcacgctcgtcgtttggatggccttcatcagct
ccggttcccaacgatcaaggcgcagttacatgatccccatgttgtgcaaaaaagcgggttagctc
cttcggctcctccgatcgttgtcagaagtaagttggccgcagtggttatcactcatgggtatggca
gcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtaact
caaccaagtcattctgagaatagtgatgcggcgcagccagttgctcttgcccggcgtcaatac
ggataataaccgcgccacatagcagaactttaaagtgtcatcattggaaaacgcttcttcgggg
cgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaaccactcgtgcaccca
actgatcttcagcatcttttactttaccagcgttctggtgagcaaaaaacaggaaggcaaaa
tgccgcaaaaaagggaataagggcgcacacggaaatggtgaataactcatactcttcttttcaa
tattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttaga
aaaataaacaataggggttccgcgcacatttccccgaaaagtgcacacctgacgctc

**>pY016 (pcDNA3.1-hLbCpf1), secuencia de hLbCpf1 en mayúscula
negrita**

gacggatcgggagatctcccgatcccctatgggtgcaactctcagtacaatctgctctgatgccgc
atagttaagccagtatctgctccctgcttgtgtgttggaggctcgtgagtagtgcgcgagcaaa
atthaagctacaacaaggcaaggcttgaccgacaattgcatgaagaatctgcttaggggttaggc
gttttgcgctgcttcgcgatgtacgggcccagatatacgcgcttgacattgattattgactagtta
ttaatagtaatcaattacggggctcattagttcatagccatataatggagttccgcgcttacataa
cttacggtaaatggcccgcctggctgaccgccaacgacccccgcccattgacgtcaataatga
cgtatgttcccataagtaacgccaatagggactttccattgacgtcaatgggtggagatatttacg
gtaaactgcccacttggcagtacatcaagtgtatcatatgccaaagtagccccctattgacgctc
aatgacggtaaatggcccgcctggcattatgccagtagacattatgggactttcctactt
ggcagtagatctacgtattagtcacgctattaccatgggtgatgcgggttttggcagtagatcaa
tgggcgctggatagcgggttgactcacggggatttccaaagtctccacccccattgacgtcaatggg
agtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtacaactccgccccattga

cgcaaatgggcggttaggcgtgtacgggtgggaggtctatataagcagagctctctggctaactag
agaaccactgcttactggcttatcgaaattaatacgaactcactatagggagacccaagctggc
tagcgtttaacttaagcttggtagccacc**ATGAGCAAGCTGGAGAAGTTTACAACTGCTA**
CTCCCTGTCTAAGACCCTGAGGTTCAAGGCCATCCCTGTGGGCAAGACCAGGAGAACATCGAC
AATAAGCGGCTGCTGGTGGAGGACGAGAAGAGAGCCGAGGATTATAAGGGCGTGAAGAAGCTGC
TGGATCGCTACTATCTGTCTTTTATCAACGACGTGCTGCACAGCATCAAGCTGAAGAATCTGAA
CAATTACATCAGCCTGTTCCGGAAGAAAACCAGAACCAGAGAAGGAGAATAAGGAGCTGGAGAAC
CTGGAGATCAATCTGCGGAAGGAGATCGCCAAGGCCTTCAAGGGCAACGAGGGCTACAAGTCCC
TGTTTAAGAAGGATATCATCGAGACAATCCTGCCAGAGTTCCTGGACGATAAGGACGAGATCGC
CCTGGTGAACAGCTTCAATGGCTTTACCACAGCCTTCACCGGCTTCTTTGATAACAGAGAGAAT
ATGTTTTCCGAGGAGGCCAAGAGCACATCCATCGCCTTCAGGTGTATCAACGAGAATCTGACCC
GCTACATCTCTAATATGGACATCTTCGAGAAGGTGGACGCCATCTTTGATAAGCACGAGGTGCA
GGAGATCAAGGAGAAGATCCTGAACAGCGACTATGATGTGGAGGATTTCTTTGAGGGCGAGTTC
TTTAACTTTGTGCTGACACAGGAGGGCATCGACGTGTATAACGCCATCATCGGCGGCTTCGTGA
CCGAGAGCGGCGAGAAGATCAAGGGCTGAACGAGTACATCAACCTGTATAATCAGAAAACCAA
GCAGAAGCTGCCTAAGTTAAGCCACTGTATAAGCAGGTGCTGAGCGATCGGGAGTCTCTGAGC
TTCTACGGCGAGGGCTATACATCCGATGAGGAGGTGCTGGAGGTGTTTGAACACCCCTGAACA
AGAACAGCGAGATCTTCAGCTCCATCAAGAAGCTGGAGAAGCTGTTCAAGAATTTTGACGAGTA
CTCTAGCGCCGGCATCTTTGTGAAGAACGGCCCCGCCATCAGCACAATCTCCAAGGATATCTTC
GGCGAGTGGAACGTGATCCGGGACAAGTGAATGCCGAGTATGACGATATCCACCTGAAGAAGA
AGGCCGTGGTGACCGAGAAGTACGAGGACGATCGGAGAAAGTCTTCAAGAAGATCGGCTCCTT
TTCTCTGGAGCAGCTGCAGGAGTACGCCGACGCCGATCTGTCTGTGGTGGAGAAGCTGAAGGAG
ATCATCATCCAGAAGGTGGATGAGATCTACAAGGTGTATGGCTCCTCTGAGAAGCTGTTGACG
CCGATTTTGTGCTGGAGAAGAGCCTGAAGAAGAACGACGCCGTGGTGGCCATCATGAAGGACCT
GCTGGATTCTGTGAAGAGCTTCGAGAATTACATCAAGGCCTTCTTTGGCGAGGGCAAGGAGACA
AACAGGGACGAGTCTTCTATGGCGATTTTGTGCTGGCCTACGACATCCTGCTGAAGGTGGACC
ACATCTACGATGCCATCCGCAATTATGTGACCCAGAAGCCCTACTCTAAGGATAAGTTCAGCT
GTATTTTTCAGAACCCTCAGTTCATGGGCGCTGGGACAAGGATAAGGAGACAGACTATCGGGCC
ACCATCCTGAGATACGGCTCCAAGTACTATCTGGCCATCATGGATAAGAAGTACGCCAAGTGCC
TGCAGAAGATCGACAAGGACGATGTGAACGGCAATTACGAGAAGATCAACTATAAGCTGCTGCC
CGGCCCTAATAAGATGCTGCCAAAGGTGTTCTTTTCTAAGAAGTGGATGGCCTACTATAACCCC
AGCGAGGACATCCAGAAGATCTACAAGAATGGCACATTCAGAAGGGCGATATGTTTAACTGA
ATGACTGTCACAAGCTGATCGACTTCTTTAAGGATAGCATCTCCCGGTATCCAAAGTGGTCCAA
TGCCTACGATTTCAACTTTTCTGAGACAGAGAAGTATAAGGACATCGCCGGCTTTTACAGAGAG
GTGGAGGAGCAGGGCTATAAGGTGAGCTTCGAGTCTGCCAGCAAGAAGGAGGTGGATAAGCTGG
TGGAGGAGGGCAAGCTGTATATGTTCCAGATCTATAACAAGGACTTTTCCGATAAGTCTCACGG
CACACCCAATCTGCACACCATGTACTTCAAGCTGCTGTTTGACGAGAACAATCACGGACAGATC
AGGCTGAGCGGAGGAGCAGAGCTGTTTCATGAGGCGCGCCTCCCTGAAGAAGGAGGAGCTGGTGG
TGCACCCAGCCAACCTCCCCTATCGCCAACAAGAATCCAGATAATCCCAAGAAAACCAACCCT
GTCTACGACGTGTATAAGGATAAGAGGTTTTCTGAGGACCAGTACGAGCTGCACATCCCAATC
GCCATCAATAAGTGCSCCAAGAACATCTTCAAGATCAATACAGAGGTGCGCGTGTCTGCTGAAGC
ACGACGATAACCCCTATGTGATCGGCATCGATAGGGGCGAGCGCAATCTGCTGTATATCGTGGT
GGTGGACGGCAAGGGCAACATCGTGGAGCAGTATTCCCTGAACGAGATCATCAACAACTTCAAC
GGCATCAGGATCAAGACAGATTACCACTCTCTGCTGGACAAGAAGGAGAAGGAGAGGTTTCGAGG
CCCGCCAGAACTGGACCTCCATCGAGAATATCAAGGAGCTGAAGGCCGGCTATATCTCTCAGGT
GGTGCACAAGATCTGCGAGCTGGTGGAGAAGTACGATGCCGTGATCGCCCTGGAGGACCTGAAC
TCTGGCTTTAAGAATAGCCGCGTGAAGGTGGAGAAGCAGGTGTATCAGAAGTTCGAGAAGATGC
TGATCGATAAGCTGAACTACATGGTGGACAAGAAGTCTAATCCTTGTGCAACAGGGCGGCCCT
GAAGGGCTATCAGATACCAATAAGTTCGAGAGCTTTAAGTCCATGTCTACCCAGAACGGCTTC
ATCTTTTACATCCCTGCCTGGCTGACATCCAAGATCGATCCATCTACCGGCTTTGTGAACCTGC
TGAAAACCAAGTATAACCAGCATCGCCGATTCGAAGAAGTTCATCAGCTCCTTTGACAGGATCAT
GTACGTGCCCGAGGAGGATCTGTTCCGAGTTTGCCCTGGACTATAAGAATTTCTCTCGCACAGAC
GCCGATTACATCAAGAAGTGAAGCTGTACTCCTACGGCAACCGGATCAGAATCTTCCGGAATC
CTAAGAAGAACAACGTGTTCCGACTGGGAGGAGGTGTGCCTGACCAGCGCCTATAAGGAGCTGTT
CAACAAGTACGGCATCAATTATCAGCAGGGCGATATCAGAGCCCTGCTGTGCGAGCAGTCCGAC
AAGGCCTTCTACTCTAGCTTTATGGCCCTGATGAGCCTGATGCTGCAGATCGGGAACAGCATCA

**CAGGCCGCACCGACGTGGATTTTCTGATCAGCCCTGTGAAGAACTCCGACGGCATCTTCTACGA
TAGCCGGAACCTATGAGGCCAGGAGAATGCCATCCTGCCAAAGAACGCCGACGCCAATGGCGCC
TATAACATCGCCAGAAAGGTGCTGTGGGCCATCGGCCAGTTCAAGAAGGCCGAGGACGAGAAGC
TGGATAAGGTGAAGATCGCCATCTCTAACAAGGAGTGGCTGGAGTACGCCAGACCAGCGTGAA
GCACAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCCAAAAAGAAAAAGGGATCCTACCCA
TACGATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCC
CCGACTATGCC**TAAG**aattctgcagatatccagcacagtgggcgccgctcgagctctagagggcc
cgtttaaaccgctgatcagcctcgactgtgccttctagttgccagccatctggttgttggccc
tccccgctgccttccttgacctggaaggtgccactcccactgtccttctctaataaaatgagg
aaattgcatcgcattgtctgagtaggtgtcattctattctggggggtgggggtggggcaggacag
caagggggaggattgggaagacaatagcaggcatgctggggatgcggtgggtctatggcttct
gaggcggaaagaaccagctggggctctagggggtatccccacgcgcctgtagcggcgcattaa
gCGCGGGCGGGTGTGGTGGTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGC
tcctttcgctttcttcccttcccttctcgccacgttcgcccggctttccccgtcaagctctaaat
cgggggctccctttagggttccgatttagtgctttacggcacctcgacccccaaaaaacttgatt
agggatgaggttcacgtagtgggccatcgccctgatagacggtttttgccttttgacgttggg
gtccacgttctttaatagtgactcttgttccaaactggaacaacactcaaccctatctcggtc
tattcttttgattataagggttttggcgtttcggcctattgggttaaaaaatgagctgattt
aacaaaaatttaacgcgaattaattctgtggaatgtgtgtcagttaggggtgtggaaagtcccc
ggctccccagcaggcagaagtatgcaaagcatgcatctcaattagtcagcaaccagggtgtggaa
agtccccaggctccccagcaggcagaagtatgcaaagcatgcatctcaattagtcagcaaccat
agtccccgcccctaactccgcccattcccgccttaactccgcccagttccgcccattctccgccc
catggctgactaatttttttatttatgcagaggccgaggccgcctctgcctctgagctattcc
agaagtagtgaggaggcttttttggaggcctaggcttttgcaaaaagctccccgggagcttgtat
atccattttcggatctgatcaagagacaggatgaggatcgtttcgcatgattgaacaagatgga
ttgcacgcaggttctccggccgcttgggtggagaggctattcggctatgactgggcacaacaga
caatcggtctgctctgatgcccgcgtgttccggctgtcagcgcaggggCGCCCGGTTCTTTTGT
caagaccgacctgtccgggtccctgaatgaactgcaggacgaggcagcggctatctggctg
gccacgacggggcttcccttgcgcagctgtgctcgacgttgtcactgaagcgggaagggactggc
tgctattgggCGAAGTgcccgggagggatctcctgtcatctcaccttgtcctgcccagaaaagt
atccatcatggctgatgcaatgcccgggctgcatacgttgatccggctacctgcccattcgac
caccaagcgaacatcgcacgagcagcagcactcggatggaagcgggtcttgtcgatcagg
atgatctggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgg
catgcccgacggcgaggatctcgtcgtgacctatggcgatgctgcttggcgaatatcatgggtg
gaaaatggccgcttttctggattcatcgactgtggccggctgggtgtggcggaccgctatcagg
acatagcgttggctaccgctgatattgctgaagagcttggcggcgaatgggctgaccgcttcc
cgtgctttacggtatcgcgcctcccgattcgcagcgcacatgccttctatcgccttcttgacgag
ttcttctgagcgggactctgggttcgaaatgaccgaccaagcgcacgcccacactgccatcacg
agatttcgattccaccgccccttctatgaaagggtgggcttcggaatcgtttccgggacgcc
ggctggatgatcctccagcgcggggtctcatgctggagttcttcgcccacccccacttgttta
ttgcagcttataatgggttacaataaaagcaatagcatcacaatttcacaataaaagcattttt
ttactgcatctagttgtgggtttgtccaaactcatcaatgtatcttatcatgtctgtataaccg
tcgacctctagctagagcttggcgtaatcatgggtcatagctgtttcctgtgtgaaattgttatc
cgctcacaattccacacaacatacagagccggaagcataaagtgtaaagcctgggggtgcctaag
agtgagctaaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcg
tgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggCGCTCTT
ccgcttccctcgctcactgactcgctgcgctcggctcggttcggctgoggcgagcggatcagctca
ctcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagca
aaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgcgtggcggtttttccataggctcc
gccccctgacgagcatcacaataatcgacgctcaagtcagaggtggcgaaaccccgacaggact
ataaagataccaggcgtttccccctggaagctccctcgctgcgctctcctgttccgacctgccc
cttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgct
gtaggtatctcagttcgggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgt
tcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgac
ttatcgccactggcagcagccactggtaacaggattagcagagcagggatgttagggcggtgcta
cagagttcttgaagtgggtggcctaactacggctacactagaagaacagtatttgggtatctgcgc**

tctgctgaagccagttaccttcggaaaaagagttggttagctcttgatccggcaaaacaaaccacc
gctggttagcggtttttttgtttgcaagcagcagattacgcgcagaaaaaaggatctcaagaag
atcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggat
ggatcatgagattatcaaaaaggatcttcacctagatccttttaattaaaaatgaagt
ttaaatacaatctaaagtataatgatgagtaaacttggctctgacagttaccaatgctta
atcagtgaggcacctatctcagcgatctgtctatcttcggtccatccatagttgctgact
ccccgctcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaat
gataccgcgagacccacgctcacggctccagattatcagcaataaaccagccagccgga
agggccgagcgcagaagtggctctgcaactttatccgcctccatccagtctattaattg
ttgcccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacggtggttgc
cattgctacaggcatcgtgggtgtcacgctcgtcgttggatggcttcattcagctccg
gttcccaacgatcaaggcagttacatgatccccatggtgtgcaaaaaagcggttagct
ccttcggtcctccgatcgttgtcagaagtaagttggccgagtggtatcactcatgggt
tatggcagcactgcataattctcttactgtcatgccatccgtaagatgctttctgtgact
ggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctctt
gccccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcat
catggaaaaacgttcttcggggcgaaaactctcaaggatcttaccgctggtgagatccag
ttcgatgtaaccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgt
ttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaaggaataagggcgacacg
gaaatggtgaatactcatactcttcctttttcaatattattgaagcatttatcagggt
tattgtctcatgagcggatacatatattgaatgtatttagaaaaataaacaatagggg
gttccgcgcacattttccccgaaaagtgcacctgacgctc