

**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL**

UNIDAD ZACATENCO

DEPARTAMENTO DE BIOTECNOLOGÍA Y BIOINGENIERÍA

**Diseño y generación de Soya resistente a herbicida y con menor
contenido de ácido fítico mediante edición de genomas**

Tesis que presenta

Ana María Leyva Téllez

Para obtener el grado de
MAESTRA EN CIENCIAS
EN LA ESPECIALIDAD DE BIOTECNOLOGÍA

Directora de la Tesis:
Dra. Guadalupe Beatriz Xoconostle Cázares

Ciudad de México

Octubre, 2018

AGRADECIMIENTOS

Agradezco al Consejo de Ciencia y Tecnología (CONACyT) por su apoyo económico durante dos años para la realización de este proyecto de tesis.

Agradezco a mi directora de tesis Dra. Beatriz Xoconostle Cázares durante el desarrollo de este proyecto, por sus acertados comentarios y su valioso apoyo a mi desarrollo académico.

A mis asesores: Dr. Roberto Ruíz Medrano por su entusiasta contribución a este proyecto. Por contagiarme las ganas de seguir experimentando y conociendo. Al Dr. Jesús Hinojosa Moya por la disponibilidad que mostró para poder concluir el proceso que represento este proyecto.

Agradezco a los auxiliares de laboratorio M. en C. Leandro Alberto Núñez Muñoz y M. en C. Brenda Yazmín Vargas Henrnández por su apoyo técnico en la realización de este proyecto.

Al personal y estudiantes que forman parte del laboratorio de Biología Molecular de Plantas por su apoyo y colaboración durante mi estancia.

INDICE DE CONTENIDO

AGRADECIMIENTOS.....	I
INDICE DE CONTENIDO.....	II
RESUMEN.....	IV
ABSTRACT.....	V
INTRODUCCIÓN.....	1
1. Edición de Genomas.....	2
1.1. Nucleasas dedos de Zinc.....	3
1.2. TALENs.....	3
1.3. CRISPR/Cas9.....	4
2. Mecanismos de reparación de DNA.....	7
2.1. Unión de extremos no homólogos.....	8
1.2.2. Recombinación de homólogos.....	8
3. La semilla de la Soya.....	9
4. Metabolismo de AHAS e IPK.....	10
3.2. Inositol Pentakisfosfato Cinasa.....	10
3.1. Acetolactato Sintasa.....	12
JUSTIFICACIÓN.....	14
HIPÓTESIS.....	15
OBJETIVOS.....	16
ESTRATEGIA EXPERIMENTAL.....	17
MATERIALES Y MÉTODOS.....	18
1. Material biológico.....	18
2. Condiciones de crecimiento vegetal.....	18
2.1. Germinación de semillas de soya.....	18
3. Análisis de características fenotípicas antes de la edición.....	19
3.1. Cuantificación de ácido fítico.....	19
4. Análisis <i>in silico</i> del sitio de edición de los genes a modificar.....	19
4.1. Diseño de gRNAs.....	19
5. Determinación de la secuencia parcial de los genes IPK y AHAS en <i>G. max</i>	20
6. Construcción del vector de edición de genomas.....	23

6.1. Reacción de Ligación para inserción de inserto de gRNA.....	23
7. Transformación de <i>Glycine max</i>	25
7.1. Transformación mediante <i>Agrobacterium tumefaciens</i>	25
7.2. Transformación del embrión de la soya mediante biobalística	26
8. Análisis de transformantes.....	27
8.1. Extracción de DNA genómico vegetal.....	27
8.2. PCR punto final.....	28
8.3. Análisis de edición mediante la endonucleasa T7.....	28
8.4. Análisis de edición mediante secuenciación.	28
RESULTADOS Y DISCUSIÓN	29
1. Análisis in silico genes de edición.....	29
1.1. Inositol Fosfato Cinasa.....	29
1.2. Acetolactato Sintasa	30
2. Cuantificación del Ácido Fítico de la Soya.....	31
3. Determinación de la secuencia parcial de los genes de interés	32
3.1. IPK	33
3.2. AHAS	33
4. Diseño de RNA guía.....	34
4.1. IPK	34
4.2. AHAS	35
5. Construcción de los vectores para edición	37
5.1. Clonación del gRNA en el vector de edición.....	37
5.2. Construcción del vector DONOR-AHAS	38
6. Selección de plantas transformadas y análisis.....	41
6.1. IPK.....	42
7. Identificación de edición	45
CONCLUSIONES.....	49
REFERENCIAS	50

RESUMEN

La soya (*Glycine max*) es un cultivo de gran importancia a nivel mundial por su alto contenido proteico y oleaginoso. Esta semilla es ampliamente utilizada en el consumo tanto humano como ganadero. Sin embargo, el alto contenido de ácido fítico presente de manera natural en la semilla, además de ser altamente contaminante, reduce la asimilación de iones metálicos en los consumidores, así como la biodisponibilidad de estos en suelos agrícolas. Por otro lado, el rendimiento de la cosecha se ve reducido de un 25 a 30% debido al mal manejo de malezas durante su siembra. La ingeniería genética ha permitido durante los últimos años el desarrollo y producción de plantas de soya mejoradas. No obstante, la liberación comercial de estos cultivos se encuentra regulada por la LBOGM. Para lograr una liberación no regulada del cultivo de la soya se propone el uso de la técnica CRISPR-Cas9 para realizar cambios puntuales en dos genes que le confieran a la semilla de la soya un menor contenido de ácido fítico y que reduzca la susceptibilidad de la planta a los herbicidas utilizados para el control de malezas. Entre los resultados obtenidos destacan la obtención de una plataforma para edición de genoma en los genes IPK y AHAS de la soya, empleando la cepa hipervirulenta de *Agrobacterium tumefaciens* AGL1, el diseño de gRNAs para editar los genes mencionados y la obtención de plantas transgénicas que cuentan con el cassette de edición de plantas dicotiledóneas.

ABSTRACT

Soybean (*Glycine max*) is an important crop employed worldwide, due to its high protein and oil content. This seed is widely used for both human and animal consumption. However, its high phytic acid content in seeds poses a problem, since it is a highly polluting agent, reduces metallic ion assimilation and bioavailability in agricultural soil. On the other hand, the crop yield is reduced from 25 to 30% due to competition with weeds during its sowing. Genetic engineering has allowed the development and production of improved soybean plants. Nevertheless, the commercial release of these crops is regulated by the LBOGM. To achieve a non-regulated release of the soybean crop, we propose the use of CRISPR-Cas9 to induce point mutations in two genes that would decrease the phytic acid content and induce herbicide resistance for weed management. Among the results obtained we highlight obtaining a platform for genome editing in *Glycine max* genes *IPK* and *AHAS*, employing the hypervirulent *Agrobacterium tumefaciens* strain AGL1. The gRNA design to edit the mentioned genes and the transgenic plants obtained which have the editing cassette for dicot plants.

INTRODUCCIÓN

Los cultivos vegetales de interés agrícola se encuentran expuestos a agentes estresantes ya sean bióticos o abióticos. Naturalmente, las plantas que se encuentran sometidos a cualquier tipo de estrés logran responder de manera positiva a estos, desde sequías e inundaciones, hasta suelos salinos y patógenos. Producto de este comportamiento, los cultivos tanto comerciales como no comerciales han sido objeto de estudio en las diversas ramas de la biotecnología vegetal, con la finalidad de entender los mecanismos por los cuales estos se han vuelto tolerantes.

El primer paso para realizar el estudio y desarrollo de variedades de interés agro biotecnológico consiste en la identificación de fenotipos de interés para el cultivo de variedades resistentes y/o tolerantes a diversos tipos de estrés, resultantes de mutaciones génicas (Oladosu et al., 2016).

Particularmente, se han desarrollado estrategias para el análisis genómico funcional en plantas. Además, el desarrollo de técnicas de mejoramiento tradicional y no tradicional en plantas, han sido ampliamente estudiados para su aplicación. En conjunto, todo lo anterior se ha vuelto fundamental para el desarrollo de nuevas variedades de especies de interés agrícola, las cuales se espera, cuenten con características capaces de mejorar su potencial económico (Lakhanpaul, 2006).

En ese sentido, la selección de características deseadas se puede llevar a cabo mediante el uso de diferentes marcadores, ya sean puede morfológicos, bioquímicos o fenotípicos. Sin embargo, lo anterior puede conllevar a diversas implicaciones genotípicas. Debido a ello, durante las últimas décadas se han desarrollado marcadores de DNA, los cuales son una herramienta prometedora para identificar en el genoma regiones en las cuales un único nucleótido difiere en los alelos de una especie llamados Polimorfismos de nucleótidos únicos (SNPs) (Yin, Gao, & Qiu, 2017) Los SNPs se presentan como una estrategia prometedora para la generación de mutaciones (sustituciones) las cuales representarán en la especie vegetal de interés una mejora durante su cultivo y cosecha. Para

cumplir el objetivo de mejoramiento agrobiotecnológico, se han desarrollado estrategias que permiten la mutación de sitios específicos en una secuencia dada de DNA, esta estrategia es denominada modificación genética dirigida (TGM) o edición genómica (Chen & Gao, 2015).

1. Edición de Genomas

La edición del genoma se define como el conjunto de técnicas desarrolladas para producir modificaciones específicas y dirigidas en secuencias nucleotídicas, permitiendo la modificación de un determinado locus mediante sustituciones, deleciones o inserciones en sitios determinados del genoma (Belhaj, Chaparro-Garcia, Kamoun & Nekrasov, 2013).

Una de las estrategias que se han utilizado a lo largo de los años para la investigación funcional de genes y las aplicaciones biotecnológicas es la modificación genómica dirigida (TGM). Dicha estrategia ha sido ampliamente aplicada a organismos como la mosca de la fruta, ratones, células humanas y plantas (K. Chen & Gao, 2015).

La modificación genética, sin embargo, encuentra en el desarrollo de estrategias confiables y eficientes para introducir cambios dirigidos en el genoma, su mayor desafío (Guha, Wai, & Hausner, 2017a). Para lograr la modificación o edición genéticas cuales se basan en el uso de nucleasas compuestas de dominios específicos de unión a DNA, se han desarrollado principalmente tres plataformas de edición: (1) Nucleasas dedos de Zinc (ZFN), (2) Nucleasas de Efectores Activadores de la Transcripción (TALENs) y (3) el sistema Repeticiones Palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR) y la proteína asociada a CRISPR (Cas). Estas nucleasas inducen la ruptura de DNA de doble cadena (DSBs) en una región específica dirigida (Gaj, Gersbach, & Barbas, 2013; Guha, Wai, & Hausner, 2017b).

1.1. Nucleasas dedos de Zinc

Son proteínas quiméricas, compuestas de un dominio de factores de transcripción y una endonucleasa. Los dedos de Zinc son aproximadamente 30 aminoácidos que se pueden unir en una combinación limitada de tripletes. Mediante el uso de estas combinaciones, una única secuencia de DNA en el genoma puede ser su objetivo (Figura 1a) (Yin, Kauffman, & Anderson, 2017).

Para el uso de los dedos de Zinc, dos quimeras son producidas. Cada una de ellas se une a una secuencia específica de DNA, las cuales se encuentran cercanas una de la otra. En estas secuencias, la nucleasa realiza un corte que libera un inserto, estimulando los mecanismos de reparación, que en este documento se explican más adelante (Figura 4).

A pesar de todas las ventajas que esta técnica ofrece, una de las principales desventajas de esta técnica es su alta probabilidad de off-targets dentro del genoma. Debido a lo anterior, es de vital importancia que, para asegurar (Urnov, Rebar, Holmes, Zhang, & Gregory, 2010)

1.2. TALENs

Los TALENs son nucleasas unidas por ingeniería genética a regiones activadoras de transcripción altamente conservados (TALE) aislados de una proteobacteria. Estos activadores de transcripción son introducidos a la célula de manera natural mediante el sistema de secreción Tipo II, uniéndose a la célula y así facilitando la colonización bacteriana (Figura 1b) (Bortesi & Fischer, 2015; Liang, Zhang, Chen, & Gao, 2014)

Los TALENs se unen a DNA mediante un arreglo de aminoácidos conservados, ubicados en la posición 33-35, flanqueados por extremos amino y carboxi-terminal. Cada TALEN, de manera individual, se une específicamente a un nucleótido de una secuencia de DNA. La secuencia específica de unión a DNA es determinada por la identidad de dos residuos variables en la posición 12 y 13 dentro de cada repetición. Desde su descubrimiento, se han estudiado la combinación de

diversos residuos para la estandarización del reconocimiento de los nucleótidos específicos siendo los más aceptados N(N o K), NI, HD and NG para la unión sitio específico de G, A, C y T, respectivamente (K. Chen & Gao, 2015; Guha et al., 2017b)

El uso de TALENs asociados a nucleasas para el reconocimiento y corte de secuencias dadas se ha acrecentado ampliamente durante los últimos años. Esta estrategia resulta favorable para la generación de variedades de plantas con nuevas características fenotípicas en un periodo de tiempo más corto. Su uso se ha aplicado a plantas modelo como *Arabidopsis thaliana*, así como de consumo humano particularmente maíz y tabaco (Q.-T. Li et al., 2017).

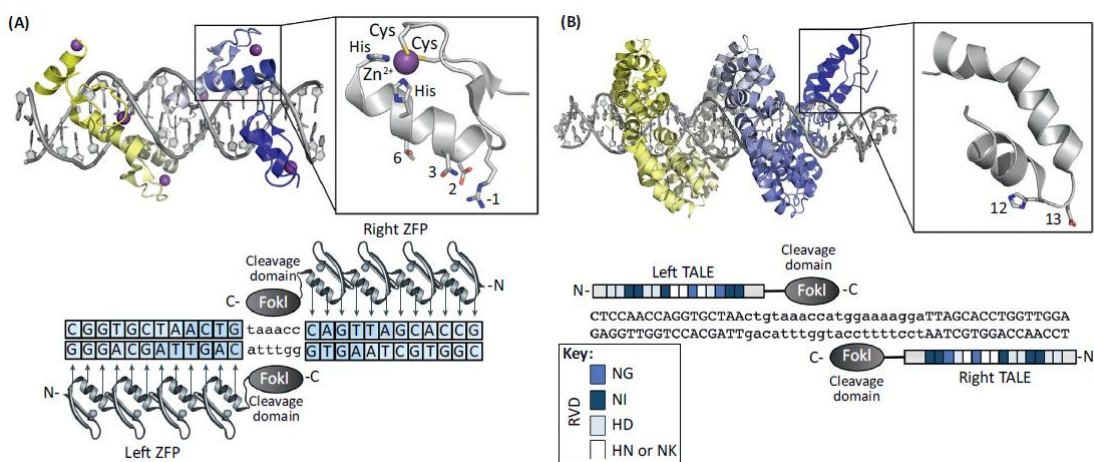


Figura 1. Representación esquemática de las nucleasas utilizadas para edición. a) ZFN; b) TALENs (Gaj et al., 2013)

1.3. CRISPR/Cas9

Una estrategia nueva y diferente de las nucleasas sitio específico, es el sistema CRISPR/Cas. Este nuevo sistema emergió en el año 2012 como una tecnología de reconocimiento específico modular y de fácil y rápida aplicación en diversos sistemas (Bortesi & Fischer, 2015).

CRISPR/CAS es un sistema de tres componentes, los cuales consisten en: una endonucleasa (Cas9), una secuencia específica como elemento objetivo (crRNA) y una secuencia que se une a

crRNA y a Cas9 (tracrRNA). Una parte fundamental en la acción de este sistema es un Motif Adyacente al Protoespaciador (PAM, por sus siglas en inglés), esta es una secuencia determinada por la proteína Cas de estudio, la cual es la que determinará la unión de dicha proteína (Sander & Joung, 2014).

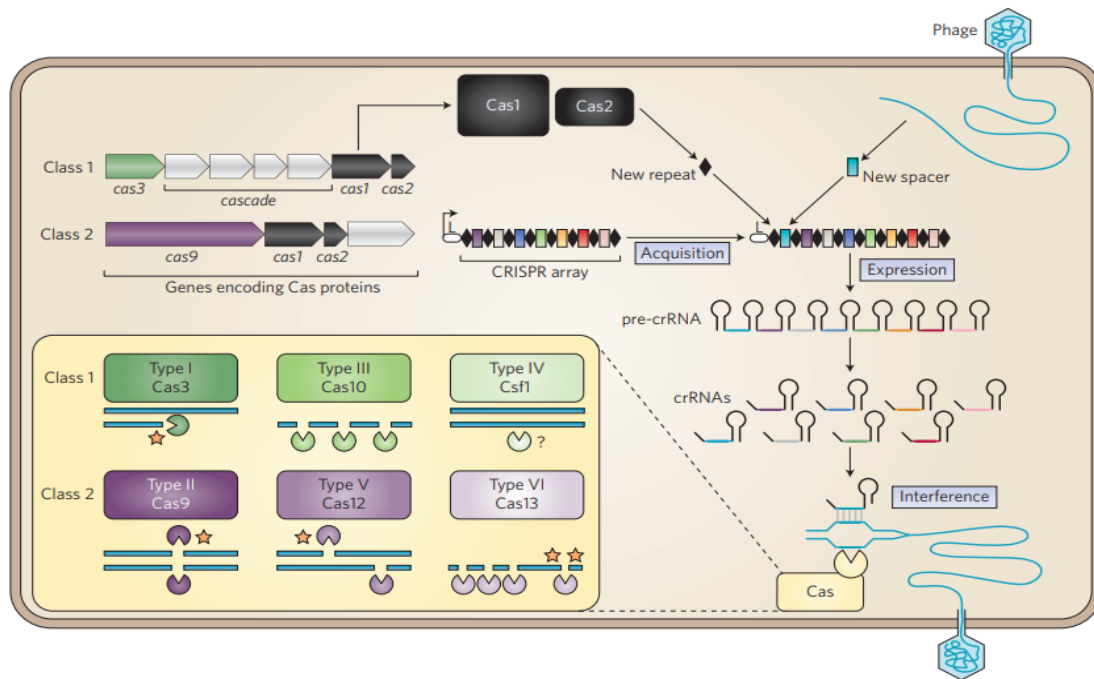


Figura 2. Representación esquemática del sistema CRISPR/Cas9 proveniente de *Streptococcus pyogenes* (Barrangou & Horvath, 2017)

Es un sistema por primera vez encontrado en *Streptococcus pyogenes*, este sistema es conocido por ser parte del sistema inmunológico de la bacteria. CRISPR son secuencias cortas de DNA divididas espaciadores, cuando un organismo exógeno es introducido a la célula las secuencias se sintetizan como RNA. El RNA se procesa dando lugar a secuencias de RNA independientes unas de otras (crRNAs), mientras el tracrRNA se encarga de dirigir al crRNA al sitio de unión de Cas9, el cual será dirigido por cada uno de los nucleótidos que forman la secuencia de RNA. Finalmente, Cas9 realizará la inactivación del gen endógeno mediante su actividad como endonucleasa (Figura 2) (Hruscha et al., 2013; Lowder et al., 2015).

Mediante estrategias de ingeniería genética, las secuencias crRNA y tracrRNA se combinan en una sola molécula llamada RNA guía (gRNA). La proteína Cas9 es dirigida por una molécula corta de RNA (20 pb), la cual se aparea perfectamente con la secuencia de DNA, la endonucleasa realiza en la cadena de DNA un corte 3-6 pb río debajo de la secuencia PAM. Lo anterior, como todas las estrategias de edición descritas, dará lugar a los mecanismos de reparación de DNA (Lowder et al., 2015).

La proteína Cas9 de *Streptococcus pyogenes* (SpCas9): es una enzima multifuncional monomérica con un peso de 160 KDa, compuesta por 6 diferentes dominios, de los cuales dos tienen actividad de endonucleasa: El dominio HNH es responsable de la digestión de la hebra complementaria tres nucleótidos adyacentes al PAM y el dominio similar a RuvC que rompe la hebra no complementaria en una región de 3 a 8 nucleótidos adyacentes al PAM (Anders, Niewoehner, & Jinek, 2015; Doudna & Charpentier, 2014) (Fig. 3)

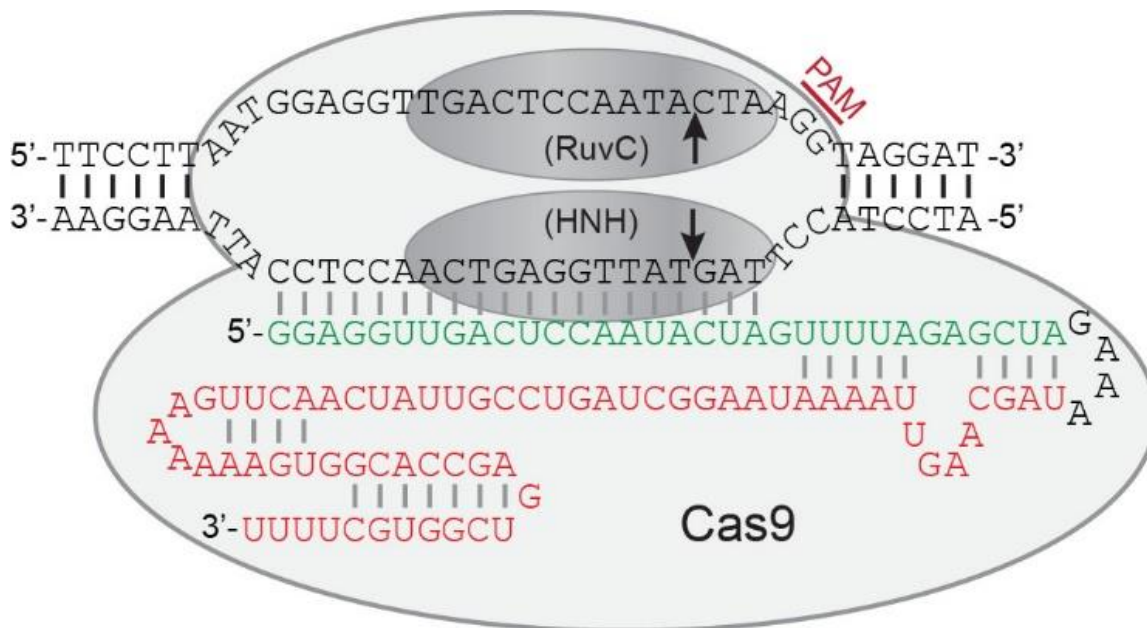


Figura 3. Estructura y actividad endonucleasa de la proteína Cas9 (Jinek et al., 2014)

Existen otras proteínas similares a SpCas9 en diferentes microorganismos. Ejemplos de ello son los ortólogos de Cas9 y Cpf1 de *Francisella novocida*, y C2c2 de *Leptotrichia sashii* que tienen una actividad similar a la de Cas9, pero PAMs, longitudes de gRNA diferentes (East-Seletsky et al., 2016; Zetsche et al., 2015), por lo cual las opciones de diseño para los sitios de edición, seguirán incrementándose en los próximos años con el descubrimiento de proteínas similares en otros microorganismos (Kingdom et al., 2017; Zetsche et al., 2015).

2. Mecanismos de reparación de DNA

Una vez que las nucleasas han realizado su actividad, cortando la doble cadena de DNA en una región específica, la célula inicia con un proceso de reconocimiento de los cortes en el genoma. Los primeros transductores de señales de la ruptura en el DNA genómico son dos proteínas similares a cinasas: ATM y ATR. Ambas cinasas inician un proceso de fosforilación en cascada en la histona H2AX en un dominio de la cromatina al rededor del corte. La fosforilación anterior acumulará otros factores de respuesta a daño que, en conjunto, desencadenará la reparación de doble cadena (Amiard, Depeiges, Allain, White, & Gallego, 2011).

Una vez que las proteínas reconocen la ruptura de la doble cadena de DNA, se lleva a cabo la reparación principalmente por dos rutas: (1) Unión de extremos no homólogos (NHEJ) o (2) reparación homóloga (HDR) (Gaj et al., 2013).

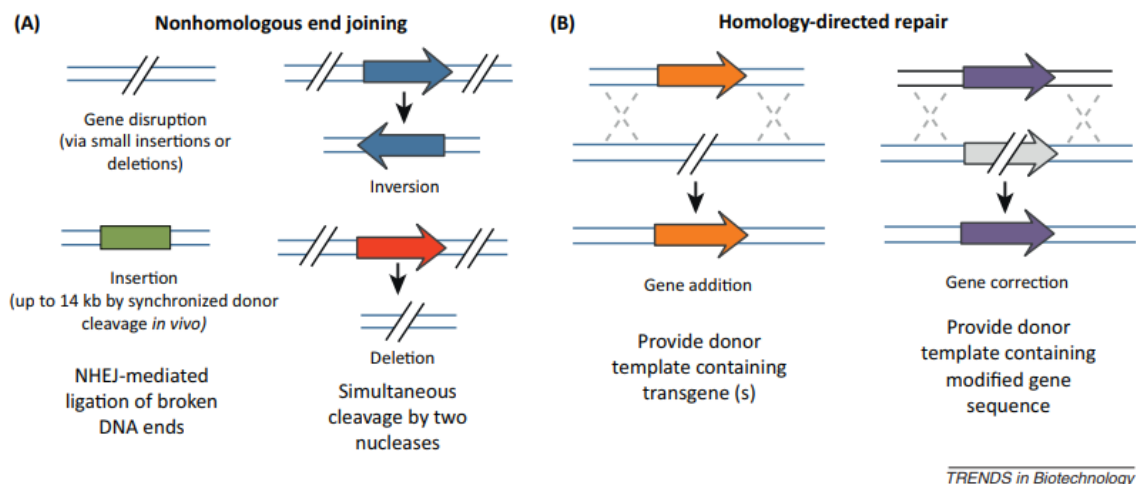


Figura 4. Vías de reparación del DNA por cortes en la doble cadena. a) Unión de extremos homólogos. b) Reparación dirigida por homología

2.1. Unión de extremos no homólogos

La unión de extremos no homólogos puede ser subdividido en dos grupos ocurriendo de manera convencional y alternativa. En ambos casos, este proceso ocurre en la célula sin el uso de un templado como molde, dando como resultado la inserción o deleción (indels) de fragmentos de DNA de una longitud desde 1 a 1000 pares de base en la cadena con la ruptura. El mecanismo de NHEJ ocurre, en plantas, con mucha mayor frecuencia que la recombinación homóloga, y la reparación por este mecanismo provoca generalmente una pérdida de función de un gen de interés (knockout) (Chen & Gao, 2015; Urnov et al., 2010).

1.2.2. Recombinación de homólogos

La recombinación homóloga, es un mecanismo de reparación de mucho menor eficiencia en la reparación de cortes de doble cadena. Generalmente, se encuentra presente con mucho mayor actividad durante las fases S y G del ciclo celular (Goh, Ko, Koh, Kim, & Bae, 2012).

Con la finalidad de introducir en el genoma cambios puntuales en un gen de interés, se ha introducido en la ingeniería genética el uso de endonucleasas con actividad dirigida a sitios

específicos dentro del genoma de un organismo. De esta manera, el conocimiento de los mecanismos de reparación, podrán verse aplicados como una herramienta para la mutagénesis “natural” de algunos organismos, ya sea mediante NHEJ o HMR (Gaj et al., 2013).

3. La semilla de la Soya

El cultivo de soya (*Glycine max*), dentro de las oleaginosas, tiene una importancia fundamental debido a su alto contenido de proteínas y aceites. Se estima que entre el año 2015/2016 la producción mundial fue de 313.553 millones de toneladas (USDA, 2016). En México, la producción estimada total se estima en 341 y 509 de Toneladas en los años 2015 y 2016, respectivamente de acuerdo con los datos reportados por la SIAP, SAGARPA. A nivel internacional, Estados Unidos es el principal productor, seguido de Brasil, Argentina, China e India (ASERCA, 2015).

Debido a esta gran importancia a nivel tanto nacional como internacional, la soya ha sido objeto de estudio para la generación de nuevas variedades con características de interés en el comercio. Y es también debido a esto, que es importante solucionar las problemáticas más emblemáticas de la misma. El cultivo de soya tiene dos problemáticas importantes durante su producción y aplicación en la industria de los alimentos: el mal manejo de malezas durante su cultivo y el alto contenido de ácido fítico de la semilla. Debido a la competencia a la que la soya se ve sometida por presencia de malezas, este cultivo puede llegar a perderse desde un 30 a un 50% (FAO; ADAMA FAUBA, 2012). Además, el ácido fítico que este contiene, si es consumido en altas cantidades, puede llegar a actuar como anti nutriente en el organismo, así como disminuir la biodisponibilidad de algunos iones metálicos en el suelo (Humer & Zebeli, 2015). Por otro lado, este mismo compuesto tiene efectos negativos en el medioambiente, debido a que favorece la eutrofización en corrientes de agua al no poder ser asimilado por mamíferos no rumiantes (Redekar et al., 2015).

Además de las estrategias de transgénesis normales, se han buscado la mejora en cultivo mediante ingeniería genética, la cual no represente en la planta la presencia de un gen exógeno. Un ejemplo importante es el anteriormente mencionado: la edición del genoma mediante el sistema CRISPR/Cas9.

Para la aplicación de este sistema, sin embargo, hay consideraciones a considerar. Lo más importante es el diseño del experimento al elegir el gen a editar de acuerdo los objetivos que se pretenden alcanzar.

4. Metabolismo de AHAS e IPK

A continuación, se describen los genes involucrados en las rutas de biosíntesis del ácido Fítico, así como la enzima a la cual la herbicida sulfonil urea efecta alostéricamente.

3.2. Inositol Pentakisfosfato Cinasa

La biosíntesis del ácido fítico en plantas se da primordialmente mediante dos rutas: la lípido dependiente y la lípido independiente (Figura 5) (Stevenson-Paulik, Bastidas, Chiou, Frye, & York, 2005). Durante la ruta hay diversas enzimas que colaboran para su síntesis, dentro de las más importantes destacan la mio-inositol fosfato sintasa (MIPS) que cataliza la formación de un metabolito secundario muy importante para el desarrollo de la semilla: Inositol-3-fosfato. Los pasos subsecuentes son dados por una serie de cinasas, dentro de las cuales se encuentra la enzima: inositol Pentakisfosfato cinasa o también llamada IPK1 (IPK) (Rasmussen, Ingvarlsen, & Torp, 2010)

Mediante la mutación de este gen en diversas especies se ha demostrado que la semilla reduce la cantidad de ácido fítico en la semilla tales como maíz, soya, avena, entre otros (Shi et al., 2003; Yuan et al., 2012; Zhao et al., 2013)

3.1. Acetolactato Sintasa

La primera enzima en la síntesis de aminoácidos de cadena ramificada (valina, isoleucina y leucina) es la Acetohidroxiácido sintasa, también conocida como Acetolactato sintasa (AHAS), la cual cataliza la descarboxilación a partir de piruvato (Figura 6). Al igual que en bacterias, la enzima AHAS de las plantas se conforma de dos subunidades, la primera corresponde a la subunidad menor y la segunda a la subunidad mayor, la cual cuenta con actividad catalítica. Particularmente en las plantas esta enzima se encuentra realizando su catálisis en cloroplastos (K. Y. Lee et al., 1988; Y. T. Lee & Duggleby, 2001). Además, de ser dependiente de tiamina difosfato y flavin adenina dinucleótido (FAD) como cofactores.

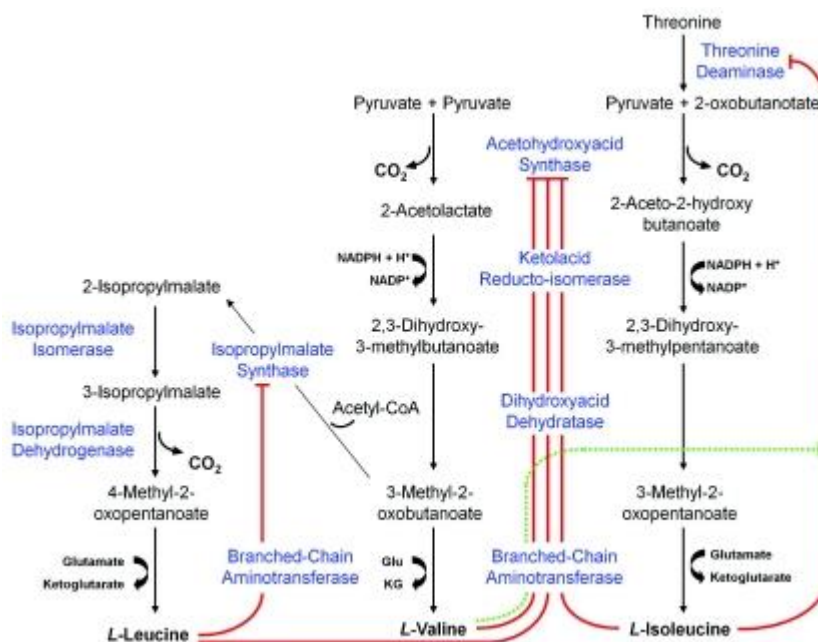


Figura 6. Síntesis de aminoácidos de cadena ramificada (valina, isoleucina, leucina)

La enzima Acetolactato sintasa ha sido estudiada ampliamente debido a que es el objetivo de diferentes herbicidas comerciales, uno de los más utilizados es el clorsulfuron, perteneciente al grupo de la sulfonilurea. Su mecanismo de acción se basa en la inactivación de dicha enzima, evitando la síntesis de los aminoácidos de cadena ramificada, actuando sobre la subunidad grande. La falta de aminoácidos tan indispensables para la planta provocará en la misma el cese de división celular, crecimiento vegetativo, clorosis en hojas y necrosis en los nódulos axilares (Brown, 1990; Ray, 1982).

Bioquímicamente los herbicidas no se acoplan al sitio de unión al sustrato por ser homólogos a los sustratos naturales de la planta, pero si se unen al canal de acceso de los sustratos de esta enzima, bloqueando la entrada de estos al sitio activo. Este modo de acción, independiente del sustrato, explica porque muchas plantas expuestas al herbicida generan rápidamente mutaciones que les confieren de resistencia al mismo. Hasta la fecha se sabe que existen 17 modificaciones en la cadena polipeptídica las cuales confieren de alta tolerancia a las sulfonil ureas (Binder, 2010; Chang & Duggleby, 1998). Se ha demostrado que, en la mayoría de los casos, un polimorfismo en la subunidad catalítica en la posición 197 (*Arabidopsis thaliana*) confiere en las plantas menor susceptibilidad a la sulfonil urea.

JUSTIFICACIÓN

El cultivo de soya presenta diversas problemáticas como el alto contenido de ácido fítico y susceptibilidad a herbicidas que generan un impacto al consumidor y productor.

La edición de genomas contribuirá de forma rápida y fácil a la implementación de un sistema para la obtención de plantas de soya con disminución de ácido fítico y proveerá tolerancia a herbicidas al cultivo de la misma, cultivos que se considerarán libres de transgénicos.

HIPÓTESIS

El cambio en el marco de lectura del gen que codifica para IPK generará una enzima cinasa no funcional interrumpiendo la ruta de biosíntesis del ácido fítico y la acumulación de InositolP5, permitiendo la disminución del contenido de ácido fítico.

La edición del gen AHAS permitirá generar resistencia a herbicidas del grupo Sulfonilurea mediante la modificación de un sitio de unión importante para el funcionamiento del herbicida.

OBJETIVOS

OBJETIVO GENERAL

Mejorar genéticamente a plantas de soya a través del uso de herramientas de edición de genomas.

OBJETIVOS PARTICULARES

- Determinar mediante estudios *in silico* el mejor gen para llevar a cabo la edición.
- Determinar las secuencias de gRNA que permitan la edición de los genes involucrados en la síntesis de ácido fítico y la resistencia a herbicidas en *Glycine max*.
- Producir, analizar y seleccionar genómicamente las plantas con potencial edición en los genes.

ESTRATEGIA EXPERIMENTAL

La estrategia experimental consta de cinco etapas (Figura 7) correspondientes a cada uno de los objetivos propuestos incluyendo el análisis *in silico* para determinar el sitio de edición para de los genes *IPK* y *AHAS*, la construcción de plásmidos para la edición de genomas vegetales, la transformación y selección de transformantes, la corroboración de secuencias editadas, así como análisis fenotípicos de las plantas.

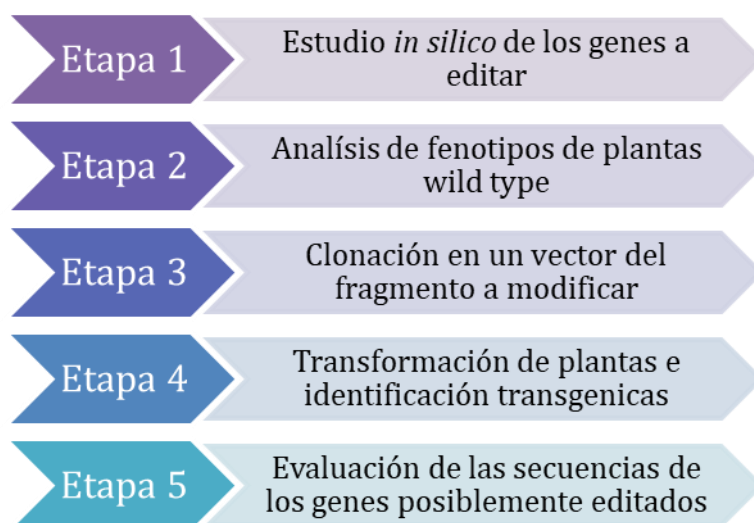


Figura 7. Estrategia experimental del proyecto

MATERIALES Y MÉTODOS

1. Material biológico

En este trabajo se emplearon semillas de soya pertenecientes a las variedades Nainari. Asimismo, se contó con la cepa hipervirulenta de *Agrobacterium tumefaciens* AGL1 y la cepa de conservación de *Escherichia coli* MACH1TM-T1 (Invitrogen).

2. Condiciones de crecimiento vegetal

2.1. Germinación de semillas de soya

Las semillas de la soya fueron esterilizadas mediante gas de cloro, los cuales fueron obtenidos por una mezcla de 100 ml de hipoclorito de sodio comercial y 3.0 ml de ácido clorhídrico (HCl) concentrado. Las semillas fueron expuestas a estos gases durante 16 horas. Una vez esterilizadas, estas fueron transferidas a un medio de crecimiento (MS basal, adicionado con Sacarosa 3 % y 1 mg/L de 6-bencilaminopurina (BAP) para su germinación. Las plantas se mantuvieron en una cámara de crecimiento con una temperatura aproximada de 25°C y un fotoperiodo de 16 horas luz: 8 horas oscuridad. Las semillas germinadas fueron transferidas semanalmente a medio fresco durante un mes y medio.

Posteriormente, las plántulas obtenidas *in vitro* fueron transferidas a una mezcla estéril de suelo: turba: agrolita, cubiertas con papel plástico e incubadas en una cámara de crecimiento (Conviron A1000, Winnipeg, Canadá) con humedad relativa de 70%, una temperatura de 22°C y con fotoperiodo de 16 horas luz: 8 horas oscuridad, regadas con una solución nutritiva de MiracleGro. Tras 2 semanas de aclimatación, la cubierta plástica fue retirada mientras las condiciones de temperatura y luminosidad se mantuvieron. Dos semanas después, las semillas germinadas fueron llevadas a invernadero con una temperatura de 27°C.

3. Análisis de características fenotípicas antes de la edición

3.1. Cuantificación de ácido fítico.

Se realizó la determinación indirecta del contenido de ácido fítico en 7 variedades de la semilla de soya mediante el protocolo propuesto por Latta y Eskin (1980). Este ensayo se realizó mediante una prueba colorimétrica de precipitación del hierro con una longitud de onda de 480 nm. La cuantificación se realizó siguiendo una curva patrón mediante la ecuación $y = 115081x + 1977$.

4. Análisis *in silico* del sitio de edición de los genes a modificar

El análisis *in silico* de los genes objetivo fue hecha usando como referencia el genoma de la soya variedad Williams 82, disponible en las bases de datos del NCBI (<https://ncbi.nlm.nih.gov>) y Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>). Asimismo, se hizo uso de la plataforma UniProt (<http://www.uniprot.org/>) para el análisis estructural de las proteínas estudiadas en el presente proyecto.

Se evaluaron los sitios de unión más importantes para ambas proteínas mediante el análisis de secuencias conservadas en organismos del Phylum Viridiplantae. Posteriormente se realizó un alineamiento de secuencias mediante el software MEGA 7.0 empleando el método de alineamiento de ClustalW.

4.1. Diseño de gRNAs.

El diseño de gRNAs para la edición de este dominio se realizó empleando el software en línea Benchling (<https://benchling.com>) a fin de identificar los mejores candidatos para gRNAs en el gen codificante para Inositol Pentafosfato Cinasa de *Glycine max* (GCF_000004515.4) y Acetolactato Sintasa (GCF_000004515.4). El diseño consideró para ello la mayor proximidad al dominio de unión a editar, la cercanía a aminoácidos de interés de edición, la mayor actividad on-traget del

gRNA, la menor actividad off target y excluyendo aquellos gRNAs cuya potencial actividad off target se encuentre en sitios codificantes del genoma.

5. Determinación de la secuencia parcial de los genes IPK y AHAS en *G. max*

Con la finalidad de poder iniciar estudios de diseño de gRNAs en las variedades de estudio de la soya, se realizó la secuenciación y comparación del genoma de referencia *Glycine max* var. Williams 82. Para la amplificación de la región del gen IPK y AHAS que incluyera el dominio de unión a sustrato en variedades de soya cuya secuencia nucleotídica es desconocida; se diseñaron oligos para amplificar mediante PCR punto final el sitio potencial de edición en dichas variedades. Los oligos empleados para dicha amplificación (Tabla i) fueron diseñados utilizando como referencia la variedad de soya Williams 82, los cuales flaqueaban secuencias de 443 y 393 pb de los genes de interés IPK y AHAS, respectivamente.

Tabla i. Secuencias de los oligonucleótidos sintetizados para la amplificación parcial de los genes a editar

Nombre	Oligonucleótidos 5'-3'
GmALS fw	GTGGTGCTTCCATGGAGATCCACCAGG
GmALS rv	CACAGGGCCAGGTCTACCCGAAGTGCC
GmIPK fw	CCCAAATGCGGATTTCTTCCTCTTC
GmIPK rv	GTCCTCCAAGTATGAGAGAGCC

Los lineamientos generales seguidos en este trabajo, para la amplificación de productos de PCR mediante PCR punto final, se indican en las Tablas ii (mezcla de reacción) y iii (programa de temperaturas para amplificación).

Tabla ii. Mezcla de reacción para PCR punto final.

Reactivo	Volumen (µL)
ddH ₂ O estéril	8.70
Buffer 10X	1.25
dNTPs (10 µM)	1.00
Primer F (100 µM)	0.25
Primer R (100 µM)	0.25

Takara Ex Tak HS	0.05
DNA (30-100 ng/ μ L)*	1.00
Volumen final	12.50

*=Para PCR de DNA plasmídico ajustar el DNA a concentración de 30 ng/ μ L. Para PCR de DNA genómico de *G. max*, ajustar el DNA a concentración de 100 ng/ μ L.

Tabla iii. Programa de temperaturas para amplificación de fragmentos por PCR punto final.

Etapa	Temperatura (°C)	Tiempo	
		DNA plasmídico	DNA genómico
Desnaturalización inicial	94.00	03:00	05:00
Desnaturalización	94.00	00:30	00:30
Alineamiento	50.0-68.0	00:30	00:30
Extensión	72.0	1 kb/minuto	
Extensión final	72.0	5:00	10:00
Enfriamiento	16.0	∞	

*=La temperatura de alineamiento dependerá de cada PCR en particular y deberá ser optimizada en función de cada par de oligos con base en su T_m .

Una vez amplificados los fragmentos por PCR punto final y visualizados en gel agarosa al 1%, dichos productos de PCR fueron clonados en el vector pCR8/GW/TOPO (Fig. 8), siguiendo la mezcla de reacción señalada en la Tabla iv e incubados a 22°C toda la noche. Posteriormente, la mezcla de clonación fue empleada para la transformación de células comerciales de *E. coli* MACH1TM-T1 (Invitrogen) mediante choque térmico siguiente las indicaciones del fabricante seleccionados. Las células se inocularon en placas de medio LB suplementado con espectinomicina (100 μ g/mL).

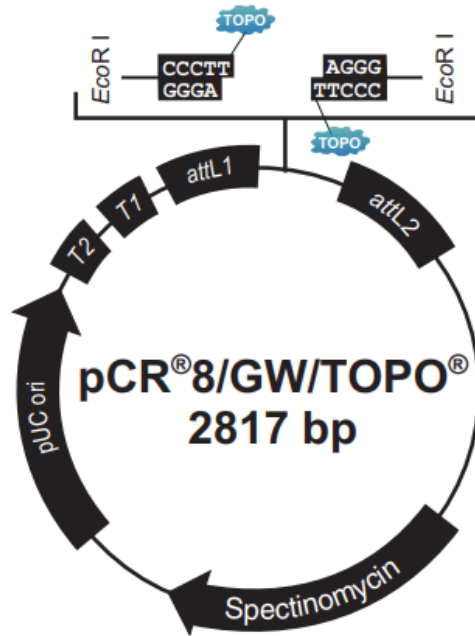


Figura 8. Mapa del vector pCR8, empleado para la clonación de una región del gen IPK y AHAS de *G. max*

Tabla iv. Mezcla de reacción para clonación en el vector pCR8

Componente	Volumen (μL)
Producto de PCR purificado (~100 ng)	1.00 – 4.00*
Solución Salina	1.00
pCR TM 8/GW/TOPO [®] vector	0.5
H ₂ O Milli-Q estéril	cbp vol. final 6 μL
Volumen final	6.0

*En función de la concentración del producto de PCR

La presencia del inserto en las colonias obtenidas fue verificada por PCR punto final y estas clonas fueron secuenciadas para determinar su homología de secuencias nucleotídicas respecto al genoma de referencia a partir de secuenciación por el método de Sanger en el Servicio de Secuenciación de DNA del Laboratorio de Bioquímica Molecular, UBIPRO, FES-Iztacala, UNAM. Las secuencias obtenidas fueron comparadas a través de un alineamiento con BLASTN para regiones genómicas de organismos de la familia Fabaceae (<https://blast.ncbi.nlm.nih.gov>).

6. Construcción del vector de edición de genomas.

El vector de edición fue obtenido por el equipo de trabajo del Laboratorio de Biología Molecular en Cinvestav, I.P.N. Para la construcción de dicho vector se utilizó como base el plásmido pBUN411 (Addgene plasmid # 50581), al cual se adicionó un cassette de espectinomicina y el cassette del promotor U6 para dirigir la transcripción de gRNA.

6.1. Reacción de Ligación para inserción de inserto de gRNA

Para la construcción de los plásmidos pBUN4U6SM, se alinearon los oligonucleótidos de 24 pares de bases con regiones codificantes para los protoespaciadores de los genes IPK y AHAS de *Glycine max* en un volumen final de 20 μ L, añadiendo 1 μ L de cada oligo a 100 μ M y 2 μ L de Buffer de alineamiento 10X (Tris-HCl 100 mM pH 8.0, NaCl 500 mM y EDTA 10 mM). La mezcla de reacción para los alineamientos se incubó de acuerdo al programa descrito en la Tabla V.

Tabla v. Rampa de temperaturas para el alineamiento de secuencias complementarias

Temperatura ($^{\circ}$ C)	Tiempo (min)	Δ T ($^{\circ}$ C/s)
94	5	-3.0
94-84	-	-2.0
84-25	-	-0.2
16	∞	-3.0

Ambos productos de alineamiento se visualizaron en un gel con agarosa Nusieve al 5%. Posteriormente, se realizó la digestión del vector pBUN4U6SM con la enzima BsaI-HF (NEB). La banda de 13 Kb fue purificada directamente del gel de agarosa. El producto del plásmido linealizado, así como el alineamiento de 24 pares de bases fue ligado mediante Rapid DNA Ligation Kit (ThermoFisher Scientific) a una proporción molar de 10:1 (vector:inserto) respecto al gRNA alineado. Las clonas de *E. coli* obtenidas se verificaron mediante PCR punto final específica para el promotor AtU6-F con gRNA-R y digestión con la enzima HindIII, finalmente la correcta

inserción del gRNA fue corroborada mediante secuenciación por Sanger. Posteriormente 1 µg de plásmido fue transformado en células de *A. tumefaciens* AGL1 mediante electroporación. Se seleccionaron 5 colonias y se verificaron mediante PCR punto final para *Agrobacterium* (100ng), específica para U6 con gRNA-R.

Para el análisis de colonias de *Agrobacterium*, se crecieron 5 colonias candidatas transformantes en 3 mL de medio LB con los respectivos antibióticos (carbenicilina 50 mg/L, kanamicina 25 mg/L, espectinomicina 50 mg/L) por separado, a 28°C durante 2 días a 180 rpm. Los cultivos se centrifugaron a 13000 g durante 5 min y cada pastilla celular se lavó dos veces con 500 µL de agua estéril, centrifugando a 13000 rpm durante 5 min y desechando el sobrenadante. Cada pastilla celular se resuspendió en 20 µL de agua estéril, se incubó a 94°C durante 10 min y luego en hielo por 10 min. Se realizó una última centrifugación durante 5 min a 13000 g a 4°C y del sobrenadante se tomó 1 µL para la realización de PCR punto final.

1.1. Transformación de *Agrobacterium*. Una carga de 100-150 ng del vector binario verificado fue utilizado para la transformación de 50 µL de células electrocompetentes de *A. tumefaciens* AGL1 utilizando una celda de 0.1 cm a 1800 V, 15 µF 335 R en un equipo electroporador EquiBio EasyJect Optima (Ashford, UK). Tras incubar 3 h a 28°C a 225 rpm, se sembraron 5 µL de la mezcla de transformación en placas de LB con carbenicilina (100 mg/L), kanamicina (50 mg/L) y espectinomicina (100 mg/L) durante 2-3 días a 28 °C. Las colonias fueron transferidas a medio LB líquido con carbenicilina (50 mg/L), kanamicina (25 mg/L) y espectinomicina (50 mg/L) e incubadas por 2 días a 28 °C, 180 rpm. Posteriormente las posibles transformantes fueron analizadas mediante PCR. Una vez verificadas las colonias de *A. tumefaciens* AGL1 se emplearon para la transformación de plantas de acuerdo al procedimiento especificado en cada caso.

7. Transformación de *Glycine max*

7.1. Transformación mediante *Agrobacterium tumefaciens*

La cepa de *A. tumefaciens* AGL1 con el plásmido pBUN4U6-IPK fue cultivada en medio LB líquido suplementada con carbenicilina (50 mg/L), kanamicina (25 mg/L) y espectinomicina (50 mg/L) hasta alcanzar O.D.₆₂₀ 0.8-1.0. El cultivo (50 ml) fue centrifugado durante 10 minutos a 5000 rpm. El pellet obtenido fue resuspendido en 25 ml medio de infección (0.1X Sales Gamborg B5, 0.1X vitaminas GB5, 3 % Sacarosa, 3.9 g/L MES, pH 5.4. Adicionado con GA3 (0.25 mg/L), BAP (1.67 mg/L), y 40 mg/L de acetosiringona, esterilizados por filtración). La bacteria se incubó durante 25 minutos a temperatura ambiente. Las semillas previamente embebidas durante 16 horas fueron colocadas en grupos de 10 sobre una toalla absorbente. Con ayuda de una navaja se realizó un corte a través del hilum para remover la testa y remover uno de sus cotiledones. El cotiledón que aun contenía el embrión fue sometido a un corte en el eje radicular.

Los cotiledones fueron expuestos al medio de infección durante 20 minutos en agitación suave (60 rpm). Posteriormente, el exceso de medio fue eliminado al colocar las semillas en papel estéril. Las semillas fueron colocadas en medio de Co-cultivo (0.1X Sales Gamborg B5, 0.1X vitaminas GB5, 3% de Sacarosa, 3.9 g/L MES y 4.25 g/L Noble agar, pH 5.4. Adicionado con GA3 (0.25 mg/L), BAP (1.67 mg/L), Cisteína (400 mg/L), Dithiothrietol (154.2 mg/L) y 40 mg/L acetosiringona esterilizados por filtración) cubierto con papel filtro durante 3-5 días.

Las semillas fueron después traspasadas a medio de inducción de brotes (1X Sales Gamborg B5, 1X vitaminas GB5, 3% Sacarosa, 0.59 g/L Mes y 7 g/L Noble agar. Adicionado con BAP (1.11 mg/L), Meropenem (100 mg/L), Cefotaxima (100 mg/L) y Vancomicina (50 mg/L) esterilizados por filtración). Dichos explantes permanecieron en el medio durante 2-3 semanas hasta alcanzar un tamaño aproximado de 3-5 cm. Posteriormente, fueron traspasados a medio de elongación de

brotos (1X Sales Gamborg B5, 1X vitaminas GB5, 3% Sacarosa, 0.59 g/L MES y 7 g/L Noble agar, pH 5.7. Adicionado con BAP (1.11 mg/L), Meropenem (100 mg/L), Cefotaxima (100 mg/L), Vancomicina (50 mg/L) y Glufosinato de amonio (4 mg/L).

7.2. Transformación del embrión de la soya mediante biobalística

La bacteria *E. coli* conteniendo el vector de edición y el vector pCR8-DONADOR fue incubada durante 16 horas a 37°C en medio LB con espectinomicina 100 mg/L a una D.O._{600nm}=1.0. Tras este periodo se realizó la extracción de DNA plasmídico de ambas bacterias. El plásmido de edición pBUN4U6-AHAS se linealizó mediante la enzima SspI. El plásmido pCR8-DONADOR se digirió con la enzima PvuI para obtener el plásmido de forma lineal. Ambos productos linealizados se purificaron mediante QIAquick(R) PCR Purification Kit, se cuantificaron y se visualizaron en un gel de agarosa 0.8%.

Las semillas de soya se embebieron durante 16 horas en agua destilada estéril. A las semillas previamente embebidas se extrajeron los embriones y se colocaron con la cara abaxial en contacto con el medio, dispuestas de manera circular manteniendo un diámetro de 3.5 cm. A 60mg as partículas de oro (0.6 uM) resuspendidas en una agua se les adicionaron 3.5 ug de DNA por disparo en una proporción 1:3 por cada par de base (pBUN4U6-AHAS:pCR8-DONADOR). A la reacción se le adicionaron 50 uL de CaCl₂ 2.5 M y 20 uL de espermidina 0.1 M. La mezcla se homogeneizó y se recuperaron las partículas de oro mediante centrifugación. Estas, se resuspendieron en 75 uL de Etanol absoluto grado biología molecular y se alicuotaron 10 uL en cada acarreador para realizar los disparos. Los disparos se llevaron a cabo a 1100 psi impulsados con Helio comercial con una atmósfera de vacío a 7 psi.

Los explantes bombardeados fueron transferidos a un medio de regeneración (MS basal 1X, Vitaminas MS 1X, 3 % Sacarosa, 5 mg/L BAP, pH 5.7, 0.4% fitagel) en oscuridad durante dos

semanas. Una vez transcurrido ese tiempo, los embriones fueron transferidos a medio fresco de regeneración y expuestos a la luz durante dos semanas. Posteriormente, los embriones regenerados fueron transplantados a un medio de regeneración adicionado con 4 mg del herbicida glufosinato de amonio.

8. Análisis de transformantes.

8.1. Extracción de DNA genómico vegetal.

A las plantas que no presentaron amarillamiento tras la aplicación del herbicida se les realizó extracción de DNA genómico total. Se recuperaron 100 mg de tejido foliar, los cuales fueron congelados en nitrógeno líquido y pulverizados en un homogeneizador de tejidos TissueLyser a 50 Hz durante 2 min. Tras la pulverización las muestras se incubaron en 600 μ L de CTAB Buffer (Tris-HCl 100 mM pH 8.0, EDTA 20 mM pH 8.0, NaCl 1.4 M, CTAB 2%, PVP 4000 1%) a 65°C durante 20 min a 300 rpm. A esta mezcla se adicionó un volumen de fenol: cloroformo: alcohol isoamílico (12:12:1), se centrifugó y el sobrenadante se mezcló con un volumen de cloroformo: alcohol isoamílico (24:1). Se centrifugó nuevamente y al sobrenadante se le añadieron 0.7 volúmenes de isopropanol frío, mezclando por inversión e incubando a -20°C durante 30 min. La pastilla obtenida posterior a la centrifugación se lavó dos veces con etanol frío al 70% y se secaron los tubos a 55°C, resuspendiendo el precipitado en agua Milli-Q estéril con RNAsa A a 20 μ g/mL. La integridad del DNA fue verificada en un gel de agarosa al 0.8% y teñido con bromuro de etidio (0.1 μ g/mL); mientras que la concentración fue determinada con un Nanodrop (Thermo Scientific, WA).

8.2. PCR punto final.

Las plantas se analizaron por PCR punto final para detectar un fragmento de 424 pb correspondiente al gen codificante para Cas9 (zCas9-F: 5'-CGGCCTCGATATTGGGACTAACTCT-3', zCas9-R: 5'-CTTATCTGTGGAGTCCACGAGCTTC-3'), 5 µL de cada producto de PCR fue visualizado en un gel de agarosa al 1.5 %.

8.3. Análisis de edición mediante la endonucleasa T7.

Para la identificación de cambios puntuales en la secuencia de DNA en el gen de interés se realizó un ensayo con la endonucleasa T7, la cual reconoce y realiza el corte en secuencias mal apareadas a partir de 1 pb. Para ello, se realizó el alineamiento de dos productos de PCR del gen de interés utilizaron; el primero correspondiente a la planta de soya Nainari silvestre y el segundo a la planta de soya transformada. El alineamiento se realizó siguiendo el programa anteriormente señalado en la Tabla v.

El alineamiento se incubó con 1 uL de Endonucleasa T7 a 37°C durante 1 hora. Posteriormente, el producto se visualizó en un gel de agarosa 2 % para observar si existía desapareo entre las secuencias.

8.4. Análisis de edición mediante secuenciación.

A las plantas que resultaron positivas se les realizó una reacción de amplificación del gen IPK. El producto de PCR se utilizó para la clonación en el vector TOPO pCR2.1, posteriormente se realizó la transformación de células de *E. coli* cepa MACH1-T1. A las colonias resultantes se le realizó la extracción de DNA plasmídico, se corroboró la presencia del inserto y el DNA que resultó positivo al inserto se envió a secuenciar con los oligonucleotidos M13-F (5' ACTGGCCGTCGTTTTAC

3') y M13-R (5' GGAAACAGCTATGACCATG 3'). Los datos obtenidos de la secuenciación se analizaron mediante un alineamiento con la secuencia del gen en estudio obtenida de la planta silvestre, de este modo se evaluó la existencia de ediciones en la secuencia de interés.

RESULTADOS Y DISCUSIÓN

1. Análisis in silico genes de edición

1.1. Inositol Fosfato Cinasa

El gen de IPK es de vital importancia para la síntesis del ácido fítico, debido a que es el encargado de realizar el último paso de fosforilación en la reacción. Shi y colaboradores (2017) encontraron que un silenciamiento del gen IPK en cultivares de maíz reducía el contenido de ácido fítico de la semilla F1 provenientes de las plantas transformadas.

De acuerdo a las bases de datos analizadas, el dominio más importante dentro de este gen pertenece a la familia de Ins_P5_fosfato, el cual abarca los aminoácidos 24-339. Con ayuda de los resultados obtenidos del BLAST de proteína del dominio Ins_P5_Fosfato con diversos organismos (Tabla vi), se pudieron delimitar las secuencias mayormente conservadas y, por homología, determinar los sitios más importantes para la enzima catalítica.

Tabla vi. Especies empleadas para realizar la búsqueda de sitios conservados de IPK

<i>Arabidopsis thaliana</i>	<i>Phaseolus angularis</i>
<i>Cucumis sativus</i>	<i>Phaseolus vulgaris</i>
<i>Glycine gracilis</i>	<i>Sorghum bicolor</i>
<i>Glycine max</i>	<i>Trifolium pratense</i>
<i>Glycine soja</i>	<i>Vigna angularis</i>
<i>Gossypium hirsutum</i>	<i>Vigna radiata</i>
<i>Jatropha curcas</i>	<i>Zea mays</i>
<i>Mendicago truncatula</i>	

Se encontraron cuatro secuencias de importancia en el dominio Ins_P5_fosfato (Tabla vii).

Considerando los resultados del BLAST, se realizó la representación gráfica de la firma de unión

a sustrato del gen IPK. Debido a que en la posición 206 de esta proteína se encuentran los sitios importantes para la actividad catalítica, se decidió realizar el diseño del gRNA cerca de este sitio, esperando un cambio en el marco abierto de lectura que le confiriera a la proteína una conformación no funcional dentro de dicha región.

Tabla vii. Regiones de relevancia para la funcionalidad del gen IPK

Sitio	Secuencia
Unión a sustrato	KXXK
Unión a sustrato	KXXXXEIS
Unión ATP	RXF
Sitio activo	DHSXF

Adicionalmente, en estudios anteriores, Yuan y colaboradores (2012) se encontró una mutación puntual al final exón 5 del gen IPK la cual confería a la semilla de la soya una baja cantidad de ácido fítico. Esta mutación corresponde al aminoácido 206 de la cadena polipeptídica, el cual se encuentra dentro al sitio de unión a sustrato. Lo anterior, corresponde al sitio de edición elegido en este proyecto.



Figura 9. Representación gráfica de la secuencia consenso para el dominio de unión a sustrato de la enzima IPK en plantas (<http://weblogo.berkeley.edu/>).

1.2. Acetolactato Sintasa

Esta enzima, ha sido ampliamente estudiada debido a representar la proteína objetivo de los herbicidas sulfonilurea. Se evaluó mediante alineamiento si existía en plantas de *Arabidopsis thaliana* que se han reportado como resistentes al herbicida Sulfonilurea, algún cambio en la cadena

primaria de la proteína AHAS. El análisis bioinformático arrojó diferencias importantes dentro de los dominios de unión a tiamina pirofosfato, las cuales se basan en el cambio de un aminoácido (P→S). Con el cambio de este aminoácido es posible obtener una planta tolerante a herbicidas.

Mediante la inserción de AHAS modificado a diversas especies vegetales se han logrado obtener plantas resistentes a sulfonil urea tales como tabaco, *Arabidopsis thaliana*, arroz, tabaco, entre otros (Haughn & Somerville, 1986; Z. Li, Hayashimoto, & Murai, 1992).

Más recientemente se han obtenido plantas de *Arabidopsis thaliana* resistentes a Clorsulfuron. Lo anterior se logró mediante el cambio en el aminoácido 197 del gen AtAHAS de Prolina a Serina. Dicha modificación fue obtenida a partir de la estrategia de edición de genomas CRISPR/Cas9 (Y. Chen et al., 2017). Así mismo esta estrategia se ha probado en plantas de maíz (*Zea mays*) y soya (*Glycine max*) (Zhongsen Li et al., 2015; Liang, Zhang et al., 2014). Debido a lo anterior, en este proyecto se realizará la sustitución del triplete CCC a TCT, para modificar el aminoácido Prolina a Serina, de acuerdo a lo reportando en bases de datos y los autores anteriormente descritos.

2. Cuantificación del Ácido Fítico de la Soya

Se realizó la cuantificación del ácido fítico de la soya mediante el protocolo propuesto por Latta y Eskin (1980). Lolas (1976) reportan que el ácido fítico se encuentra en cantidades de 0.5 a 1 g PA/ 100 g de muestra. Lo anterior corresponde a los resultados obtenidos en este proyecto, presentados en la Tabla viii:

Tabla viii. Cuantificación del ácido fítico en las diversas variedades utilizadas en este proyecto.

Semilla	Promedio	[Fe] ug/L	[PA] g/100 g muestra
H200	0.015	3645.67	1.220
H400	0.021	4393.70	1.350
Nainari	0.010	3070.26	0.880
Oil	0.009	3041.49	0.935
Protein	0.004	2379.78	0.522
Rapid Shoot	0.009	3041.49	0.990
Suaqui	0.011	3185.35	1.06

3. Determinación de la secuencia parcial de los genes de interés

Antes de realizar el diseño del gRNA se analizaron las secuencias obtenidas para evaluar la homología del gen a editar. Se analizaron un total de 7 variedades (Tabla ix) por dos genes de estudio. Se esperaba que los resultados obtenidos de la secuenciación tuvieran un alto porcentaje de homología en la región a editar de manera que el vector de edición de cada uno de los genes fuera funcional si estos se aplicaban de manera independiente en un evento de transformación vegetal

Tabla ix. Variedades empleadas en el análisis de secuencia de una región parcial de los genes de estudio

Código	Nombre de la variedad
H200	Huasteca 200
H400	Huasteca 400
Nainari	Nainari
Oil	Soya con alto contenido de aceite
Protein	Soya con alto contenido de proteína
RapidShoot	Soya de rápida producción de brotes
Suaqui	Suaqui

3.1. IPK

El gen de IPK consta de 7 exones, el sitio de unión a sustrato se encuentra en el exón aproximándose al final del exón 5. Para realizar el análisis de secuencia se amplificaron 443 pb, las cuales se extienden sobre la región del exón 5 hasta la mitad del exón 6 (Figura 10).



Figura 10. Estructura del gen de IPK en Glycine max

Las flechas rojas indican la posición donde se diseñaron oligos para amplificación por PCR punto final y la flecha en color azul indica el sitio de interés para la realización de la edición genética

De la amplificación y clonación se obtuvieron 5 clonas, las cuales se analizaron mediante secuenciación, obteniendo como secuencia final la mostrada a continuación:

```
>Gm_IPK1
TGCTATCAAAAAGGAGAATAACCCGATTTGAAATGCACCAAACCTCTGAAATTGCTTCAAGGAGAGGTACATCCTACAAA
ATTTAACCTTTTATGGTGATCTAGTGGCTGCATAGTGCATACTTCTGTTGTCATATGTTCTAGGCTCAGCATTTAAT
TCTCTGTCTTATGATTGATGTTGTAATAGTCAATGTCCATTATTTATTTGGCATTAAATCTGAGTTAGAAAATCTTTG
AATATGTTTGTAGATATCGCAACTAAGTGAGTACAATCCACTTGATCTGTTCTCTGGATCCAAGGAAAGAATTCTGAA
AGCTATTAAAGGTCTCCTTACAACCTCCTCAAAACAATTTTCGTGTATTTTTGAATGGCTCTCTCATACTTGGAGGGAC
TGGGAAGA
```

La secuencia anterior es resultado de un alineamiento realizado para las 7 variedades. Dicha cadena de DNA tiene un grado alto de homología en todas las variedades que aquí fueron analizadas.

3.2. AHAS

El gen AHAS es una secuencia de 1962 pb, las cuales corresponden al marco abierto de lectura de la proteína con el mismo nombre de 653 aminoácidos. El fragmento amplificado cubría el marco desde la posición 308-1018, los resultados obtenidos de la secuenciación se presentan como un alineamiento en la Figura 11.



Figura 11. Estructura del gen de AHAS subunidad mayor en Glycine max

Las flechas rojas indican la posición donde se diseñaron oligos para amplificación por PCR punto final y la flecha en color azul indica el sitio de interés para la realización de la edición genética

Las secuencias parciales obtenidas del proceso de secuenciación para este gen fueron alineadas, producto de ello se obtuvo una cadena de DNA consenso para las variedades cuyo gen se estudió.

Dicha secuencia se muestra a continuación:

```
>Gm_AHAS
CTCCCTCGCCACGAACAGGGCGGCGTCTTTCGCCGCCGAGGGCTACGCCCGCTCTTCCGGCCTCCCCGGCGTCTGCATC
GCCACCTCCGGCCCCGGCGCCACCAACCTCGTCTCCGGCCTCGCCGACGCCTTGCTTGACAGCGTCCCCCTCGTCGCC
ATCACCGGCCAGGTCCCCCGCCGCATGATCGGGCACAGACGCCTTCCAAGAAACCCCATCGTCGAGGTAACACGTTCC
ATCACTAAGCATAACTATCTCGTTCTCGATGTTGATGACATTCTAGAAATCGTTAATGAAGCGTTTTTCTTAGCCACT
TCGGGTAGACCTGG
```

Dichas secuencias de DNA mostraron una homología del 100 % en las variedades analizadas en el proyecto a partir del nucleótido 308 hasta el 694. Sin embargo, en el genoma de referencia en comparación con la secuencia de las variedades estudiadas, se observaron polimorfismos a lo largo de la misma, lo cual nos podría indicar una diversidad fenotípica.

4. Diseño de RNA guía

Se hizo uso de la plataforma Benchling, para el análisis de secuencia de DNA de los genes a editar. La plataforma se alimentó con la secuencia de nucleótidos obtenida de la secuenciación. Se evaluaron los mejores parámetros arrojados por la plataforma.

4.1. IPK

Para el gen IPK, la plataforma se alimentó de una secuencia de 443 pb. Dentro de esta secuencia, se buscó una secuencia de 20 pb, cuyo on target se encuentre arriba de 50 %, de la misma manera se esperaba un off target con valores arriba de 40 %. A pesar de que los valores on target fueron adecuados para los parámetros que se buscaban, el off target resultó con un porcentaje más bajo. Para evaluar el efecto que el off target podría tener dentro del genoma de la soya, se evaluaron las secuencias arrojadas por la página como target off, se descartaron aquellas con un valor <10%. En

en el caso de IPK, no existieron secuencias con un puntaje de off target alto, por lo cual se procedió a elegir la secuencia con mayor on target. La secuencia de gRNA se posicionó del nt 1417 al 1436.

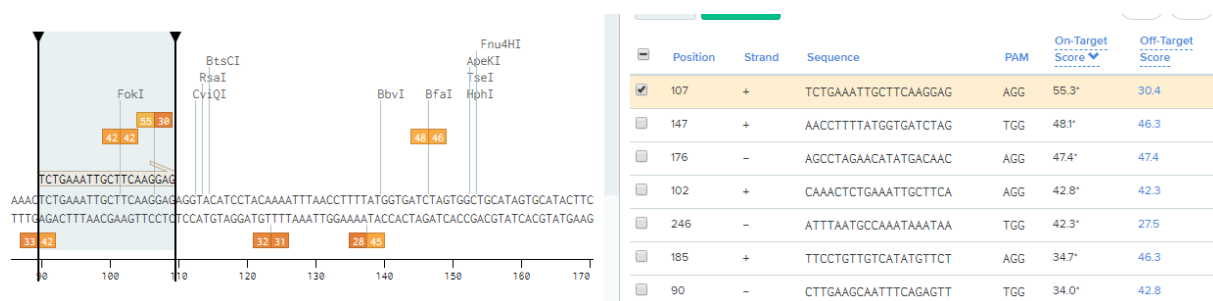


Figura 12. Diseño de gRNA para la edición del gen IPK en Glycine max

Se analizaron las secuencias parciales obtenidas del gen IPK para confirmar que en todas las variedades existiera la presencia de los 20 pb que darán lugar al gRNA. Como se observa en la Figura 13, las secuencias se conservaron a lo largo de la región de 20 pb y asimismo en la posición de la secuencia PAM.

Species/Abbrv	Group Name	* * * * *
1. Nainari_IPK		A A C T C T G A A A T T G C T T C A A G G A G A G G T A C A T C C
2. WilliAMS82_IPK		A A C T C T G A A A T T G C T T C A A G G A G A G G T A C A T C C
3. Oil_IPK		A A C T C T G A A A T T G C T T C A A G G A G A G G T A C A T C C
4. H400_IPK		A A C T C T G A A A T T G C T T C A A G G A G A G G T A C A T C C
5. H200_IPK		A A C T C T G A A A T T G C T T C A A G G A G A G G T A C A T C C
6. Suaqui_IPK		A A C T C T G A A A T T G C T T C A A G G A G A G G T A C A T C C
7. RapidShooting_IPK		A A C T C T G A A A T T G C T T C A A G G A G A G G T A C A T C C
8. Protein_IPK		A A C T C T G A A A T T G C T T C A A G G A G A G G T A C A T C C

Figura 13. Comprobación de la presencia del gRNA de IPK en las secuencias de las variedades de soya

4.2. AHAS

En el caso de la secuencia de AHAS, la región seleccionada para el diseño de los gRNA se ubica entre las posiciones 516-552. El gRNA, de una longitud de 20 pb, se diseñó para que la proteína Cas9 realice el corte de secuencia en el codón CCC que codifica para una Prolina. Los valores de on target se encuentran cercanos al 50 %. Por otro lado, a pesar de que el off target es bajo, la

evaluación del mismo indico una probabilidad del 50 % de realizar el corte en una secuencia redundante (Figura 14).

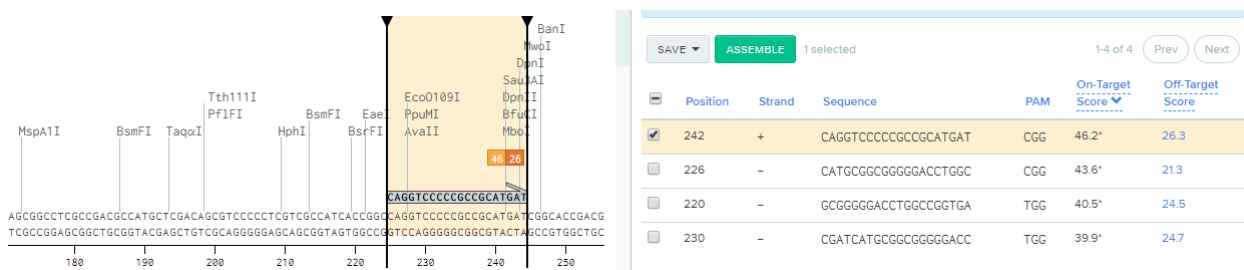


Figura 14. Diseño de gRNA para la edición del gen AHAS en Glycine max

Debido a que los genes de AHAS son redundantes y la estrategia de edición es recombinación homóloga, se corroboró que los genes con más alto target off se encontraran únicamente dentro de la familia de las Acetolactato Sintasa en la subunidad mayor, de esta manera si Cas9 actuara sobre dichos genes, la planta podría adquirir resistencia a herbicidas.

Posteriormente, las secuencias parciales se evaluaron mediante un alineamiento en el programa MEGA 7. Se corroboró la presencia de la secuencia de 20 pb que dará lugar a gRNA, de esta manera se comprobó que el vector resultante de este proyecto podrá ser aplicado en las variedades secuenciadas en esta posición en particular (Figura 15).

Species/Abbrv	Group Name	* * * * *
1. H200		C G T C G C C A T C A C C G G C C A G G T C C C C G C C G C A T G A T C G G C A C A G A C
2. H400		C G T C G C C A T C A C C G G C C A G G T C C C C G C C G C A T G A T C G G C A C A G A C
3. Nainari		C G T C G C C A T C A C C G G C C A G G T C C C C G C C G C A T G A T C G G C A C A G A C
4. Oil		C G T C G C C A T C A C C G G C C A G G T C C C C G C C G C A T G A T C G G C A C A G A C
5. Protein		C G T C G C C A T C A C C G G C C A G G T C C C C G C C G C A T G A T C G G C A C A G A C
6. Rapid_Shooting		C G T C G C C A T C A C C G G C C A G G T C C C C G C C G C A T G A T C G G C A C A G A C
7. Suaqui		C G T C G C C A T C A C C G G C C A G G T C C C C G C C G C A T G A T C G G C A C A G A C
8. Williams82		C G T C G C C A T C A C C G G C C A G G T C C C C G C C G C A T G A T C G G C A C C G A C

Figura 15. Comprobación de la presencia del gRNA de AHAS en las secuencias de las variedades de soya

Los gRNA obtenidos del programa Benchling fueron sintetizados como oligonucleótidos de DNA, tal cual se describe en la Tabla x.

Tabla x. Secuencias de los oligonucleótidos que darán lugar al gRNA de ambos genes

Nombre	Secuencia
gRNA IPK F	ATTGTCTGAAATTGCTTCAAGGAG
gRNA IPK R	AAACCTCCTTGAAGCAATTCAGA
gRNA AHAS F	ATTGCAGGTCCCCCGCCGCATGAT
gRNA AHAS R	AAACATCATGCGGCGGGGACCTG

5. Construcción de los vectores para edición

5.1. Clonación del gRNA en el vector de edición

Una vez que se tenían los oligonucleótidos sintetizados, se realizó la clonación en el vector pBUN4U6SM. Este vector fue construido por el grupo de trabajo utilizando como base el plásmido comercial pBUN411.

Los oligos fueron diseñados para ser complementario a lo largo de 20 pb, y dejar extremos cohesivos complementarios al vector abierto en los extremos 5' y 3'. El alineado de se aprecia en un tamaño de 25 pb como se aprecia en la figura 16.

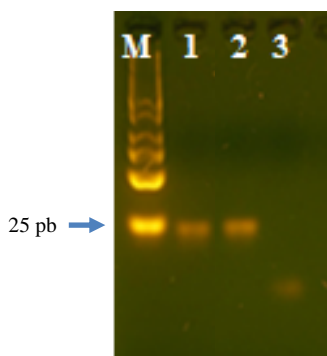


Figura 16. Alineamiento de los oligonucleótidos descritos en la Tala iv. M) Marcador de bajo peso molecular; 1) Alineamiento correspondiente al gRNA del gen IPK; 2) Alineamiento correspondiente al gRNA del gen AHAS; 3) Oligonucleótido sin alinear como control

La correcta clonación se evaluó mediante la digestión y PCR de las colonias obtenidas de la ligación como se muestra en la figura 17. Las clonas positivas se corroboraron mediante secuenciación.

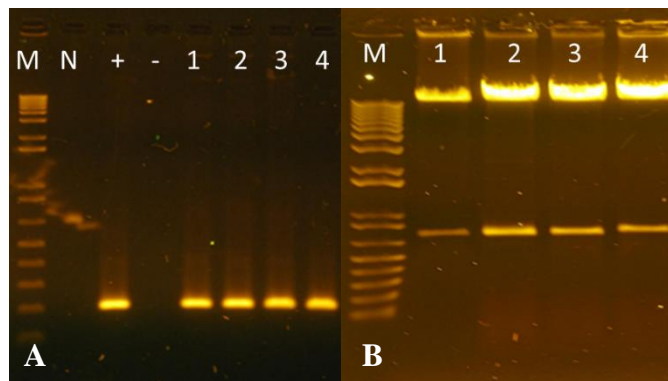


Figura 17. Corroboración de la correcta clonación del inserto. A) PCR punto final U6p-gRNA: N, NTC; +, control positivo; -, control negativo; 1-4, muestra a analizar. B) Digestión mediante HindIII para liberación del inserto: 1-4) muestras a analizar

Se seleccionaron dos clonas las cuales se encontraban ya analizadas por los resultados de secuenciación. De este análisis resultaron dos plásmidos que se aplicaron durante este trabajo. Aquel cuyo gRNA se dirigirá a la secuencia de la soya de Inositol Pentakisfosfato Cinasa (pBUN4U6-IPK) y el clonado con el gRNA que dirigirá a la proteína Cas hacia el gen que codifica para la acetolactato Sintasa (pBUN4U6-AHAS).

5.2. Construcción del vector DONOR-AHAS

De acuerdo con varios reportes, los experimentos de knock in dirigidos por CRISPR/Cas9 se han realizado con un DONOR como inserto en plásmido de longitud total variando desde 0.47 hasta 5.2 kb, para modificaciones de fragmentos variando desde 46 pb hasta 1.6 kb, respectivamente (Collonnier et al., 2017).

Por otro lado, Shi y colaboradores (2017) corroboraron, aunque no en el mismo sitio de edición, que el DNA de cadena simple proveniente de un producto de PCR comparado con el DONOR

circular presenta menores porcentajes de eficiencia de edición de las plantas transformadas. En este mismo estudio se comprobó que, con un 2.05-4.1% de edición, la eficiencia del sistema aumento ampliamente con un DONOR de longitud total de 1 kb en cada extremo y una inserción de 3.1 kb. Por lo anterior se realizó la construcción del DONOR de AHAS en un vector pCR8/TOPO. Se diseñaron oligonucleótidos para obtener las secuencias flanqueantes al sitio de edición, marcando el cambio de secuencia de C a T en las posiciones 537 y 539 para lograr el cambio de Prolina a Serina como se indica en la tabla xi.

Tabla xi. Secuencia para la obtención de los brazos homólogos para realizar la recombinación homóloga en explantes de soya. Indicadas en dirección 5'-3'

1	DON ALS-short Left Fw	GTTCACCGCCCTCCCTTCATCAT
2	DON ALS-short Right R	CATAAACAGTCCCGTGCATCCCAA
3	DON ALS-Fw Ser-Ala	CCCTCGTCGCCATCACCGGCCAGGTCTCTCGCCGCATGATCGCC
4	DON ALS-Rv Ser-Ala	GGGGTTTCTTGGAAGGCGTCGGTGCCGATCATGCGGCGAGAGACCTG
5	DON ALS-Fw Ser-Gly	CCCTCGTCGCCATCACCGGCCAGGTCTCTCGCCGCATGATCGGC
6	DON ALS-Rv Ser-Gly	GGGGTTTCTTGGAAGGCGTCGGTGCCGATCATGCGGCGAGAGACCTG

La consideración más importante para el diseño del DONOR fue el cambio de la región PAM (CGG). Para esto, se observó la secuencia de aminoácidos para la cual codificaba esta secuencia en la soya.

La secuencia PAM de este gen se encuentra posicionado dentro del sitio de traducción para dos aminoácidos. Debido a que el codón para Glicina es el triplete GGN, el cambio de secuencia PAM significaría un cambio en la cadena de aminoácidos codificada en este gen. Para estudiar si la región a editar AITGQVPRRMIGTDAFQETPIVE se encontraba altamente conservada se realizó un BLAST de proteína en una amplia selección de organismos, el BLAST arrojó un cambio en la secuencia proteica del gen AHAS en *Bacillus aeruginosus*.

Con los resultados obtenidos, se observó que existía un cambio en la secuencia de Glicina a Alanina, por lo cual se diseñaron los oligonucleótidos para la obtención del locus con un cambio de CGG a CGC de la posición 552 a la 554.

Se obtuvieron los brazos homólogos del LOCUS del gen AHAS mediante amplificación por PCR como se observa en la Figura 18. La primera construcción DONOR consistió en el cambio del codón CCC→TCT manteniendo la secuencia silvestre para la región PAM, a este plásmido le denominamos pCR8-AHAS-Gly. La segunda construcción corresponde al cambio del codón que codifica a Prolina por un codón codificante para Serina y, adicionalmente, el cambio en la región PAM que codificará para Alanina, este plásmido se llamó pCR8-AHAS-Ala.

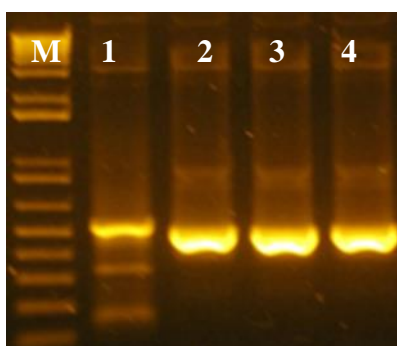


Figura 18. Amplificación por PCR punto final de los brazos homólogos con cambio en el ORF M, Marcador 1 Kb plus; 1, amplificación de extremo 5' cambios P→S; 2, amplificación de extremo 3' cambios P→S; 3, amplificación de extremo 5' cambios P→S y G→A; 4, amplificación de extremo 3' cambios P→S y G→A

Los brazos homólogos, con los cambios esperados en la secuencia se alinearon mediante una rampa de temperatura, realizando también una reamplificación para su posterior clonación. El tamaño esperado del alineado completo se esperaba en 989 pb, para su evaluación se observó en un gel como se muestra en la Figura 19.

6.1. IPK

El vector de edición pBUN4U6-IPK se entregó a las células de la planta mediante la estrategia de *A. tumefaciens*. Esta estrategia ha sido probada ampliamente en soya mediante diversas técnicas. En la literatura, Paz y colaboradores (2006) reportaron una relativa alta de transformación en embriones de soya siendo esta de 4 %.

Los explantes de soya colocados en medio con agente de selección (Glufosinato de amonio DL₅₀) que lograron sobrevivir después de 30 días, fueron evaluados mediante PCR punto final para corroborar la presencia del transgén Cas9 de 400 pb (Figura 21a). De la misma manera, para realizar la doble corroboración de la presencia del cassette de edición se realizó la amplificación de un fragmento de 232 pb el cual corresponde al tamaño desde el promotor U6 hasta el extremo 3' del scaffold del gRNA (Figura 21b).

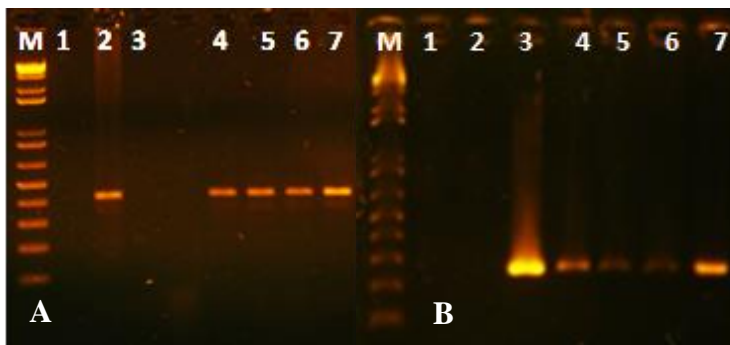


Figura 21. PCR punto final de detección Cas9 M, marcador 1 kb plus; 1, NTC; 2, control plasmídico positivo; 3, Control negativo silvestre; 4, control positivo At; 5-7, muestras a analizar. b) PCR punto final de detección promotor U6 a scaffold. M, marcador 1 kb plus; 1, NTC; 2, control negativo silvestre; 3, control plasmídico positivo; 4, control positivo At; 5-7, muestras a analizar.

Con esta corroboración, se pudieron delimitar las plantas que se encuentran efectivamente transformadas debido a la inserción de T-DNA en el genoma. En contraste con los resultados de Paz (2006) en este proyecto los porcentajes de transformación no se observan altos. Sin embargo, se lograron obtener líneas estables que contienen el cassette de edición.

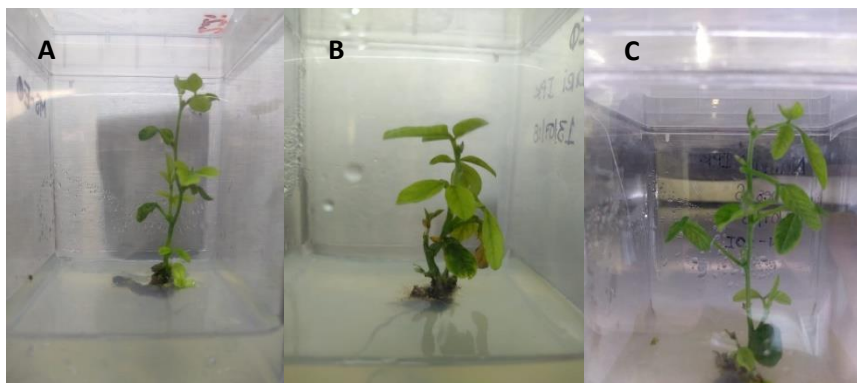


Figura 22. Líneas transformadas mediante *Agrobacterium tumefaciens* en medio de selección con glufosinato de amonio.

El resumen de las eficiencias de transformación mediante *Agrobacterium tumefaciens* se muestra en la tabla. A pesar de que la eficiencia de transformación es baja comparado con Paz en el 2006, se obtuvieron líneas estables expresoras de Cas9.

Tabla xii. Resumen de las eficiencias de transformación de los explantes sometidos al proceso de transformación por *A. tumefaciens* AGL1

pBUN4U6-IPK	Eficiencia
Explantes transformados	650 (100%)
Explantes regenerados	210 (32%)
Explantes en selección	25 (3.8%)
Explantes positivos	3 (0.5%)

6.2.AHAS

En el caso de la transformación del plásmido, la cual se transformó mediante biobalística está reportada como una estrategia con una mayor eficiencia de transformación aproximándose al 9 % de acuerdo a lo reportado por Rech y colaboradores (2008) y Soto (2016). Se realizó el bombardeo con la aplicación de dos diferentes donadores (pCR8-AHAS-Gly y pCR8-AHAS-Ala) para la recombinación homóloga. Se realizó la confirmación de presencia del T-DNA mediante PCR del ORF de Cas9 (Figura 23).

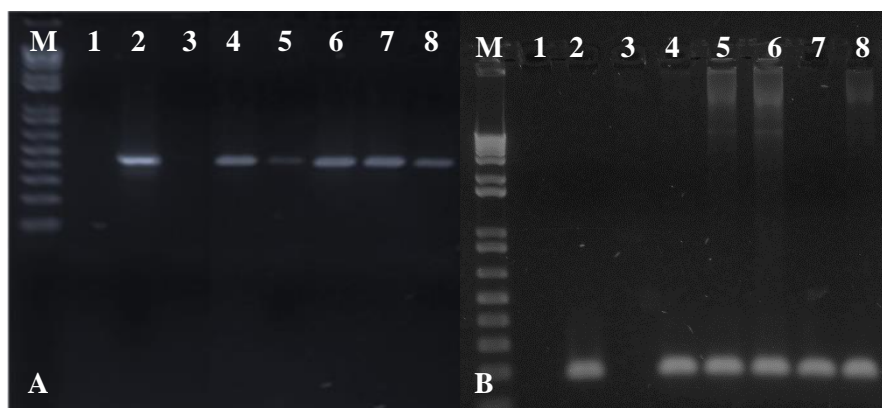


Figura 23. a) PCR punto final de detección Cas9 para el gen AHAS. M, marcador 1 kb plus; 1, NTC; 2, control positivo; 4-8, muestras a analizar. b) PCR punto final de detección promotor U6 a scaffold. M, marcador 1 kb plus; 1, NTC; 2, control positivo; 3, control negativo (silvestre); 4-7, muestras a analizar.

El resumen de las eficiencias obtenidas se muestra a continuación:

Tabla xiii. Resumen de eficiencias obtenidas mediante transformación por biobalística

pBUN4U6-ALS	pCR8-AHAS-Gly	pCR8-AHAS-Ala
Explantes transformados	470 (100%)	500 (100%)
Explantes regenerados	58 (12.3%)	50 (10%)
Explantes en selección	10 (2.12%)	7 (1.4%)
Explantes positivos	3 (0.70%)	2 (0.40%)

Como se puede observar en la Tabla xii, existe mediante este método una baja eficiencia (0.4 – 0.7 %) de transformación en comparación con lo reportado por Soto y Rech. Las eficiencias generales de este protocolo son bajas, esto se puede deber en primer lugar a la variedad con la que se trabajó en este proyecto, pues se ha demostrado que la eficiencia de transformación depende muchas veces de los varietales. Por otro lado, la baja eficiencia de transformación en el marco de este proyecto puede deberse al gen elegido para realizar la modificación. Ya que este gen (AHAS) es de vital importancia en la síntesis de aminoácidos de cadena ramificada (Isoleucina, Leucina y Valina), una mutación en AHAS puede resultar para la planta de estudio. Lo anterior explicaría el bajo porcentaje de regeneración después de la transformación.

7. Identificación de edición

Una vez que se seleccionaron las plantas transformadas para ambos genes. Se procedió a analizar si existía un cambio en la secuencia de interés. Se implementó un ensayo de digestión mediante la aplicación de la Endonucleasa T7 propuesto en el 2010 por Guschin y colaboradores. Esta endonucleasa tiene como actividad reconocer cortes en DNA de doble cadena cuyo apareamiento no sea perfecto, con una cobertura desde 1 pb.

Amplificación y alineamiento se realizó considerando que las ediciones se dan generalmente en un único alelo, por lo tanto, si existen modificaciones de uno de los alelos, el alineamiento podrá darse de manera no específica en el sitio de edición, permitiendo la acción de la enzima.

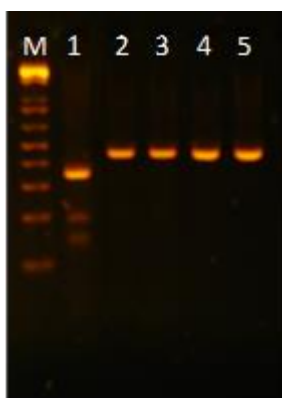


Figura 24. Ensayo de la endonucleasa T7 para el gen IPK. M, marcador 1 kb plus; 1, control positivo de At editada; 2; control negativo silvestre; 3-5, muestras a analizar

En la figura 24 se muestran los resultados de la reacción de digestión por la endonucleasa T7. Como se puede observar en dicha imagen, no existe edición en ninguno de los alelos de la planta. Para corroborar lo anterior, se realizó la clonación de las secuencias de interés en un vector de conservación pCR2.1 para la secuenciación.

Se evaluaron 5 colonias con el inserto de IPK de cada una de las plantas transformadas. Los resultados se observan en la Figura 25.

Species/Abb.	Group Name	* *
1. Gm_IPK_wt		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
2. IPKL1-C1		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
3. IPKL1-C2		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
4. IPKL1-C3		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
5. IPKL1-C4		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
6. IPKL1-C5		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
7. IPKL2-C1		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
8. IPKL2-C2		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
9. IPKL2-C3		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
10. IPKL2-C4		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
11. IPKL2-C5		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
12. IPKL3-C1		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
13. IPKL3-C2		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
14. IPKL3-C3		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
15. IPKL3-C4		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A
16. IPKL3-C5		A C T C T G A A A T T G C T T C A A G G A G A G G T A C A


 PAM

Figura 25. Alineamiento de secuencias de las líneas transformadas con el plásmido pBUN4U&-IPK.

El cuadro indica la secuencia del gRNA. La flecha señala la región PAM

En el caso del gen AHAS se realizó también el ensayo de la endonucleasa T7, para este gen tampoco se observó la actividad de la enzima (Figura 26), por lo cual se estima que en las plantas transformadas no existió modificación en la secuencia. Debido a lo anterior, es necesaria la corroboración mediante secuenciación de las plantas transformadas obtenidas mediante el protocolo de Biobalística.

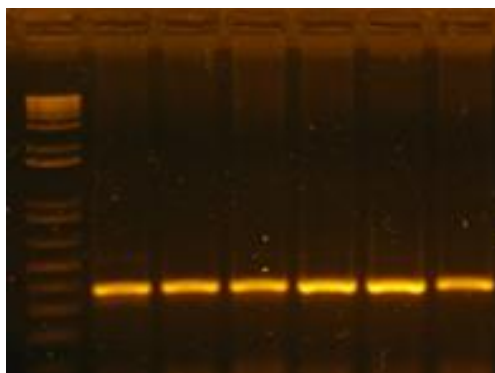


Figura 26. Ensayo de la endonucleasa T7 para el gen IPK. M, marcador 1 kb plus; 1, control positivo de At editada; 2; control negativo silvestre; 3-5, muestras a analizar

Como se puede apreciar, para este caso particular, no existieron modificaciones en las secuencias de interés. Esto puede deberse a diversas razones. Como reporta Xu y colaboradores (2015) el porcentaje de edición utilizando como estrategia knock out alcanza porcentajes desde 2.4 a 6.25% de las plantas transformadas totales. Lo anterior, podría justificar la nula tasa de edición en los explantes de soya, debido a que se debieron tener al menos 100 explantes de soya transformados para obtener de 2 a 6 plantas modificadas. Por otro lado, la baja expresión de la proteína Cas9 y/o gRNA en los explantes de soya podrían evitar el funcionamiento de la maquinaria de edición. De manera general, se ha reportada que, con el uso de promotores de RNA pequeño (U6) se utilice para el inicio de la transcripción una Guanina. Sin embargo, considerando que durante el diseño se seleccionaron los gRNA con los mejores puntajes y más cercanos al sitio de edición, en ambos casos el nucleótido inicial no corresponde a G. A pesar de que la guanina no es el codón de preferencia para el promotor U6, Zhu y colaboradores (2015) reportaron que la expresión de los RNA de tamaño pequeño en maíz esta potenciada, mas no limitada por el nucleótido G como iniciador de la transcripción. De acuerdo con lo reportado, el gRNA diseñado se le puede adicionar

una G al inicio de la secuencia nucleotídica, con lo cual se espera una eficiencia de modificación más alta.

En casos más generales, Ma y colaboradores (2014) determinaron que con el uso del promotor U6, la expresión del RNA no inicia en todos los casos de la guanina en la posición uno, sino que puede saltarse hasta dos nucleótidos de la cadena de DNA, cualquiera que fuera el nucleótido.

CONCLUSIONES

- Los genes de la soya más viables para realizar la edición fueron IPK, gen involucrado en la ruta de síntesis del ácido fítico, y AHAS, gen involucrado en el funcionamiento del herbicida sulfonilurea.
- Las secuencias de gRNA seleccionadas son las óptimas para realizar la edición al encontrarse cerca del sitio de unión a sustrato de ambas proteínas.
- Para realizar la transformación de soya, a metodología de transformación por *Agrobacterium tumefaciens* suponen una ventaja para la obtención de plantas transgénicas.
- Se obtuvieron 3 líneas independientes positivas al cassette de edición pBUN4U6-IPK con una eficiencia del 0.5 % en los explantes de soya. Las líneas positivas para el cassette de edición pBUN4U6-AHAS fue del 0.4 al 0.7 %.

REFERENCIAS

1. Amiard, S., Depeiges, A., Allain, E., White, C. I., & Gallego, M. E. (2011). Arabidopsis ATM and ATR Kinases Prevent Propagation of Genome Damage Caused by Telomere Dysfunction. *The Plant Cell*, 23(12), 4254–4265. <https://doi.org/10.1105/tpc.111.092387>
2. Anders, C., Niewoehner, O., & Jinek, M. (2015). In Vitro Reconstitution and Crystallization of Cas9 Endonuclease Bound to a Guide RNA and a DNA Target. In *Methods in Enzymology* (Vol. 558, pp. 515–537). Elsevier.
3. Barrangou, R., & Horvath, P. (2017). A decade of discovery: CRISPR functions and applications. *Nature Microbiology*, 2(June), 1–9. <https://doi.org/10.1038/nmicrobiol.2017.92>
4. Belhaj, K., Chaparro-Garcia, A., Kamoun, S., & Nekrasov, V. (2013). Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods*, 9(1), 39. <https://doi.org/10.1186/1746-4811-9-39>
5. Binder, S. (2010). Branched-Chain Amino Acid Metabolism in *Arabidopsis thaliana*. *The Arabidopsis Book*, 8, e0137. <https://doi.org/10.1199/tab.0137>
6. Bortesi, L., & Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, 33(1), 41–52. <https://doi.org/10.1016/j.biotechadv.2014.12.006>
7. Brown, H. M. (1990). Mode of action, crop selectivity, and soil relations of the sulfonylurea herbicides. *Pesticide Science*, 29(3), 263–281. <https://doi.org/10.1002/ps.2780290304>
8. Chang, A. K., & Duggleby, R. G. (1998). Herbicide-resistant forms of Arabidopsis thaliana acetohydroxyacid synthase: characterization of the catalytic properties and sensitivity to inhibitors of four defined mutants. *Biochem. J.*, 333, 765–777.

<https://doi.org/10.1042/bj3330765>

9. Char, S. N., Neelakandan, A. K., Nahampun, H., Frame, B., Main, M., Spalding, M. H., ... Yang, B. (2017). An Agrobacterium-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. *Plant Biotechnology Journal*, *15*(2), 257–268. <https://doi.org/10.1111/pbi.12611>
10. Chen, K., & Gao, C. (2015). Targeted Gene Mutation in Plants. In X.-Q. Li, D. J. Donnelly, & T. G. Jensen (Eds.), *Somatic Genome Manipulation: Advances, Methods, and Applications* (pp. 253–272). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4939-2389-2_12
11. Chen, Y., Wang, Z., Ni, H., Xu, Y., Chen, Q., & Jiang, L. (2017). CRISPR/Cas9-mediated base-editing system efficiently generates gain-of-function mutations in Arabidopsis. *Science China Life Sciences*, *60*(5), 520–523. <https://doi.org/10.1007/s11427-017-9021-5>
12. Collonnier, C., Guyon-Debast, A., Maclot, F., Mara, K., Charlot, F., & Nogu e, F. (2017). Towards mastering CRISPR-induced gene knock-in in plants: Survey of key features and focus on the model *Physcomitrella patens*. *Methods*, *121–122*, 103–117. <https://doi.org/10.1016/j.ymeth.2017.04.024>
13. Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213), 1258096–1258096. <https://doi.org/10.1126/science.1258096>
14. East-Seletsky, A., O’Connell, M. R., Knight, S. C., Burstein, D., Cate, J. H. D., Tjian, R., & Doudna, J. A. (2016). Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature*, *538*(7624), 270–273. <https://doi.org/10.1038/nature19802>
15. Gaj, T., Gersbach, C. A., & Barbas, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based

- methods for genome engineering. *Trends in Biotechnology*, 31(7), 397–405.
<https://doi.org/10.1016/j.tibtech.2013.04.004>
16. Goh, C. H., Ko, S. M., Koh, S., Kim, Y. J., & Bae, H. J. (2012). Photosynthesis and Environments: Photoinhibition and Repair Mechanisms in Plants. *Journal of Plant Biology*, 55(2), 93–101. <https://doi.org/10.1007/s12374-011-9195-2>
17. González, B., Baños-Sanz, J. I., Villate, M., Brearley, C. A., & Sanz-Aparicio, J. (2010). Inositol 1,3,4,5,6-pentakisphosphate 2-kinase is a distant IPK member with a singular inositide binding site for axial 2-OH recognition. *Proceedings of the National Academy of Sciences USA*, 107(21), 9608–9613. <https://doi.org/10.1073/pnas.0912979107>
18. Guha, T. K., Wai, A., & Hausner, G. (2017a). Programmable Genome Editing Tools and their Regulation for Efficient Genome Engineering. *Computational and Structural Biotechnology Journal*, 15, 146–160. <https://doi.org/10.1016/j.csbj.2016.12.006>
19. Guha, T. K., Wai, A., & Hausner, G. (2017b). Programmable Genome Editing Tools and their Regulation for Efficient Genome Engineering. *Computational and Structural Biotechnology Journal*, 15, 146–160. <https://doi.org/10.1016/j.csbj.2016.12.006>
20. Haughn, G. W. G. W., & Somerville, C. (1986). Sulfonylurea-resistant mutants of *Arabidopsis thaliana*. *MGG Molecular & General Genetics*, 204(3), 430–434.
<https://doi.org/10.1007/BF00331020>
21. Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C., & Schmid, B. (2013). Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development*, 140(24), 4982–4987. <https://doi.org/10.1242/dev.099085>
22. Humer, E., & Zebeli, Q. (2015). Phytate in feed ingredients and potentials for improving the utilization of phosphorus in ruminant nutrition. *Animal Feed Science and Technology*, 209, 1–15. <https://doi.org/10.1016/j.anifeedsci.2015.07.028>

23. Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., ... Lin, S. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, *343*(6176), 1247997.
24. Kingdom, U., Kingdom, U., London, F., Health, N., Foundation, S., Kingdom, U., & Kingdom, U. (2017). Two Distinct RNase Activities of CRISPR-C2c2 Enable Guide RNA Processing and RNA Detection. *Nature*, *21*(2), 129–139. <https://doi.org/10.5588/ijtld.16.0716.Isoniazid>
25. Lakhanpaul, S. (2006). Single nucleotide polymorphism (SNP)–Methods and applications in plant genetics: A review DNA FINGERPRINTING OF NATIVE BIODIVERSITY View project. *Article in Indian Journal of Biotechnology*.
26. Lee, K. Y., Townsend, J., Tepperman, J., Black, M., Chui, C. F., Mazur, B., ... Bedbrook, J. (1988). The molecular basis of sulfonylurea herbicide resistance in tobacco. *The EMBO Journal*, *7*(5), 1241–8. <https://doi.org/16453837>
27. Lee, Y. T., & Duggleby, R. G. (2001). Identification of the regulatory subunit of *Arabidopsis thaliana* acetohydroxyacid synthase and reconstitution with its catalytic subunit. *Biochemistry*, *40*(23), 6836–6844. <https://doi.org/10.1021/bi002775q>
28. Li, Q.-T., Lu, X., Song, Q.-X., Chen, H.-W., Wei, W., Tao, J.-J., ... Zhang, J.-S. (2017). Selection for a Zinc-Finger Protein Contributes to Seed Oil Increase during Soybean Domestication. *Plant Physiology*, *173*(4), 2208–2224. <https://doi.org/10.1104/pp.16.01610>
29. Li, Z., Hayashimoto, A., & Murai, N. (1992). A Sulfonylurea Herbicide Resistance Gene from *Arabidopsis thaliana* as a New Selectable Marker for Production of Fertile Transgenic Rice Plants. *Plant Physiology*, *100*(2), 662–668. <https://doi.org/10.1104/pp.100.2.662>
30. Li, Z., Liu, Z.-B., Xing, A., Moon, B. P., Koellhoffer, J. P., Huang, L., ... Cigan, A. M. (2015). Cas9-Guide RNA Directed Genome Editing in Soybean. *Plant Physiology*, *169*(2),

- 960–970. <https://doi.org/10.1104/pp.15.00783>
31. Liang, Z., Zhang, K., Chen, K., & Gao, C. (2014). Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. *Journal of Genetics and Genomics*, *41*(2), 63–68. <https://doi.org/10.1016/j.jgg.2013.12.001>
32. Lowder, L. G., Zhang, D., Baltus, N. J., Paul, J. W., Tang, X., Zheng, X., ... Qi, Y. (2015). A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. *Plant Physiology*, *169*(2), 971–985. <https://doi.org/10.1104/pp.15.00636>
33. Ma, H., Wu, Y., Dang, Y., Choi, J. G., Zhang, J., & Wu, H. (2014). Pol III promoters to express small RNAs: Delineation of transcription initiation. *Molecular Therapy - Nucleic Acids*, *3*(May), e161. <https://doi.org/10.1038/mtna.2014.12>
34. Oladosu, Y., Rafii, M. Y., Abdullah, N., Hussin, G., Ramli, A., Rahim, H. A., ... Usman, M. (2016). Principle and application of plant mutagenesis in crop improvement: A review. *Biotechnology and Biotechnological Equipment*, *30*(1), 1–16. <https://doi.org/10.1080/13102818.2015.1087333>
35. Paz, M. (2009). *Agrobacterium*-mediated transformation of soybean and recovery of transgenic soybean plants, 1–6.
36. Rasmussen, S. K., Ingvarsen, C. R., & Torp, A. M. (2010). Mutations in genes controlling the biosynthesis and accumulation of inositol phosphates in seeds. *Biochemical Society Transactions*, *38*(2), 689–694. <https://doi.org/10.1042/BST0380689>
37. Ray, T. B. (1982). The mode of action of chlorsulfuron: A new herbicide for cereals. *Pesticide Biochemistry and Physiology*, *17*(1), 10–17. [https://doi.org/10.1016/0048-3575\(82\)90120-1](https://doi.org/10.1016/0048-3575(82)90120-1)
38. Rech, E. L., Vianna, G. R., & Aragão, F. J. L. (2008). High-efficiency transformation by biolistics of soybean, common bean and cotton transgenic plants. *Nature Protocols*, *3*(3),

- 410–418. <https://doi.org/10.1038/nprot.2008.9>
39. Redekar, N. R., Biyashev, R. M., Jensen, R. V., Helm, R. F., Grabau, E. A., & Maroof, M. A. S. (2015). Genome-wide transcriptome analyses of developing seeds from low and normal phytic acid soybean lines. *BMC Genomics*, *16*(1), 1–16. <https://doi.org/10.1186/s12864-015-2283-9>
40. Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology*, *32*(4), 347–350. <https://doi.org/10.1038/nbt.2842>
41. Shi, J., Gao, H., Wang, H., Lafitte, H. R., Archibald, R. L., Yang, M., ... Habben, J. E. (2017). ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnology Journal*, *15*(2), 207–216. <https://doi.org/10.1111/pbi.12603>
42. Shi, J., Wang, H., Wu, Y., Hazebroek, J., Meeley, R. B., & Ertl, D. S. (2003). The maize low-phytic acid mutant *lpa2* is caused by mutation in an inositol phosphate kinase gene. *Plant Physiology*, *131*(2), 507–15. <https://doi.org/10.1104/pp.014258>
43. Soto, N., Delgado, C., Hernández, Y., Rosabal, Y., Ferreira, A., Pujol, M., ... Enríquez, G. A. (2017). Efficient particle bombardment-mediated transformation of Cuban soybean (INCASoy-36) using glyphosate as a selective agent. *Plant Cell, Tissue and Organ Culture*, *128*(1), 187–196. <https://doi.org/10.1007/s11240-016-1099-x>
44. Stevenson-Paulik, J., Bastidas, R. J., Chiou, S.-T., Frye, R. a, & York, J. D. (2005). Generation of phytate-free seeds in *Arabidopsis* through disruption of inositol polyphosphate kinases. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(35), 12612–12617. <https://doi.org/10.1073/pnas.0504172102>
45. Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome

- editing with engineered zinc finger nucleases. *Nature Reviews Genetics*, *11*(9), 636–646.
<https://doi.org/10.1038/nrg2842>
46. Xu, R. F., Li, H., Qin, R. Y., Li, J., Qiu, C. H., Yang, Y. C., ... Yang, J. B. (2015). Generation of inheritable and “transgene clean” targeted genome-modified rice in later generations using the CRISPR/Cas9 system. *Scientific Reports*, *5*(May), 1–10.
<https://doi.org/10.1038/srep11491>
47. Yin, H., Kauffman, K. J., & Anderson, D. G. (2017). Delivery technologies for genome editing. *Nature Reviews Drug Discovery*, *16*(6), 387–399.
<https://doi.org/10.1038/nrd.2016.280>
48. Yin, K., Gao, C., & Qiu, J.-L. (2017). Progress and prospects in plant genome editing. *Nature Plants*, *3*(8), 17107. <https://doi.org/10.1038/nplants.2017.107>
49. Yuan, F. J., Zhu, D. H., Tan, Y. Y., Dong, D. K., Fu, X. J., Zhu, S. L., ... Shu, Q. Y. (2012). Identification and characterization of the soybean IPK1 ortholog of a low phytic acid mutant reveals an exon-excluding splice-site mutation. *Theoretical and Applied Genetics*, *125*(7), 1413–1423. <https://doi.org/10.1007/s00122-012-1922-7>
50. Zetsche, B., Gootenberg, J., Abudayyeh, O., Slaymaker, I., Makarova, K., Essletzbichler, P., Zhang, F. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*, *163*(3), 759–771. <https://doi.org/10.1016/j.cell.2015.09.038>
51. Zhao, H. J., Cui, H. R., Xu, X. H., Tan, Y. Y., Fu, J. J., Liu, G. Z., ... Shu, Q. Y. (2013). Characterization of OsMIK in a rice mutant with reduced phytate content reveals an insertion of a rearranged retrotransposon. *Theoretical and Applied Genetics*, *126*(12), 3009–3020. <https://doi.org/10.1007/s00122-013-2189-3>
52. Zhu, J., Song, N., Sun, S., Yang, W., Zhao, H., Song, W., & Lai, J. (2016). Efficiency and Inheritance of Targeted Mutagenesis in Maize Using CRISPR-Cas9. *Journal of Genetics*

and Genomics, 43(1), 25–36. <https://doi.org/10.1016/j.jgg.2015.10.006>