



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

UNIDAD ZACATENCO

DEPARTAMENTO DE BIOTECNOLOGÍA Y BIOINGENIERIA

**Inducción de resistencia a roya asiática causada por
Phakopsora pachyrhizi en soya (*Glycine max*)**

Tesis que presenta:

Edgar Yebrán Villegas Vázquez

Para obtener el grado de

DOCTOR EN CIENCIAS

EN LA ESPECIALIDAD DE BIOTECNOLOGÍA

Director de la Tesis: **Dr. Roberto Ruiz Medrano**

Ciudad de México

junio de 2022

Este trabajo se realizó en el laboratorio de Biología Molecular del departamento de Biotecnología y Bioingeniería bajo la dirección del Dr. Roberto Ruiz Medrano y la Dra. Beatriz Xoconostle Cázares.

Agradecimientos

Al consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo financiero mediante el proyecto número FC-1234 y la beca para el cumplimiento del programa de doctorado.

A los doctores Roberto Ruiz Medrano y Beatriz Xoconostle por abrimme las puertas de su laboratorio, por su tiempo, dedicación, paciencia y conocimiento que me brindaron.

A mis asesores Dra. Berenice Calderón, Dra. Rosalía Lira, Dra. María del Carmen Montes y Dr. Jesús Hinojosa gracias por su apoyo, conocimiento y tiempo.

A los auxiliares Brenda Vargas y Leandro Núñez por su asesoría, apoyo técnico y logístico.

A José Luis por su apoyo en el laboratorio y en cualquier cosa que se necesitara, así como sus platicas y consejos.

A todos mis compañeros de laboratorio por su apoyo incondicional.

A mis amigos, Gabriel Pérez, Micaela Pucheta, Mario Rojas y Domingo por su amistad, apoyo y por todos los momentos que vivimos juntos.

A la Dra. Laura Itzel Quintas Granados por ser mi compañera de vida, por su apoyo incondicional, sus consejos, su amor, por ser mi maestra de paciencia y por compartir tantos momentos, infinitamente gracias.

A mis papás Angelica Vázquez y Heriberto Villegas, eternamente agradecido por su amor incondicional, su apoyo, su confianza, consejos y educación que han sido parte importante en el desarrollo de mi vida.

A mi hermana Brisna Villegas por ser un ejemplo de carácter y templanza, a mi hermano Ronald Jardel por su fortaleza emocional e intelectual.

Resumen

La soya (*Glycine max*) tiene importancia a nivel mundial en la industria alimentaria por su alto contenido en aceites y proteínas. Sin embargo, los cultivos de soya son afectados por plagas y enfermedades, como la roya causada por el hongo *Phakopsora pachyrhizi*, generando pérdidas de hasta 90% del cultivo. Las proteínas AvrPphB susceptible1 (PBS1) y la de resistencia a *Pseudomonas syringae* 5 (RPS5) están implicadas en la transducción de señales para evocar la respuesta inmunitaria innata de las plantas. En *Arabidopsis thaliana*, PBS1 es escindida por la proteasa AvrPphB (del inglés *Pseudomonas phaseolicola* avirulence protein B) activando RPS5 y dando lugar a una respuesta hipersensible (HR). Una estrategia que se ha propuesto para combatir la roya es utilizando estas dos proteínas.

En este trabajo de investigación se planteó la generación de soya con resistencia a hongos patógenos. La estrategia propuesta fue la edición de genes de resistencia, en particular, la edición del gen que codifica para la proteína PBS1 mediante el uso del sistema CRISPR/Cas9. De esta manera, la cinasa ortóloga PBS1 funcionaría como “señuelo” para distintos efectores, permitiendo mayor resistencia. Los resultados sugieren que el método de transformación fue efectivo; sin embargo, no se logró la regeneración de explantes. Adicionalmente, se buscaron ortólogos de PBS1 para rastrear su origen y evolución. El análisis evolutivo sugiere que PBS1 se originó en el linaje que dio lugar a las embriofitas, con las secuencias de las angiospermas formando un clado separado de la proteína de las pteridofitas. En resumen, el uso de la proteína PBS1 para generar resistencia a diversos patógenos es una estrategia prometedora empleando un método novedoso como CRISPR/Cas9.

Abstract

Soybean (*Glycine max*) is a worldwide important crop for the food industry due to its high protein and oil contents. However, the crops are affected by pests and diseases, such as soybean rust caused by the fungus *Phakopsora pachyrhizi* which has the potential to cause up to 90% of the yield losses. A control strategy for this disease is modulating the AvrPphB susceptible1 (PBS1) protein. PBS and the resistance to *Pseudomonas syringae* 5 (RPS5) proteins are involved in signal transduction to elicit the innate immune response of plants. In *Arabidopsis thaliana*, PBS1 is cleaved by the AvrPphB (*Pseudomonas phaseolicola* avirulence protein B) protease activating RPS5 and leading to a hypersensitive response.

The main objective of this work was the generation of a pathogenic fungi resistant soybean line by the genetic edition of the *pbs1* gene using the CRISPR/Cas9 system. Thus, the orthologous PBS1 kinase would function as a "decoy" for different effectors, allowing pathogen resistance. Our research suggested that soybean transformation method was effective. However, we could not be able to achieve explant regeneration. Additionally, we searched for orthologs of PBS1 to analyze its origin and evolution. Our findings suggested that PBS1 originated in the lineage that gave rise to embryophytes, with the angiosperm sequences forming a separate clade from the pteridophytes. Finally, the use of the PBS1 protein to generate resistance to various pathogens using gene editing with CRISPR/Cas9 is promising.

Índice de Figuras	10
Índice de tablas	12
1.Introducción	13
1.1 Soya	16
1.1.2 Origen y distribución	16
1.1.3 Importancia nacional e internacional de la soya	18
1.1.4 Fases de desarrollo de la soya	20
1.1.4.1 Etapa vegetativa	20
1.1.4.2 Etapa reproductiva	21
1.1.5 Taxonomía	23
1.2 Phakopsora pachyrhizi	23
1.2.1 Clasificación taxonómica	25
1.3 CRISPR/Cas: edición de genomas	25
1.3.1 Sistema CRISPR/Cas	27
1.3.2 Mecanismo de CRISPR/Cas9	29
1.3.3 CRISPR/Cas9 en plantas	31
1.3.4 CRISPR para mejora de cultivos	31
1.4 Proteína cinasa de serina/treonina PBS1	32
2. Justificación	36
3. Hipótesis	36
4.Objetivo general	37
4.1.1Objetivos Particulares	37
5. Materiales y Métodos	38
5.1 Metodología general	38
5.1.2 Metodología del análisis bioinformático	38
5.2 Material biológico	39

5.3 Siembra de soya	39
5.4 Extracción de DNA	39
5.5 Análisis in silico para identificar el gen de PBS1 en soya	40
5.6 Secuencia del gen PBS1 de soya	40
5.7 Purificación del producto de PCR correspondiente a PBS1 en soya	42
5.8 Análisis de secuencia y alineación	42
5.9 Extracción de DNA plasmídico	43
5.10 Linealización del plásmido PBUN4116	43
5.11 Diseño de gRNA	44
5.12 Alineamiento de gRNA	44
5.13 Purificación del producto linealizado	45
5.14 Ligación del vector PBUN411U6 con el gRNA-PBS1	45
5.15 Diseño de secuencia DONOR	46
5.16 Obtención de secuencia DONOR en el gen PBS1 de soya	47
5.17 Ensamblaje por PCR	48
5.18 Transformación de E.coli Mach-T1 químicamente competentes	50
5.19 Selección de colonias transformadas	50
5.20 Análisis de restricción	50
5.21 PCR y secuenciación para corroborar presencia de inserto	51
5.22 Transformación de A. tumefaciens por electroporación	52
5.22.1 Preparación de A. tumefaciens para transformación de soya	53
5.23 Transformación genética de soya	53
5.23.1 Esterilización de las semillas	54
5.23.2 Imbibición de semillas	54
5.23.3 Preparación de explantes para la infección	54
5.23.4 Cocultivo	55
5.23.5 Inducción de brotes	55
5.23.6 Elongación de Brotes	56
5.23.7 Enraizamiento de plantas	56

5.24 Transformación de soya por biobalística	57
5.24.1 Extracción de embriones de soya	57
5.24.2 Preparación de los explantes de soya	57
5.24.3 Preparación y recubrimiento de micropartículas de oro	58
5.24.4 Bombardeo de embriones de soya por biobalística	58
5.25 Análisis molecular de las plantas transformadas de soya	60
5.25.1 Análisis con la endonucleasa T7	61
5.26 Identificación de proteínas PBS1 en especies vegetales	62
5.26.1 Identificación de subdominios en secuencias ortólogas de PBS1	63
5.26.2 Análisis filogenético de PBS1	63
5.26.3 Reconstrucción ancestral de secuencias ortólogas de PBS1	64
5.26.4 Modelado de estructura PBS1	65
6.Resultados	66
6.1 Cultivo de soya	66
6.2 Extracción de DNA de plantas de soya	66
6.3 Secuencia in silico del gen PBS1	67
6.4. Presencia del gen PBS1 en variedades de soya y análisis de la secuencia correspondiente al gRNA	69
6.5 Diseño de gRNA para el gen PBS1 de soya	71
6.6 Construcción del vector para edición de genoma	72
6.7 Obtención de secuencia DONOR	76
6.8 Obtención de plantas de soya transgénicas	80
6.9 Transformación por biobalística	83
6.10 Análisis molecular del gen cas9 T0	85
6.11 Análisis molecular del sitio de edición	87
	90
7. Análisis bioinformático de PBS1	90
7.1 Identificación de proteínas PBS1, subdominios de cinasa conservados y motivos involucrados en la respuesta inmunitaria indirecta en PBS1 en el reino vegetal	91
7.2 Árbol filogenético de ortólogos de PBS1 en clorofitas	97

7.3 Análisis filogenético de ortólogos de PBS1 de diferentes linajes de plantas	100
7.4 Reconstrucción ancestral de ortólogos de PBS1 en plantas terrestres	102
7.5 Modelado de estructuras proteicas de ortólogos de PBS1	107
8. Discusión	109
9. Conclusiones	120
10. Bibliografía	121
11. Anexos	126
Artículo	126

Índice de Figuras

Figura 1. Principales países productores de soya 2022 _____	18
Figura 2. Principales Países importadores de soya 2022 _____	19
Figura 3. Superficie y producción de soya _____	20
Figura 4. Estados vegetativos de soya _____	21
Figura 5. Etapa reproductiva de la soya _____	22
Figura 6. Infección por el hongo de la roya asiática en soya _____	24
Figura 8. Fases del sistema CRISPR/Cas _____	28
Figura 9. Edición de genes mediada por CRISPR / Cas9 _____	30
Figura 10. Ingeniería de señuelo PBS1 _____	34
Figura 11. Siembra de soya _____	66
Figura 12. Extracción de DNA _____	67
Figura 13. Secuencia de aminoácidos de PBS1 de <i>A. thaliana</i> . Sustrato, sitio de corte y proteasa. _____	67
Figura 14. Secuencia de aminoácidos de soya (<i>Glycine max</i> cv. Williams 82) _____	68
Figura 15. Secuencia nucleotídica del exón 5 del gen <i>PBS1</i> _____	68
Figura 16. Perfil electroforético del gen <i>PBS1</i> _____	69
Figura 17. Secuencia del fragmento del gen <i>PBS1</i> que incluye el sitio de corte en diferentes variedades de soya _____	70
Figura 18. Diseño de gRNA para el gen <i>PBS1</i> de Soya _____	72
Figura 19. Linealización del Vector PBUN411U6 _____	72
Figura 20. Hibridación del gRNA _____	73
Figura 21. Corroboración de la presencia del gRNA- <i>PBS1</i> en el vector PBUN411U6 _____	74
Figura 22. Secuenciación del vector PBUN411U6-gRNA- <i>PBS1</i> _____	75
Figura 23. Esquema del vector de expresión pBUN411U6-gRNA- <i>PBS1</i> _____	75
Figura 24. Diseño de secuencia DONOR _____	76
Figura 25. Producto de PCR del fragmento del gen <i>PBS1</i> y secuencia DONOR _____	77
Figura 26. Amplificación y purificación de los productos 1.1 y 1.2 _____	77
Figura 27. Ensamblaje de la secuencia DONOR y el gen <i>PBS1</i> _____	78

Figura 28. Figura esquemática del fragmento del gen <i>PBS1</i> con la secuencia DONOR_____	79
Figura 29. Cotiledones de soya en medio de cocultivo_____	80
Figura 30. Explantes en medio de inducción de brotes_____	81
Figura 31. Explantes en medio de elongación (1 semana). _____	81
Figura 32. Explantes en medio de elongación (2 semanas). _____	82
Figura 33. Embriones de soya _____	83
Figura 34. Embriones en medio de elongación_____	85
Figura 35. Detección del gen <i>cas9</i> en plantas transformadas _____	86
Figura 36. Análisis con endonucleasa T7 _____	87
Figura 37. Esquema de edición del gen <i>PBS1</i> de soya mediante CRISPR-Cas9 _____	88
Figura 38. Análisis molecular del fragmento del gen <i>PBS1</i> con la secuencia DONOR_____	88
Figura 39. Análisis de secuenciación del sitio de edición en <i>PBS1</i> _____	89
Figura 40. Detección del gen <i>cas9</i> en plantas de soya transformadas_____	90
Figura 41. Representación esquemática de la activación del efector <i>AvrPphB</i> de <i>P. syringae</i> de <i>RPS5</i> mediada por <i>PBS1</i> que conduce a la respuesta de defensa. _____	91
Figura 42. Análisis de subdominios de cinasa conservados en ortólogos de <i>PBS1</i> en clorofitas. _____	93
Figura 43. Análisis de subdominios de cinasa conservados y motivos importantes en ortólogos de <i>PBS1</i> en plantas terrestres _____	95
Figura 44. Comparación de subdominios y motivos de cinasa conservados en ortólogos de <i>PBS1</i> en plantas terrestres_____	97
Figura 45. Relación filogenética entre <i>PBS1</i> de <i>A. thaliana</i> y proteína cinasas ortólogas de clorofitas. _____	99
Figura 46. Análisis filogenético de ortólogos de <i>PBS1</i> en plantas terrestres _	101
Figura 47. Reconstrucción de secuencias de ortólogos ancestrales de <i>PBS1</i> en plantas terrestres _____	104
Figura 49. Reconstrucción del estado ancestral en ortólogos de <i>PBS1</i> basados en el motivo <i>SEMPH</i> _____	106
Figura 50. Estructuras predichas de <i>PBS1</i> de varias especies de plantas ____	108

Índice de tablas

Tabla 1. Superficie sembrada de soya por entidad federativa _____	19
Tabla 2. El género <i>Glycine</i> _____	23
Tabla 3. Mezcla de reacción PCR punto final _____	41
Tabla 4. Programa de amplificación del gen <i>PBS1</i> _____	41
Tabla 5. Mezcla de alineamiento del gRNA _____	44
Tabla 6. Programa de alineamiento _____	45
Tabla 7. Ligación del gRNA _____	46
Tabla 8. Reacciones de PCR para secuencia DONOR _____	48
Tabla 9. Reacción de alineamiento 1 _____	49
Tabla 10. Reacción de alineamiento 2 _____	49
Tabla 11. Programa de alineamiento _____	49
Tabla 12. Reacción de digestión _____	50
Tabla 13. Reacción de PCR para gRNA _____	51
Tabla 14. Programa de amplificación de gRNA _____	52
Tabla 15. Programa de amplificación. _____	60
Tabla 16. Programa de amplificación del gen <i>pbs1</i> con la modificación ____	61
Tabla 17. Mezcla de reacción de alineamiento para la digestión con endonucleasa T7 _____	62
Tabla 18. Condiciones para el alineamiento de oligonucleótidos _____	62
Tabla 19. Eventos de transformación de embriones de soya mediante el sistema CRISPR/Cas9. _____	84

1.Introducción

La soya (*Glycine max*) es una leguminosa herbácea, erecta, anual y ramificada que tuvo su origen en la región oriental del continente asiático (China), posteriormente se expandió a otros países de Asia, Europa y el continente americano (Agricultura y Rural, 2006; Grassini et al., 2021).

La soya tiene gran importancia a nivel mundial por su demanda en la industria alimentaria. México es uno de los principales importadores de soya aportando internacionalmente un 3% de la comercialización (Shapouri H et al.,2001; SAGARPA, 2022). No obstante, la productividad de este cultivo es impactada por estrés biótico (plagas y enfermedades) y abiótico (frío, sequía y suelos alcalinos) (SAGARPA, 2011). En este sentido, la soya es susceptible a infección por virus, bacterias, hongos y nemátodos, que están en continua evolución y generando nuevas razas de mayor virulencia.

Las plantas en general, y en particular las terrestres, han coexistido con una variedad de microorganismos durante cientos de millones de años a través de diferentes tipos de interacciones (Han, 2019). La coexistencia de plantas y microorganismos ambientales llevó a la evolución de los sistemas vegetales involucrados en la interacción con estos últimos, incluyendo patógenos (Jones J. D. y Dangl J. L., 2006). Tal interacción resultó en el aumento de la complejidad del genoma tanto en plantas como en los microorganismos interactuantes (Chisholm et al., 2006).

Las plantas poseen diversas estrategias de defensa contra patógenos. Una de éstas incluye proteínas receptoras de reconocimiento de patrones (PRR) (Jones J. D. y Dangl J. L., 2006; Schwessinger y Zipfel, 2008), que reconocen patrones moleculares asociados a microbios o patógenos (MAMP o PAMP), como flagelina bacteriana o fragmentos de pared celular en hongos que son PAMP's que activa la inmunidad por PAMP (PTI) (Gómez-Gómez y Boller, 2000; Jones J. D. y Dangl J. L., 2006; Kaku et al., 2006). Las plantas contienen diferentes cinasas similares a

receptores (RLK) que funcionan como PRR (Boller y Felix, 2009; Han, 2019). Las RLK consisten principalmente en un dominio N-terminal de unión a ligando, un dominio transmembranal y un dominio C-terminal de cinasa (Shiu S. H. y Bleecker, 2001; Sun et al., 2017). Éstos tienen un papel importante en el estrés abiótico, las interacciones planta-microbio y el desarrollo de la planta (Shiu S.-H. et al., 2004; Tang et al., 2017). Los receptores citoplásmicos similares a cinasas (RLCK) son un subconjunto de miembros de la familia RLK, sin embargo, carecen de dominios extracelulares y transmembranales. La proteína AvrPphB Susceptible1 (PBS1) y una serie de proteínas similares a PBS1 (PBL), como BIK1 y PBL2, pertenecientes a la subfamilia de RLCK VII participan en la PTI. Las plantas contienen una diversidad de genes denominados de resistencia (R), cuyos productos son capaces de reconocer efectores de patógenos de manera específica, induciendo la respuesta inmune innata de la planta. Existen genes de resistencia que codifican para proteínas que contienen sitios de unión a nucleótidos (NBS) y repetidos ricos en leucina (LRR) denominados NBS-LRR (Jones Jonathan DG y Dangl Jeffery L, 2006). Los productos de estos genes presentan especificidad contra genes de patógenos; sin embargo, se ha observado que mutaciones en dominios LRR permiten cambios en la resistencia a patógenos (Ashfield et al., 2012). Por este hecho, se hipotetiza que la edición de genes de resistencia y su activación podrían generar variedades con un aumento en la mitigación o resistencia al ataque por patógenos.

La roya asiática es una enfermedad causada por el hongo biotrófico *Phakopsora pachyrhizi*. Esta enfermedad se encuentra presente en regiones productoras de soya causando pérdidas severas (Goellner et al., 2010). La respuesta fenotípica de la soya ante el hongo es la presencia de lesiones esporulantes bronquiales, lesiones marrón-rojas con escasas o ninguna esporulación y lesiones no visibles indicativas de una respuesta inmune (Goellner et al., 2010). La roya asiática es controlada mediante el uso de fungicidas, los cuales están perdiendo eficacia. Por lo tanto, existe interés en la generación de plantas con resistencia a enfermedades (Whitham et al., 2016). A pesar de que existen variedades con resistencia a

algunos patógenos, ésta es parcial. Por lo anterior, se justifica el uso de biotecnología que permita ampliar el rango de resistencia a un mayor número de patógenos que inciden en la productividad de este importante cultivo (Whitham et al., 2016).

Los microorganismos patógenos son capaces de introducir efectores en las células huésped; un ejemplo notable de PRR que se ha reclutado en algunas especies para la inmunidad mediada por efectores (ETI), es la cinasa PBS1 y varias proteínas similares a PBS1 (PBL). PBS1 es una proteína cinasa de serina/treonina conservada en angiospermas que está anclada a la membrana plasmática por S-acilación en el extremo N-terminal (Caldwell y Michelmore, 2009). PBS1 de *Arabidopsis thaliana* (AthPBS1) es cortada por AvrPphB, un efector tipo cisteín proteasa de *Pseudomonas syringae*, internalizado por un sistema de secreción de tipo tres (TTSS) (Afzal et al., 2008). La escisión se produce en el motivo GDK, lo que resulta en un cambio conformacional; esto expone otro motivo (la secuencia SEMP, localizada en el bucle C-terminal) que es reconocido y unido por el miembro de la familia de proteínas NLR RPS5. Esta interacción activa RPS5, que a su vez conduce a la posterior inducción de respuesta hipersensible (HR) (Ade et al., 2007; Qi et al., 2014). RPS5 se localiza en la membrana plasmática y se asocia a PBS1 a través de su dominio amino terminal (CC) (Ade et al., 2007). Por lo tanto, PBS1 puede servir como señuelo, protegido por RPS5 para detectar la actividad efectora (Kim S. H. et al., 2016). De acuerdo con investigaciones recientes, muchas proteínas de resistencia podrían ser activadas de manera indirecta por efectores provenientes de los patógenos (Jones Jonathan DG y Dangl Jeffery L, 2006). Por ejemplo, en tomate, la proteína PBS1 es una cinasa que al ser hidrolizada por la proteasa AvrPphB induce la expresión de genes de la respuesta inmune innata (Kourelis et al., 2016).

En este trabajo de investigación se planteó la generación de soya-con resistencia a hongos patógenos que causan enfermedades e inciden negativamente en su productividad. La estrategia propuesta fue la edición de genes de resistencia que

fueran capaces de reconocer efectores diversos para aumentar la resistencia al patógeno. En particular, la edición del gen que codifica para la proteína PBS1, sustrato de la proteasa del patógeno, permitiría ampliar el rango de proteasas que activan el sistema de defensa. De esta manera, la cinasa ortóloga PBS1 funcionaría como “señuelo” para distintos efectores, permitiendo un espectro de resistencia más amplio. El objetivo del proyecto fue analizar los ortólogos de PBS1 en Viridiplantae a través de la caracterización de subdominios y motivos de cinasa en PBS1 en clorofitas y plantas terrestres, así como realizar una reconstrucción ancestral de los motivos de PBS1 y un análisis estructural predictivo de las proteínas de los ortólogos de PBS1 en diferentes linajes de plantas terrestres.

1.1 Soya

La soya (*G. max*) es una planta herbácea, anual y ramificada. Contiene ramas que proceden del tallo principal, presenta hojas, tallos y vainas; su altura puede variar entre 30 cm y 2 m. De acuerdo con su variedad y condiciones ambientales, su ciclo de vida puede ir desde 80 hasta 200 días (Rosas y Young, 1989). Esta oleaginosa del género *Glycine* pertenece a la familia Fabaceae, subfamilia Papilionoideae y tribu Phasoleae. Tuvo su origen en la región oriental del continente asiático (China) y posteriormente se expandió a otros países de Asia, Europa y el continente americano. El genoma de soya es un tetraploide diploidizado con $2n = 40$ cromosomas y un tamaño del genoma de aproximadamente 1115 Mb (Schmutz et al., 2010).

1.1.2 Origen y distribución

En la antigua China, la soya era considerada una de las cinco plantas sagradas junto con la cebada, el trigo, frijoles y el arroz, por lo que cada año se realizaban ceremonias en el que el emperador participaba en la siembra de los frijoles. De acuerdo con la literatura antigua, Sheng -Nung realizó el primer registro de la soya en su obra médica (2,838 a.a.c) describiéndola como una planta de alto valor

nutritivo y con propiedades curativas. Desde sus orígenes en el pueblo chino, la soya se ha empleado como alimento humano ya que era su principal fuente de proteína (Torres y Torres y Tovar-Palacio, 2009).

La soya se desplazó a las naciones de Asia y Europa desde 1950 (Hymowitz, 2008). En EUA la aceptación fue lenta; sin embargo, en 1938 hubo un avance en el desarrollo de la infraestructura para su procesamiento y comercialización. Posteriormente, la soya se convirtió en el cultivo de mayor importancia económica para EUA. En Brasil, la soya no tenía importancia económica; sin embargo, entre los años 1965 y 1975 aumentó la superficie sembrada con un rendimiento del 28%, siendo los principales factores para su expansión los subsidios gubernamentales, y el desarrollo de infraestructura para su procesamiento y comercialización (Hymowitz, 2008).

La introducción de la soya en algunos países de América Latina se debió, en parte, a una lucha en contra de la desnutrición de niños de familias que no podían adquirir fuentes de proteína, como la leche y el huevo (Bourges, 1978; Torres y Torres y Tovar-Palacio, 2009). En 1970, México fue uno de los primeros países latinoamericanos que desarrolló diferentes productos que han utilizado la proteína de soya (PS) en un porcentaje de 20 a 30% para enriquecer o sustituir a la proteína de origen animal de varios productos y hacerlos de esta manera más económicos (Bourges, 1978).

1.1.3 Importancia nacional e internacional de la soya

La soya presenta un alto contenido en proteína y aceite, por lo que tiene gran importancia a nivel mundial; su producción se ha incrementado considerablemente debido a la demanda de la industria alimentaria (Intagri, 2015).

El departamento de Agricultura de Estados Unidos (USDA) estima que la producción de soya en el periodo 2021/2022 será de 350.72 millones de toneladas. El año pasado la producción fue de 367.76 millones por lo que este año podría significar una disminución del 4.63% alrededor del mundo. Los principales productores a nivel internacional son Brasil, EUA, Argentina, China e India. México se posiciona en el lugar 17 (Figura 1) (SAGARPA, 2022).

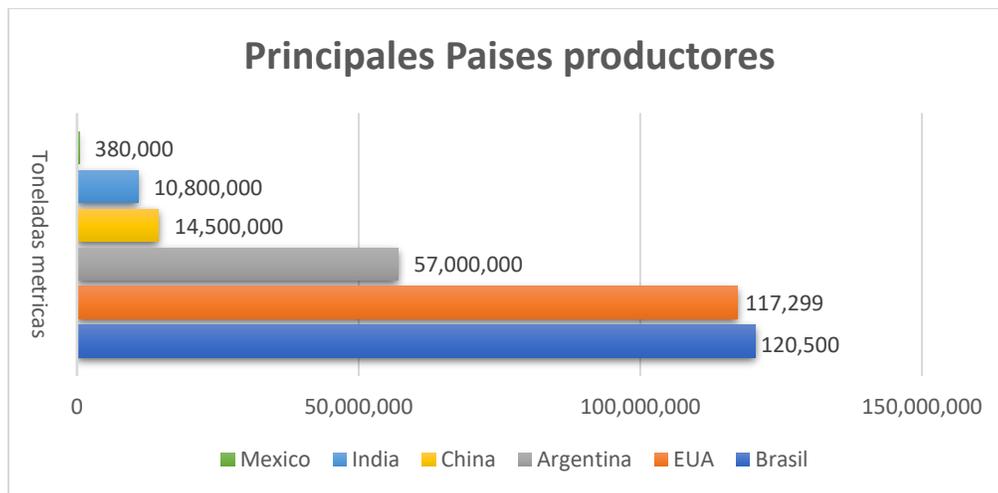


Figura 1. Principales países productores de soya 2022. El principal productor de soya es Brasil, seguido de EUA, Argentina China y en el lugar 17, México (SAGARPA, 2022).

El aumento del comercio internacional por el consumo de la soya ha mejorado el crecimiento de las importaciones. Los principales países importadores son China, la Unión Europea, México, Japón y Tailandia (Shapouri et al.,2001)(SAGARPA, 2022)(Figura 2).

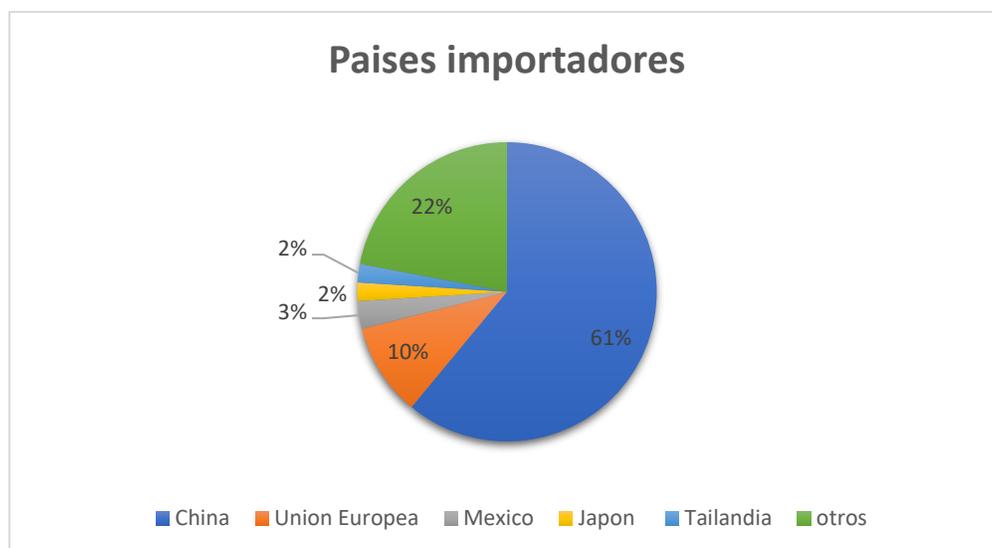


Figura 2. Principales Países importadores de soya 2022. China fue el principal importador de soya seguido de la Unión Europea. (SAGARPA, 2022).

México es uno de los principales importadores de soya a nivel mundial, aportando un 3% de la comercialización a nivel mundial. En Latinoamérica, México es el principal importador de aceite y productos de soya. La Asociación América de la soya despertó el interés para la preparación y usos de la soya en México. En el año 2021 se obtuvieron 4,694 toneladas más que el año anterior. Los estados con mayor participación respecto al volumen nacional obtenido son: Yucatán, Tamaulipas, Chiapas, Campeche (Tabla1).

Tabla 1. Superficie sembrada de soya por entidad federativa
Ciclo otoño-invierno 2019-2020. Avance julio 2021

Entidad	Volumen	obtenido	Var. % anual	Participación % 2021
	2020	2021 ¹		
Nacional	960	5654	489	100
Yucatán	0	4170	ns	73,8
Tamaulipas	591	1142	93,2	20.2
Chiapas	249	258	3.6	4.6
Campeche	0	84	ns	1.5
Nuevo León	120	0	ns	0.0

¹ Cifras preliminares; ns no significativo (SIAP, 2021)

La superficie sembrada en 2021 fue de 3025 más de hectáreas respecto al año anterior; el mismo patrón se observa para la superficie cosechada y la producción con 2541 hectáreas y 4,694 toneladas más comparado con 2020 (Figura 3). Para el año 2022 los principales estados productores son: Tamaulipas (34.3%), Campeche (21.9%), San Luis Potosi (16.7%), Sinaloa (7%), Chiapas (5.6%) y Sonora (4.5%) (SAGARPA).

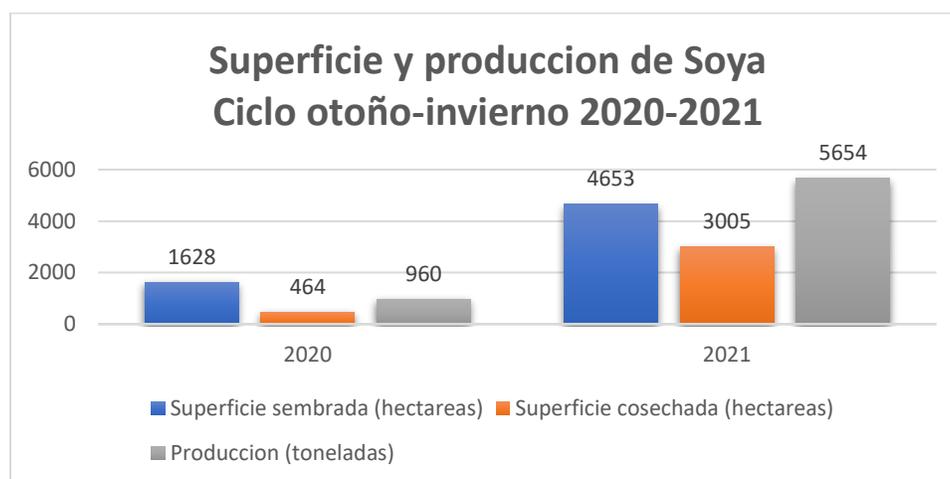


Figura 3. Superficie y producción de soya en México. (SIAP, 2021).

1.1.4 Fases de desarrollo de la soya

El desarrollo de la planta se divide en vegetativo (V) y reproductivo (R). Con excepción de las primeras dos etapas, las etapas V son designadas numéricamente. Los 8 estados R son designados numéricamente (Wright y Lenssen, 2013).

1.1.4.1 Etapa vegetativa

VE: se produce la emergencia de la plántula, los cotiledones están sobre la superficie del suelo.

VC: el hipocótilo se endereza, los cotiledones se encuentran desplegados y las hojas unifoliadas se encuentran separadas.

V1: primer nudo. Hojas unifoliadas expandidas, primera hoja trifoliada desplegada
V2: segundo nudo. Primer trifolio expandido, segunda hoja trifoliada desplegada.
V3: tercer nudo. Segundo trifolio expandido, tercera hoja trifoliada abierta desenrollada.
V4: enésimo nudo. Enésimo nudo con trifolio abierto sobre tallo principal, hojas trifoliadas desplegadas, antes de la floración (Figura 4).

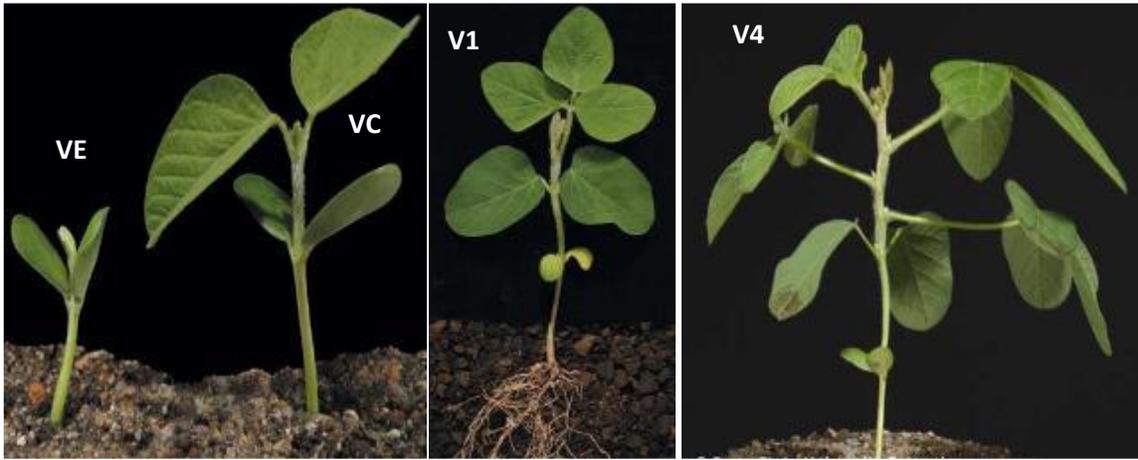


Figura 4. Estados vegetativos de soja. Fotografías de 4 etapas vegetativas VE, VC, V1 y V4.

Imagen tomada de Iowa State

University(https://crops.extension.iastate.edu/soybean/production_growthstages.html).

1.1.4.2 Etapa reproductiva

Las etapas de crecimiento pueden variar debido a las diferencias que pueden existir en el campo, causados por la temperatura y el estrés biótico.

R1: comienzo de floración. La planta presenta una flor en cualquier nudo.

R2: floración completa. Floración completa en uno de los nodos superiores del tallo de la planta con hojas completamente desenrolladas.

R3: formación de vainas. Formación de la vaina en uno de los 4 nodos superiores del tallo, mide 5 mm de largo con hojas totalmente desplegadas.

R4: vainas completas. La vaina mide 2 cm en uno de los 4 nudos superiores del tallo con hojas totalmente desplegadas.

R5: formación de semillas. Las semillas miden 3 mm de largo se encuentra dentro de la vaina de uno de los 4 nudos superiores.

R6: semilla completa. Una de las vainas de uno de los 4 nudos superiores contiene una semilla que llena la cavidad de la vaina.

R7: comienzo de maduración. Una vaina en el tallo principal ha alcanzado el color de madurez (bronceado, marrón o negro).

R8: maduración completa. El 95% de la planta ha alcanzado su madurez (Figura 5).

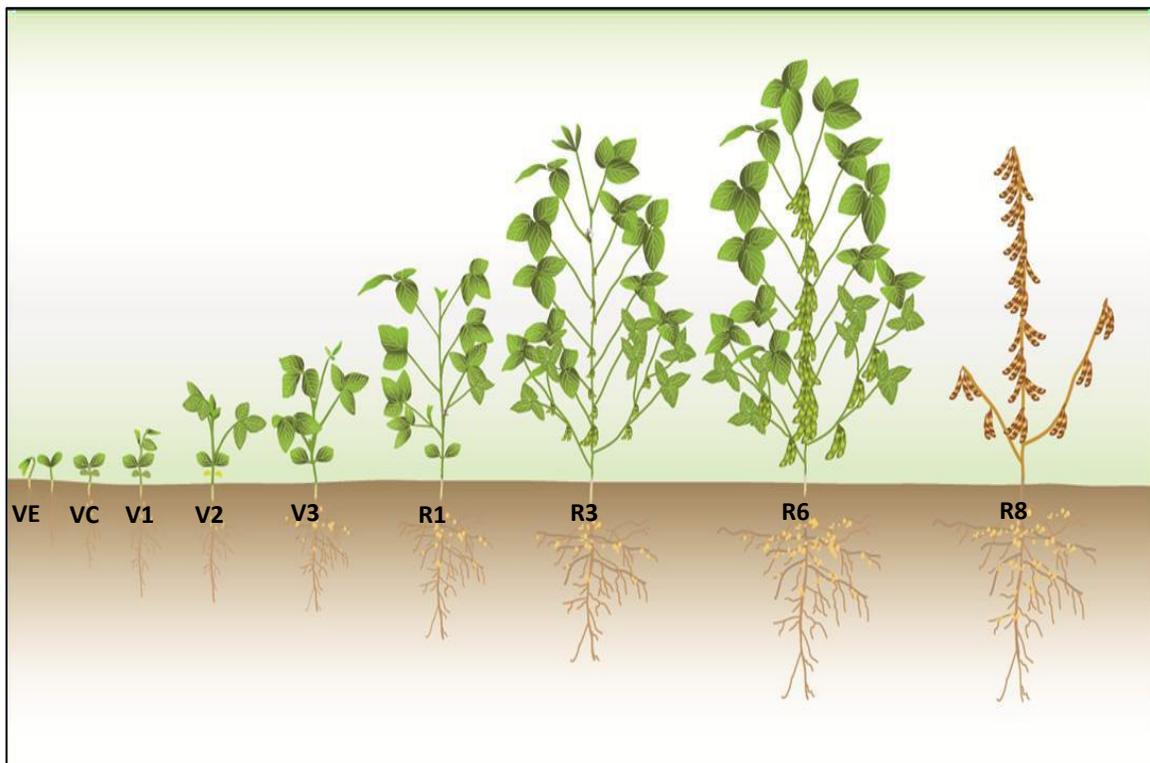


Figura 5. Etapa reproductiva de la soja. Etapas de desarrollo en soja. Imagen tomada de Iowa State University (https://crops.extension.iastate.edu/soybean/production_growthstages.html).

1.1.5 Taxonomía

Se han reportado más de 280 especies, subespecies y variedades taxonómicas del género *Glycine* (Rosas y Young, 1989). El género *Glycine* se subdivide en 3 subgéneros, como se muestra en la Tabla 2:

Tabla 2. El género *Glycine*

Nombre	Distribución
Subgénero <u>Glycine</u>	
<i>G. clandestina</i>	Australia, Pacífico Sur
<i>G. falcata</i>	Australia
<i>G. latrobeana</i>	Australia
<i>G. canescens</i>	Australia, S.E. Asia
<i>G. tometella</i>	Australia, S.E. Asia
Subgénero <u>Bracteata</u>	
<i>G. wightii</i> (soya perenne)	África, S.E. Asia
Subgénero <u>Soja</u>	
<i>G. soja</i> (soya silvestre)	Este de Asia
<i>G. max</i> (frijol soya)	Cultivada

1.2 *Phakopsora pachyrhizi*

P. pachyrhizi es el agente causal de la roya asiática; tuvo su origen en Asia-Australia, posteriormente colonizó el norte y sur de América. *P. pachyrhizi* provoca una de las principales enfermedades que más pérdidas económicas genera en la agricultura (Yorinori et al., 2005).

Este organismo biotrófico es considerado un parásito obligado ya que es incapaz de completar su ciclo de vida alimentándose de materia orgánica muerta. La presencia de este hongo se manifiesta a través de pústulas sobre distintas partes aéreas de la planta como hojas y espigas. Estas pústulas contienen esporas redondas y unicelulares denominadas uredinosporas, que son dispersadas de manera rápida y extensa, expandiendo así la enfermedad.

La aparición de soros oscuros corresponde a la presencia de esporas bicelulares y su forma invernante, que se encuentran envueltas en una membrana espesa;

estas telidsporas germinan en primavera, infectan a la planta y mediante las pústulas generadas salen esporas que pueden infectar a los cultivos, continuando de esta forma su ciclo evolutivo (Bromfield, 1984).

El ciclo de vida de *P. pachyrhizi* es típico en las royas de los hongos. Las uredosporas son esporas que son transportadas fácilmente por el aire diseminándose en los cultivos. Una vez que inicia la germinación, la uredospora germina a través de un tubo (GT) que crece por la superficie de la hoja hasta formar un apresorio (AP) en la superficie. La penetración ocurre de 7 a 12 h después de que la espora se encuentre en la superficie adaxial de la hoja. La penetración es directa a través de aberturas naturales o heridas en la hoja; 20 h después, la hifa de penetración (PH) pasa a través de la cutícula emergiendo en el espacio intercelular donde se produce la hifa de la infección primaria (IH) (Koch y Hoppe, 1987). Entre 24 y 48 h después de posicionarse la espora, la hifa de infección crece para llegar a las células del mesófilo donde forma el haustorio (H) (Figura 6).

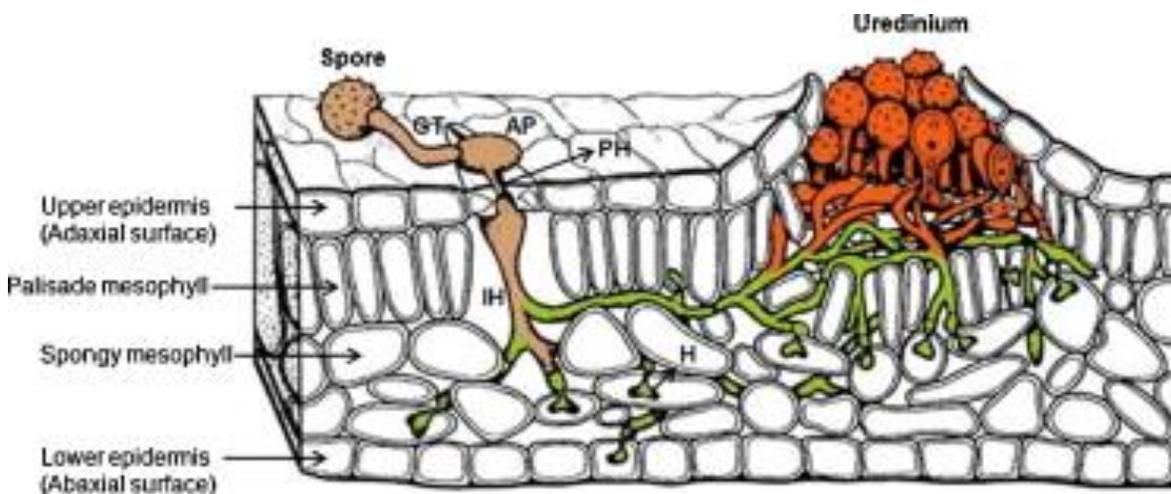


Figura 6. Infección por el hongo de la roya asiática en soya. Estructura interna de una hoja dicotiledónea típica de soya que muestra las diferentes capas celulares y la infección por el hongo de la roya. GT, tubo germinativo; AP, apresorio; PH, penetración hifas; IH, hifas de infección; H, haustorio. Tomado de (Tremblay et al., 2010).

1.2.1 Clasificación taxonómica

Dominio: Eukaryota
Reino: Fungi
Phylum: Basidiomycota
Subphylum: Pucciniomycotina
Clase: Pucciniomycetes
Orden: Pucciniales
Familia: Phakopsoraceae
Género: *Phakopsora*
Especie: *Phakopsora pachyrhizi*

1.3 CRISPR/Cas: edición de genomas

CRISPR-Cas es una herramienta molecular de gran importancia en la actualidad ya que ofrece la capacidad de modificar el genoma, por lo que tiene un gran potencial en medicina, medio ambiente y agricultura. Interesantemente, la historia de CRISPR (repeticiones palindrómicas agrupadas regularmente inter espaciadas) se remonta a 1987, cuando Ishino y colaboradores reportaron un conjunto de secuencias homólogas de 29 nucleótidos espaciados por 32 nucleótidos en *Escherichia coli*; sin embargo, no supieron explicar el significado biológico de estas secuencias. En estudios posteriores se confirmó la presencia de estas secuencias en cepas de *Shigella dysenteria* y *Salmonella enterica* (Nakata et al., 1989), *Mycobacterium tuberculosis* y en las arqueas *Haloferax volcanii* y *Haloferax mediterranei* (Ishino et al., 2018).

CRISPR/Cas (del inglés CRISPR- associated proteins), acrónimo acuñado por el microbiólogo español Mojica (Mojica y Almendros, 2017), fueron identificadas en bacterias y arqueas, como un tipo de inmunidad adaptativa de células procariontes que se caracteriza por la degradación de secuencias de DNA exógenas provenientes de fagos o plásmidos. Este sistema inmunológico mediado por CRISPR/Cas previene la infección por fagos en procariontes mediante la memoria en forma de DNA viral del huésped bacteriano (Makarova et al., 2011). Este DNA está rodeado de secuencias repetidas de nucleótidos, alrededor de éstas se

encuentran secuencias que codifican para las proteínas Cas. Cas es una nucleasa que reconoce como blanco al DNA exógeno desactivándolo mediante cortes específicos. Con estas características, este sistema fue utilizado y modificado de forma artificial para editar el genoma en distintas especies mediante la reprogramación de endonucleasas guiadas hacia su gen objetivo. Por lo tanto, la técnica de edición del genoma modifica el DNA/RNA interno de manera específica y reprogramable (Mojica y Almendros, 2017). La historia de CRISPR/Cas se resume en la Figura 7.



Figura 7. Cronología del sistema CRISPR/Cas. Eventos importantes de 1987 hasta 2013 en el desarrollo de edición de genomas.

Esta herramienta molecular se ha aplicado en diferentes campos, como la industria alimentaria, la agricultura y en el desarrollo de agentes terapéuticos. La investigación en CRISPR continúa avanzando en laboratorios de todo el mundo, y la tecnología resultante se ha traducido en artículos científicos y diversas aplicaciones (Nidhi et al., 2021).

1.3.1 Sistema CRISPR/Cas

Existen una gran variedad de sistemas CRISPR/Cas en diversas especies de arqueas y bacterias con diferentes componentes y mecanismos de acción. El grupo dirigido por Rodolphe Barrangou y colaboradores descubrieron que, después de una infección viral, se insertan nuevas secuencias espaciadoras del bacteriófago invasor en el genoma del huésped (Nidhi et al., 2021). Esto llevó a deducir la función de tales inserciones. Estos pasos funcionales del sistema CRISPR son la adaptación, expresión e interferencia (Gleditzsch et al., 2019; Makarova et al., 2020; McGinn y Marraffini, 2019).

La adaptación es un proceso en el que un elemento portador de material genético se reconoce como agente invasivo, del que se identifica una secuencia protoespaciadora (o para mayor simplicidad espaciador) que se corta e integra en el locus CRISPR en el extremo 5'. La integración de un nuevo espaciador está mediada por un complejo proteico a la secuencia líder de la matriz de CRISPR (McGinn y Marraffini, 2019). Esta secuencia se transcribe generando un precursor de RNA (pre-crRNA) que se procesa a un CRISPR RNA (crRNA) maduro. El procesamiento consiste en el corte del pre-crRNA en la secuencia repetida directa para el crRNA. Por último, la interferencia utiliza la formación de un complejo proteico o una sola proteína efectora con múltiples dominios perteneciente a la clase 2 dentro de las diferentes clases de proteínas Cas que utiliza el sistema CRISPR (Hochstrasser y Doudna, 2015). El paso de interferencia en algunos sistemas requiere de la presencia de secuencias 5'NGG-3 o 5'NAG-3 llamadas motivo adyacente al protoespaciador (PAM) en el genoma invasor. Esta secuencia PAM es una secuencia corta de DNA que no se encuentra en el genoma bacteriano. La secuencia PAM sirve para que algunas proteínas Cas puedan unirse y generar el corte (Figura 8) (Nidhi et al., 2021). El propósito es reconocer secuencias iguales o similares en el genoma del virus o plásmido invasor. Posterior al reconocimiento, el genoma invasor es escindido e inactivado por el complejo proteico; estos acontecimientos permiten al microorganismo identificar

elementos extraños portadores de material genético y, de esta forma, neutralizarlos.

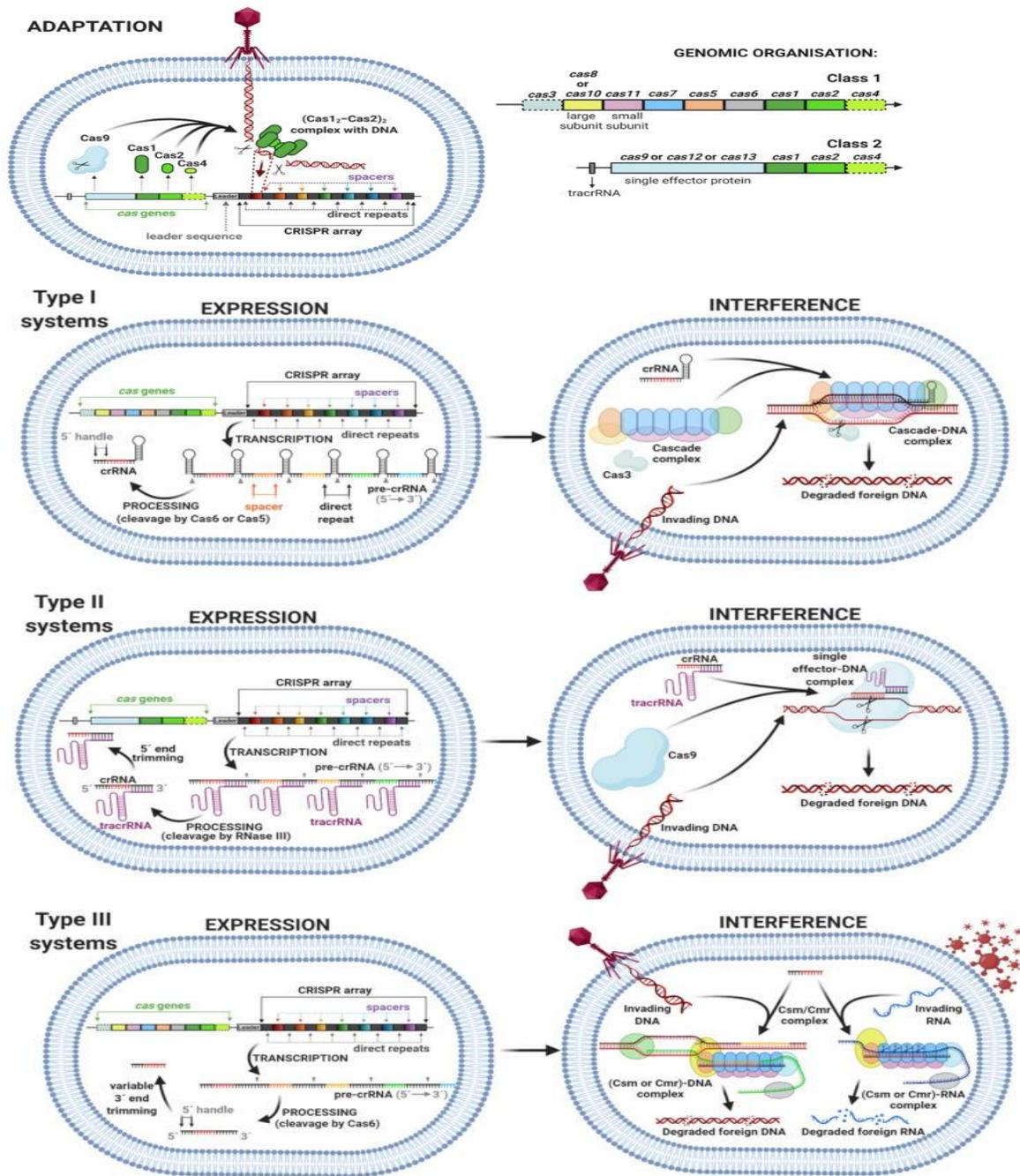


Figura 8. Fases del sistema CRISPR/Cas. Lado izquierdo: paso de adaptación CRISPR. Lado derecho: organización genómica generalizada en los sistemas de clase 1 y clase 2. Un contorno discontinuo indica que el gen es prescindible o falta en algunos subtipos o variantes. Paso de expresión e interferencia CRISPR en los sistemas de tipo I arriba), tipo II (centro) y tipo III (abajo). Tomado de (Nidhi et al., 2021).

1.3.2 Mecanismo de CRISPR/Cas9

La gran versatilidad de CRISPR/Cas9 en el campo de la edición de genomas se debe a su eficiencia, simplicidad y robustez. La herramienta CRISPR/Cas9 consta principalmente de dos componentes. El primero, una proteína endonucleasa Cas9 bacteriana que contiene el dominio RuvC encargado de cortar la cadena de DNA no complementaria a la secuencia espaciadora y el dominio HNH que escinde la hebra complementaria generando un corte bicatenario (DBS) en el DNA. El segundo, un RNA guía (gRNA), una secuencia nucleotídica diseñada de manera específica que va de entre 17 a 20 pb homólogo al DNA blanco (Jiang y Doudna, 2017). La presencia de secuencias PAM es un requisito para el corte del DNA blanco. Por lo tanto, el gRNA lleva al complejo CRISPR/Cas9 a su ubicación genómica prevista. Esto desencadena las dos vías de reparación del DNA endógeno: la unión de extremos no homólogos (NHEJ) o reparación dirigida por homología (HDR) (Nidhi et al., 2021). NHEJ implica la inserción y eliminación aleatoria de pares de bases (indels) en el sitio de corte, lo que da como resultado mutaciones en el marco de lectura, creando en muchas ocasiones un codón de paro o un polipéptido no funcional. La otra vía es la HDR, que implica el uso de regiones homologas de la cadena de DNA sin editar, siendo una secuencia donante o plantilla para corregir el DNA dañado, lo que resulta en una reparación sin errores (Thurtle-Schmidt y Lo, 2018) (Figura 9).

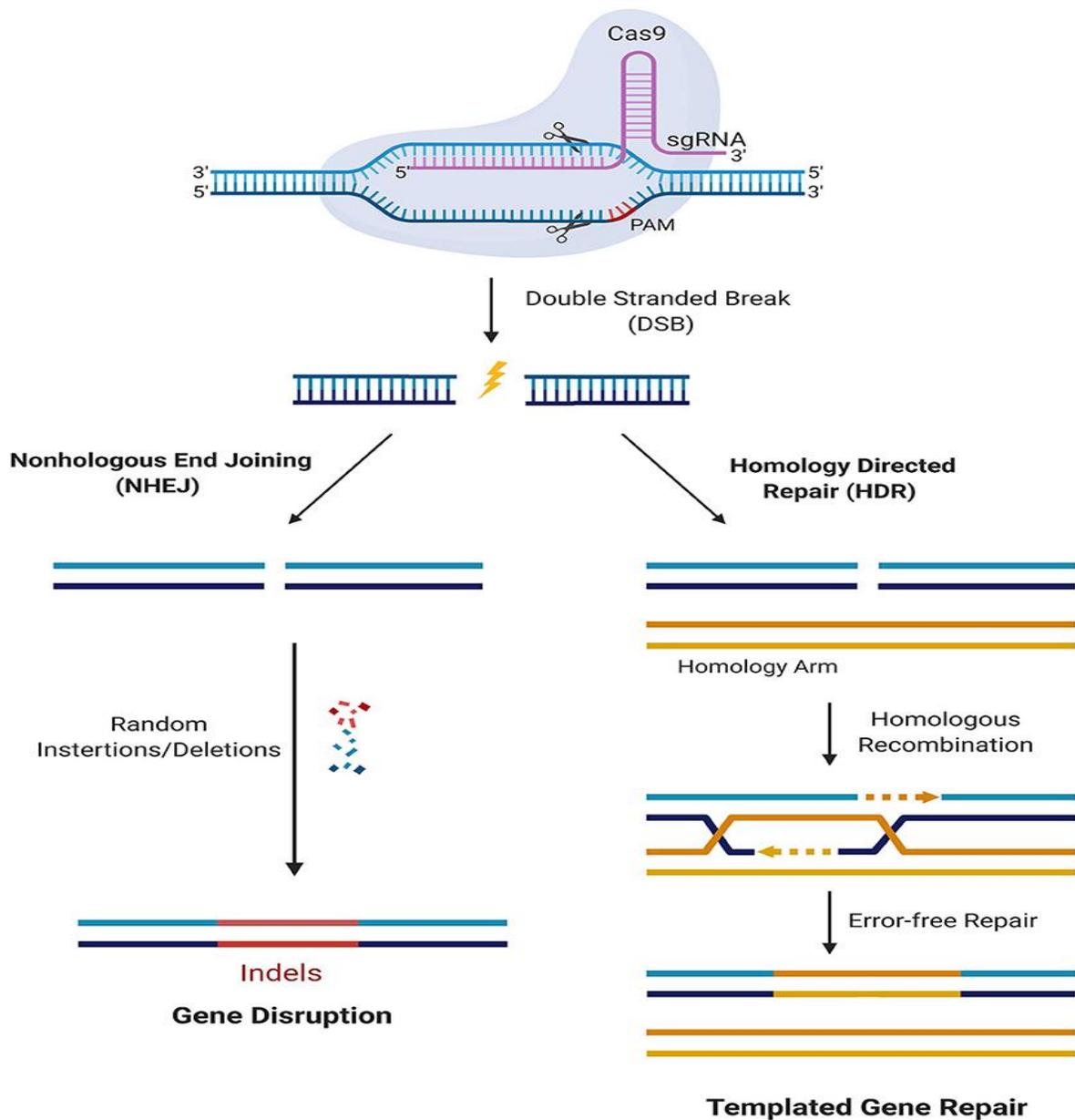


Figura 9. Edición de genes mediada por CRISPR / Cas9. Cas9 en complejo con el gRNA se dirige al gen respectivo y crea DSB cerca de la región PAM. La reparación del daño del DNA procede a través de la vía NHEJ o HDR. En la vía NHEJ, se introducen inserciones y deleciones (indeles) aleatorias en el lado del corte y se ligan, lo que da como resultado una reparación propensa a errores. En la vía HDR, el DNA cromosómico homólogo sirve como molde para el DNA dañado durante la reparación, lo que resulta en una reparación sin errores. Tomado de (Thurtle-Schmidt y Lo, 2018).

1.3.3 CRISPR/Cas9 en plantas

El uso del sistema CRISPR/Cas9 en plantas se ha aplicado con éxito. Entre los modelos más estudiados está *A. thaliana*, pero también en cultivos como soya, tabaco, trigo, maíz, tomate, papa, manzana, sorgo y plátano, entre otros (Liu et al., 2017). El material vegetal más empleado son protoplastos, embriones, hojas y flores. El propósito de las aplicaciones incluye la mejora de la resistencia al estrés abiótico o biótico, el aumento en la producción, así como la ingeniería en vías metabólicas.

El uso de vectores binarios que expresen Cas9 y gRNA dentro de una célula blanco permite la modificación del genoma en las plantas. El uso de promotores de RNA polimerasa III específicos para plantas como AtU6 (*Arabidopsis*), TaU6 (trigo), OsU6 u OsU3 (arroz) son utilizados para la expresión de Cas9 y el gRNA en sistemas vegetales (Ma X. et al., 2016). Una vez efectuada la transformación, las mutaciones introducidas se heredan a la siguiente generación, indicando la edición del genoma de la planta y permitiendo su uso para la investigación y la producción de plantas para un propósito específico (Belhaj et al., 2015).

1.3.4 CRISPR para mejora de cultivos

La edición de genes mediada por CRISPR/Cas9, para la obtención de mejores características relacionadas con respuesta a estrés abiótico, biótico y rendimiento, ha sido introducida en casi 20 especies de cultivos (Jaganathan et al., 2018). Uno de los problemas más grandes de los diferentes cultivos es el estrés biótico causado por microorganismos patógenos, lo que hace necesario el desarrollo de cultivos resistentes a enfermedades. Esto representa más del 42% de pérdida del rendimiento y un 15% en la disminución mundial de alimentos (Oerke, 2006). Por lo tanto, la edición de genomas se utiliza para aumentar la resistencia a las

diversas enfermedades que afectan cultivos, así como una respuesta más eficiente al estrés abiótico.

Uno de los cultivos más importantes por su alto contenido proteico y oleaginoso es la soya (Bourges, 1978). Este cultivo ha sido editado mediante el uso del sistema CRISPR/Cas9. Cai y colaboradores (2015) utilizaron 6 gRNA dirigidos contra dos genes endógenos (GmFEI2 y GmSHR), lo que resultó en supresiones e inserciones en dos sitios genómicos del cromosoma 4 de la soya (Li Z. et al., 2015). El promotor GmU6-16-1 resultó ser más eficiente en la edición simultánea en relación con el promotor de *A. thaliana* AtU6-26 (Hu et al., 2019). Asimismo mediante CRISPR se pudo editar el gen de avirulencia (AvR4/6) de *Phytophthora sojae* resultando en pérdida de virulencia por parte del patógeno (Fang y Tyler, 2016).

1.4 Proteína cinasa de serina/treonina PBS1

Las plantas están en constante exposición a un ambiente con una variedad de microorganismos, por lo tanto, han desarrollado mecanismos sofisticados de defensa. Uno de los sistemas de defensa contra patógenos microbianos más estudiados es el PBS1-RPS5. El inicio de la inmunidad innata comienza con el reconocimiento de microorganismos patógenos mediante receptores transmembranales e intracelulares (Chisholm et al., 2006). Los primeros receptores son los PRR que reconocen a los PAMP provenientes de los microorganismos, lo que activa la inmunidad activada por PAMP; sin embargo, muchos microorganismos exitosos son capaces de secretar una diversidad de proteínas efectoras que anulan el efecto de la inmunidad innata. Para contrarrestar este ataque proteico, las plantas han adquirido otro mecanismo de defensa utilizando proteínas de resistencia (R) para detectar efectores patógenos y así

activar la inmunidad por efectores (ETI) (Cui et al., 2015; Jones J. D. y Dangl J. L., 2006; Kim S. H. et al., 2016).

Hace algunos años, no era posible realizar modificaciones de forma específica en proteínas receptoras de plantas. El descubrimiento de la ingeniería de señuelos realizada por Kim S.H. et al. (2016) puede ampliar el reconocimiento de una diversidad de proteínas efectoras de manera específica hacia su receptor. El objetivo principal es la modificación de una proteína receptora (PBS1) para ser reconocida de manera específica por diversas proteasas derivadas de patógenos (Kim S. H. et al., 2016).

RPS5 es una proteína NLR que se activa tras el corte de la proteína PBS1 a través de la proteasa AvrPphB que hidroliza los enlaces alfa-peptídicos. Dicho corte expone un bucle de 5 aminoácidos que se cree activa a RPS5, lo que a su vez activa la respuesta inmune caracterizada por la HR (Zhang et al., 2010). Interesantemente, este tipo de inmunidad requiere de ambos fragmentos de PBS1, el motivo de secuencia GDKSH que es requerido para la hidrolisis, así como el motif SEMP (A. *thaliana*) de unión a la proteína RPS5. Ésta se activa después del corte y el cambio conformacional de PBS1. Por lo tanto, la percepción de AvrPphB sigue un mecanismo trampa donde el corte de PBS1 (cebo) activa la trampa y de esta forma activa a RPS5, activando la inmunidad (Kim S. H. et al., 2016). Debido a esto, se ha inferido que la respuesta inmune mediada por RPS5 podría modificarse mediante el cambio del motivo de secuencia de corte de AvrPphB por otro que pueda ser reconocido y cortado por proteasas secretadas por diversos patógenos (Kourelis et al., 2016) (Figura 10).

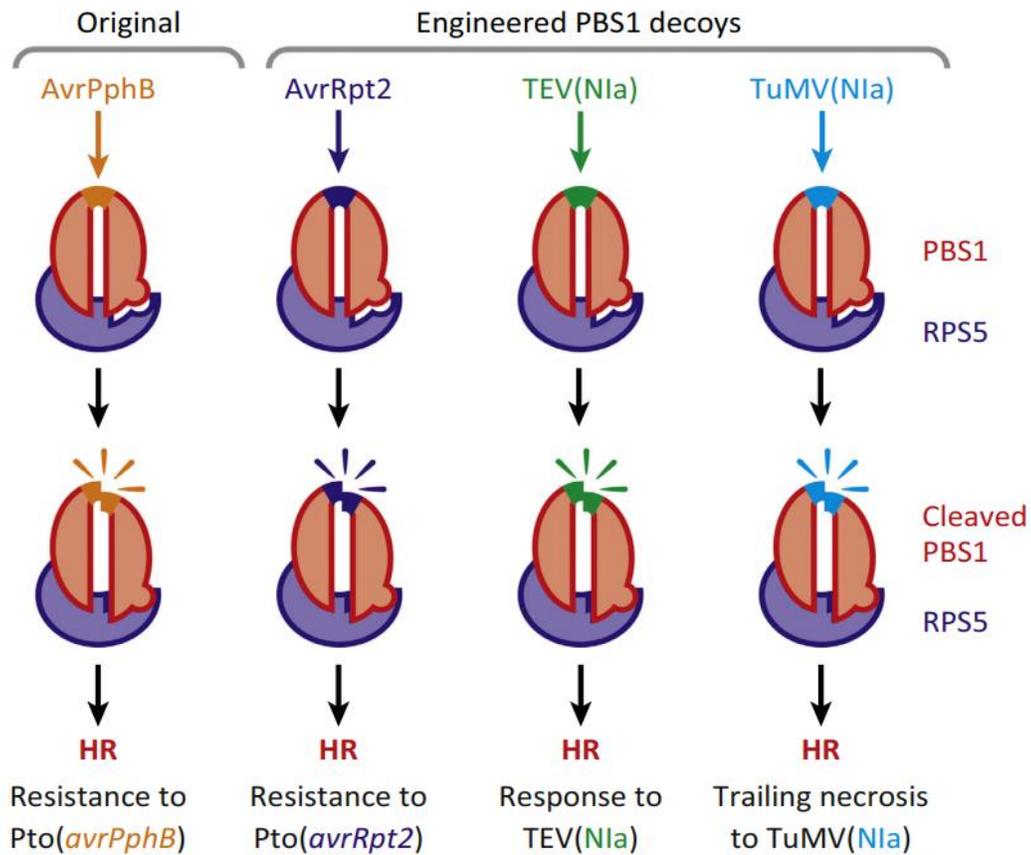


Figura 10. Ingeniería de señuelo PBS1. Mediante la introducción de sitios de corte para proteasas no relacionadas da como resultado nuevas especificidades de reconocimiento del receptor inmune RPS5. La proteína PBS1 nativo (izquierda) fue diseñado reemplazando el sitio de corte AvrPphB (naranja) por sitios de corte dirigidos por proteasas derivadas de patógenos no relacionados: AvrRpt2 (azul) de *Pseudomonas syringae* (Pto); Nla proteasa (verde) de Tobacco Etch Virus (TEV) y la proteasa Nla (cian) del Virus del Mosaico del Nabo (TuMV). Tomado de (Kourelis et al., 2016).

De manera interesante, se ha comenzado a estudiar el motivo de secuencia de reconocimiento a RPS5; el sistema más estudiado es en *A. thaliana* en el que el motivo de reconocimiento es SEMPH y se ha demostrado la activación del sistema inmune mediante este mecanismo (Qi et al., 2014). Sin embargo, estudios en un homólogo PBS1 de trigo (TaPBS1) identificaron el motivo STRPH y se observó que la asociación de éste con RPS5 fue más débil, por lo que no se pudo desencadenar una respuesta hipersensible. La diferencia entre los motivos SEMPH y STRPH radica en los aminoácidos EM considerando que el ácido

glutámico (E) tiene carga negativa necesaria para el reconocimiento de PBS1 y RPS5 (Sun et al., 2017). Por tanto, el análisis de proteínas PBS1 homólogas de diferentes especies de plantas a nivel de los motivos de corte y reconocimiento es importante para comprender el mecanismo de la inmunidad en plantas.

2. Justificación

El cultivo de soya es impactado por enfermedades producidas por diferentes patógenos afectando su productividad como la roya asiática, enfermedad causada por *P. pachyrhizi*, la cual es controlada con el uso de fungicidas que están perdiendo eficacia. Por lo tanto, es importante la generación de nuevas variedades con resistencia a este hongo y otros patógenos.

3. Hipótesis

La modificación del gen PBS1 de soya (*G. max*) mediante el sistema CRISPR/Cas9 generará plantas resistentes a la roya asiática causada por *P. pachyrhizi*.

4.Objetivo general

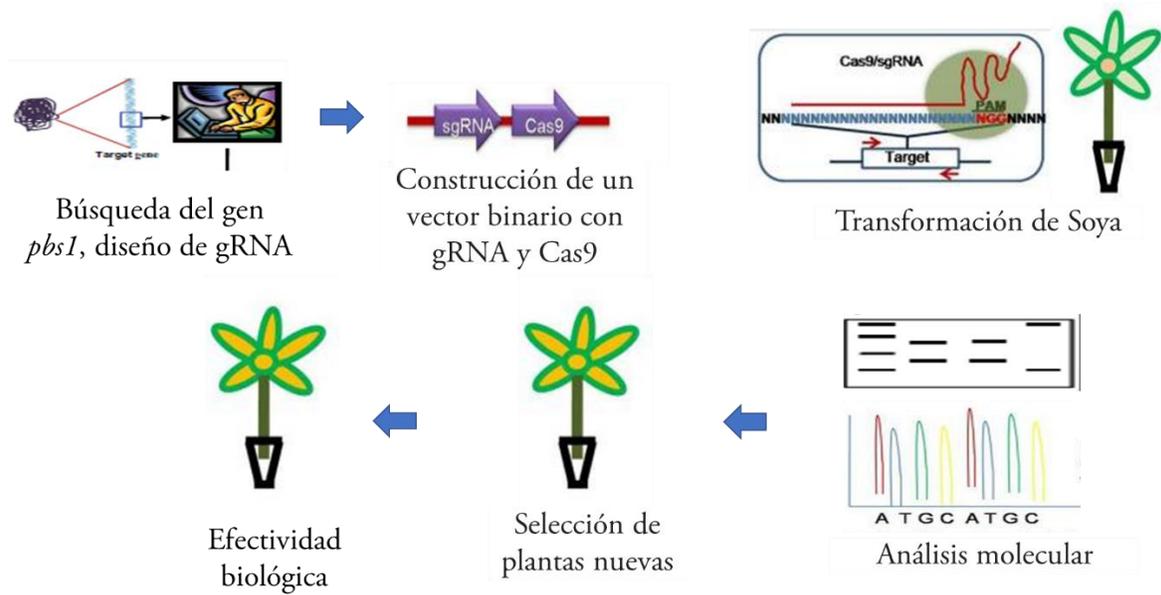
Editar el gen *PBS1* para generar potencial resistencia a la roya asiática de la soya.

4.1.1Objetivos Particulares

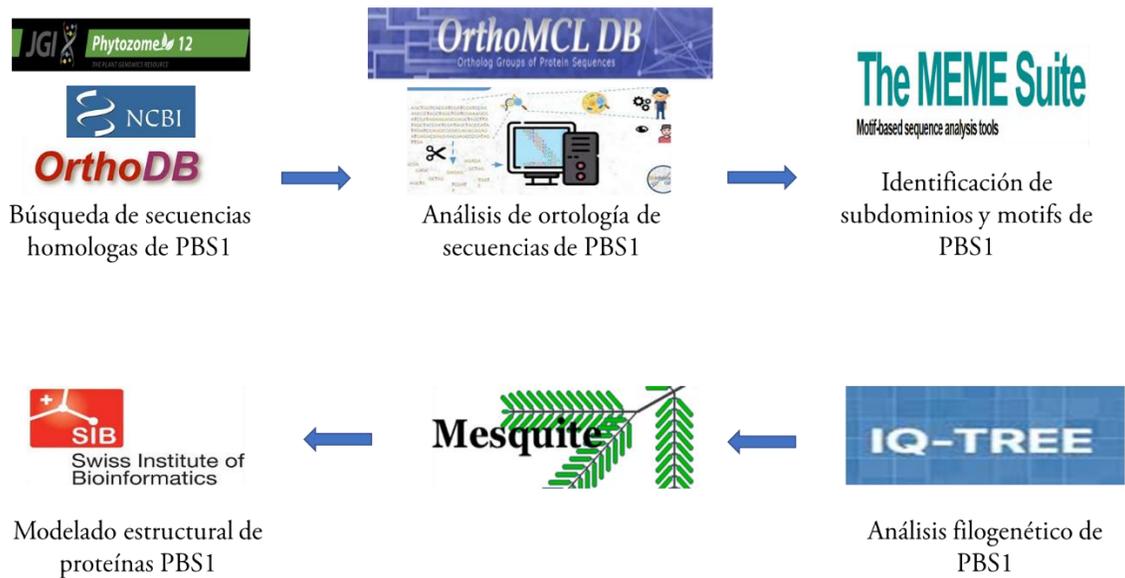
- Analizar evolutivamente proteínas ortólogas *PBS1* de *A. thaliana*.
- Identificar el gen ortólogo *PBS1* de *A. thaliana* en soya.
- Transformar genéticamente soya con la construcción que contiene *PBS1* mediante vectores de edición.
- Analizar la progenie de la planta para identificar las variantes.
- Evaluar la resistencia mediante ensayos de efectividad biológica con esporas de *P. pachyrhizi*.

5. Materiales y Métodos

5.1 Metodología general



5.1.2 Metodología del análisis bioinformático



5.2 Material biológico

En la presente investigación se utilizaron plantas de soya variedades Nainari y Suaqui, rápido crecimiento y alta proteína, como fuentes para la extracción de productos biológicos. Además, se utilizaron *Agrobacterium tumefaciens* para el proceso de transformación y *E. coli* MACH1TM-T1 (Invitrogen, La Jolla CA) como fuente de extracción de DNA plasmídico.

5.3 Siembra de soya

A las semillas de soya seleccionadas (Nainari, Suaqui, rápido crecimiento ó alta proteína) se les realizó lavados con cloro: agua en proporción 1:4 durante 20 min en agitación; posteriormente, se lavaron con agua destilada por 5 min y se dejaron secar. Las semillas se sembraron en pellets de turba comprimida los cuales se mantuvieron en cámara de crecimiento (Conviron, A1000, Winnipeg, Canadá) con humedad relativa al 70% a 24°C. Después de 1 semana, las semillas germinadas fueron sembradas en macetas con una mezcla de suelo:turba:agrolita en proporción 2:2:1 con humedad de 20-40% a temperatura de 22-32°C con riego inter diario y semanal con solución nutritiva Miracle Grow.

5.4 Extracción de DNA

Se tomaron 100 mg de hoja de cada una de las variedades de soya, congelados en nitrógeno líquido y pulverizados con pistilo. Posteriormente se incubaron con 600 µl de CTAB (Tris-HCl 100 mM pH 8.0, EDTA 20 mM pH 8.0, NaCl 1.4M, CTAB 2%, PVP 4000 1%) precalentado por 20 min a 65°C. Una vez transcurrido el tiempo se agregó 600 µl de fenol:cloroformo:isoamílico (50:48:2), se mezcló con

vórtex por 1 min y se centrifugó; se colectó el sobrenadante y se mezcló con 600 µl de cloroformo:isoamílico (24:1). Se centrifugó y se añadieron 0.7 volúmenes de isopropanol frío al sobrenadante, se mezcló por inversión y se incubó a -20°C por 20 min. La pastilla obtenida fue lavada dos veces con 200 µl de etanol al 70% y se dejó secar. La pastilla se resuspendió en agua Mili-Q estéril con RNasa a 20µg/ml. La concentración de DNA fue determinada con un Nanodrop (Thermo Scientific, Waltham MA) y su integridad se verificó en gel de agarosa al 0.8% teñido con bromuro de etidio (0.1 µg/ml).

5.5 Análisis *in silico* para identificar el gen de *PBS1* en soya

Se realizó una búsqueda en la base de datos de UNIPROT (<https://www.uniprot.org/>) y MEROPS (<https://www.ebi.ac.uk/merops/>) del gen *pbs1* de *A. thaliana*. Una vez que se obtuvo la secuencia y se identificó el sitio de corte en secuencia aminoacídica, se realizó la búsqueda del gen ortólogo de *PBS1* en soya en la base de datos Phytozome v12.1 (<https://phytozome.jgi.doe.gov/>) mediante alineamiento.

5.6 Secuencia del gen *PBS1* de soya

Para comenzar con los estudios de gRNA en las 4 variedades de soya se realizó secuenciación del gen *pbs1* así como la comparación con el gen de referencia para soya (*G. max Wm82.a2.v1*). La amplificación del gen *PBS1* se realizó considerando la secuencia nucleotídica que codifica para el sitio de corte de *PBS1*. Se diseñaron oligonucleótidos para amplificar mediante PCR punto final la secuencia que contiene el sitio potencial de edición en todas las variedades. Los oligonucleótidos que se emplearon fueron los siguientes:

For-Ex5-PBS1 G max 5'-GGCAAATCCTCCTGTCATCTAC-3'

Rev-Ex5-PBS1 G max 5'-CGGCCAGTAATCAGCTCTAAG-3'.

La mezcla de reacción para la amplificación de los productos de PCR se indica en la Tabla 3 y el programa de amplificación en la Tabla 4.

Tabla 3. Mezcla de reacción PCR punto final

Reactivo	Volumen (µl)
<i>Buffer 10x</i>	1.25
<i>dNTPs (10µM)</i>	1.00
<i>Primer F (10µM)</i>	0.25
<i>Primer R (10µM)</i>	0.25
<i>Takara Ex Taq</i>	0.05
<i>H2O esteril</i>	8.70
<i>DNA (100 ng/µl)</i>	1.00
<i>Volumen final</i>	12.50

Tabla 4. Programa de amplificación del gen *PBS1*

<i>Etapa</i>	Temperatura °C	Tiempo
<i>Desnaturalización inicial</i>	94.0	3:00 min
<i>Desnaturalización</i>	94.0	0:45 s
<i>Alineamiento</i>	58.0	0:30s
<i>Elongación</i>	72.0	0:45s
<i>Elongación final</i>	72.0	3:00 min

Nota: se amplificó por 30 ciclos a partir del paso de desnaturalización hasta elongación

Una vez que se obtienen los productos de PCR se verifica su integridad en gel de agarosa de 1.5%.

5.7 Purificación del producto de PCR correspondiente a *PBS1* en soya

El producto de PCR del gen *PBS1* de soya fue purificado utilizando el kit comercial QIAquick PCR Kit (Qiagen). A partir de 50 µl de reacción de PCR se agregaron 250µl del buffer PB, una vez que se obtuvo la mezcla se colocó en una columna QIAquick y se centrifugó por 1 min. Se realizó un lavado agregando 750µl de buffer PE a la columna, se centrifugó por 1 min y se desechó el sobrenadante. Para eluir el producto de PCR, la columna se colocó en un tubo limpio de 1.5 ml y se agregó 20µl de buffer EB al centro de la membrana de la columna QIAquick y se centrifugó por 1min. Se cuantificó la concentración del producto de PCR mediante un Nanodrop (Thermo Scientific) se verificó la purificación en gel de agarosa al 1.5%.

5.8 Análisis de secuencia y alineación

Para determinar la homología de las 4 variedades de soya con el gen de referencia *Glycine max Wm82.a2.v1*, se utilizó la secuenciación por el método Sanger en el servicio de secuenciación del Laboratorio de Bioquímica Molecular, UBIPRO, FES-Iztacala UNAM en el cual las muestras fueron procesadas. Las secuencias obtenidas fueron analizadas a través de un alineamiento en ClustalW 2.1 con el gen de referencia.

5.9 Extracción de DNA plasmídico

El vector PBUN4116 fue utilizado para la edición de genoma mediante el uso de la proteína Cas9 y sugRNA. A partir de una cepa de *E.coli* se realizó la extracción de DNA plasmídico (PBUN4116). La bacteria se cultivó en medio LB con espectinomicina (100µg/ml) y se colocó en agitación orbital a 200 rpm a 37°C por 16h. Transcurrido el tiempo a partir de 1.5ml del cultivo bacteriano se centrifugo a 10000 g por 1min obteniéndose una pastilla celular a la cual se le agregó 100 µl de solución 1 (Tris HCL 25 mM pH 8.0, EDTA a 10mM pH 8.0) homogenizando la pastilla. Posteriormente se agregaron 200µl de solución II (NaOH 0.2N, SDS 1%), se mezcló por inversión y se incubó en hielo por 2 min. Después se agregaron 150 µl de la solución III (acetato de potasio 3M, pH 4.5), se mezcló por inversión y se incubó en hielo por 5 min y se centrifugo por 5 min, el sobrenadante se transfirió a un tubo nuevo y se añadieron 2µl de RNAsa (10 mg/ml) se incubó a 37° C por 20 min. Se añadieron 450µl de fenol:cloroformo:isoamílico (25:24:1) y se centrifugó por 5 min. La fase acuosa se transfirió a un tubo nuevo y el DNA se precipitó con 240µl de isopropanol y se mezcló por inversión, se centrifugó por 10 min, se descartó el sobrenadante y el precipitado se lavó con 500µl de etanol al 70% en dos ocasiones, la pastilla se secó a temperatura ambiente y se resuspendió en 30µl de agua MiliQ estéril. Se cuantificó la concentración del DNAP mediante un Nanodrop (Thermo Scientific) y se verificó su integridad en gel de agarosa al 0.8%.

5.10 Linealización del plásmido PBUN4116

Se realizó una digestión para linealizar el plásmido PBUN4116, en una mezcla de reacción conteniendo 1.5 µl de buffer 10xNEB, 1.0 µl de enzima Bsal, 2.5 µl de H₂O MilliQ estéril y 10 µl DNA (1000 ng/µl). La reacción se incubó a 37°C por 5h. La digestión se verificó en gel de agarosa al 0.8%.

5.11 Diseño de gRNA

El diseño de gRNAs se realizó utilizando el software CRISPR-P2.0 (<http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR>) con el objetivo de obtener los gRNAs con mejor registro y menor número de blancos inespecíficos potenciales (off target) del gen *PBS1* de soya considerando la cercanía a la secuencia nucleotídica que codifica para el sitio de corte.

Una vez que se tiene la secuencia de *PBS1* de soya, se realizó el diseño del gRNA en el software, seleccionando: PAM (NGG), SnoRNA promotor (U6), andamio o esqueleto de RNA (Scaffold predeterminado), longitud de secuencia (20nt), genoma blanco (soya) y por último la secuencia nucleotídica del gen *PBS1*. Los gRNAs seleccionados fueron:

F-gRNA-PBS1 5' ATTGGTTTCTACCCGTGTCAT 3'

R-gRNA-PBS1 5' AAACATGACACGGGTAGAAAC 3'

F-gRNA PBS1-20nt 5' ATTGGATGTTTCTACCCGTGTCAT 3'

R-gRNA PBS1-20nt5' AAACATGACACGGGTAGAAACATC 3'.

5.12 Alineamiento de gRNA

Una vez que se obtienen los gRNAs sentido y antisentido se realizó su ensamblaje preparando la siguiente mezcla de alineamiento (Tabla 5).

Tabla 5. Mezcla de alineamiento del gRNA

Componente	Volumen (µL)
<i>Oligonucleótido F (100µM)</i>	1.0
<i>Oligonucleótido R (100µM)</i>	1.0
<i>Buffer de alineamiento 10X*</i>	2.0
<i>ddH₂O estéril</i>	16.0
<i>Volumen final</i>	20.0

*Buffer de alineamiento (10X): Tris-HCl pH8.0 100 mM, NaCl 500 mM y EDTA 10mM. Esterilizar por filtración en membrana de acetato de celulosa y almacenar a -20 °C

Se sigue el siguiente programa de alineamiento en el termociclador (Tabla 6).

Tabla 6. Programa de alineamiento

Temperatura (°C)	Tiempo	ΔT (°C/s)
94	5min	3.0
94-84	10s	2.0
84-25	10s	0.2
16	∞	3.0

Se cuantificó la concentración del alineamiento de gRNAs mediante un Nanodrop (Thermo Scientific). El correcto alineamiento se verificó mediante un gel de agarosa nusieve 3:1 (Lonza Biotech, Suiza) al 5%.

5.13 Purificación del producto linealizado

El producto linealizado (PBUN4116) se purificó utilizando el sistema comercial Zymoclean Gel DNA Recovery (Zymo; Irvine CA) siguiendo las instrucciones del fabricante. Se cortaron las bandas de DNA plasmídico del gel de agarosa, se colocaron en un tubo de 1.5ml y se agregaron 3 volúmenes de la solución ADB, se incubó por 10 min a 55°C y se mezcló por vórtex. La mezcla se transfirió a una columna Zymo-Spin TM en un tubo de recolección y se centrifugó por 1 min a 13000 rpm y se desechó el sobrenadante. A la columna se le agregaron 200 μ l de buffer DNA Wash, se centrifugó por 30 seg a 13000 rpm, se desechó el sobrenadante y se repitió este paso. La columna se pasó a un tubo nuevo y se agregaron 10 μ l de buffer de elución para DNA directamente a la columna y se centrifugó por 1 min a 13000 rpm para eluir.

5.14 Ligación del vector PBUN411U6 con el gRNA-PBS1

Los gRNAs alineados se ligaron al vector linealizado PBUN4116. Se realizó la siguiente mezcla de reacción (Tabla 7):

Tabla 7. Ligación del gRNA

Reactivo	Volumen (µL)
	Ligasa T4
<i>Buffer de reacción de ligasa</i>	4.0
<i>Plásmido linearizado</i>	1.0 (97 ng/ µl)
<i>Inserto</i>	1.0 (124 ng/µl)
<i>ddH₂O estéril</i>	13.5
<i>Ligasa T4</i>	0.5

La mezcla de ligación se homogenizó y se incubó por 1 h a 22 °C, se enfrió en hielo y la mezcla de ligación se utilizó para transformar en células *E. coli* Mach-TI químicamente competentes.

5.15 Diseño de secuencia DONOR

Se realizó una búsqueda de proteasas de *P. pachyrhizi* reportadas en bases de datos UNIPROT (<https://www.uniprot.org/>) y Genbank, usando la herramienta NCBI BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Se realizó un alineamiento y búsqueda de una proteasa caracterizada que tuviera el potencial de procesar proteolíticamente a PBS1; posteriormente, se accedió a la base de datos MEROPS (<https://www.ebi.ac.uk/merops/>) y se identificó el sustrato, el sitio de corte y la proteasa potenciales. Se realizó posteriormente un alineamiento de la secuencia nucleotídica de la proteasa de *P. pachyrhizi* con la secuencia de referencia de Genbank en la que se identificó la triada catalítica.

Una vez que se identificó la secuencia de corte se empleó el software de Addgene (<https://www.addgene.org/crispr/plant/>) para el diseño de la secuencia DONOR (secuencia a insertar). De acuerdo al gRNA diseñado el software predice el sitio donde la secuencia DONOR se va a insertar considerando la homología del largo de los sitios flanqueantes de ambos sentidos 5´y 3´. Posterior a este análisis se diseñaron oligonucleótidos con las secuencias DONOR, que corresponde a una secuencia de 24 nt que se traduce en el sitio de corte; a esta secuencia se le

anexó una secuencia de 26 nt a cada oligonucleótido con homología al gen *PBS1*. Así se obtuvieron:

F-DONOR-FFTEGGHD

5' TTCTTCACTGAAGGTGGTCACGATAGAGTTATGGGAACTTATGGTTACTG 3'

R-DONOR-FFTEGGHD

5' ATCGTGACCACCTTCAGTGAAGAAAACAGGACCAAGCTTCGCAAG 3'

5.16 Obtención de secuencia DONOR en el gen *PBS1* de soya

Para aumentar la longitud de los sitios flanqueantes se diseñaron oligonucleótidos 150 pb río arriba y 150 pb río abajo del gen *PBS1* considerando el sitio de inserción de la secuencia DONOR y se realizó una PCR como se describió anteriormente (Tabla 3). Se utilizó el programa de amplificación descrito en el apartado (Tabla 4). Se diseñaron y utilizaron los siguientes oligonucleótidos:

F-PBS1-arm- 5'CCTTCACGATAAGGCAAATC 3'

R-PBS1-arm- 5' TTCTGTTCTCCCTGCGGCTG 3'

Se verificó la amplificación del producto en un gel de agarosa al 1.5%. (producto 1). Se realizó una PCR punto final utilizando los oligonucleótidos F-PBS1-DONOR y R-PBS1-DONOR como de describe en el apartado 5.17 (Tabla 7) sin muestra de DNA. El producto amplificado se verificó en un gel de agarosa nusieve 3:1 al 5% (producto 2). Ambos productos de PCR se purificaron utilizando el sistema comercial QIAquick PCR Kit (250) descrito en el apartado 5.7. Una vez que se obtuvieron los productos purificados, se cuantificaron y se realizó una PCR punto final utilizando ambos productos. Las reacciones se realizaron de acuerdo con la Tabla 8. Se utilizó el programa de amplificación de la Tabla 4. Los productos de PCR se verificaron en un gel de agarosa al 2.5%.

Tabla 8. Reacciones de PCR para secuencia DONOR

Reacción 1		Reacción 2	
Reactivo	Volumen (µl)	Reactivo	Volumen (µl)
<i>Buffer 10x</i>	1.25	<i>Buffer 10x</i>	1.25
<i>MgCl₂(50nM)</i>	0.375	<i>MgCl₂(50nM)</i>	0.375
<i>dNTPs (10µM)</i>	0.25	<i>dNTPs (10µM)</i>	0.25
<i>F-pbs1-arm (10µM)</i>	0.25	<i>F-pbs1-DONOR (10µM)</i>	0.25
<i>R-pbs1-DONOR (10µM)</i>	0.25	<i>R-pbs1-arm (10µM)</i>	0.25
<i>Taq Platinum</i>	0.05	<i>Taq Platinum</i>	0.05
<i>H2O esteril</i>	8.10	<i>H2O esteril</i>	8.10
<i>*Producto 1 (100 ng/µl)</i>	1.00	<i>*Producto 1 (100 ng/µl)</i>	1.00
<i>*Producto 2 (100 ng/µl)</i>	1.00	<i>*Producto 2 (100 ng/µl)</i>	1.00
<i>Volumen final</i>	12.50	<i>Volumen final</i>	12.50

*Producto 1: amplicón del gen *pbs1* de 300 pb

*Producto 2: amplicón de la secuencia DONOR de 68pb

5.17 Ensamblaje por PCR

Los productos obtenidos en el paso anterior se purificaron utilizando el sistema comercial Zymoclean Gel DNA Recovery descrito anteriormente (5.13). Una vez que se obtiene el producto de la reacción 1 (producto 1.1) y el producto de la reacción 2 (producto 2.1) de interés, se realizó una reacción de alineamiento de ambos productos como se indica en la Tabla 9.

Tabla 9. Reacción de alineamiento 1

Componente	Volumen (µL)
<i>Producto 1.1</i>	2.0
<i>Producto 2.1</i>	2.0
<i>Buffer de Alineamiento 10X</i>	2.0
<i>ddH₂O estéril</i>	14.0
<i>Volumen final</i>	20.0

Una vez que se obtiene la mezcla de reacción se sigue el programa de alineamiento descrito anteriormente. Posteriormente se realizó una reacción para complementar el producto alineado (Tabla 10).

Tabla 10. Reacción de alineamiento 2

Componente	Volumen (µL)
<i>Reacción de alineamiento</i>	5.0
<i>Buffer 10x</i>	2.0
<i>dNTPs (10µM)</i>	1.0
<i>Takara Ex Taq</i>	0.1
<i>ddH₂O estéril</i>	11.9
<i>Volumen final</i>	20.0

Se utilizó el siguiente programa en el termociclador (Tabla 11):

Tabla 11. Programa de alineamiento

Temperatura (°C)	Tiempo
50	2min
60	2min
72	5min
80	2min
72	5min
16	∞

Se realizó PCR punto final del producto obtenido y se verificó en gel de agarosa al 2%. El producto obtenido se utilizó para transformar en células *E. coli* Mach-TI químicamente competentes.

5.18 Transformación de *E.coli* Mach-T1 químicamente competentes

Las células químicamente competentes se descongelaron en hielo, se colocaron 5µl de la ligación al vial de células *E. coli*, se incubó por 30 min en hielo, posteriormente se realiza choque térmico a 42 °C por 50 seg, se incubó en hielo por 2 min y se agregó 250µl de medio SOC a temperatura ambiente. Después se incubó a 37°C por 1 h con agitación a 200 rpm. Una vez transcurrido el tiempo, el cultivo se distribuyó en medio LB solido + espectinomicina (Spc) 100µg/ml. Las cajas se incubaron a 37 °C toda la noche, posteriormente se seleccionaron las colonias transformadas.

5.19 Selección de colonias transformadas

Del crecimiento obtenido en el paso anterior se picaron 10 colonias con palillos estériles y se estriaron en otra caja Petri con medio LB sólido + Spc 100 µg/ml para obtener una caja maestra, la cual fue incubada a 37 °C. Los palillos se colocaron en tubos de ensayo con 3 ml de medio LB liquido (Spc 100 µg/ml) y se incubaron a 37°C durante toda la noche en agitación. La extracción de DNA plasmídico se realizó de acuerdo con el procedimiento descrito en el apartado 5.9.

5.20 Análisis de restricción

Una vez obtenido el DNA plasmídico se realizó el análisis de restricción mediante una digestión utilizando la enzima HindIII (Tabla 12).

Tabla 12. Reacción de digestión

Reactivos	Volumen (µl)
<i>DNAp</i>	4.0
<i>HindIII</i>	0.25
<i>Cutsmart 10x</i>	1.5
<i>H2O MiliQ esteril</i>	9.25

5.21 PCR y secuenciación para corroborar presencia de inserto

Para corroborar la presencia del gRNA de *PBS1* en el vector PBUN4116 se realizó una PCR como se describe en la Tabla 13.

Tabla 13. Reacción de PCR para gRNA

Reactivo	Volumen (µl)
<i>Buffer 10x</i>	1.25
<i>MgCl₂(50nM)</i>	0.375
<i>dNTPs (10µM)</i>	0.25
<i>Primer F (10µM)</i>	0.25
<i>Primer R (10µM)</i>	0.25
<i>Taq Platinum</i>	0.05
<i>H2O esteril</i>	9.10
<i>DNA (100 ng/µl)</i>	1.00
<i>Volumen final</i>	12.50

Se emplearon los siguientes oligonucleótidos:

F-gRNA-PBS1- 5'ATTGGTTTCTACCCGTGTCAT 3'

R-U6t 5'CCGAAGAACAGAGGAAGAAG-3'

Se utilizó el programa de amplificación descrito en la Tabla 14.

Tabla 14. Programa de amplificación de gRNA

<i>Etapas</i>	Temperatura °C	Tiempo
<i>Desnaturalización inicial</i>	94.0	3:00 min
<i>Desnaturalización</i>	94.0	0:35 s
<i>Alineamiento</i>	60.0	0:30s
<i>Elongación</i>	72.0	0:35s
<i>Elongación final</i>	72.0	7:00 min

Se eligieron 3 muestras de DNA plasmídico las cuales fueron purificados utilizando el sistema comercial Minipreps (ZR Plasmid Miniprep) siguiendo las indicaciones del fabricante. Se cuantificó la concentración del DNA plasmídico mediante un Nanodrop (Thermo Scientific) y se verificó su integridad en gel de agarosa al 0.8%. Las muestras fueron procesadas en el Laboratorio de Bioquímica Molecular, UBIPRO, FES-Iztacala UNAM mediante secuenciación por el método Sanger.

5.22 Transformación de *A. tumefaciens* por electroporación

Una vez se obtuvo el plásmido con el inserto de interés, se realizó la transformación de *A. tumefaciens* (AGL1) electrocompetentes. Se descongelaron las células AGL1 en hielo y se les agregó 2 µl de DNA plasmídico. Las celdas de electroporación se enfriaron previamente, ya descongeladas las células se mezclaron por pipeteo con el DNA plasmidico. Se tomó todo el volumen y se depositó en la celda de electroporación de 1 ml y se colocó en el electroporador, se le dio un pulso eléctrico (1800 V); posteriormente se le agregó 300µl de medio SOC (triptona 2%, extracto de levadura 0.5% NaCl 10nM, KCl 2.5mM, MgCl₂ 10mM y glucosa 20 mM pH 7.0). Se recuperó el volumen y se transfirió a un tubo

nuevo de 1.5ml el cual se incubó por 3h a 28°C en agitación. Posteriormente se distribuyó homogéneamente en medio LB solido con antibióticos (kanamicina 50 µg/ml, espectinomicina 100 µg/ml y carbenicilina 100 µg/ml) y se incubaron a 28°C de 3-5 días. Una vez transcurrido el tiempo se picaron colonias y se obtuvieron cajas maestras y crecimiento en medio líquido. Posterior al crecimiento celular se realizó extracción de DNA plasmídico como se describió anteriormente (5.9). La corroboración de la presencia del inserto se realizó mediante PCR punto final como se describió anteriormente.

5.22.1 Preparación de *A. tumefaciens* para transformación de soya

Se preparó medio líquido LB (que contiene espectinomicina, kanamicina y carbenicilina), se agregaron 3 ml de medio LB a tubos de ensayo. Las cepas de *Agrobacterium* fueron retiradas del ultracongelador, con un palillo se tomaron alícuotas de los stocks de glicerol e inmediatamente se colocaron en los tubos de ensayo con medio LB. Estos se incubaron a 28°C por 36h con agitación a ~250 rpm. Del cultivo anterior se tomó un inóculo de 1ml, que se agregó a un matraz de 500ml con 100ml de medio LB, y se incubó toda la noche a 28°C ~250 rpm (hasta llegar a una OD₆₂₀=0.8-1.0). Posteriormente se colectó en tubos Falcon de 50ml y se centrifugó a 4,000 rpm por 15 min a TA (22°C); posteriormente se resuspendió la pastilla bacteriana en 25 ml de medio de infección. Las células para transformación se agitaron suavemente a 60 rpm por 30 min.

5.23 Transformación genética de soya

La transformación genética se realizó empleando el protocolo de la Universidad de Iowa (Paz et al., 2006). El método consiste en transformar el embrión adherido a uno de los cotiledones para regenerar a la soya vía embriogénesis directa. Se transformó la variedad Suaqui proveniente del norte del país. Primero se describe

el método de transformación utilizando *Agrobacterium*, y posteriormente se describe el método de transformación por biobalística.

5.23.1 Esterilización de las semillas

Las semillas se colocaron en cajas Petri, que se colocaron a su vez en un desecador que contenía un vaso de precipitados de 250ml con 100ml de hipoclorito de sodio comercial puro; posteriormente se agregó lentamente 3.5 ml de HCl, enseguida se cerró el desecador. Las semillas fueron expuestas al gas cloro por 16h. Una vez cumplido el tiempo se cerraron las cajas y se retiraron del desecador. Las semillas se dejaron en aeración por 30 min en campana de flujo laminar. Pasado el tiempo las semillas estaban listas para el proceso de transformación.

5.23.2 Imbibición de semillas

En una campana de flujo laminar se colocaron las cajas Petri que contienen las semillas de soya y se les agregó agua a $\frac{3}{4}$ de la capacidad de la caja, se cerraron las cajas y se cubrieron con papel aluminio para mantenerlas en oscuridad toda la noche a temperatura ambiente.

5.23.3 Preparación de explantes para la infección

De las semillas de soya embebidas anteriormente, se tomaron grupos de 20 y se colocaron sobre una toalla de papel absorbente estéril para eliminar el exceso de humedad. Posteriormente, se realizó un corte longitudinal a lo largo del hilum de la semilla para separar los cotiledones y remover la testa. Una vez separados los

cotiledones se tomó aquél que tiene el nodo axilar y se le realizó una incisión al eje radicular. Se colocaron 30 cotiledones por caja Petri y se agregaron 30 ml de medio de infección (Medio Gamborg solución inorgánica 1/10X, Vitaminas B5 medio Gamborg 1/10X, sulfato ferroso 2.8 mg, NaEDTA 3.8 mg, sacarosa 30g, ácido 2-(N - morfolino) etanosulfónico (MES) 3.9 g, pH 5.7, ácido giberélico (GA₃) 0.25 mg, bencilaminopurina (BAP) 1.11 mg, acetosiringona 40 mg) con *Agrobacterium GL1*. Se incubó por 30 min a temperatura ambiente con agitaciones ocasionales.

5.23.4 Cocultivo

Posterior a la infección, se removió el exceso de medio de infección con una toalla de papel estéril. Enseguida se transfirieron los explantes a medio de cocultivo (Medio Gamborg solución inorgánica 1/10X, vitaminas B5 medio Gamborg 1/10X, sulfato ferroso 2.8 mg, NaEDTA 3.8 mg, sacarosa 30g, MES 3.9 g, noble agar 4.5 g pH 5.4 GA₃ 0.25 mg, BAP 1.67 mg, cisteína 400 mg, DTT 154.2 mg acetosiringona 140 mg), se colocó papel filtro estéril sobre el medio con la finalidad de reducir crecimientos bacterianos durante el cocultivo. Se sellaron las cajas con Parafilm y se incubaron a 24 °C de 3 a 5 días bajo un fotoperiodo de 16 h de luz y 8 de oscuridad.

5.23.5 Inducción de brotes

Después de los 3-5 días de cocultivo, se removieron los explantes del medio de cocultivo y se pasaron a una caja Petri vacía a la que se le agrego 50ml de medio de lavado (Medio Gamborg solución inorgánica 1X, Vitaminas B5 medio Gamborg 1X, sulfato ferroso 2.8 mg, NaEDTA 3.8 mg, sacarosa 30g, MES 0.59 g, pH 5.7,

BAP 1.11 mg, meropenem 50mg, cefotaxima 100mg, vancomicina 50mg); el exceso de humedad se eliminó con una toalla de papel estéril. Se colocaron los explantes con el lado adaxial en contacto con el medio de inducción de brotes (Medio MS solución inorgánica 1X, vitaminas B5 medio Gamborg 1X, sulfato ferroso 28 mg, NaEDTA 38 mg, sacarosa 30g, MES 0.59 g, agar noble 7 g pH 5.7, BAP 1.67 mg, meropenem 50mg, cefotaxima 100 mg, vancomicina 50 mg, glufosinato de amonio 4 mg), en un ángulo de 30° a 40°. Las cajas se sellaron con Parafilm y se incubaron a 24 °C, bajo fotoperiodo de 16 h de luz y 8 de oscuridad. El medio de inducción de brotes se cambió pasados 14 días a medio fresco; los explantes se mantuvieron en este medio por 28 días.

5.23.6 Elongación de Brotes

Después de los 28 días, los explantes se cambiaron de medio, y a cada uno se le realizó una incisión en la base. Los explantes se cambiaron a medio de elongación de brotes (Medio MS solución inorgánica 1X, vitaminas B5 medio Gamborg 1X, sulfato ferroso 28 mg, NaEDTA 38 mg, sacarosa 30g, MES 0.59 g, agar noble 7 g pH 5.7, asparagina 50 mg, ácido glutámico 100mg IAA 0.1 mg GA3 0.5 mg, zeatina-R 1 mg, meropenem 50 mg, cefotaxima 100 mg, vancomicina 50 mg, glufosinato de amonio 4 mg). Los explantes se cambiaron de medio cada dos semanas.

5.23.7 Enraizamiento de plantas

Después de que los explantes alcanzaron 8 cm de altura se pasaron a medio de enraizamiento (Medio MS solución inorgánica 1/2X, vitaminas B5 medio Gamborg 1/2X, sulfato ferroso 28 mg, NaEDTA 38 mg, sacarosa 20g, MES 0.59 g, agar

noble 7 g pH 5.6, IBA 1 mg, glufosinato de amonio 3mg). Posteriormente se les realizó una incisión en la base y se pasaron al nuevo medio de cultivo.

Las cajas se sellan con Parafilm y se incubaron a 24°C bajo fotoperiodo de 16 h de luz y 8 de oscuridad, de 1 a 2 semanas.

5.24 Transformación de soya por biobalística

El método de transformación por biobalística presenta algunas modificaciones con respecto al método anterior. El método de esterilización e imbibición se llevó a cabo como se describió en los puntos 5.23.1 y 5.23.2.

5.24.1 Extracción de embriones de soya

Las semillas de soya embebidas anteriormente se procesaron en la campana de flujo laminar, a cada semilla se le realizó un corte longitudinal para quitar la testa. Posteriormente se separaron los cotiledones de la semilla y se realizó un corte sobre el hipocótilo para llevar a cabo la separación del embrión del cotiledón. Los embriones se colocaron en medio MS basal (sales MS 1X, 3 % sacarosa, 0.4 % fitagel) durante toda la noche en oscuridad a 25-26°C.

5.24.2 Preparación de los explantes de soya

Posterior a la extracción de embriones, éstos son dispuestos de manera circular, equidistantes entre 6 y 12 mm del centro de la placa (5 mm de diámetro) con la región meristemática apical dirigida hacia arriba (de acuerdo con el protocolo descrito por (Rech et al., 2008). Los embriones fueron colocados sobre medio de cultivo MS

osmótico (sales MS 1X, 8 % sacarosa, 0.4 % fitagel) durante 6 h. Una vez transcurrido este tiempo se procedió al proceso de bombardeo por biobalística.

5.24.3 Preparación y recubrimiento de micropartículas de oro

Para la preparación de micropartículas de oro se siguió el protocolo descrito por (Tomes et al., 1995); modificado por (Lule-Chávez et al., 2021). Se pesaron 15 mg de partículas de 0.6 μm de diámetro y se transfirió a un tubo estéril de 1.5 ml. Posteriormente se adicionaron 500 μl de etanol al 100% frío (campana) y se sonicaron las partículas hasta que estuvieron completamente dispersas; se precipitaron y centrifugaron a 3000 rpm durante 1 minuto y se removió el sobrenadante (lavado dos veces). Se adicionó 1 ml de agua Milli Q fría y se resuspendieron las partículas con vórtex para su posterior precipitación. Se centrifugó a 5000 rpm por 15-30 segundos y se removió el sobrenadante, este proceso se repitió dos veces más. Por último, las partículas se resuspendieron en 500 μL de agua Milli Q, se sonicaron por 30-45s para hacer alícuotas de 50 μL = 1.5mg = 1X. Se almacenaron a -20°C.

5.24.4 Bombardeo de embriones de soya por biobalística

Para comenzar con el proceso de bombardeo se recubrieron las micropartículas con DNA y se tomaron 50 μl de partículas de oro después de mezclar. Se adicionó de 3.5-5 μg de DNA (máximo volumen 9 μl) y se homogenizó suavemente. Una vez se tiene esta mezcla, se adicionaron 50 μL de cloruro de calcio a una concentración de 2.5 M [6.93625 g/0.05L (25 mL)] y 20 μL de espermidina base libre 0.1M, se sónico y centrifugó a 5000 rpm por 15 segundos. Se retiró el sobrenadante y se comenzó la limpieza adicionando 600 μL de etanol absoluto frío.

Posteriormente se sónico y centrifugó a 5000 rpm por 15 segundos. Finalmente, se retiró el sobrenadante y se adicionaron 75 μ L de etanol absoluto, para realizar un total de 7 disparos.

Para obtener la transformación deseada en los embriones de soya se emplearon dos plásmidos, uno correspondiente al vector pCR8 GW TOPO (Invitrogen) que contiene el molde para la recombinación homóloga del gen *PBS1* de soya, y el segundo plásmido correspondiente al vector binario PBUN4116 que contiene el gen *cas9*, así como el gRNA de *PBS1*.

De los 75 μ L de la mezcla final se colocaron 10 μ l en el centro de la membrana de kaptón (macroacarreador) y se procedió a bombardear los embriones de soya con partículas de oro (0.6 μ m) que contienen ambos plásmidos a una relación molar de 3:1 o 5:1 de plásmido de recombinación homóloga: plásmido PBUN4116.

Para el proceso de bombardeo, se utilizó un equipo de biobalística PDS 100/He (Biorad; Hercules CA) con discos de ruptura de 1110 psi, a una distancia de 3 cm y 6 cm para algunos lotes en una cámara de vacío de a una presión de 27.5 mm de Hg. Se procesaron lotes de entre 100 y 140 embriones de soya.

Los embriones bombardeados se transfirieron a medio de inducción de brotes (sección 5.23.5) durante 15 días. Se realizó un proceso de cambio de temperatura durante los primeros 7 días. El primer día se colocaron los embriones alternándose la temperatura de incubación a 37°C y 28°C un día en cada caso, hasta cumplir los 7 días, posterior a ese tiempo los embriones continuaron incubándose a 28°C. Posteriormente los embriones fueron transferidos a medio de elongación (sección 5.23.6) durante 15-20 días. Cabe señalar que algunos de los medios de inducción y de elongación no se les agregó glufosinato para permitir el desarrollo del explante y procesarlo para su análisis molecular.

5.25 Análisis molecular de las plantas transformadas de soya

El proceso de análisis moleculares se realizó mediante la extracción de DNA genómico del tejido foliar obtenido a partir de los embriones desarrollados después del bombardeo por biobalística. Esto se realizó con todos los lotes de embriones de soya transformados. Las muestras fueron analizadas por PCR. El primer análisis molecular fue la amplificación del gen *cas9* como se indica a continuación (Tabla 15):

Tabla 15. Programa de amplificación.

<i>Etapas</i>	Temperatura °C	Tiempo
<i>Desnaturalización inicial</i>	94.0	3:00 min
<i>Desnaturalización</i>	94.0	0:30 s
<i>Alineamiento</i>	60.0	0:30s
<i>Elongación</i>	72.0	0:40s
<i>Elongación final</i>	72.0	5:00 min

Nota: se amplifico por 35 ciclos a partir del paso de desnaturalización hasta elongación

Posteriormente, las muestras que amplificaron *cas9* se utilizaron para identificar plantas editadas. Para ello se utilizaron los siguientes oligonucleótidos y programa de amplificación descrito en la Tabla 16.

F-PBS1-arm- 5'CCTTCACGATAAGGCAAATC 3'

F- donor-pbs1 5' TTCTTCACTGAAGGTGGTCACGAT 3'

R-ORF-NONSTOP 5'AGCATCAAGACTGCCCTGCTGCAA 3'

Tabla16. Programa de amplificación del gen *pbs1* con la modificación

<i>Etapas</i>	Temperatura °C	Tiempo
<i>Desnaturalización inicial</i>	94.0	5:00 min
<i>Desnaturalización</i>	94.0	0:35 s
<i>Alineamiento</i>	60.0	0:35s
<i>Elongación</i>	72.0	0:45s
<i>Elongación final</i>	72.0	5:00 min

Nota: se amplificó por 30 ciclos a partir del paso de desnaturalización hasta elongación

El oligonucleótido F-donor-pbs1 corresponde a la secuencia específica de la modificación, por lo que la amplificación de este producto de PCR confirmó la presencia de la edición. Para realizar el análisis de T7 endonucleasa se realizó la amplificación utilizando los oligonucleótidos, F-PBS1-arm y R-ORF-NONSTOP utilizando el programa de amplificación indicado en la Tabla 16. Una vez que se obtuvieron los productos de PCR se visualizaron en un gel de agarosa de 1.5%.

5.25.1 Análisis con la endonucleasa T7

El producto de PCR anterior se purificó empleando QIAquick PCR Purification Kit (QIAGEN; Hilden, Alemania), y se llevó a cabo la restricción con la endonucleasa T7 con los componentes y volúmenes indicados en la Tabla 17 y después se realizó un alineamiento del producto de PCR con las condiciones mostradas en la Tabla 18.

Tabla 17. Mezcla de reacción de alineamiento para la digestión con endonucleasa T7

Componente	Vf=19 µL
DNA	200 ng
10X NEBuffer 2	2 µL
Agua destilada	a 19 µL

Tabla 18. Condiciones para el alineamiento de oligonucleótidos

Etapa	Temperatura	Variación de temperatura	Tiempo
Desnaturalización Inicial	95 °C		5 minutos
Alineamiento	95-85 °C	-2°C/s	
	85-25 °C	-0.1 °C/s	
Enfriamiento	4 °C		

Posteriormente se añadió 0.5 µL de endonucleasa T7 al producto de PCR alineado. La mezcla se incubó por 1 h a 37°C y después se procedió a terminar la reacción añadiendo 1.5 µL de EDTA 0.25 M. Por último, las muestras se visualizaron en un gel de agarosa al 1.5%. Las muestras positivas fueron analizadas por PCR y se procesaron en el Laboratorio de Bioquímica Molecular, UBIPRO, FES-Iztacala UNAM mediante secuenciación por el método Sanger.

5.26 Identificación de proteínas PBS1 en especies vegetales

Se utilizó la secuencia proteica PBS1 de *A. thaliana* (ID: Q9FE20) para identificar proteínas homólogas en las bases de datos: UNIPROT (<https://www.uniprot.org/>), OrthoDB (<https://www.orthodb.org/>) y Phytozome12 (**¡Error! Referencia de hipervínculo no válida.**) mediante la herramienta BLASTP. Las secuencias redundantes fueron eliminadas por la herramienta BlastClust usando Biolinux (se usó la siguiente línea de comando: blastclust -i infile -o outfile -p T -L 1 -b T -S 100). Posteriormente se utilizó OrthoMCL para la identificación de proteínas ortólogas utilizando secuencias de aminoácidos homólogas de PBS1 en el límite del exponente del valor $E = 1 \times 10^{-5}$ y el límite de coincidencia porcentual = 50 [46]. Un total de 881 secuencias de aminoácidos ortólogas de PBS1 se

encontraron en una amplia gama de especies, incluyendo monocotiledóneas, dicotiledóneas, briofitas, pteridofitas, gimnospermas, licofitas, hepáticas, carofitas y clorofitas, que fueron identificadas y utilizadas en los siguientes análisis.

5.26.1 Identificación de subdominios en secuencias ortólogas de PBS1

Las secuencias PBS1 de diferentes especies se sometieron a múltiples alineamientos, que se realizaron utilizando ClustalW (<https://www.genome.jp/tools/bin/clustalw>) con la configuración predeterminada [alineación por pares: Fast/Approximate: K-tuple (word) tamaño: 1, tamaño de ventana: 5, penalización por hueco entre secuencias (gap): 3] para identificar los subdominios de cinasa conservados, así como los motivos GDK y SEMPH. Las regiones conservadas obtenidas mediante alineamiento múltiple se analizaron utilizando el programa MEME (<https://meme-suite.org/meme/index.html>) (Bailey et al., 2015)) con las siguientes configuraciones: modo discriminativo, secuencia primaria PBS1 (ID: Q9FE20), cualquier número de repeticiones, y 10 motivos.

5.26.2 Análisis filogenético de PBS1

Las relaciones evolutivas entre las proteínas PBS1 se investigaron utilizando un total de 881 secuencias de aminoácidos ortólogos de PBS1 correspondientes a plantas terrestres y especies de clorofitas. De un total de 881 secuencias, 182 secuencias de ortólogos de PBS1 fueron seleccionados de acuerdo con el registro de homología más alto dentro de cada género y especie de plantas terrestres. En el caso de las clorofitas, se utilizaron 116 secuencias de ortólogos de PBS1 para construir la filogenia. Se realizaron múltiples alineaciones utilizando el algoritmo

muscle (Edgar, 2004), implementado en Seaview (Gouy et al., 2010; Larsson, 2014), y editado usando Aliview (Larsson, 2014); las regiones que muestran gaps fueron eliminadas.

Los árboles filogenéticos se elaboraron utilizando máxima probabilidad de alineaciones múltiples en el programa IQ-TREE (<http://www.iqtree.org/>) con parámetros -m TESTNEWONLY -b 1000 y con bootstrapping remuestreado 1000 veces con reemplazo de mejor evolución según el modelo Corrected Akaike Information Criterion VT + F + R9 (modelo de conversión, AG = CT y frecuencia base desigual, y frecuencias base empíricas; este es el valor predeterminado si el modelo tiene una frecuencia base desigual) (Trifinopoulos et al., 2016). Este método da como resultado el árbol con mayor probabilidad de acuerdo con los datos. Las filogenias se visualizaron utilizando el programa Figtree y fueron enraizados por punto medio.

5.26.3 Reconstrucción ancestral de secuencias ortólogas de PBS1

Para las reconstrucciones ancestrales de PBS1, 182 secuencias de aminoácidos ortólogos de PBS1 de plantas terrestres fueron utilizadas. Los árboles filogenéticos se produjeron utilizando estadísticas bayesianas (<https://nbisweden.github.io/MrBayes/download.html>). Se utilizaron 5 millones de generaciones para construir la filogenia, con parámetros chains=4, printfreq=100, samplefreq=100 y burnin=200. El árbol se visualizó en Mesquite (<https://www.mesquiteproject.org/>), y la reconstrucción de caracteres ancestrales se llevó a cabo utilizando el algoritmo de máxima parsimonia.

5.26.4 Modelado de estructura PBS1

Para generar un modelo proteico de PBS1, las siguientes proteínas ortólogas fueron seleccionadas: A003— *A. thaliana*, A635—*T. aestivum*, A654—*G. max*, A179—*S. polyrhiza*, A352—*P. patens* y A773—*C. reinhardtii*. El modelado estructural se realizó utilizando Swiss Model (<https://swissmodel.expasy.org/interactive>) y fue visualizado y manipulado con el software Pymol (<http://pymol.sourceforge.net/newman/user/S0200start.html>). Las estructuras 3D de los homólogos de PBS1 de las especies mencionadas se analizaron utilizando Swiss Model (<https://swissmodel.expasy.org/>). Las estructuras proteicas conocidas que se utilizaron en este estudio fueron la serina / treonina proteína cinasa denominada cinasa 1 inducida por *Botrytis* (BIK1) para las especies A003, A635 y A654; receptor cinasa 1 asociado a *brassinosteroid insensitive 1* (BAK1) para A179; proteína *brassinosteroid insensitive 1* (BR1) para A352; y proteína cinasa (PK) para A773. Los números de acceso de Swiss Model para las estructuras conocidas son los siguientes: 5tos.1 (A003, a635 y A654), 3tl8.1 (A179), 5lpz.1. A (A352) y 6j5t.1 (A773).

6.Resultados

6.1 Cultivo de soya

Se sembraron 4 variedades de soya, rápido crecimiento, alta proteína, Nainari y Suaqui en el invernadero de Biotecnología vegetal del CINVESTAV Zacatenco con el objetivo de obtener plántulas para la obtención de DNA (Figura 11).

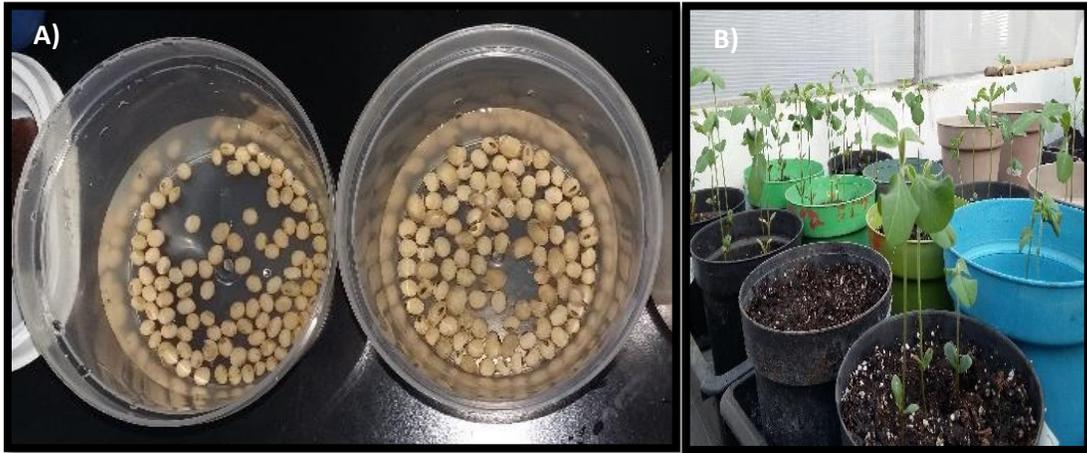


Figura 11. Siembra de soya. **A)** Esterilización de semillas de soya. **B)** Plántulas de soya.

6.2 Extracción de DNA de plantas de soya

Una vez se obtuvieron las plántulas se realizó la extracción de DNA del tejido foliar de la planta. En la Figura 12 se muestra el perfil electroforético de DNA de las 4 variedades de soya, así como la concentración obtenida.

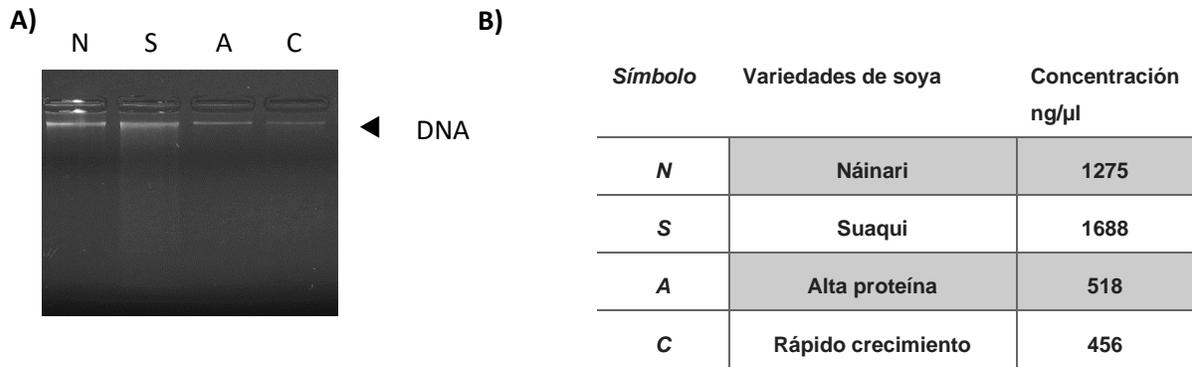


Figura 12. Extracción de DNA. **A)** Perfil electroforético de las 4 variedades de soya. **B)** Cuantificación de DNA de las 4 variedades de soya.

6.3 Secuencia *in silico* del gen *PBS1*

Se realizó la búsqueda del gen *PBS1* en la base de datos de UNIPROT obteniendo el número de referencia Q9FE20 (*PBS1_ARATH*) de *A. thaliana*, posteriormente se accedió a la base MEROPS proporcionando información sobre la proteasa, el organismo que la secreta, el sustrato (*PBS1*) y el sitio de corte de la proteasa como lo muestra la Figura 13.

Substrate	Definition: Serine/threonine-protein kinase <i>PBS1</i>
	Symbol: <i>Arabidopsis thaliana</i>
Substrate	NCBI Accession: Q9FE20
	SubstrateDB: 16387
	UniProt Accession: Q9FE20
	[Sequence on/off]
	<p>MGCFSCFDSSDDEKLNIPVDESNHGQKQSQPTVSNINISGLPSGGKLSKKTNGGSKRELLLPDGLGQIA AHTFAFRELAATMNFHPDFTFLGEGGFRVYKGRLDSTGQVAVKQLDRNGLQGNREFLVEVLMLSLLHH PNLVNLIGYCADGDQRLVYEFMPLGSLDHLHDLPPDKEALDWMRMKIAAGAAKGLEFLHDKANPPVI YRDFKSSNILLDEGFHPKLSDFGLAKLGPTGDKSHVSTRVMGTGYGCAPEYAMTGLTVKSDVYSFGVVF LELITGRKAIDSEMPHGEQNLVAVARPLFNDRRKFIKLADPRLKGRFPTRALYQALAVASMCIQEAATR PLIADVVTALSYLANQAYDPSKDDSRNRDERGARLITRNDDGGGSGSKFDLEGSEKEDSPRETARILNR DINRERAVAEAKMNGESLREKRRQSEQGTSESNSTG 1.....2.....3.....4.....5.....6.....7</p>
Cut-site	Sequence: TGDK-SHVS Position: 243-244
Protease	Definition: AvrPph8 g.p.
	Organism: <i>Pseudomonas syringae</i>
	MEROPS ID: MER019812
	PMAP ID: pmap.11787
	MEROPS code: C58.002

Figura 13. Secuencia de aminoácidos de *PBS1* de *A. thaliana*. Sustrato, sitio de corte y proteasa.

Posteriormente la secuencia de PBS1 de *A. thaliana* se utilizó para realizar un BLAST en la base de datos Phytozome 12.1 la cual contiene la secuencia de Soya (*G. max* cv Williams 82) obteniendo la siguiente información de la proteína PBS1 en soya:

Organismo: *Glycine max*

Nombre del transcrito: Glyma.08G360600.1

Localización: Chr08:47242629..47246912 reverse

Descripción: K13430 - serine/threonine-protein kinase PBS1 (PBS1)

>Glyma.08G360600.1

MSCFCEFHCKATFLCLGINKMGCFCFDSSSKEDHNLRPQHQPPLPSQISRLPSGADKLRSRNNGGSKREL
 QQPPPTVQIAAQTFTFRELAATKNFRPESFVGGGGFGRVYKGRLETTAQIVAVKQLDKNGLQGNREFLVEVL
 MSLLLHHPNLVNLIGYCADGDQRLLVYEFMPLGSLEDHLHDLPPDKEPLDWNTRMKIAVGAAGKLEYLHDKAN
 PPVIYRDFKSSNILLDEGYHPKLSDFGLAKLGPV**GDKSHVS**TRVMGTYGYPEYAMTGQLTVKSDVYSFGVV
 FLELITGRKAIDSTQPQGEQNLVTWARPLFNDRRKFSLADPRLQGRFPMRGLYQALAVASMCIQESAATRPL
 IGDVVTALSYLANQAYDPNGYRGSDDKRNRRDDKGGRI SKNDEAGGSGRRWDLEGSEKDDSPRETARILNRDL
 DRERAVAEAKMWGENLRQKRKQSLQQGSLDA*

Figura 14. Secuencia de aminoácidos de soya (*Glycine max* cv. Williams 82). Información de la proteína PBS1: Organismo, transcrito, localización, descripción y secuencia de aminoácidos en amarillo se indica el sitio de corte ortólogo a la proteína PBS1 de *A. thaliana*.

En la secuencia anterior se señala en amarillo el sitio de corte que contiene la proteína PBS1 en soya siendo el ortólogo de la proteína PBS1 de *A. thaliana*. Posteriormente se obtuvo la secuencia nucleotídica del gen *pbs1* localizando la secuencia nucleotídica que traduce al sitio de corte de la proteína PBS1 en el exón 5 enmarcada en rojo. Las secuencias enmarcadas en morado indican los oligonucleótidos utilizados para la amplificación del gen *PBS1* y la secuencias en rojo indica la secuencia que codifica para el sitio de corte de PBS1 Figura 16.

ATCTTCCCCCTGATAAGGAACCACTAGATTGGAACACTAGAATGAAAATTGCTGTTGGTGCTGCAAAAAGGATT
 AGAATACCTTCACGATAAG**GGCAAATCCTCCTGTCATCTAC**AGAGACTTCAAGTCATCTAACATATTACTTGAT
 GAAGGATAACCACCCAAAACCTTCTGACTTTGGTCTTGCGAAGCTTGGTCCTGTT**GGTGACAAATCACATGTTT**
CTACCCGTGTCATGGGAACCTTATGGTTACTGTGCTCCTGAGTATGCTATGACTGGACAGCTGACTGTGAAGTC
 TGATGTATATAGTTTTGGGGTGGTCTT**CTTAGAGCTGATTACTGGCCG**TAAAGCAATTGACAGCACCCAGCCG
 CAGGGAGAACAGAACCTTGTCCACATGG

Figura 15. Secuencia nucleotídica del exón 5 del gen *PBS1*. La secuencia génica en rojo codifica para el sitio de corte presente en PBS1.

6.4. Presencia del gen *PBS1* en variedades de soya y análisis de la secuencia correspondiente al gRNA

Se realizó un análisis molecular del gen *PBS1* en las 4 variedades de soya, con el fin de extraer DNA. En la figura 16 se muestra el perfil electroforético de los productos de PCR de 249 pb del gen *PBS1* de dichas variedades.

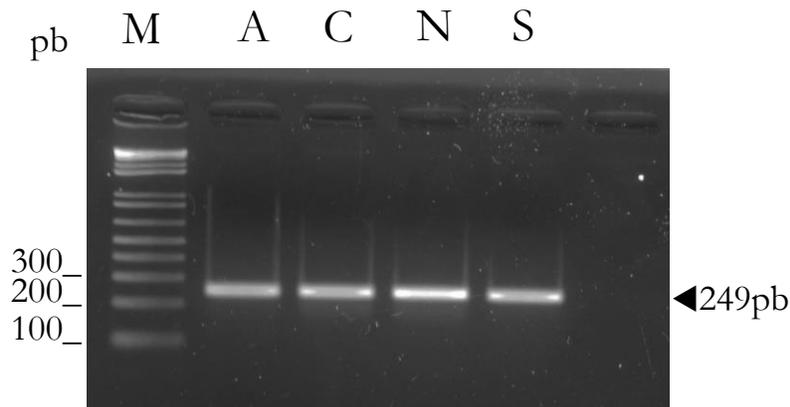
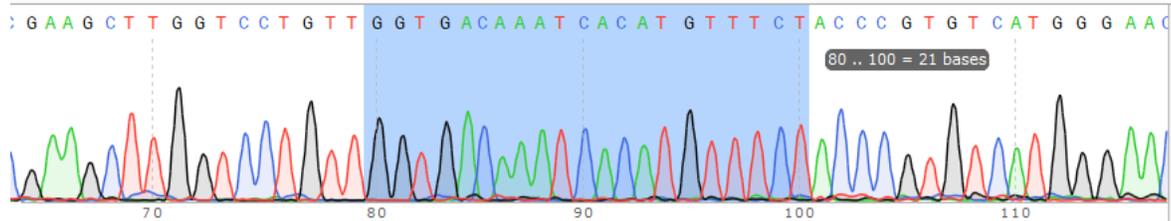


Figura 16. Perfil electroforético del gen *PBS1*. Producto de PCR de las cuatro variedades de soya. M: marcador, A: Soya con alto contenido de proteína C: Soya con alto contenido de aceite, N: Nainari S: Suaqui.

Antes de diseñar la secuencia de gRNA se secuenció el producto de PCR que contiene la región correspondiente a la secuencia de interés del gen *PBS1* de cada una de las variedades los cuales fueron analizados mediante un alineamiento bioinformático en Clustal 2.1. Se observó un elevado grado de semejanza con la secuencia nucleotídica de interés (sitio de corte) en todas las variedades respecto a la secuencia de referencia (*G. max* cv. William 82); para el diseño del gRNA se consideró el mismo gRNA para todas las variedades (Figura 17).

A)



B)

GmR.crecimiento	TGGCAAATCCTCCTGTCATCTACAGACACTTCAAGTCATCTAACATATTACTTGTATGAAG
GmAltaproteína	-GGCAAATCCTCCTGTCATCTACAGAGACTTCAAGTCATCTAACATATTACTTGTATGAAG
GmSuaqui	AGGCAAATCCTCCTGTCATCTACAGAGACTTCAAGTCATCTAACATATTACTTGTATGAAG
GmNainari	TGGCAAATCCTCCTGTCATCTACAGAGACTTCAAGTCATCTAACATATTACTTGTATGAAG
Gm.08G360600.1	-GGCAAATCCTCCTGTCATCTACAGAGACTTCAAGTCATCTAACATATTACTTGTATGAAG

GmR.crecimiento	GATACCACCCAAAACTTTCTGACTTTGGTCTTGCGAAGCTTGGTCTGTTGGTGACAAAT
Gmaltaproteína	GATACCACCCAAAACTTTCTGACTTTGGTCTTGCGAAGCTTGGTCTGTTGGTGACAAAT
GmSuaqui	GATACCACCCAAAACTTTCTGACTTTGGTCTTGCGAAGCTTGGTCTGTTGGTGACAAAT
GmNainari	GATACCACCCAAAACTTTCTGACTTTGGTCTTGCGAAGCTTGGTCTGTTGGTGACAAAT
Gm.08G360600.1	GATACCACCCAAAACTTTCTGACTTTGGTCTTGCGAAGCTTGGTCTGTTGGTGACAAAT

GmR.crecimiento	CACATGTTTCTACCCGTGTCATGGGAACTTATGGTACTGTGCTCCTGAGTATGCTATGA
GmAltaproteína	CACATGTTTCTACCCGTGTCATGGGAACTTATGGTACTGTGCTCCTGAGTATGCTATGA
GmSuaqui	CACATGTTTCTACCCGTGTCATGGGAACTTATGGTACTGTGCTCCTGAGTATGCTATGA
GmNainari	CACATGTTTCTACCCGTGTCATGGGAACTTATGGTACTGTGCTCCTGAGTATGCTATGA
Gm.08G360600.1	CACATGTTTCTACCCGTGTCATGGGAACTTATGGTACTGTGCTCCTGAGTATGCTATGA

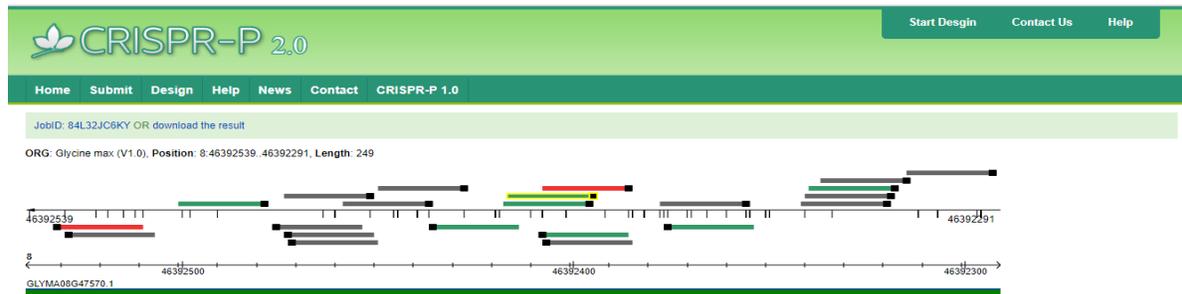
GmR.crecimiento	CTGGACAGCTGACTGTGAAGTCTGATGTATATAGTTTGGGGTGGTCTTCTTA-AGCTGA
GmAltaproteína	CTGGACAGCTGACTGTGAAGTCTGATGTATATAGTTTGGGG-----
GmSuaqui	CTGGACAGCTGACTGTGAAGTCTGATGTATATAGTTTGGGGTGGTCTTCTTAGAGCTGA
GmNainari	CTGGACAGCTGACTGTGAAGTCTGATGTATATAGTTTGGGGTGGTCTTCTTAAG-CTGA
Gm.08G360600.1	CTGGACAGCTGACTGTGAAGTCTGATGTATATAGTTTGGGGTGGTCTTCTTAGAGCTGA

GmR.crecimiento	TTACTGGCCGA
GmAltaproteína	-----
GmSuaqui	TTACTGGCCGA
GmNainari	TTACTGGCCGA
Gm.08G360600.1	TTACTGGCCG-

Figura 17. Secuencia del fragmento del gen PBS1 que incluye el sitio de corte en diferentes variedades de soya. **A)** Electroferograma de la reacción de secuenciación en las 4 variedades de soya. **B)** Alineamiento de dicho fragmento en las cuatro variedades de soya mostrando identidad respecto a la secuencia de referencia.

6.5 Diseño de gRNA para el gen *PBS1* de soja

Una vez que se confirmó la conservación de la secuencia de interés se utilizó el software CRISP-P 2.0. Con ello se identificó el gRNA idóneo y el más cercano al sitio de corte. Los resultados del análisis se muestran en la Figura 18.



	<i>On-score</i>	<i>Secuencia</i>	<i>Región</i>	<i>%GC</i>
<i>guide1</i>	0.5855	TCTCTGTAGATGACAGG AGG	exón	50%
<i>guide2</i>	0.5039	CGTGTCATGGGAACCTA TGG	exón	50%
<i>guide3</i>	0.4043	GTTTCTACCCGTGTCATGGG	exón	55%
<i>guide4</i>	0.3614	CCAGTCATAGCATACTC AGG	exón	50%
<i>guide5</i>	0.2944	AACATATTACTTGATGA AGG	exón	30%

Sitios target-off:

<i>Secuencia</i>	<i>Off-score</i>	<i>MMs</i>	<i>Locus</i>	<i>Gen</i>	<i>Región</i>
<i>GTTTCACCCGTGTCATGGG</i>	0.909	1MMs	10:+50901219	GLYMA10G44580	exón
<i>GTTTCAACCCGTGTCATGGG</i>	0.867	1MMs	20:+46701982	GLYMA20G39370	exón
<i>GATTCATCCGTGTCATAGG</i>	0.642	4MMs	15:+34099566	GLYMA15G30830	exón
<i>TTTTCTACCAAGATCATTGG</i>	0.640	4MMs	1:-3722911		Intergénica
<i>GTTTCACCAAGATCATGGG</i>	0.630	3MMs	17:-41826084	GLYMA17G38150	Exón
<i>GTTTCAACCAAGATCATGGG</i>	0.601	3MMs	9:+41237971	GLYMA09G34980	Exón
<i>GTATCTACCCGAGTCATGGG</i>	0.571	2MMs	9:-45528285	GLYMA09G40650	Exón
<i>GTTTCAACTAGAGTCATGGG</i>	0.526	4MMs	17:-3988530	GLYMA17G05660	Exón
<i>GTTTCAACTAGAGTCATGGG</i>	0.526	4MMs	13:+20873533	GLYMA13G17050	Exón
<i>GTTTTACTTGTGTCATTGG</i>	0.524	3MMs	10:-9998866		Intergénica
<i>GTTGCTACTTATGTCATTGG</i>	0.515	4MMs	18:-61414259	GLYMA18G53005	Exón
<i>GTTTCTACCCGCGTCAAAGG</i>	0.505	2MMs	18:+62181464		Intergénica
<i>TTTTCTAATAGTGCATAGG</i>	0.493	4MMs	7:-17729400	GLYMA07G17940	Exón
<i>TTTTCTAATAGTGCATTGG</i>	0.493	4MMs	18:-10688012		Intergénica
<i>ATATCAACTCGTGCATGGG</i>	0.487	4MMs	17:+4440437	GLYMA17G06251	Exón
<i>ATTTATACCTGTGTCATGGG</i>	0.484	3MMs	5:+41010871	GLYMA05G37400	Exón
<i>ATATCAACACCGTGCATGGG</i>	0.478	4MMs	13:-20340152	GLYMA13G16280	Exón

<i>GGTTCCACGCATGTCATCGG</i>	0.473	4MMs	18:-27091095	Intergenica
<i>GTTGTATTCGTGTCATGGG</i>	0.459	3MMs	17:+2513395	Intergenica
<i>TTTTGAACTCGTGTTCATGG</i>	0.455	4MMs	13:-20732559	Intergenica

Figura 18. Diseño de gRNA para el gen *PBS1* de Soya. En amarillo se indica el gRNA más cercano a la secuencia de interés, así como el score y los target-off.

El resultado del análisis con el software mostró los mejores gRNAs candidatos en la secuencia insertada; asimismo proporcionó información de target-off, score, secuencia, secuencia PAM, región en donde se localizan y % de GC. Finalmente, se seleccionó un gRNA de 20 nt (remarcado en amarillo) cercano al sitio de corte en el gen *PBS1* con 55% de contenido en GC, con un score de 0.4043 y 20 secuencias target-off.

6.6 Construcción del vector para edición de genoma

El vector PBUN411U6 es utilizado en el laboratorio para la edición de genoma mediante el uso de la proteína Cas9 y el gRNA seleccionado. Una vez que se obtuvo la cepa de *E. coli* con el vector PBUN411U6 se realizó extracción de DNA plasmídico y posteriormente se realizó una digestión con la enzima BsaI para linealizar el plásmido, lo que se verificó en gel de agarosa al 0.8% (Figura 19).

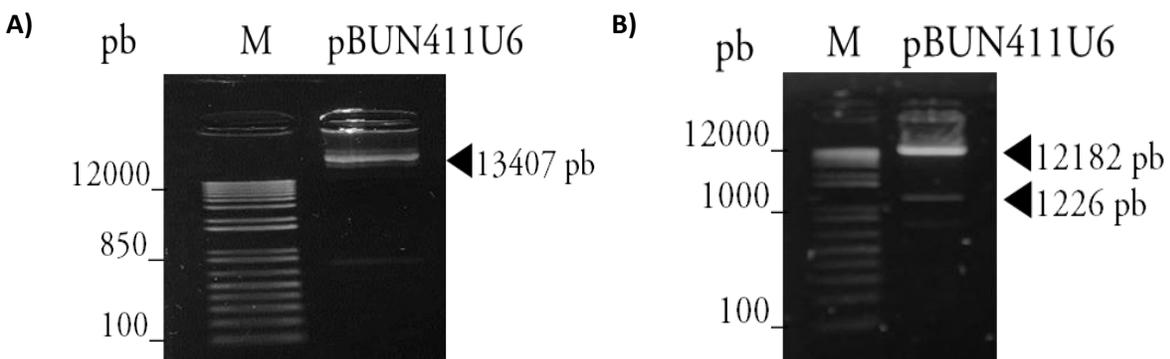


Figura 19. Linealización del Vector PBUN411U6. **A)** Perfil electroforético del vector PBUN411U6 **B)** Perfil electroforético de la digestión del vector PBUN411U6 con BsaI.

Una vez que se obtuvieron los oligonucleótidos del gRNA seleccionado se realizó una hibridación con éste (Figura 20); el híbrido resultante se purificó y se utilizó para ligar en el vector PBUN411U6.

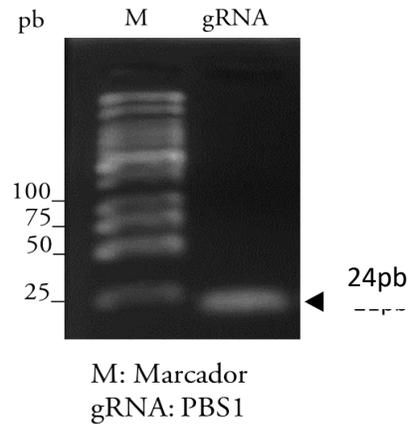


Figura 20. Hibridación del gRNA. Perfil electroforético del gRNA de *PBS1*.

Después de haber introducido el vector recombinante en *E.coli*, se seleccionaron 10 colonias a las cuales se le realizó extracción de DNA plasmídico. Para corroborar la presencia del inserto se realizó una digestión con la enzima Hind III, lo que se verificó en un gel de agarosa al 0.8%. Además, se realizó una PCR punto final utilizando los oligonucleótidos F-gRNA-PBS1, R-U6t que amplifica un fragmento de 187 pb en clonas positivas (Figura 21).

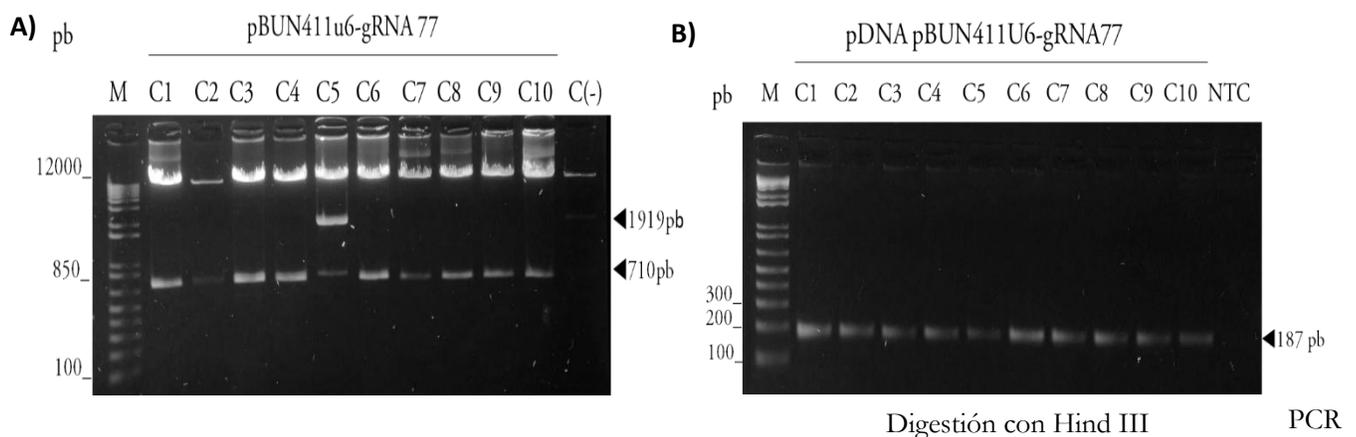
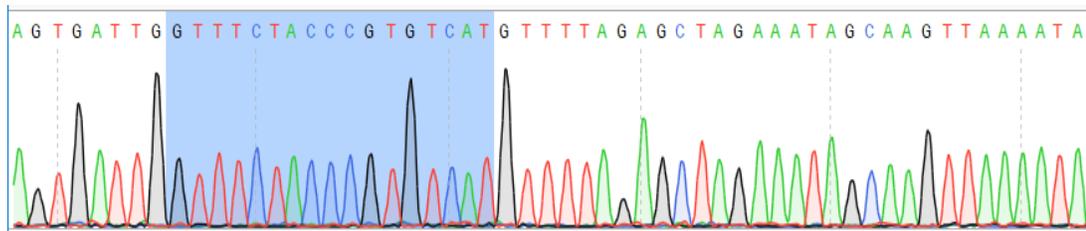


Figura 21. Corroboración de la presencia del gRNA-PBS1 en el vector PBUN411U6. **A)** Digestión con Hind III de 10 muestras de DNA plasmídico. **B)** Perfil electroforético del producto de PCR de F-gRNA-PBS1, R-U6t.

Se seleccionaron las muestras C1, C2 y C4, las cuales se secuenciaron y se corroboró la presencia del gRNA del fragmento del gen *PBS1* en el vector PBUN411U6 (Figura 22).

A)



```

gRNA-R-2 -----CCCCAATTTAAATGGGAAG--ACAATAGTATTTTC
gRNA-R-4 -----CGGAATTTTATGGGAAG--ACA-TAGTATTTTC
gRNA-R-1 -----CGCAATT-ATATGGGAAG--ACAATAGTATTTTC
gRNA-O-U6 CCTGGGAATCTGAAAGAAGAGAAGCAGGCCCATTTATATGGGAAAGAACAATAGTATTTTC
gRNA-F-1 CCTGGGAATCTGAAAGAAGAGAAGCAGGCCCATTTATATGGGAAAGAACAATAGTATTTTC
gRNA-F-2 CCTGGGAATCTGAAAGAAGAGAAGCAGGCCCATTTATATGGGAAAGAACAATAGTATTTTC
gRNA-F-4 CCTGGGAATCTGAAAGAAGAGAAGCAGGCCCATTTATATGGGAAAGAACAATAGTATTTTC
                                     ***      *
gRNA-R-2 TTATATAGGCCCATTTAAGTTGAAAACAATCTTCAAAGTCCACATCGCTTAGATAAGA
gRNA-R-4 TTATATAGGCCCATTTAAGTTGAAAACAATCTTCAAAGTCCACATCGCTTAGATAAGA
gRNA-R-1 TTATATAGGCCCATTTAAGTTGAAAACAATCTTCAAAGTCCACATCGCTTAGATAAGA
gRNA-O-U6 TTATATAGGCCCATTTAAGTTGAAAACAATCTTCAAAGTCCACATCGCTTAGATAAGA
gRNA-F-1 TTATATAGGCCCATTTAAGTTGAAAACAATCTTCAAAGTCCACATCGCTTAGATAAGA
gRNA-F-2 TTATATAGGCCCATTTAAGTTGAAAACAATCTTCAAAGTCCACATCGCTTAGATAAGA
gRNA-F-4 TTATATAGGCCCATTTAAGTTGAAAACAATCTTCAAAGTCCACATCGCTTAGATAAGA
*****
gRNA-R-2 AAACGAAGCTGAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGGTTTCTACCCGTGT
gRNA-R-4 AAACGAAGCTGAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGGTTTCTACCCGTGT
gRNA-R-1 AAACGAAGCTGAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGGTTTCTACCCGTGT
gRNA-O-U6 AAACGAAGCTGAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGGTTTCTACCCGTGT
gRNA-F-1 AAACGAAGCTGAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGGTTTCTACCCGTGT
gRNA-F-2 AAACGAAGCTGAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGGTTTCTACCCGTGT
gRNA-F-4 AAACGAAGCTGAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGGTTTCTACCCGTGT
*****
gRNA-R-2 CATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAAGGCTAGTCCGTTATCAACTTGAAAA
gRNA-R-4 CATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAAGGCTAGTCCGTTATCAACTTGAAAA
gRNA-R-1 CATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAAGGCTAGTCCGTTATCAACTTGAAAA
gRNA-O-U6 CATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAAGGCTAGTCCGTTATCAACTTGAAAA
    
```

```

gRNA-F-1      CATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAAGGCTAGTCCGTTATCAACTTGAAAA
gRna-F-2     CATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAAGGCTAGTCCGTTATCAACTTGAAAA
gRna-F-4     CATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAAGGCTAGTCCGTTATCAACTTGAAAA
*****

gRNA-R-2     AGTGGCACCGAGTCGGTGCCTTTTTTTTGCAAATTTCCAGATCGATTTCTTCTCCTCT
gRNA-R-4     AGTGGCACCGAGTCGGTGCCTTTTTTTTGCAAATTTCCAGATCGATTTCTTCTCCTCT
gRNA-R-1     AGTGGCACCGAGTCGGTGCCTTTTTTTTGCAAATTTCCAGATCGATTTCTTCTCCTCT
gRNA-O-U6    AGTGGCACCGAGTCGGTGCCTTTTTTTTGCAAATTTCCAGATCGATTTCTTCTCCTCT
gRNA-F-1     AGTGGCACCGAGTCGGTGCCTTTTTT--GCAAAATCCAAGGA-----
gRNA-F-2     AGTGGCACCGAGTCGGTGCCTTTTTT--GCAAAATACCGGT-----
gRNA-F-4     AGTGGCACCGAGTCGGTGCCTTTTTT-GCAAAATACCTTAGCCGGCCTAC-----
*****

gRNA-R-2     GTTCTTCGG
gRNA-R-4     GTTCTTCGG
gRNA-R-1     GTTCTTCGG
gRNA-O-U6    GTTCTTCGG
gRNA-F-1     -----
gRNA-F-2     -----
gRNA-F-4     -----

```

Figura 22. Secuenciación del vector PBUN411U6-gRNA-PBS1. A) electroferograma del vector PBUN411U6 con el gRNA de *PBS1* en 3 muestras de DNA plasmídico. B) Alineamiento de las secuencias analizadas mostrando la presencia del gRNA-*PBS1*.

A continuación, se muestra el esquema del vector PBUN411U6-gRNA-PBS1 que fue utilizado para los procesos de transformación en soja (Figura 23).

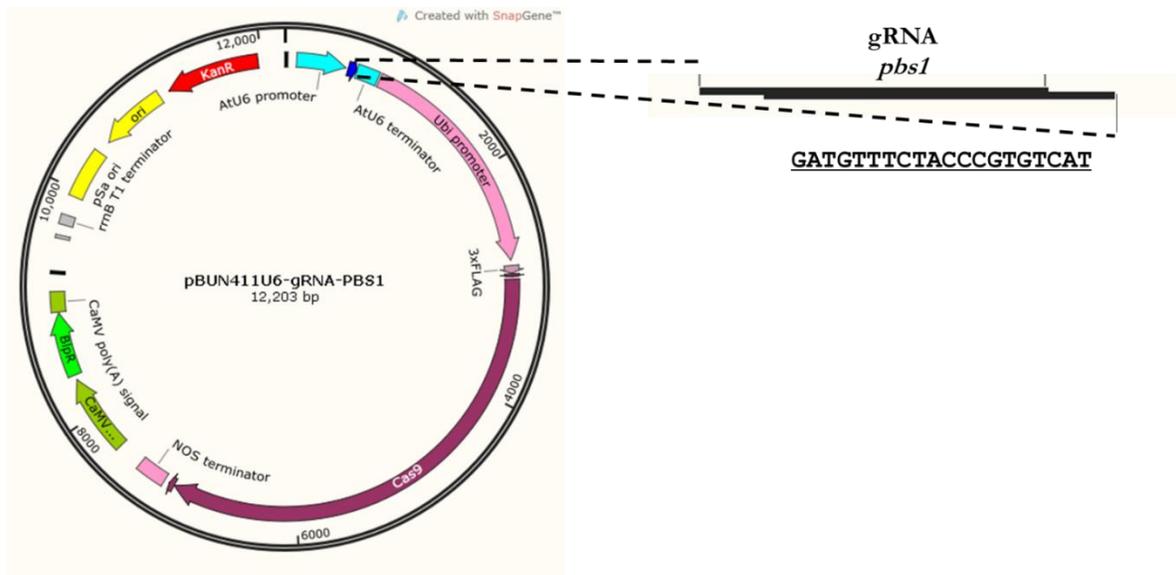


Figura 23. Esquema del vector de expresión pBUN411U6-gRNA-PBS1. AtU6: promotor U6 de *A. thaliana*, gRNA scaffold: para el sistema CRISPR / Cas9 de *Streptococcus pyogenes*, Promotor Ubi: promotor del gen de poliubiquitina de maíz, Cas9: Endonucleasa Cas9 del sistema CRISPR / Cas de *Streptococcus pyogenes* tipo II, NOS: Terminador de nopalina sintetasa, CaMV 35S: promotor 35S del virus del mosaico de la coliflor, BlpR: fosfinotricina acetiltransferasa, CaMV poli(A): señal de poliadenilación del Virus del mosaico de la coliflor.

6.7 Obtención de secuencia DONOR

Se utilizó el software de Addgene, el cual indica el sitio de inserción potencial considerando la secuencia de interés y el gRNA diseñado, de tal manera se diseñan oligonucleótidos con la secuencia DONOR (Figura 24). Para este fin se amplificaron dos productos de PCR; uno de ellos amplificó un producto de 300 pb (producto1) que corresponde al gen *PBS1* y el segundo producto amplificó un producto de 68 pb (producto 2) correspondiente a la secuencia DONOR (Figura 25).

The screenshot shows the Addgene software interface for designing a DNA donor sequence. The interface includes the following fields and options:

- Organism:** Soybean (glyMax2) Glycine Max
- Gene target:** 100305422
- Transcript:** NM_001248235.1 (-) chr8:47242074
- Guide RNA target sequence:** CATGTTTCTACCGTGTCAT
- Select modification location:** chr8:47.242.813 to chr8:47.242.743
- 5' homology arm length:** 50
- Sequence to insert:** TTTCTTCACTGAAGGAGGACACGATAC (highlighted in green)
- 3' homology arm length:** 50
- Sequence to remove:** FFTEGGHDT (highlighted in red)
- Generate DNA donor:** Button to generate the donor sequence.

Figura 24. Diseño de secuencia DONOR. Software de addgene, indica el organismo, gen blanco, transcrito, RNAg, secuencia a insertar (en color verde) sitio de corte FFTEGGHDT, homología de brazos del gen *PBS1*

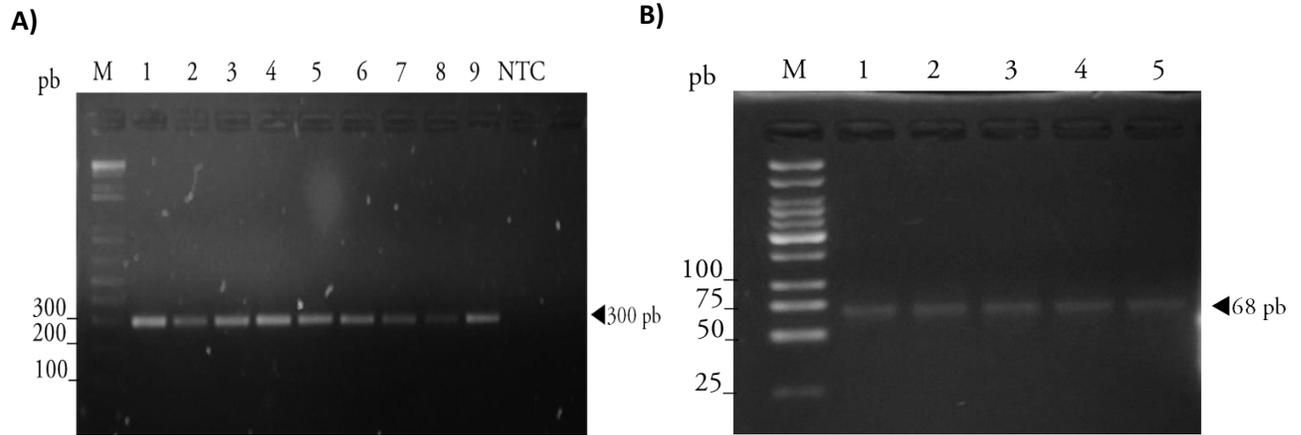


Figura 25. Producto de PCR del fragmento del gen *PBS1* y secuencia DONOR. **A)** Perfil electroforético del gen *PBS1* de 300pb. **B)** Perfil electroforético de la secuencia DONOR con la modificación.

Se purificaron ambos productos de PCR, los cuales fueron utilizados para la amplificación de dos productos de PCR de 197 pb (productos 1.1 y 1.2). El producto 1.1 contiene la secuencia inicial del gen *pbs1* y la secuencia final DONOR, y el producto 1.2 contiene la secuencia inicial DONOR y la secuencia final del fragmento del gen *PBS1* (5.16). De acuerdo con el análisis del producto de interés, se purificó la banda de 197 pb (Figura 26).

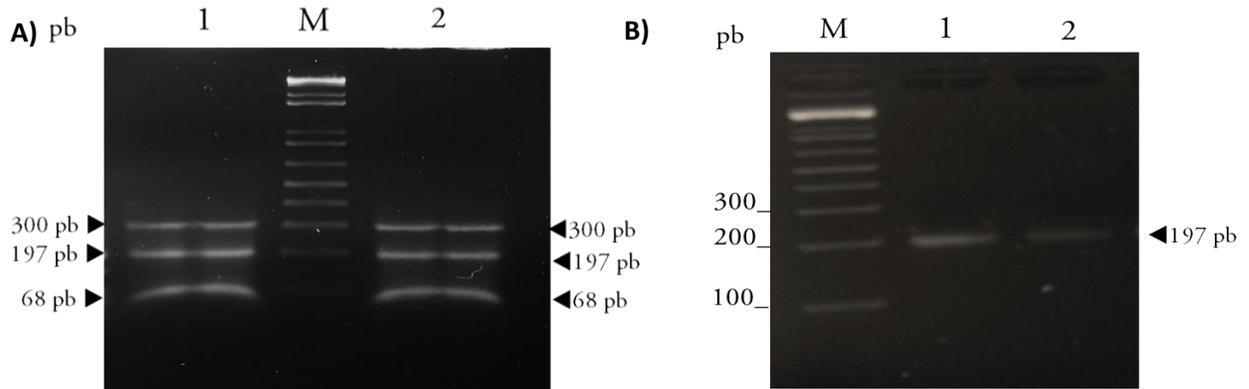


Figura 26. Amplificación y purificación de los productos 1.1 y 1.2. **A)** Perfil electroforético de la secuencia inicial del gen *PBS1* con la secuencia DONOR (carril 1) y de la

secuencia final del gen *PBS1* con la secuencia DONOR (carril 2). **B)** Purificación del producto 1.1 (carril 1) y producto 1.2 (carril 2).

Los productos purificados se utilizaron para realizar el ensamblaje como se describió en la metodología (5.17), obteniéndose un producto de PCR de 300 pb que contiene la secuencia DONOR en el fragmento del gen *PBS1* (Figura 27).

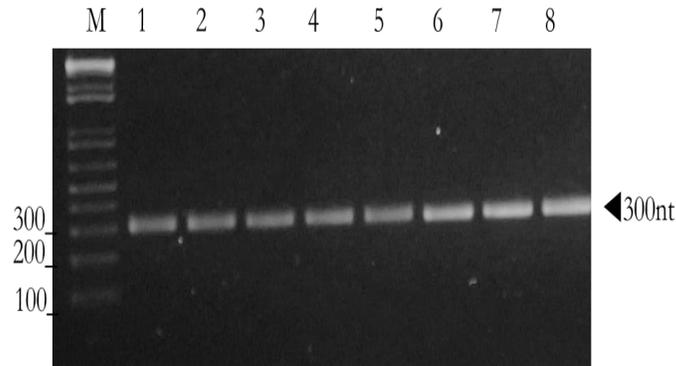


Figura 27. Ensamblaje de la secuencia DONOR y el gen *PBS1*. Perfil electroforético del gen *pbs1* con la secuencia DONOR.

Una vez que se obtuvo la secuencia DONOR se purificó el producto usando el Kit Zymoclean DNA Gel Recovery y se procedió a clonar el producto de PCR de 300 pb en el vector pCR8 GW TOPO. Posteriormente el vector se transformó en *E. coli* Mach T1. Una vez que se obtuvieron las colonias transformadas, se realizó la extracción de DNA plasmídico con la secuencia DONOR el cual fue utilizado para transformación por biobalística. A continuación, se muestra un esquema representativo de la secuencia DONOR, así como la secuencia génica de *PBS1* empleada para el ensamblaje (Figura 28A). Finalmente se esquematiza el vector obtenido de pCR8 GW TOPO con el producto de PCR que contiene la secuencia DONOR (Figura 28B).

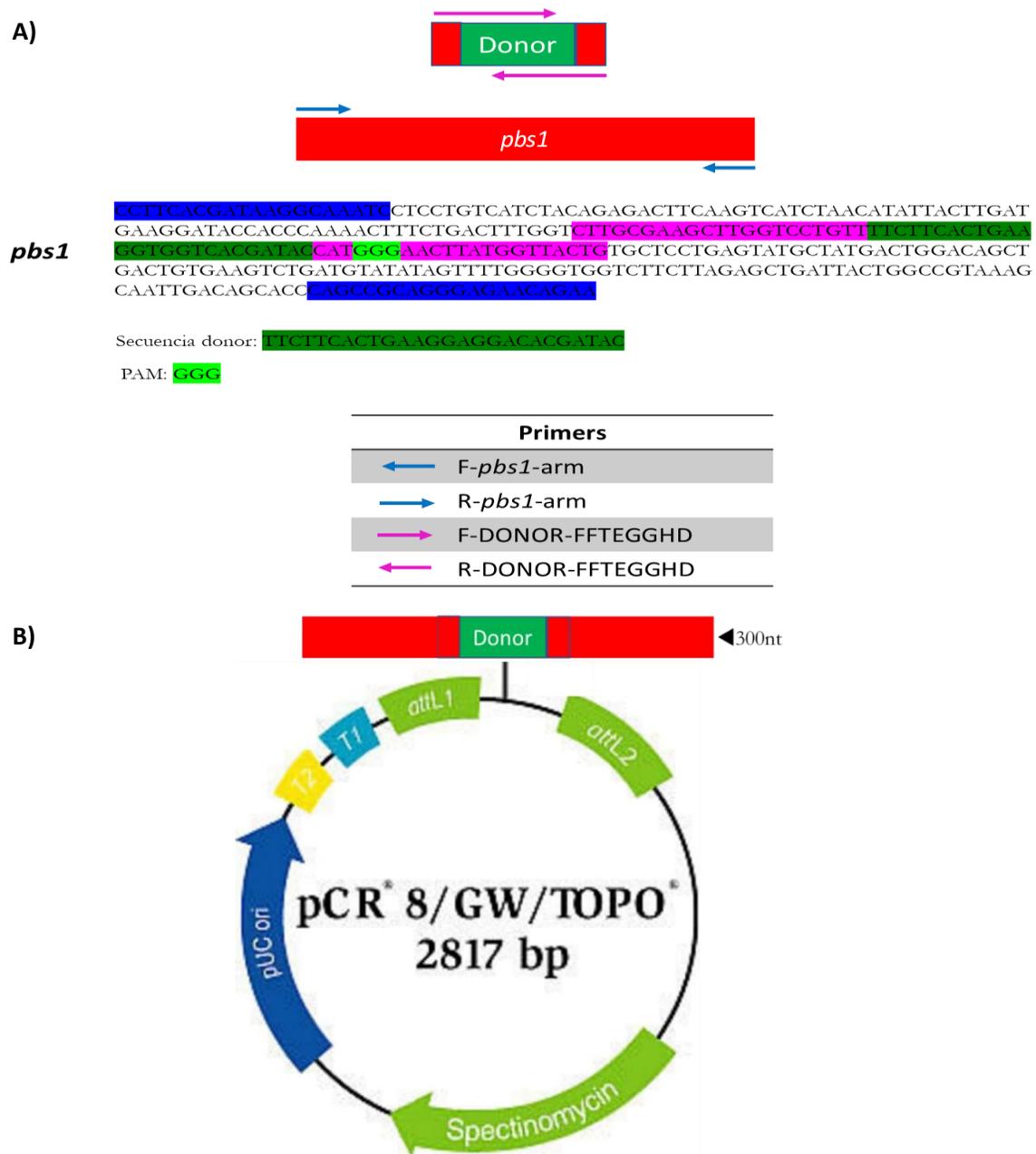


Figura 28. Figura esquemática del fragmento del gen PBS1 con la secuencia DONOR.

A) en verde se presenta la secuencia a insertar (DONOR), en rojo se representa la secuencia génica de *PBS1*, en azul se representan los oligonucleotidos F-*pbs1*-arm y R-*pbs1*-arm, en morado se representan los oligonucleotidos largos que contienen un fragmento de 21 pb correspondientes al gen *PBS1* adicionando la secuencia DONOR (verde) en ambos sentidos Fw y Rw. **B)** Vector pCR8 GW TOPO con la secuencia DONOR.

6.8 Obtención de plantas de soya transgénicas

Una vez que se corroboró la presencia del gRNA en el vector PBUN411U6 se realizó la transformación de cepas de *A. tumefaciens* por electroporación. Ya obtenido el vector de expresión en *A. tumefaciens* se realizó la transformación de soya utilizando la variedad Suaqui. Este proceso se realizó para los primeros dos lotes utilizando solo el vector PBUN411U6-gRNA-PBS1, posteriormente se empleó el método de transformación por biobalística empleando los vectores PBUN411U6-gRNA-PBS1 y DONOR-PBS1. Primero se explicará el método de transformación por *A. tumefaciens* y posteriormente por biobalística.

Se comenzó con la esterilización de semillas, la imbibición y la incisión de las semillas. Posteriormente se agro infiltraron los embriones de soya y se colocaron en medio de cocultivo (Figura 29).

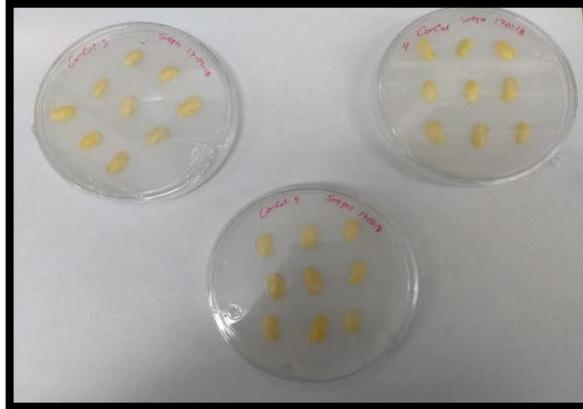


Figura 29. Cotiledones de soya en medio de cocultivo. Cajas Petri con 9 cotiledones de soya son colocados en medio de cocultivo.

Después de 5 días en medio de cocultivo los explantes se transfirieron a una caja magenta que contiene medio de inducción de brotes (Figura 30).

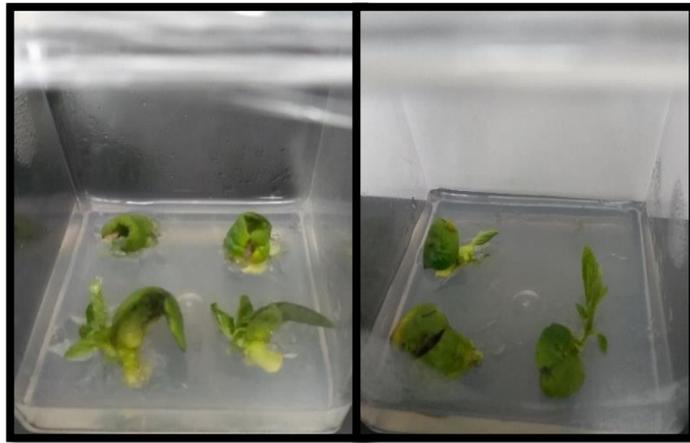


Figura 30. Explantes en medio de inducción de brotes. En cajas magenta se colocaron 4 explantes de soya en medio de inducción.

Los explantes se mantuvieron en medio de inducción de brotes por 4 semanas y se realizó un segundo cambio de medio en la segunda semana. Una vez transcurrido el tiempo, los explantes se traspasaron a medio de elongación de brotes que contiene glufosinato (medio de selección) por lo que las plantas que siguieron desarrollándose sugerían estar transformadas. Cuando las plantas alcanzaron una altura de 8 cm fueron transferidas a medio de inducción de raíces (Figuras 31 y 32).



Figura 31. Explantes en medio de elongación (1 semana). Desarrollo de explantes de soya después de 1 semana.

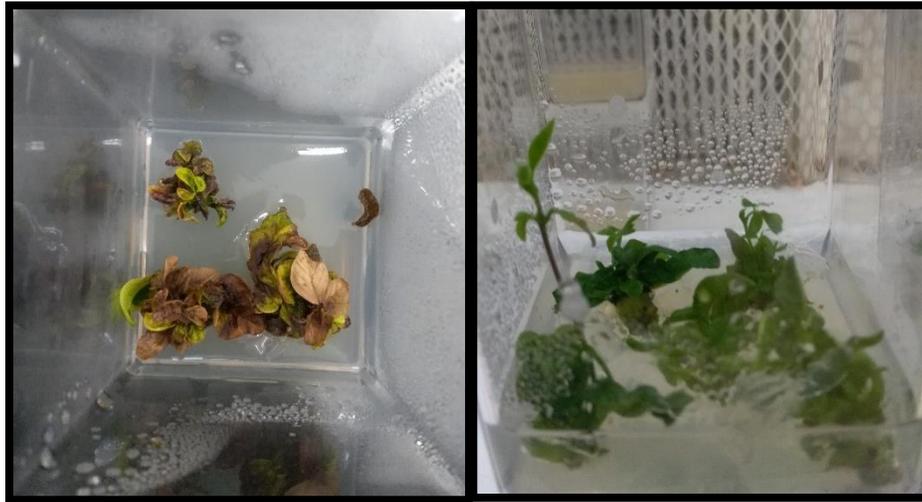


Figura 32. Explantes en medio de elongación (2 semanas). Desarrollo de explantes de soya después de 2 semanas.

Se seleccionaron plantas las cuales fueron colocadas en medio con glufosinato y sin glufosinato. De las plantas que sobrevivieron en cada uno de los procesos se les realizó extracción de DNA y se corroboró su transformación mediante la detección del gen *cas9*, sin embargo, éste no fue detectado. En esta primera etapa de transformación mediante *A. tumefaciens* se pretendió estandarizar y obtener plantas transformadas con el vector PBUN411U6-gRNA-*pbs1* con el objetivo de generar un cambio en el gen *PBS1* mediante recombinación no homóloga. De acuerdo con nuestros análisis moleculares no se detectó el gen *cas9* en el primer lote de embriones analizados (Tabla 13). Posteriormente se procedió a realizar la transformación mediante biobalística.

6.9 Transformación por biobalística

Posterior a la esterilización de las semillas se obtuvieron los embriones como se indica en la metodología. Se emplearon 110-180 semillas de soya por cada lote de las cuales se obtenía un número variable de embriones (Tabla 13). De los embriones obtenidos se dispusieron 10 embriones de manera circular en medio de cultivo MS osmótico los cuales fueron utilizados para bombardear (Figura 33).

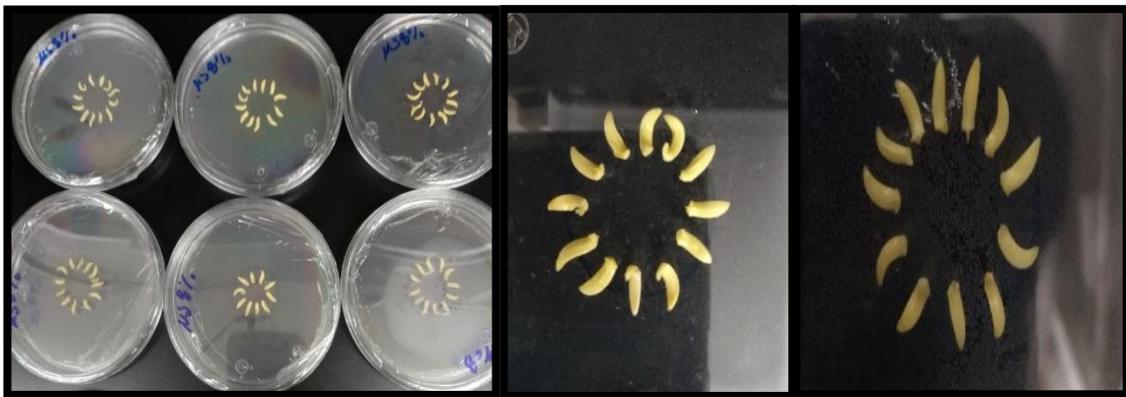


Figura 33. Embriones de soya. Colocación de embriones de soya en medio osmótico al 8%. Los embriones fueron colocados de forma circular en el centro de la caja Petri para su posterior bombardeo.

Se transformaron 12 lotes por biobalística de embriones de soya. En total se transformaron 1558 embriones (Tabla 19). Cada uno de los lotes presentó algunas modificaciones. Los primeros dos lotes fueron transformados mediante *A. tumefaciens*. Los lotes 3-12 fueron transformados por biobalística, de estos, los lotes 3-6 se bombardearon a una distancia de 6 cm y los lotes restantes a una distancia de 3 cm. También, se utilizó una relación molar diferente de los vectores en los lotes transformados; además, se empleó un cambio de temperatura (de 37°C a 28°C cada 24 h en los embriones transformados). Los lotes fueron

transformados con *A. tumefaciens* empleando el vector PBUN411U6; para los demás eventos de transformación por biobalística se empleó el plásmido de expresión, así como el plásmido de recombinación homologa. En la Tabla 19 se describen los eventos de transformación realizados.

Tabla 19. Eventos de transformación de embriones de soya mediante el sistema CRISPR/Cas9.

Lote	No. De embriones analizados	Vectores	Relación Molar	Distancia de disparo	Temperatura	% de transformación	% de edición
1	120	PBUN411U6	-	-	28°C	0	0
2	128	PBUN411U6	-	-	28°C	0	0
3	100	PBUN411U6:PCR8-DONOR	1:3	6 cm	28°C	0	0
4	110	PBUN411U6:PCR8-DONOR	1:3	6 cm	28°C	0	0
5	134	PBUN411U6:PCR8-DONOR	1:3	6 cm	28°C	0	0
6	147	PBUN411U6:PCR8-DONOR	1:3	6 cm	28°C	0	0
7	182	PBUN411U6:PCR8-DONOR	1:5	3 cm	28°C-37°C	4.9 (9)	1.6 (3)
8	140	PBUN411U6:PCR8-DONOR	1:5	3 cm	28°C-37°C	4.2(6)	2.1 (3)
9	140	PBUN411U6:PCR8-DONOR	1:5	3 cm	28°C-37°C	5.7(8)	0
10	132	PBUN411U6:PCR8-DONOR	1:5	3 cm	28°C-37°C	4.5 (6)	0
11	110	PBUN411U6:PCR8-DONOR	1:3	3 cm	28°C-37°C	4.5 (5)	0
12	115	PBUN411U6:PCR8-DONOR	1:7	3 cm	28°C-37°C	6.9 (8)	0

Una vez que se realizó el proceso de transformación los embriones se mantuvieron durante 15 días en medio de inducción, realizando un intercambio de temperaturas de 37°C a 28°C cada 24 h durante 7 días. Posteriormente se incubó a 28°C durante el desarrollo de los embriones. Después de 15 días en medio de inducción los embriones fueron colocados en medio de elongación (Figura 34).

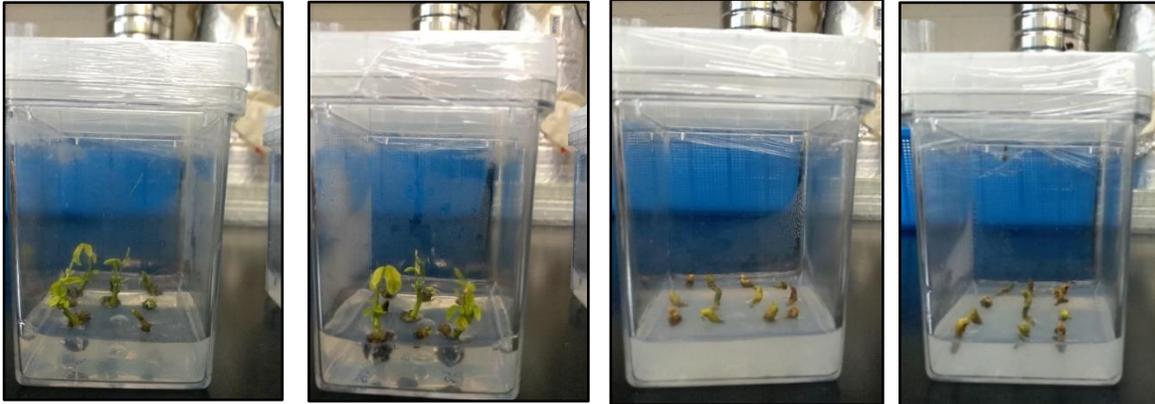


Figura 34. Embriones en medio de elongación. De 9 a 12 embriones de soja fueron colocados en medio de elongación con y sin glufosinato.

6.10 Análisis molecular del gen *cas9* T0

La transformación y edición de los explantes que generaron tejido foliar fue monitoreada por amplificación del gen *cas9*, del fragmento del gen *PBS1* con la modificación, así como con análisis por digestión con endonucleasa T7 y finalmente por secuenciación. Para ello se realizó extracción de DNA y posteriormente se realizaron los análisis citados. De todas las muestras analizadas los lotes 7-12 demostraron estar transformados ya que mostraron muestras positivas al gen *cas9* (Tabla 13). Sin embargo, solo los lotes 7 y 8 mostraron estar editados.

De acuerdo con el análisis molecular se obtuvieron 42 explantes transformados y 6 editados. Las muestras analizadas amplificaron un fragmento de 436 pb correspondientes al gen *cas9* (Figura 35).

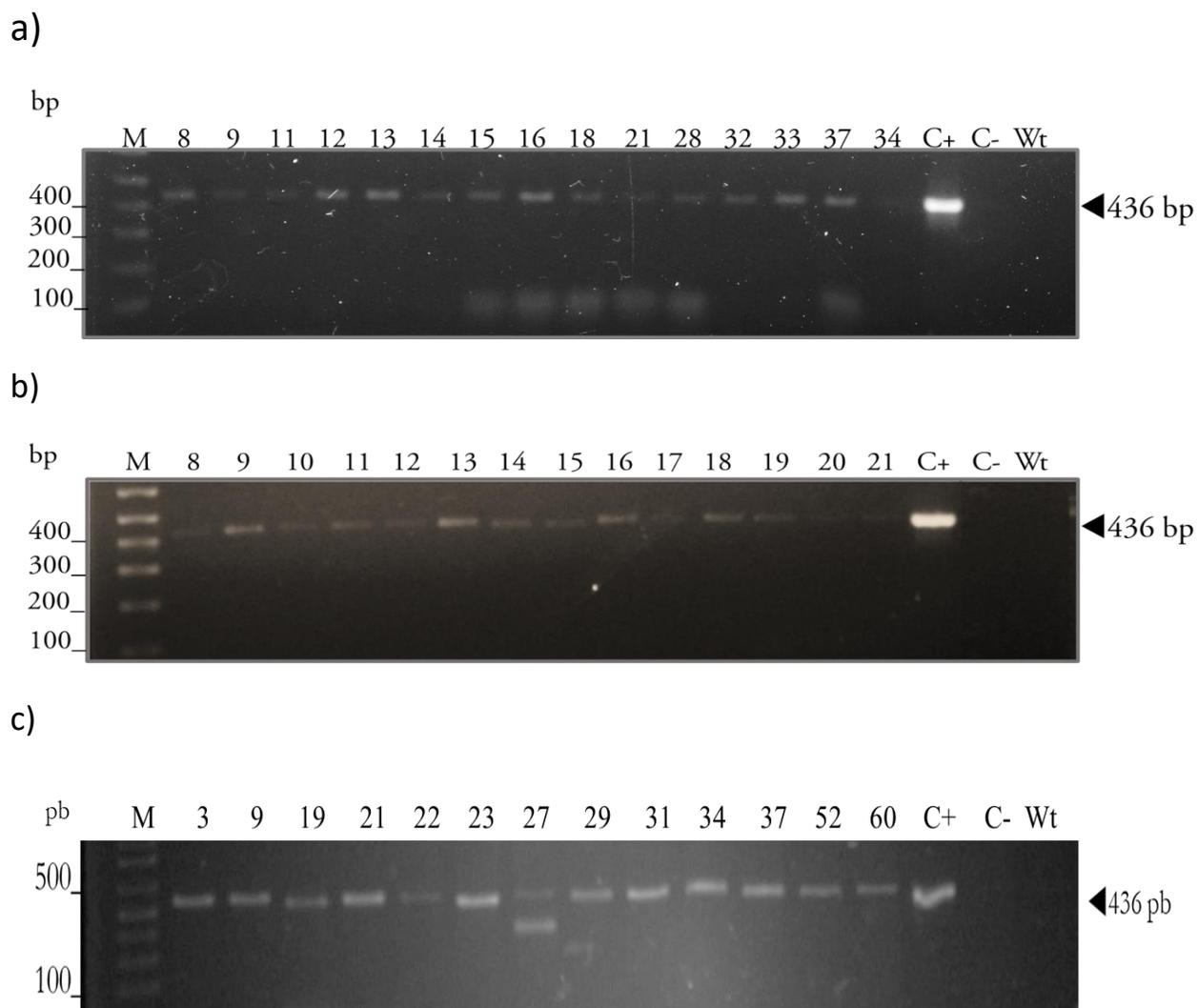


Figura 35. Detección del gen *cas9* en plantas transformadas. **A)** Perfil electroforético del gen *cas9* obtenidas del lote 7 y 8 de la generación T0. **B)** Perfil electroforético del gen *cas9* obtenidas del lote 9 y 10 de la generación T0. **C)** Perfil electroforético del gen *cas9* obtenidas del lote 11 y 12 de la generación T0.

6.11 Análisis molecular del sitio de edición

De las muestras positivas al gen *cas9* se procedió a realizar el análisis de endonucleasa T7, de acuerdo con este análisis se obtuvieron 6 muestras positivas (Figura 36).

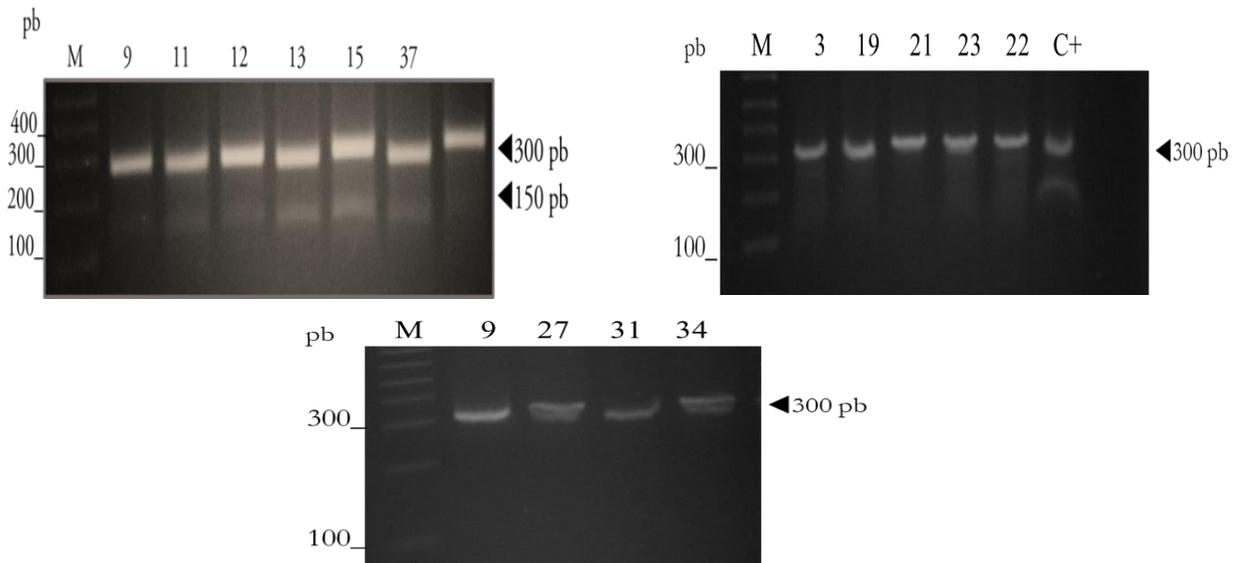


Figura 36. Análisis con endonucleasa T7. Digestión de un fragmento del fragmento del gen *PBS1* de 300 pb generando un fragmento de 150 pb indicando la presencia de una secuencia (DONOR) que no hibrida con el gen *PBS1* silvestre. La presencia de una banda de 300 pb indica la ausencia de la secuencia DONOR o alguna mutación. Las muestras de DNA se extrajeron de los lotes 7-12 de los explantes de soya.

Las muestras positivas fueron analizadas mediante amplificación del fragmento del gen *PBS1* con la modificación, como se muestra en la Figura 37.

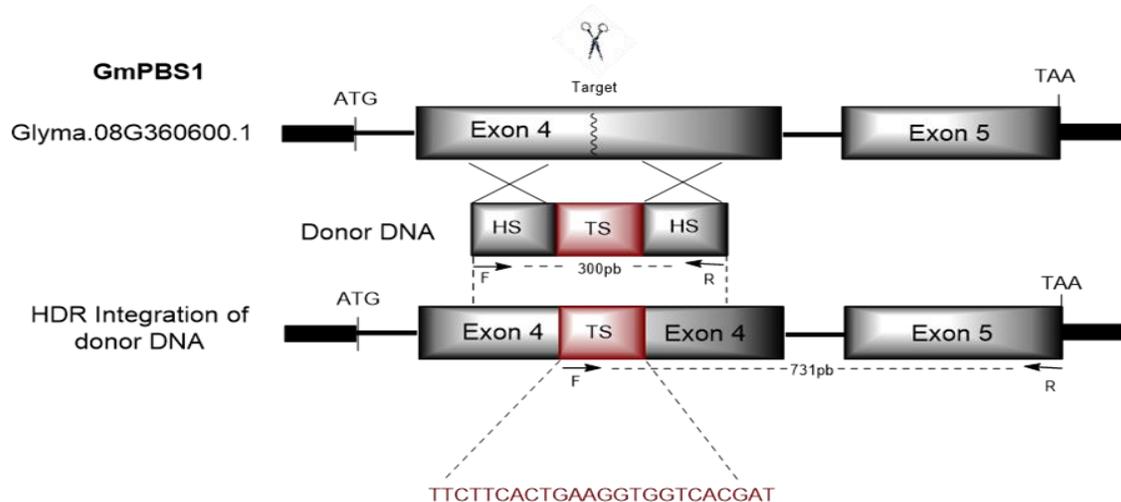


Figura 37. Esquema de edición del gen *PBS1* de soya mediante CRISPR-Cas9. Localización del sitio blanco en el exón 4 para la inserción del DNA donante HS (secuencia homóloga) en el exón 4 del gen *PBS1* TS (secuencia blanco). Las flechas indican los oligonucleótidos específicos del gen *pbs1* utilizados para la PCR.

De acuerdo con los análisis, el amplicón de 731 pb en las 6 muestras indican que contienen el sitio de edición esperado, (Figura 38), indicando la presencia del sitio de edición. Para corroborar esta observación se realizó análisis por secuenciación.

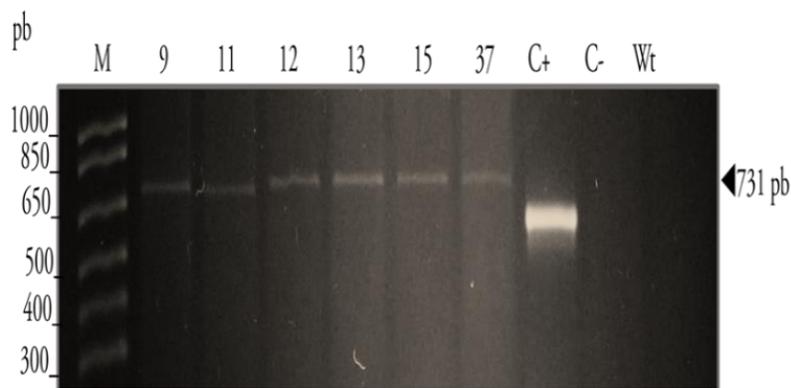


Figura 38. Análisis molecular del fragmento del gen *PBS1* con la secuencia DONOR. Perfil electroforético del fragmento del gen con la modificación, indicando la presencia de un fragmento de 731 pb. Las muestras analizadas corresponden a las muestras positivas a endonucleasa T7 provenientes de la extracción de DNA de los explantes de los lotes 7 y 8.

La secuenciación mostró que sólo la muestra 9 presentó la edición correspondiente al sitio de corte modificado en *PBS1* (Figura 39). Las 5 muestras restantes presentaron el sitio silvestre del gen *PBS1*, que puede deberse a que la

soya es una planta tetraploide, por lo que la edición pudo ocurrir en uno de los cromosomas, pero no en todos. La secuenciación de los demás alelos *PBS1* contribuirá a corroborar si este fue el caso.

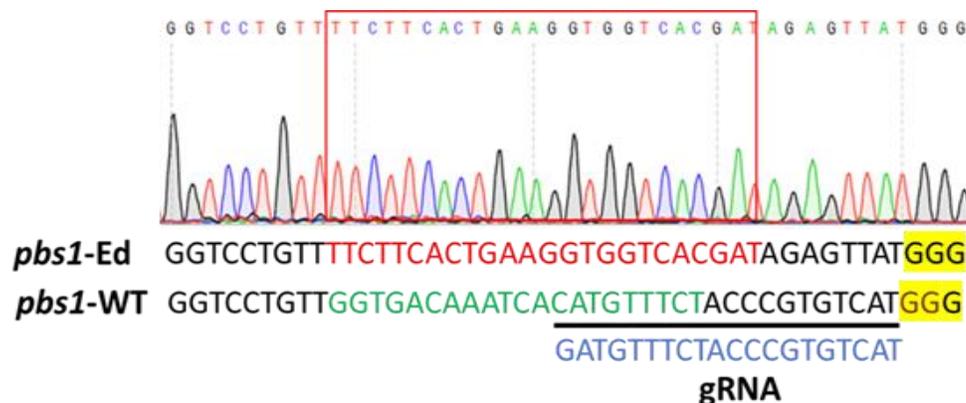


Figura 39. Análisis de secuenciación del sitio de edición en *PBS1*. Electroferograma de la región que contiene la modificación dentro del genoma de soya. El recuadro en rojo indica la secuencia DONOR. La secuencia *pbs1-Ed* indica la presencia de la edición, la secuencia *pbs1-WT* indica el genoma WT de soya. Letras en rojo indica la secuencia DONOR, letras en verde indican el fragmento génico de *PBS1*, letras en azul indican el gRNA.

Por último, las plantas transformadas de la generación T0 se mantuvieron *in vitro* hasta su enraizamiento. Posteriormente fueron trasplantadas a tierra para continuar con su desarrollo. Seis plantas de 322 alcanzaron su desarrollo completo y se obtuvo la generación T1, que fue cultivada y una vez que emergió el tejido foliar se procedió a realizar la extracción de DNA y posteriormente los análisis moleculares ya descritos. De estos análisis se obtuvo un amplicón de 436 pb correspondiente al gen *cas9* en 5 muestras (Figura 40), lo que indicó la transformación de la planta; sin embargo, al realizar la amplificación del gen *PBS1* con la modificación no se detectó. En resumen, las plantas T1 han sido transformadas, y no podemos descartar una posible edición en alguno de los alelos de *PBS1*.

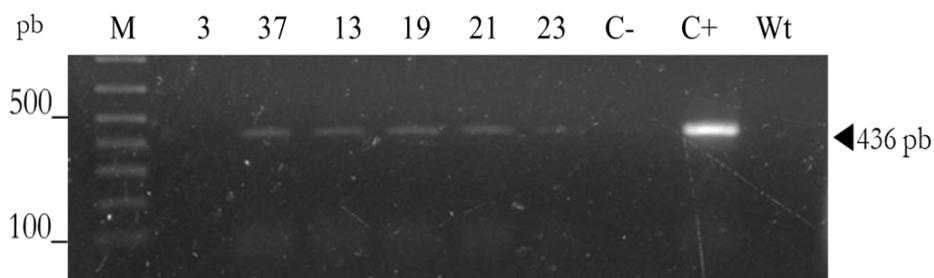


Figura 40. Detección del gen *cas9* en plantas de soya transformadas. Perfil electroforético del gen *cas9* de la generación T1. Se observa un amplicón de 436 pb correspondiente al gen *cas9*.

7. Análisis bioinformático de PBS1

El origen del sistema PBS1-RPS5 en *A. thaliana* y otras plantas no está claro. Por lo tanto, sus homólogos fueron buscados en los genomas anotados y otras bases de datos de embriofitas, briofitas, pteridofitas, algas estreptofitas y clorofitas para obtener información sobre la evolución de los genes subyacentes a este mecanismo de defensa arrojando a la luz los orígenes evolutivos de los mecanismos de defensa que implican la interacción entre un efector de patógenos y proteínas de resistencia de las plantas (Figura 41).

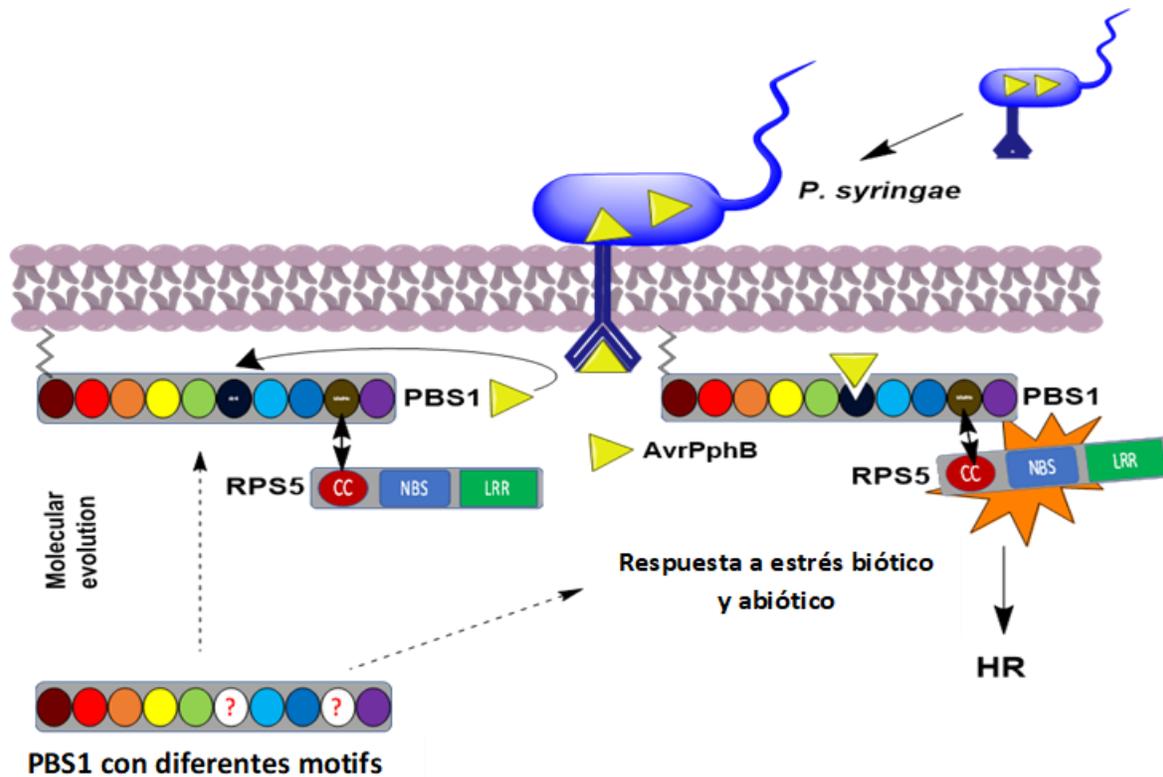


Figura 41. Representación esquemática de la activación del efector AvrPphB de *P. syringae* de RPS5 mediada por PBS1 que conduce a la respuesta de defensa. PBS1 se asocia a través de su motivo SEMPH, al dominio CC de RPS5, formando un complejo en el que este último está inactivo. Posteriormente, el efector AvrPphB (triángulo amarillo), inyectado por *P. syringae*, se une y escinde PBS1 en su motivo GDK, esto induce un cambio conformacional en RPS5, activando así la respuesta hipersensible (HR).

7.1 Identificación de proteínas PBS1, subdominios de cinasa conservados y motivos involucrados en la respuesta inmunitaria indirecta en PBS1 en el reino vegetal

Teniendo en cuenta que AthPBS1 contiene subdominios de cinasa conservados, los motivos GDK y SEMPH se utilizaron para buscar estos motivos y subdominios en proteínas homologas de PBS1 de diversos taxones de plantas. Después de eliminar las secuencias redundantes, se identificaron 881 secuencias. Con el fin de identificar las secuencias ortólogas, se realizó un análisis con la plataforma OrthoMCL, con el que se obtuvieron 881 secuencias de ortólogos

correspondientes a angiospermas, gimnospermas, briofitas, pteridofitas, licofitas, hepáticas, carofitas y clorofitas. Para obtener información sobre la evolución de los motivos y dominios antes mencionados, se realizaron tres análisis con el programa MEME. Primero, se utilizaron 116 secuencias ortólogas de PBS1 correspondientes a clorofitas; segundo, 765 secuencias ortólogas de PBS1 de plantas terrestres y finalmente 182 secuencias ortólogas de PBS1, que mostraron la mayor similitud de cada especie y género de las 765 secuencias de PBS1 analizadas.

El análisis de clorofitas mostró que estas secuencias contienen los subdominios de cinasa conservados (Figura 42). Curiosamente, los motivos GDK y SEMPH no están presentes en los ortólogos de PBS1 de clorofitas. Un análisis más detallado reveló la secuencia consenso de subdominios de cinasa en estas proteínas. Los subdominios de cinasa conservados (etiquetados con números romanos) correspondían a los subdominios previamente reportados, I: GxGxxG; II: K; III: E; VIb: DxxxxN; VII: DFG; VIII: APE; IX: DxxxxG; XI: R, algunos de los cuales se conservan en varias proteínas PBS1.

Subdominios conservados de cinasasa en ortólogos de PBS1

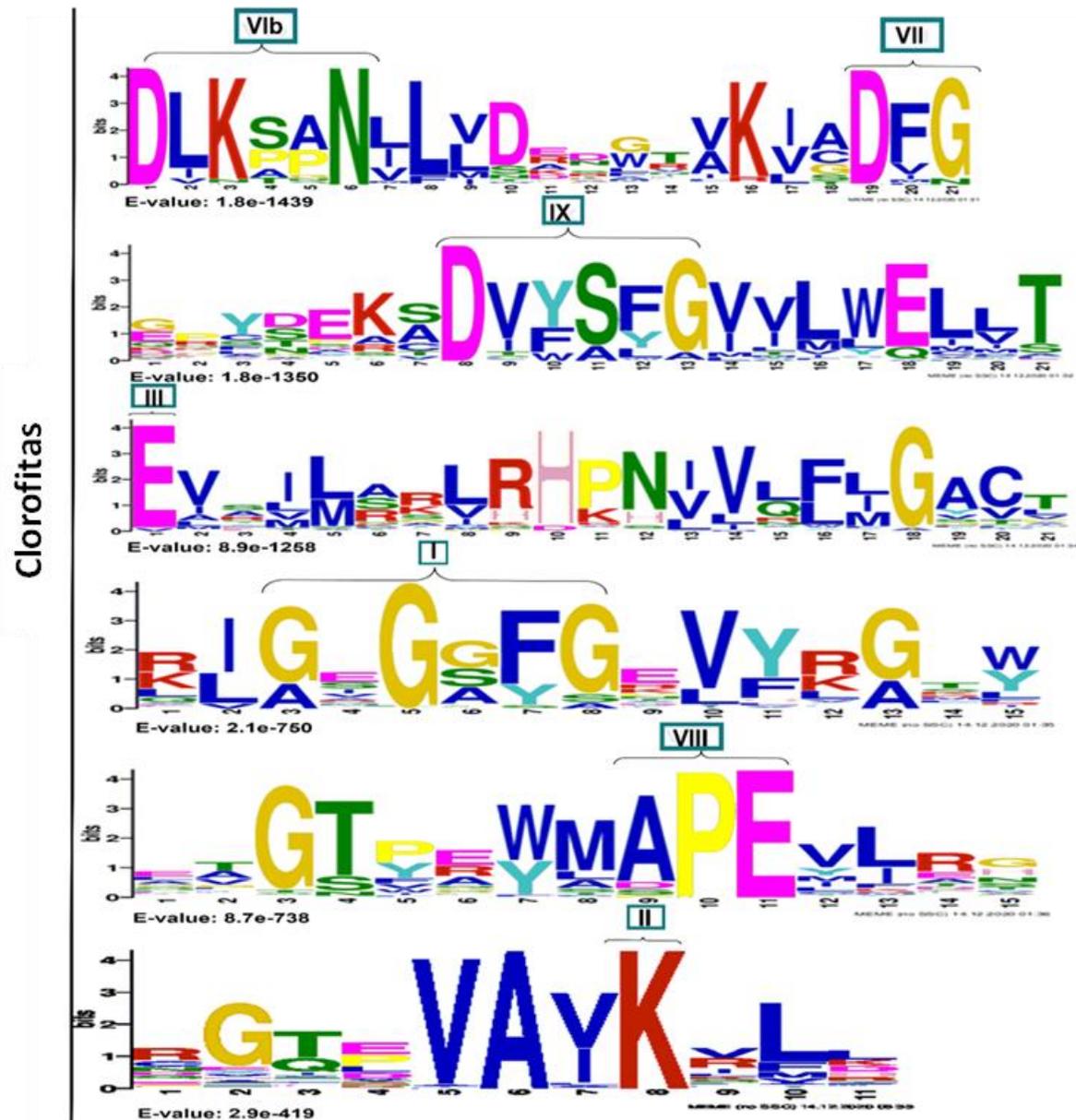


Figura 42. Análisis de subdominios de cinasasa conservados en ortólogos de PBS1 en clorofitas. La representación de logotipo de la alineación de secuencia múltiple de 116 secuencias de clorofitas con alta similitud con PBS1 de *A. thaliana* se identificó mediante análisis MEME. Los dominios están ordenados de menor a mayor de acuerdo con el valor esperado (E-value). Los subdominios de cinasasa conservados están etiquetados con números romanos. Subdominio I: GxGxxG, II: K, III: E, VIb: DxxxxN, VII: DFG, VIII: APE, IX: DxxxxG, XI: R. Los ortólogos de PBS1 en clorofitas no contienen el sitio de corte de GDK ni el motivo de reconocimiento requerido para resistencia de mediada por RPS5.

Los subdominios de cinasa, así como el motivo GDK están altamente conservado en ortólogos de PBS1 de plantas terrestres. Sin embargo, las pteridofitas y las gimnospermas carecen de estos. Se identificaron diferentes motivos en las especies analizadas, que difieren de uno a seis aminoácidos en el sitio de corte putativo en relación con el motivo canónico de AthPBS1 (Figura 43). Por un lado, secuencias de ciertas monocotiledóneas, dicotiledóneas, briofitas, licófitas, hepáticas y carófitas muestran poca variación; sin embargo, pteridofitas y gimnospermas contienen motivos completamente diferentes.

Interesantemente, las secuencias analizadas en MEME de plantas terrestres mostraron una mayor similitud entre los motivos GDKSHV y GDKTHV. El motivo SEMPH se encontró en nueve especies de dicotiledóneas. Curiosamente, diferentes motivos de reconocimiento putativos se identificaron en los ortólogos de PBS1 analizados. Los motivos de reconocimiento putativos se definieron como aquellos motivos similares en posición y composición de aminoácidos al motivo SEMPH pero con una función desconocida. Entre los motivos de reconocimiento el más conservado fue el motivo STRPH (Figura 43). El motivo GDK está presente en licofitas, hepáticas, carofitas (*Klebsormidium nitens*) y angiospermas; sin embargo, está ausente en las clorofitas. Algunas especies dentro gimnospermas, licofitas, briofitas y carofitas también carecen de este motivo (Figura 44). Los motivos de reconocimiento de resistencia son los más variables en diferentes especies. Esto puede permitir la interacción de PBS1 con una amplia gama de proteínas NBS-LRR (incluida RPS5), que a su vez generan una amplia gama de respuestas dependiendo de la naturaleza del efector.

Subdominios conservados de cinasa y ortólogos de PBS1

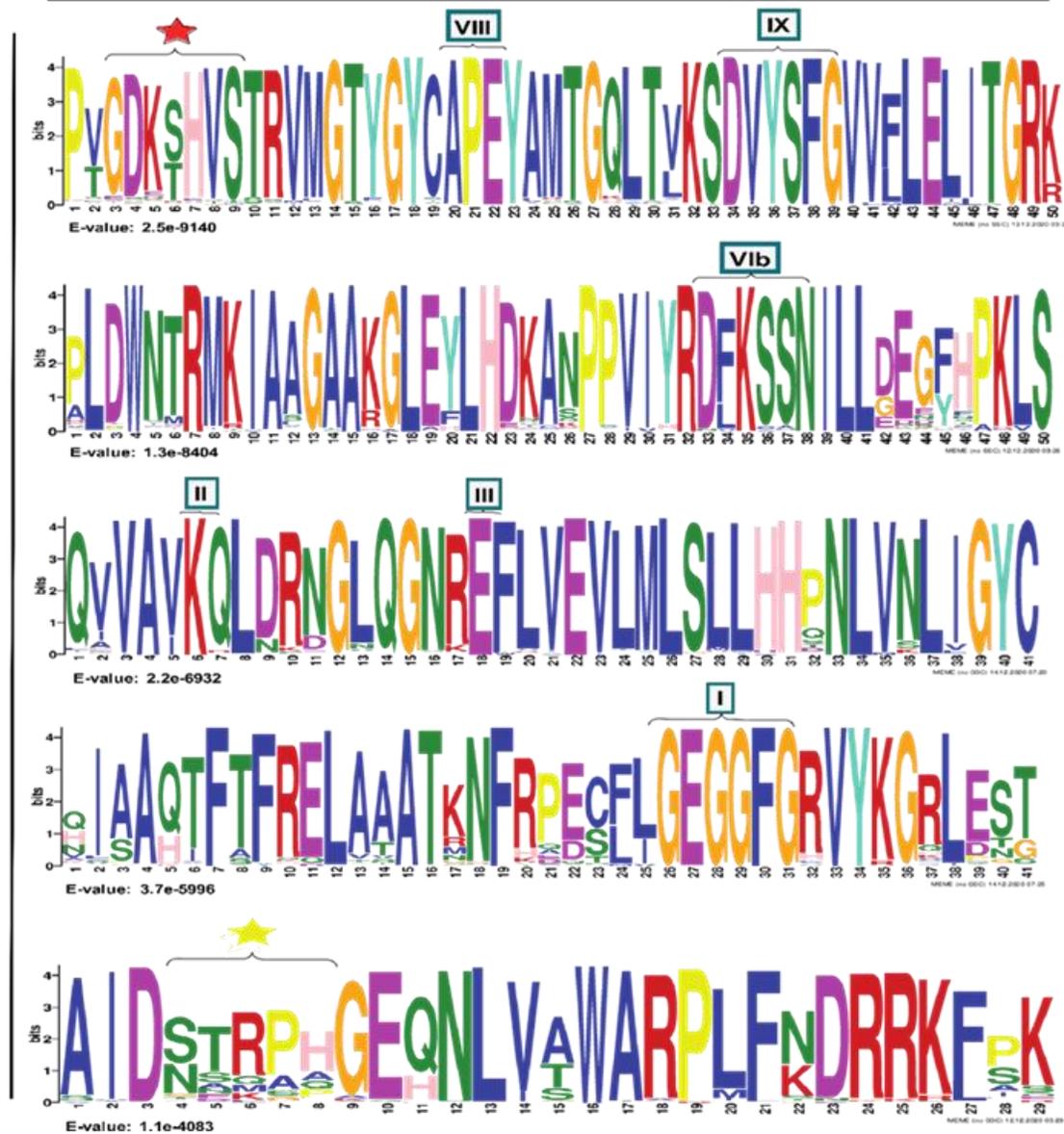


Figura 43. Análisis de subdominios de cinasa conservados y motivos importantes en ortólogos de PBS1 en plantas terrestres. Se muestra la representación del logotipo de 182 secuencias múltiples con alta similitud a PBS1 de *A. thaliana* en monocotiledóneas (37/182), dicotiledóneas (117/182), briofitas (7/182), carofitas (2/182), gimnospermas (7/182), licofitas (7/182), hepáticas (1/182) y pteridofita (4/182). Los dominios están ordenados de menor a mayor de acuerdo con el valor esperado (E-value). Los subdominios de cinasa conservados están etiquetados con números romanos. Subdominio I: GxGxxG, II: K, III: E, VIb: DxxxxN, VII: DFG, VIII: APE, IX: DxxxxG, XI: R. El sitio de corte de AvrPphB GDK está marcado con una estrella roja y el motivo de reconocimiento requerido para la resistencia mediada por RPS5 está etiquetada con una estrella amarilla.

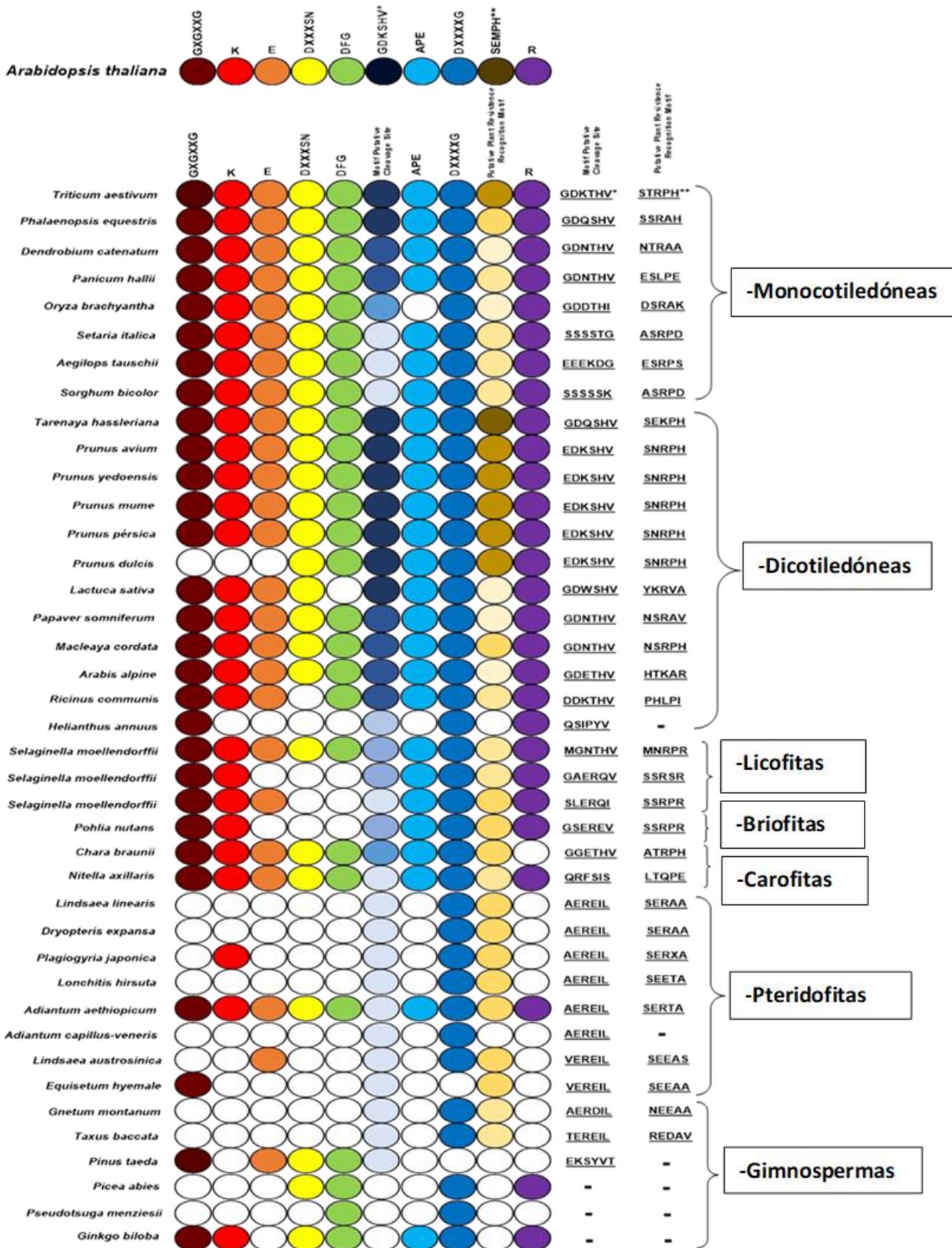


Figura 44. Comparación de subdominios y motivos de cinasa conservados en ortólogos de PBS1 en plantas terrestres. Se analizaron 40 secuencias de ortólogos de PBS1 en plantas terrestres. Las secuencias se compararon con *A. thaliana* PBS1. Las especies seleccionadas de monocotiledóneas, dicotiledóneas, briofitas, licofitas y clorofitas incluyen la mayoría de los subdominios cinasas, pero estos son más variables; en gimnospermas y pteridofitas, los subdominios cinasa y motivos están menos conservados. Las claves de color son como en la Figura 1. El sitio de corte del motivo putativo se indica con un círculo azul oscuro, el tono azul va disminuyendo si hay uno o más cambios de aminoácidos con respecto al motivo indicado en *A. thaliana*. Lo mismo se aplica al motivo de reconocimiento de resistencia putativo, pero en amarillo oscuro. Los círculos blancos indican la ausencia de un subdominio o motivo. *, ** Los motivos se toman como referencia para el análisis. +, ++ Motivos con mayor abundancia en diferentes especies.

7.2 Árbol filogenético de ortólogos de PBS1 en clorofitas

Ciento dieciséis secuencias de aminoácidos de especies clorofitas con alta similitud a AthPBS1 (*A. thaliana* ID Q9FE20) se recuperaron de diferentes bases de datos. Así, las clorofitas *Dunaliella salina*, *Chlamydomonas reinhardtii*, *Volvox carteri*, *Micromonas sp.*, *Micromonas pusillis*, *Ostreococcus luci marinus*, *Auxenochlorella protothecoides*, *Chlorella sorokiniana*, *Gonium pectorale*, *Raphidocelis subcapitata*, *Micractinium conductrix*, *Bathycoccus prasinos*, *Chlamydomonas eustigma*, *Trebouxia sp.*, *Chloropicon primus*, *Scenedesmus sp.* y *C. subellipsoidea* contienen genes que codifican para cinasas con ortología para AthPBS1. Interesantemente, una tirosina cinasa de *Coccomyxa subellipsoidea* y dos proteínas hipotéticas de *C. eustigma* y *Trebouxia sp.*, respectivamente, se agruparon con AthPBS1 (A003) en el mismo clado (B) con considerable similitud a AthPBS1 (Figura 45).

La relación filogenética entre AthPBS1 y ortólogos de clorofitas se analizó utilizando el método de máxima verosimilitud, ejecutado en IQ-TREE (Figura 45). La filogenia resultante muestra cuatro clados principales (A, B, C y D) agrupados con especies relacionadas de clorofitas y proteína cinasas o grupos estrechamente relacionados con altos valores de Bootstrap en las ramas. Un análisis preliminar identificó subdominios de cinasas GxGxxG, K, E, DxxxSN, DFG, APE, DxxxG y R en los ortólogos de PBS1 de clorofitas (Figura 42).

Además, AthPBS1 se agrupa en el clado B (Figura 45) con diferentes proteínas cinasas que contienen subdominios de cinasas (Hanks et al., 1995). PBS1 de *C. subellipsoidea* alberga dominios de cinasa clásicos y se agrupa en el mismo clado que AthPBS1 (Figura 45, clado B). Estos resultados indican que diferentes cinasas ortólogas de las clorofitas están relacionadas filogenéticamente con base en su alta similitud. Este descubrimiento apoya la hipótesis de que PBS1 tiene contrapartes evolutivas en las clorofitas, sin embargo, los motivos SEMP y GDK estaban ausentes en los ortólogos de clorofito PBS1. Por tanto, los subdominios de estas proteínas cinasas presentan sinapomorfia, ya que estos caracteres son intrínsecos a las proteínas cinasas; sin embargo, los motivos presentes en AthPBS1 pueden pertenecer a un carácter homoplástico ya que podría haber sido adquirido de forma independiente. Dada esta perspectiva, se analizaron ortólogos AthPBS1 en plantas terrestres.

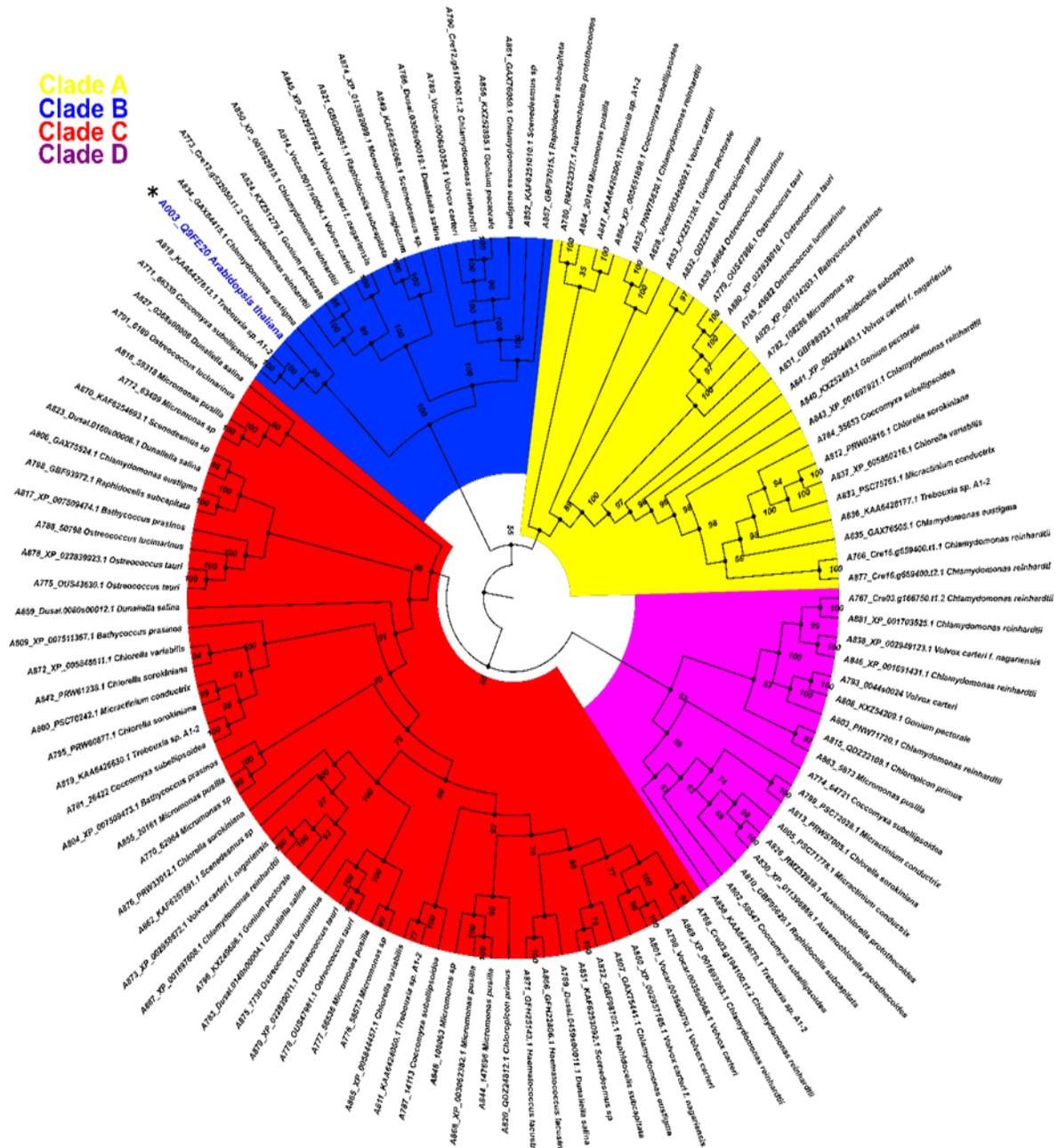


Figura 45. Relación filogenética entre PBS1 de *A. thaliana* y proteína cinasas ortólogas de clorofitas. Se analizaron un total de 116 secuencias de aminoácidos con alta similitud a AthPBS1 de diferentes bases de datos que se alinearon usando ClustalW. La filogenia se construyó utilizando máxima verosimilitud basándose en las secuencias de aminoácidos deducidas de ortólogos de PBS1 de diferentes especies de clorofitas. Los diferentes colores indican clados para diferentes ortólogos de cinasas. *A. thaliana* se incluye en el Clado B indicado con asterisco y color azul.

7.3 Análisis filogenético de ortólogos de PBS1 de diferentes linajes de plantas

Para explorar la relación evolutiva de los ortólogos de PBS1 en los principales dominios vegetales, se construyó una filogenia con 881 secuencias ortólogas de PBS1 con considerable semejanza con AthPBS1. Sin embargo, los resultados obtenidos no fueron concluyentes, porque la cantidad de politomías observadas dificultaba la resolución de sus relaciones filogenéticas. Por tanto, se utilizaron menos secuencias para construir otra filogenia y obtener un árbol optimizado. Finalmente, el análisis filogenético se realizó con 182 secuencias ortólogas de PBS1 utilizando máxima verosimilitud. Los resultados mostraron cinco clados principales (A, B, C, D y E); el clado A se subdivide en diferentes subclados; ortólogos de PBS1 en dicotiledóneas se agruparon en el clado A, que contenía AthPBS1 (A003) (Figura 46). El clado B estaba formado por ortólogos de PBS1 de monocotiledóneas, dicotiledóneas, licofitas (A295, A316), briofitas (A148, A248), gimnospermas (A344, A134, A215 y A343) y carofitas (A142). El clado C se subdivide en dos clados que incluyen solamente monocotiledóneas y dicotiledóneas. De manera interesante, una rama de *Chara brauni* (A133) surge de los tres clados analizados, sugiriendo una posible diversificación de un antepasado basal de este carófito. Por otro lado, el clado D también se subdivide en dos subclados, un grupo de ortólogos PBS1 de *Physcomitrella patens* (A164, A064, A352, A028, y A510), *Selaginella moellendorffii* (A071 y A313), *Sphagnum fallax* (A162) y *Marchantia polymorpha* (A058), mientras que el otro incluía dicotiledóneas (A758, A288 y 755), licofitas (A029 y A302) y gimnospermas (A229 y A135). Finalmente, las pteridofitos (A498, A146, A053, y A149) y una gimnosperma (A140) se agruparon en el clado E. Estos resultados indican que diferentes proteínas ortólogas de PBS1 están relacionadas filogenéticamente. Sin embargo, los clados B y D contienen proteínas de diversos taxones, lo que sugiere una evolución convergente dentro estos grupos.

- Dicotiledóneas
- Monocotiledóneas
- Gimnospermas
- Licofitas
- Pteridofitas
- Hepática
- Briófitas
- Carofitas

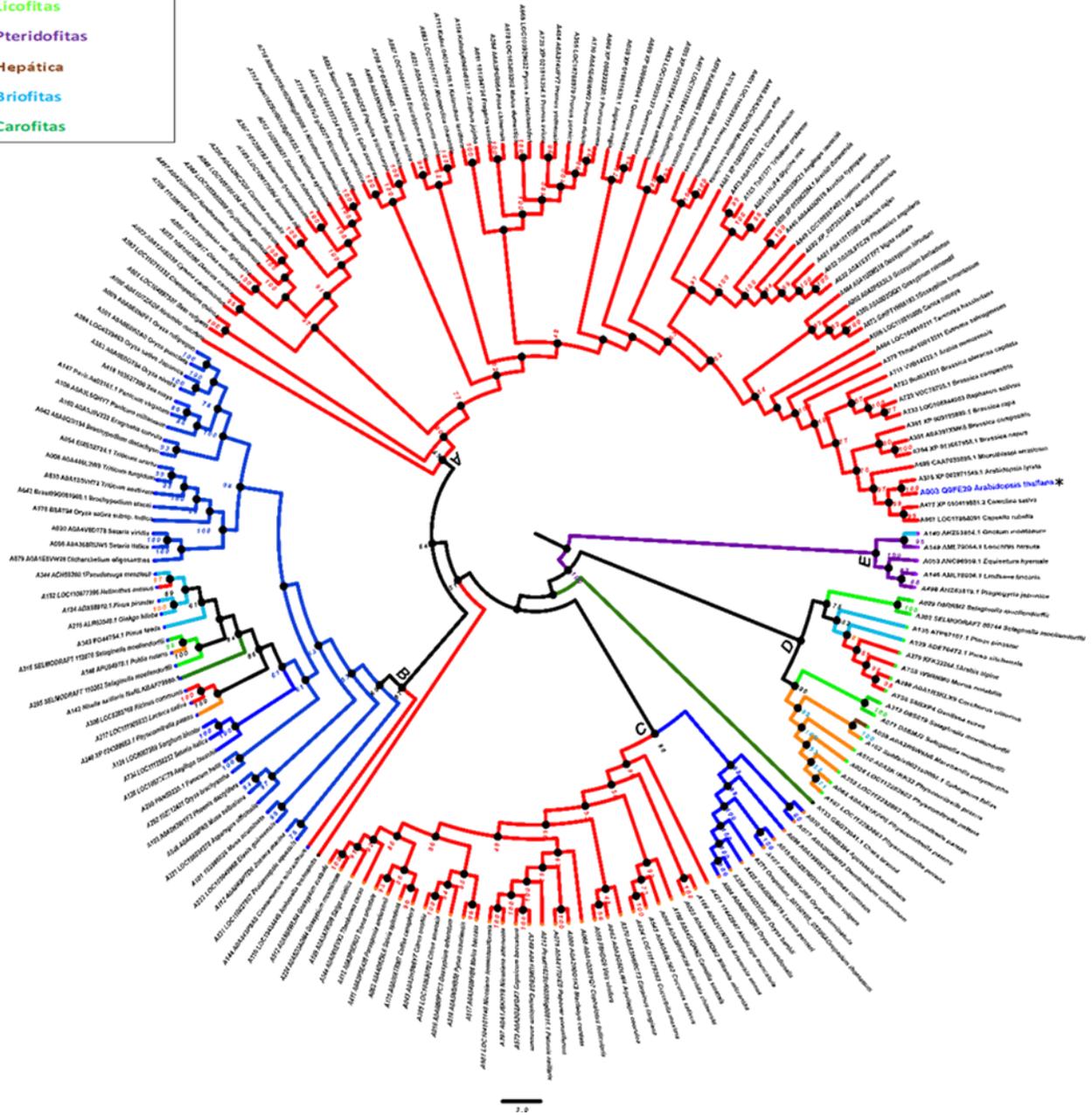


Figura 46. Análisis filogenético de ortólogos de PBS1 en plantas terrestres. Se realizaron búsquedas de 182 secuencias de aminoácidos ortólogos a AthPBS1 en diferentes bases de datos y se alinearon mediante ClustalW. El árbol filogenético de máxima probabilidad se construyó basándose en las secuencias de aminoácidos obtenidas de las proteínas PBS1 de varias plantas terrestres. El rojo (clado A) indica dicotiledóneas; azul (clado B y C), monocotiledóneas; púrpura (clado E), pteridofitas; naranja (clado B y D), briófitas; verde oscuro (clado B), carofitas; azul claro (clados B, D y E), gimnospermas; licofitas, verde (clado B y D) y hepáticas, marrón (clado D). *A. thaliana* se indica con asterisco y en color azul.

7.4 Reconstrucción ancestral de ortólogos de PBS1 en plantas terrestres

Para conocer mejor la evolución de PBS1 en plantas terrestres, se llevó a cabo su reconstrucción ancestral. En este sentido, se utilizaron 182 ortólogos de PBS1 para inferir la filogenia con el método bayesiano. La reconstrucción del estado ancestral se obtuvo mediante el método de máxima parsimonia. Con base en estos resultados, se investigó la filogenia de plantas terrestres. La filogenia (Figura 47) muestra que los ortólogos de PBS1 de pteridofitas forman una rama que se origina en el primer nodo ancestral (círculo morado). A continuación, una serie de ancestros se originó a partir de diferentes ramas, que se agruparon con diferentes especies de estreptofitos (carofitas, briofitas, hepáticas, licofitas, gimnospermas, monocotiledóneas y dicotiledóneas).

El segundo nodo ancestral (círculo azul claro) dio lugar a una rama que contiene únicamente una gimnosperma (A343), y otra que incluye un tercer nodo ancestral (círculo verde oscuro). Esto incluye una rama que contiene un carófito (A142), así como un cuarto nodo ancestral (círculo negro), que divergió en diferentes ortólogos de PBS1 de plantas terrestres, excepto pteridofitas. Finalmente, un quinto nodo ancestral (círculo rojo) dio lugar a diferentes clados que contienen exclusivamente dicotiledóneas, así como una especie de monocotiledónea (A290) (Figura 47).

Para dilucidar el posible origen de los motivos GDKSHV y SEMPH, la reconstrucción de carácter ancestral se realizó para cada posición de aminoácido. Se analizaron las posiciones 245 a 250 para el motivo GDKSHV y 296 a 300 para el motivo SEMPH. La reconstrucción evolutiva reveló un solo cambio en los estados ancestrales de los aminoácidos en las posiciones 245 (carácter ancestral glicina), 246 (carácter ancestral aspartato), 247 (carácter ancestral arginina, divergió a lisina), 249 (carácter ancestral leucina, divergió a histidina) y 250 (carácter ancestral lisina, divergió a valina). Sin embargo, el cambio más relevante

ocurrió en la posición 248, siendo el aminoácido ancestral la glicina (magenta) que más tarde divergió en treonina (rojo oscuro) y serina (amarillo) en diferentes especies (Figura 48).

Estos resultados concuerdan con el análisis realizado con MEME (Figura 43). Las distribuciones ancestrales del motivo SEMPLH fueron las más diversas. El carácter ancestral en la posición 296 era arginina (violeta claro); este aminoácido divergió principalmente para serina (amarillo) y asparagina (rojo). En la posición 297, el carácter ancestral tiene dos orígenes, alanina y valina; sin embargo, la alanina estaba más conservada. La evolución de este carácter divergió principalmente a treonina (rojo intenso). Además, un clado en el que se encuentra *A. thaliana* divergió al carácter ancestral glutamato (gris) (Figura 49). El carácter ancestral en la posición 298 tiene dos orígenes, isoleucina (azul) y arginina. (violeta claro), conservado en la mayoría de las especies. Además, un clado divergió al carácter ancestral metionina (azul claro). La arginina es el principal carácter ancestral en la posición 299 y divergió principalmente a prolina (verde claro). Finalmente, el carácter ancestral en la posición 300 era glutamato y divergió a histidina (púrpura). Este análisis sugiere que sustituciones en determinadas posiciones del motivo SEMPLH se han favorecido a lo largo de la evolución. Esto probablemente fue el resultado de la selección de aquellas variantes de PBS1 que mostraron una mayor capacidad para provocar una respuesta de defensa a través de RPS5 o proteínas funcionalmente análogas.

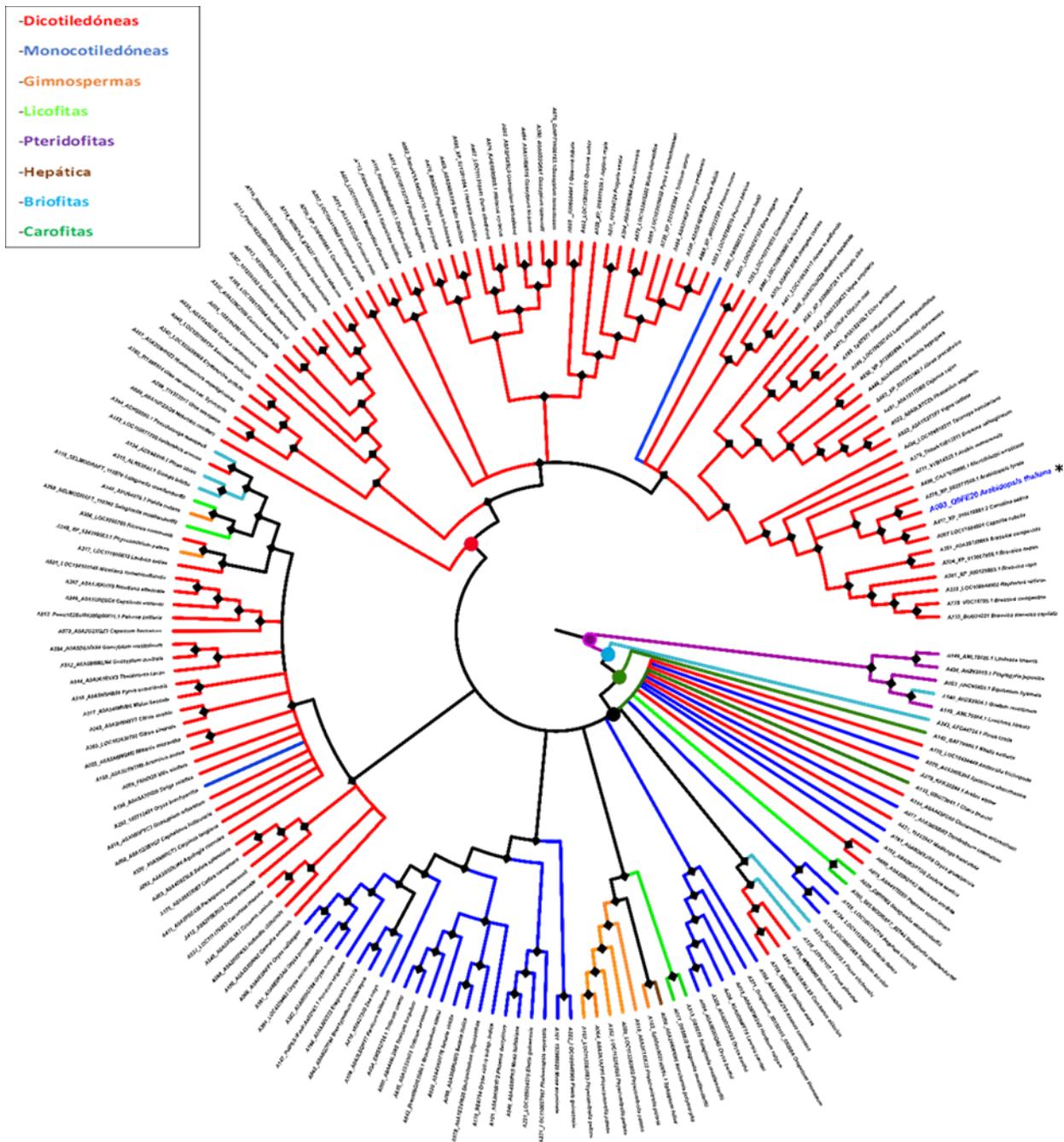


Figura 47. Reconstrucción de secuencias de ortólogos ancestrales de PBS1 en plantas terrestres. El árbol ancestral fue inferido por estadísticas bayesianas usando 182 secuencias de monocotiledóneas (azul), dicotiledóneas (rojo), briófitas (naranja), hepática (marrón), licofita (verde claro), gimnosperma (azul claro), carófitas (verde oscuro) y pteridofitas (violeta). Los círculos en los nodos indican principalmente antepasados. *A. thaliana* se indica con asterisco y color azul. El análisis de secuencias ancestrales de PBS1 se obtuvo con el software Mesquite (<http://www.mesquiteproject.org>).



Figura 48. Reconstrucción del estado ancestral en ortólogos de PBS1 basados en el motivo GDKSHV. La reconstrucción de ancestros se realizó con el algoritmo de máxima parsimonia utilizando como carácter ancestral la secuencia de aminoácidos PBS1. La reconstrucción se llevó a cabo en el aminoácido 248 de PBS1. Los colores en las ramas indican la divergencia evolutiva del aminoácido ancestral. El color en los nombres indica especies de plantas terrestres. Las transiciones la reversión se indican con asteriscos sobre los nodos. Esta figura muestra el estado ancestral de la serina (S). *A. thaliana* se indica con un asterisco y un rojo resaltado.

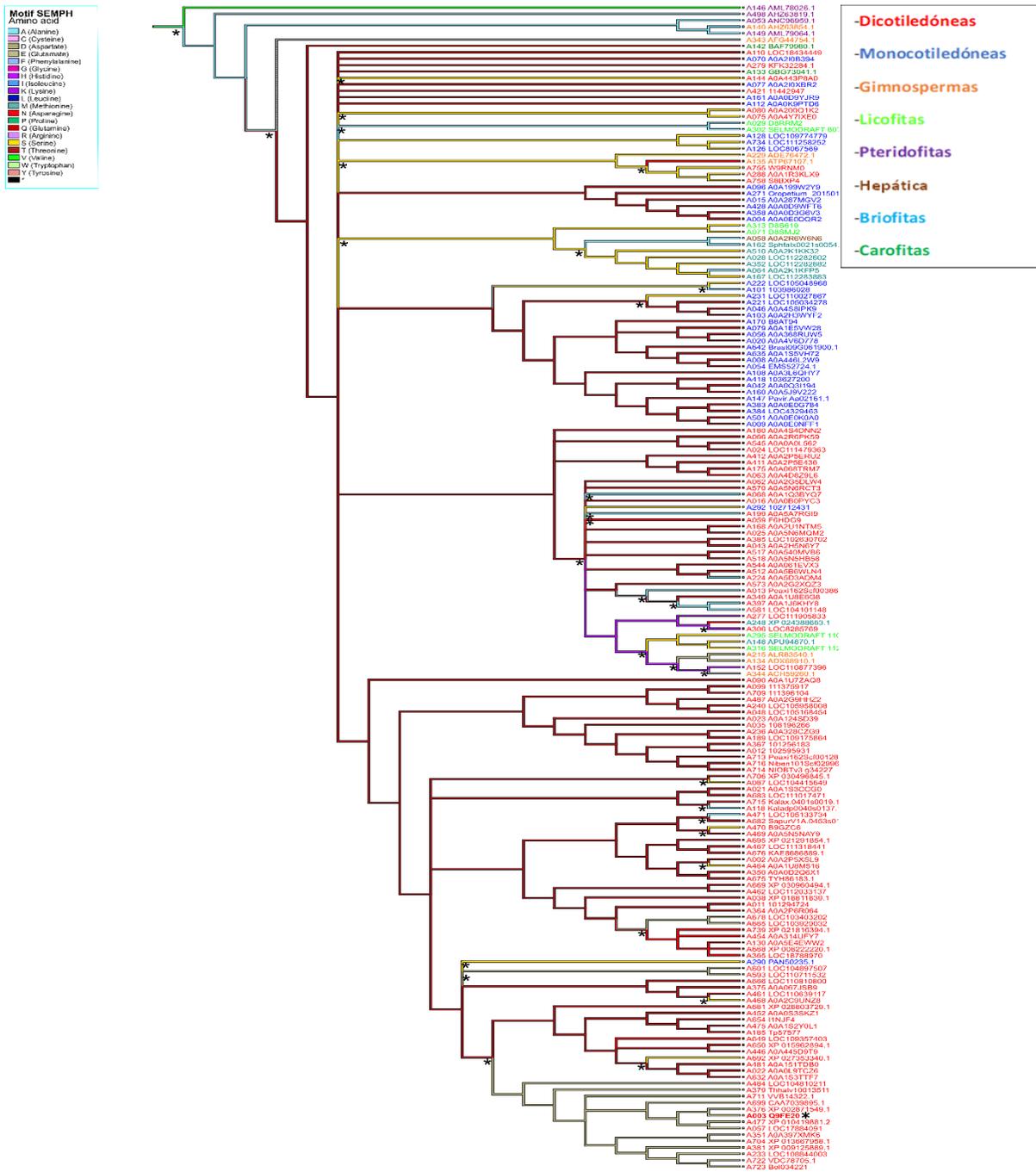


Figura 49. Reconstrucción del estado ancestral en ortólogos de PBS1 basados en el motivo SEMPH. La reconstrucción de ancestros se realizó con la máxima parsimonia utilizando como carácter ancestral la secuencia de aminoácidos de PBS1. Reconstrucción del estado ancestral en la posición del aminoácido 297 de PBS1. El color de la rama indica la divergencia evolutiva del aminoácido ancestral. Los números de acceso se muestran en rojo. Las transiciones y reversiones se indican con asteriscos sobre los nodos. Esta figura muestra el estado ancestral del ácido glutámico (E). *A. thaliana* se indica con asterisco y rojo resaltado.

7.5 Modelado de estructuras proteicas de ortólogos de PBS1

Para obtener más información sobre la conservación de los motivos GDK y SEMP, se obtuvieron modelos de la estructura de los ortólogos de PBS1 de *A. thaliana*, *T. aestivum*, *G. max*, *Spirodela polyrhiza*, *P. patens*, *C. reinhardtii*, *Chara braunii*, *Pinus pinaster* y *Lindsaea linearis*. Estos fueron seleccionados porque su función es conocida (AthPBS1 y TaPBS1), o por mayor similitud en los motivos. Además, los ortólogos de taxones filogenéticamente distantes fueron seleccionados. Según estos modelos, su estructura consiste principalmente en alfa hélices y estructuras de láminas beta (Figura 50).

Los residuos correspondientes al motivo GDK en AthPBS1, TaPBS1, GmaPBS1, SpipoPBS1-like, PpPBS1-like, y PpiHP se ubicaron dentro de un bucle (Figura 50A-E, H, bucle verde). ChbHP contiene el motivo GGETHV dentro del bucle (círculo verde). Además, no se encontró ningún motivo de corte en *C. reinhardtii* y *L. linearis* PBS1 (Figura 50F, I). Los residuos LGAPSG (*C. reinhardtii*) y AERIL (*L. linearis*) estaban en la misma posición que el motivo del sitio de corte predicho en otros ortólogos de PBS1.

Los motivos de reconocimiento de resistencia adicionales, como SEMP, STRP, STQP, ATRP y NNRAA en AthPBS1, TaPBS1, GmaPBS1, ChbHP y PpiHP, respectivamente, fueron identificados. De acuerdo con las estructuras predichas, estas secuencias están expuestas en un bucle. (Figura 50 A, C, G, H). Curiosamente, en *A. thaliana*, *T. aestivum*, *G. max* y *Ch. braunii*, los motivos analizados están próximos entre sí. Además, en SpipoPBS1-like, PpPBS1-like, y CrePKU-box de *S. polyrhiza*, *P. patens* y *C. reinhardtii*, respectivamente, contienen los motivos de reconocimiento lo que concuerda con el análisis. Por el contrario, en la misma posición donde el motivo de reconocimiento está presente en otras proteínas de PBS1, los residuos NARAA, NSRAA y QVSV (forman una hélice alfa) fueron encontrados (Figura 50 D-F). Sin embargo, en *P. pinaster*, el motivo NNRAA se ubicó dentro de un bucle (Figura 50 H).

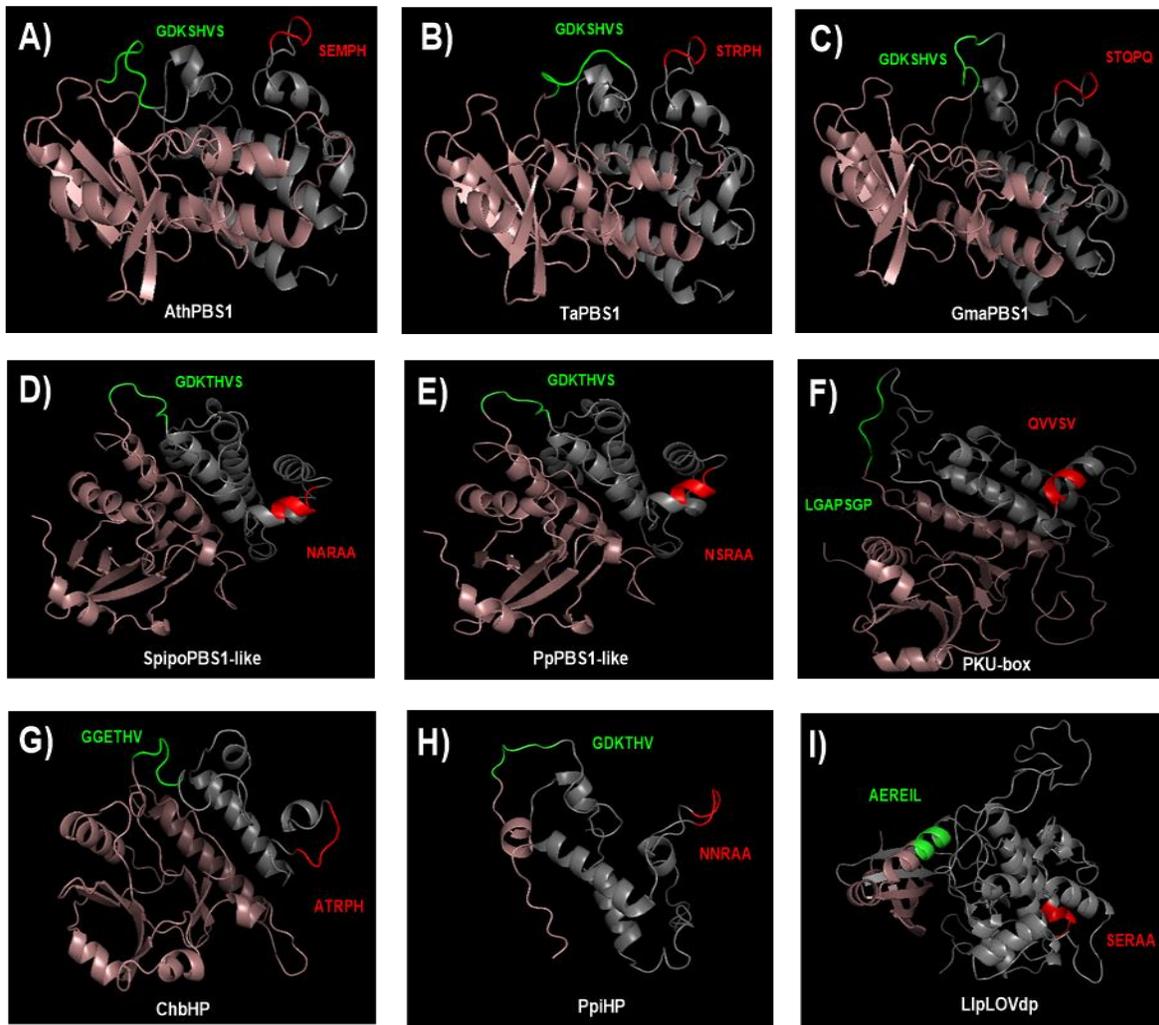


Figura 50. Estructuras predichas de PBS1 de varias especies de plantas. Se generaron estructuras para ortólogos de PBS1 predichas con Swissmodel. A) AthPBS1 de *A. thaliana* (residuos de aminoácidos 71 a 366); B) TaPBS1 de *T. aestivum* (residuos 89 a 386); C) GmaPBS1 de *G. max* (residuos 84 a 381); D) similar a SpipoPBS1 de *S. polyrhiza* (residuos 72 a 356); E) similar a PpPBS1 de *P. patens* (residuos 71 a 361); F) CrePKU de *C. reinhardtii* (residuos 474 a 786); G) ChbHP de *Ch. braunii* (residuos 16 a 249); H) PpiHP de *P. pinaster* (residuos 2 a 123); I) LlpLOVdp de *L. linearis* (residuos 346 a 689). Los modelos para AthPBS1, TaPBS1, GmaPBS1 y ChbHP se basaron en el modelo estructural BIK1 (SMTI ID: 5tos.1). Los modelos para CrePKU y PpiHP se basaron en PK (SMTI ID: 6j5t.1). Los modelos para SpipoPBS1, PpPBS1 y LlpLOVdp se basaron en BAK1 (SMTI ID: 3tl8.1), BR1 (SMTI ID: 5lpz.1) y BAC (SMTI ID: 6npz.1). A). El sitio de corte (GDKSHV) y los sitios de escisión putativos están representados en verde. El motivo de reconocimiento de RPS5 (SEMPH) y los motivos de reconocimiento de RPS5 putativos están representados en rojo. La estructura del extremo N hasta el motivo del sitio de corte putativo se muestra en rosa. La estructura de los motivos de reconocimiento putativos de RPS5 hasta el extremo C está representada en gris. BIK1, cinasa 1 inducida por Botrytis; BAK1, receptor cinasa 1 asociado a BRASSINOSTEROID INSENSITIVE 1; BR1, proteína BRASSINOSTEROID INSENSITIVE 1; PK, proteína de la superfamilia de proteínas cinasas; BAC, RAC-alfa serina / treonina-proteína cinasa. SMTI ID: número de acceso Swissmodel.

8. Discusión

La soya es considerada uno de los cultivos más nutricionales a nivel mundial, por su alto contenido de aceite y proteínas (Rosas y Young, 1989). La soya se ha convertido en un cultivo importante a nivel mundial, existe un gran interés en la mejora genética de este cultivo a través de la generación de nuevas variedades que sean capaces de enfrentar tanto estrés biótico y abiótico, cumpliendo con las demandas de la agroindustria. La soya se convierte en uno de los principales objetivos para generar cultivos modificados genéticamente mediante herramientas moleculares como CRISPR-Cas. Sin embargo, la transformación genética muestra una eficiencia muy baja. Estudios recientes han implementado la combinación de dos métodos de transformación, por biobalística y mediada por *A. tumefaciens* en el que se ha observado una eficiencia similar a protocolos anteriores en soya, sin embargo, no genera quimeras vegetales y aumenta la frecuencia de obtención de plantas transgénicas (Paes de Melo et al., 2020).

El sistema CRISPR-Cas9 se ha utilizado con éxito en plantas, como *A. thaliana*, *Nicotiana benthamiana*, arroz, tabaco, sorgo, trigo, maíz y naranja. Este sistema ha revolucionado la capacidad de modificar el genoma de las plantas, sin embargo, la soya es tetraploide y suele considerarse difícil de transformar. La mayoría de las ediciones realizadas en plantas de soya son somáticas y no son heredables cuando se utilizan promotores constitutivos (Zheng et al., 2020). La transformación de soya en raíces ha servido como un modelo rápido para el estudio del sistema CRISPR-Cas9 lo que supone un ahorro de tiempo, menor carga de trabajo, para generar plantas transformadas de soya (Do et al., 2019).

El sistema CRISPR/Cas9 puede facilitar el desarrollo de aplicaciones adicionales en la soya por de la reparación dirigida por homología (HDR) que utiliza secuencias donantes de DNA de doble cadena (dsDNA), por ejemplo, para crear inserciones genéticas dirigidas o sustituciones alélicas (Harrison et al., 2014).

En el presente estudio, se empleó el vector PBUN411U6 que contiene el gen del promotor U6 de *A. thaliana*, especie dicotiledónea como la soya, para dirigir la expresión de un gRNA de 20 nt dirigido contra el gen *PBS1* de esta última. Aunque la edición de genomas en soya enfrenta varios retos ha demostrado su viabilidad. Utilizando el sistema CRISPR-Cas9 se facilita el desarrollo de nuevas aplicaciones como la inserción genética dirigida utilizando secuencias donantes (DONOR) así como la proteína Cas9, para generar la edición deseada en soya.

La tecnología de transformación en soya se ha quedado atrás, ya que sigue dependiendo del genotipo y un número limitado de genotipos son transformables (Pareddy et al., 2020; Yamada et al., 2012). Por lo tanto, el número de plantas transgénicas de soya suele ser baja en los procesos de transformación. Interesantemente diferentes estudios han demostrado éxito en la transformación estable en la edición de genes de soya (Cai et al., 2018; Li Z. et al., 2015). Sin embargo, las tasas de mutación son bajas en plantas de soya transgénicas.

En este estudio, empleando un gRNA dirigido al gen *PBS1* de soya, se realizó una evaluación a nivel molecular de los embriones transformados por biobalística. Encontramos transformación y edición una vez que las condiciones de transformación se modificaron, considerando una mayor relación molar del vector de transformación como del vector DONOR, así como reducir la distancia de disparo y realizar cambios de temperatura (ver resultados). Interesantemente, empleando el sistema CRISPR-Cas9 encontramos un porcentaje de transformación del 4.2%-6.9% y de hasta 2.1% de edición equivalente a 3 plántulas de soya mediante detección del gen *cas9* y del fragmento editado de *PBS1* con la modificación indicando la mutación del gen blanco en la soya. Para la generación T1 sólo se obtuvieron 5 plántulas positivas al gen *cas9* indicando la transformación de la planta. De acuerdo con estudios anteriores la eficiencia fluctúa entre 2.5% y 4% empleando nodos cotiledonares, donde la inducción de brotes es el paso limitante (Paz et al., 2006) lo que pudo repercutir en una baja eficiencia de regeneración en T0, limitando el análisis en la generación T1.

Usando protocolos recientes que combinan la transformación mediada por *Agrobacterium* y biobalística demuestran una alta reproducibilidad tanto en generación como en eficiencia de transformación con una media de 9.8%. Además, algunos protocolos que usan *Agrobacterium* fluctúa entre 2 y 10% de acuerdo con el genotipo de soya, siendo similar la frecuencia de transformación que se obtiene mediante biobalística (Li S. et al., 2017; Meurer et al., 1998; Olhoft y Somers, 2001). Esto indica que se obtuvieron eficiencias elevadas de transformación y una baja eficiencia de regeneración similar a lo reportado.

Los análisis moleculares obtenidos indican la presencia de *cas9* en los explantes transformados, sin embargo, no se evaluó la forma en que las quimeras se trasladarían a la generación T0 o segregarían a la generación T1. Los eventos de HDR fueron quiméricos ya que las plantas T1 perdieron la inserción HDR de acuerdo con los análisis de secuenciación. La frecuencia de HDR revelada por PCR específica del sitio de edición es de 2.1%. La mayoría de los eventos positivos no lograron regenerar plantas T0, probablemente debido a su composición celular quimérica. La probabilidad de recuperar plantas con alguna modificación podría haber sido mayor de haberse regenerado más plantas T0 por evento. Seis plantas T0 regeneradas fueron positivas a la modificación empleando el sitio específico de la edición para realizar la PCR; esto se corroboró con análisis de T7 endonucleasa. Por último, los análisis de secuenciación comprobaron la presencia de la edición en una de las plantas T0, sin embargo, en las demás plantas no se observó dicha modificación, lo que puede ser debido a que la modificación del gen *PBS1* puede estar presente en uno de los 3 cromosomas que contienen este gen en soya. No es claro en cuál de los cromosomas ocurrió el evento de edición inducido por Cas9-RNAg. Podrían ocurrir mutaciones simultáneas si hay secuencias homólogas al blanco como *PBS1*, lo que puede favorecerse mediante el diseño de un único gRNA que se dirija a dos o más genes que compartan la secuencia objetivo común o bien el uso de dos o más gRNA para modificar de manera simultánea diferentes genes (Li J.-F. et al., 2013). Interesantemente mediante análisis moleculares se logró identificar el gen *PBS1*

de soya con la modificación, sin embargo, no conocemos en qué cromosoma se encuentra dicha modificación.

Se comprobó, por tanto, que la metodología de transformación presentó resultados positivos con un porcentaje similar a otros estudios que va del 2 al 10% (Li S. et al., 2017). No se pudo comprobar si la modificación propuesta generaría una respuesta de defensa contra *P. phachyrhizi*, ya que no se obtuvieron plantas regeneradas con el sitio de modificación esperado.

Analizando el potencial de la proteína PBS1 como blanco potencial para modificaciones genéticas se realizaron análisis bioinformáticos de distintas especies de plantas, y se investigaron proteínas relacionadas a la respuesta de defensa como RLK.

Las RLK están presentes en todas las viridiplantae (estreptofitas y clorofitas), comparten varios subdominios típicos de cinasa y son inducidos por estrés abiótico y biótico (Delaux et al., 2015; Shankar et al., 2015). Las proteínas involucradas en la respuesta de defensa tienden a ser más diversas que las proteínas involucradas en el crecimiento y desarrollo (Jose et al., 2020). Los RLCK, un subgrupo de RLK, pertenecen a la familia RLK/Pelle y son agrupados en la superfamilia de serina/treonina/tirosina cinasas en eucariotas (Shiu S. H. y Bleecker, 2001). RLCK y PBS1 en particular, evolucionaron a partir de RLK por la eliminación de un dominio extracelular (Shiu S. H. y Bleecker, 2001).

La divergencia de estos genes podría haber resultado de su eventual duplicación, con ambos tipos de proteínas que responden a conjuntos de estímulos diferentes y parcialmente superpuestos. De hecho, los RLK tienen una menor capacidad de respuesta al estrés biótico en comparación con los RLCK. La expansión de los genes RLCK fue seguida de la divergencia de las plantas terrestres ancestrales (Lehti-Shiu et al., 2009). De hecho, PBS1 y proteínas PBS1-like, como BIK1 y PBL1, juegan un papel importante en señalización de los receptores inmunitarios de la superficie celular, pero puede que no desempeñe una función esencial en la señalización de PTI, convirtiéndose en un señuelo al mimetizar objetivos de

proteínas de virulencia para desencadenar ETI (Lu et al., 2010; Zhang et al., 2010). PBS1 tiene una actividad cinasa (autofosforilación) que activa HR mediada por RPS5 en *A. thaliana*, pero también PBS1 de *N. benthamiana* puede fosforilar NbREM4 (proteína remorina) *in vitro*, permitiendo la interacción con HopZ1a (proteína efectora localizada en la membrana plasmática de la célula huésped) que desencadena la HR dependiente de ZAR1 en *N. benthamiana* (Albers et al., 2019; DeYoung et al., 2012). PBS1 de *A. thaliana* se agrupa dentro de la respuesta de estrés universal ya que el transcriptoma incluye genes que son inducidos por diversas condiciones, como choque osmótico, frío, salinidad, heridas y estrés biótico, lo que indica que PBS1 responde tanto al estrés abiótico como al biótico (Ma S. y Bohnert, 2007). Se conoce poco acerca de la filogenia y posible origen de PBS1 y otros RLCK.

Las proteínas PBS1 se caracterizan por la presencia de motivos GDK y SEMPB en *A. thaliana* (Block et al., 2008); que están ausentes en clorofitas y pteridofitas, así como en la mayoría de las plantas terrestres. En particular, el motivo GDK (y los motivos GDKSHV y GDKTHV) está presente principalmente en dicotiledóneas y monocotiledóneas, así como en algunas gimnospermas, licofitas y briófitas, lo que sugiere que este motivo (y, por lo tanto, potencialmente el mecanismo señuelo) surgió en el ancestro común de estreptofitas. Sin embargo, el motivo de reconocimiento de resistencia (RRM) es muy diverso. La secuencia de este motivo es SEMPB en PBS1 de *A. thaliana* y otros ortólogos estrechamente relacionados; sin embargo, el presente estudio mostró que el motivo STRPB está presente en diversas especies de embriofitas. Evidencia actual indica que un residuo de aminoácido cargado negativamente (glutamato en el motivo SEMPB) es requerido para el reconocimiento de PBS1 por RPS5 (Sun et al., 2017). Sin embargo, cuando el motivo STRPB se encuentra en PBS1, la unión de RPS5 se debilita, pero sigue siendo funcional (Shiu S.-H. et al., 2004). La estructura proteica tridimensional de PBS1 sugiere la importancia de este residuo para la interacción PBS1-RPS5 (Sun et al., 2017).

La función de "señuelo" de PBS1 se ha informado solo para RPS5 de *A. thaliana* y PBR1 de *Hordeum vulgare* (Kim S. H. et al., 2016). Sin embargo, en este último caso, el análisis filogenético de PBR1 indica que no está estrechamente relacionado con RPS5. Por lo tanto, es posible que las proteínas PBS1 que contienen un motivo RRM no canónico interactúan con una amplia gama de proteínas NLR. Dado que las RLCK también pueden mediar las respuestas al estrés abiótico, otros factores podrían interactuar con proteínas PBS1-like de diversos taxones.

Es concebible que, durante la evolución, no hubo restricciones a la diversificación de las RRM; sin embargo, el motivo SEMPH canónico está presente solo en *A. thaliana*, *Capsella rubella*, *Arabidopsis lyrata*, *Eutrema salsugineum*, *Camelina sativa*, *Microthlaspi erraticum*, *Arabis nemorensis*, *Brassica campestris* y *Brassica oleracea capitata* agrupadas en un clado. Cabe señalar que todas estas son especies estrechamente relacionadas. Los ortólogos de PBS1 en plantas terrestres contienen subdominios de cinasa conservados y el motivo GDK, pero el motivo RRM puede permitir la interacción con diferentes NLR. Esto está respaldado por el análisis filogenético aquí presentado.

Los RLCK evolucionaron para reconocer efectores microbianos solo en algunos taxones. En consecuencia, algunas especies contienen secuencias PBS1-like. Además, algunos de estos se encuentran en clados que incluyen taxones evolutivamente divergentes, apoyando la noción de tasas de sustitución desiguales en estos, a pesar de que PBS1 es una de las proteínas más conservadas en angiospermas (Carter et al., 2019). En clorofitas, AthPBS1 está más relacionado con una tirosina cinasa de *C. subellipsoidea* y dos proteínas hipotéticas de *C. eustigma* y *Trebouxia sp.*, respectivamente. El motivo GDK está ausente en estos ortólogos de PBS1. Además, los ortólogos de PBS1 de clorofitas y estreptofitas analizadas comparten subdominios de cinasa conservados, lo que sugiere un ancestro común para estas secuencias. Las algas verdes (es decir, Charales, Zygnematales y Coleochaetales) contienen proteínas serina/treonina cinasas, así como MAPKKK3 (del inglés mitogen-activated protein kinase kinase

kinase 3) y 4 cinasas asociadas al receptor de interleuquina-1 (IRAK4), todos pertenecientes a la familia RLK. Esto sugiere que la expansión de estos genes ha permitido una evolución acelerada en los dominios involucrados en las señales recibidas por las cinasas (Afzal et al., 2008; Kim H. S. et al., 2016).

Mediante análisis filogenético de PBS1-like se identificaron dos grandes grupos que contienen el motivo GDK. Un grupo contiene diversas cinasas y proteínas hipotéticas estrechamente relacionadas a PBS1 en diferentes estreptofitas y carófitos (angiospermas, briófitas, gimnospermas, licofitas y el alga carófito *N. axillaris*). El segundo grupo consta de diversos estreptofitos, que incluyen cinasas distantes con PBS1, como CDL1-like serina/treonina-proteína cinasas, proteínas cinasas similares a receptores LRR, familia de receptores strubbelig 8, proteínas hipotéticas que contienen dominio de proteína cinasa, y proteínas receptoras de brasinoesteroides. La proteína STRUBBELIG participa en la señalización durante la morfogénesis del tejido (Jose et al., 2020) y CDL1-like serina/treonina-proteína cinasa regula positivamente la señalización y el crecimiento de los brasinoesteroides. Esto sugiere ~~en algunos casos~~ la evolución convergente de esta proteína. Otro clado que incluye cinasas que carecen de GDK, como fototropina y el fotorreceptor neochrome, proteínas que contienen el dominio LOV y fueron identificadas en pteridofitas. Estos están relacionados con los fotorreceptores de luz azul (que, a diferencia de los RLCK, muestran un dominio transmembrana), importantes en la regulación del ciclo circadiano y la floración fotoperiódica (Christie, 2007; Ito et al., 2012). Por lo tanto, PBS1 y proteínas relacionadas han sido cooptadas durante la evolución para cumplir diversas funciones, que van desde la señalización en el sistema inmune a las respuestas relacionadas con desarrollo y crecimiento a la luz (Furukawa et al., 2020). Además, los miembros del mismo linaje según la evidencia taxonómica se encuentran en diferentes clados, lo que sugiere que la presencia de especies no relacionadas en clados comunes podría atribuirse a las características de las secuencias analizadas debido a la presión selectiva generada por el entorno.

De acuerdo con el presente análisis, las proteínas PBS1-like surgieron antes de la divergencia de embriofitas de algas estreptofitas. El hecho de que secuencias similares a PBS1 carezcan de estos motivos característicos apoya esta noción. Por lo tanto, PBS1 evolucionó a partir de cinasas preexistentes en linajes ancestrales. Para obtener más información sobre la evolución de las proteínas PBS1, se realizó un análisis ancestral de ortólogos de PBS1 en plantas terrestres.

La reconstrucción ancestral es una poderosa herramienta para adquirir conocimientos sobre la evolución y las relaciones taxonómicas modernas con sus ancestros potenciales, así como el proceso evolutivo involucrado (Garrido-Bigotes et al., 2019; Joy et al., 2016), basado en una comparación de proteínas homólogas existentes para inferir la secuencia de proteínas precursoras en especies ancestrales (Thornton, 2004). La reconstrucción ancestral de ortólogos de PBS1 se analizó con las secuencias de aminoácidos (estados de caracteres ancestrales).

Primero, el árbol ancestral mostró cinco nodos ancestrales principales, lo que indica un posible origen en pteridofitas. El siguiente nodo ancestral se origina de una gimnosperma (*Pinus taeda*) correspondiente a plantas productoras de semillas, como las pteridofitas. Curiosamente, el siguiente nodo ancestral incluía una carófito (*N. axillaris*). Finalmente, el nodo (negro) ancestral se separó en diferentes especies de plantas terrestres, incluyendo carófitos, briófitas, licofitas, gimnospermas y angiospermas; las dicotiledóneas se separaron del último nodo ancestral. Por lo tanto, proponemos que el ancestro de PBS1 contenía motivos GDK y SEMPH. Sin embargo, estos motivos no son conservados en pteridofitas, *P. taeda* o *N. axillaris*. Una especie importante que sobresale del cuarto nodo ancestral es *Amborella trichopoda*. Esta planta es considerada una especie hermana de las plantas con flores (Albert et al.). Curiosamente, Otra rama del mismo nodo ancestral corresponde a *Ch. braunii*. El ortólogo PBS1 de esta clorofita corresponde a una proteína hipotética que contiene motivos putativos modificados. Durante la evolución, el motivo GDK mostró cambios mínimos en los residuos de aminoácidos, probablemente un reflejo de presión selectiva que

mantuvo su función. Los motivos extendidos GDKSHV y GDKTHV son los más conservados en PBS1, en donde la serina y la treonina son intercambiables. Como se mencionó anteriormente, el motivo SEMPH mostró un bajo grado de conservación. Los aminoácidos que son probablemente más relevantes para la interacción con RPS5 son glutamato (posición 297) y metionina (posición 298). Según el estado ancestral, la treonina es el carácter ancestral más conservado en la posición 297, procedente de *N. axillaris*. En el caso de la posición 298, el carácter ancestral más conservado es la arginina. Su divergencia en algunas especies a un aminoácido hidrofóbico sugiere una función diferente de estos motivos (Harms y Thornton, 2013; Sun et al., 2017). Las secuencias de PBS1 que contienen motivos TDMPH, SDRPH y GDRPH se identificaron en dicotiledóneas; donde un residuo de aspartato sustituye al residuo de glutamato con carga negativa similar, lo que sugiere funciones similares de estos motivos (Shiu S.-H. et al., 2004). En resumen, aunque los motivos más conservados son GDKSHV y STRPH, algunas proteínas PBS1 tienen residuos de aminoácidos con propiedades bioquímicas diferentes o similares, lo que puede tener implicaciones en la especificidad de la función señuelo de PBS1.

El modelado de la estructura proteica de PBS1 en nueve especies diferentes permitió un análisis relevante de los motivos para su función. El motivo GDK se encuentra visible en un bucle, que sería accesible por su hendidura. Ninguna proteína ortóloga de *C. reinhardtii* y *L. linearis* alberga un motivo GDK; más bien, la secuencia LGA y AER se encuentran en su posición, respectivamente. Se desconoce la función de estas proteínas; sin embargo, basado en su similitud con las cinasas descritas, es probable que participen en la transducción de señales ambientales y de desarrollo. Esto también ilustra la diversidad de funciones de esta familia de proteínas, siendo la función señuelo en ETI una evolución tardía. En los modelos PBS1 analizados en este estudio, los motivos STQPQ están separados de el motivo GDK y están expuestos en un bucle, lo que sugiere que no hay impedimento estérico por su reconocimiento por RPS5. No se sabe si el motivo STQPQ cumple una función similar a SEMPH y STRPH, es decir, unión a

RPS5. En proteínas ortólogas PBS1 de *S. polyrhiza*, *P. patens*, *L. linearis*, y *C. reinhardtii*, se encuentra una secuencia diferente en lugar del motivo SEMPH (NARAA, NSRAA, SERAA y QVVSV, respectivamente), que forma una hélice alfa, y por lo tanto sería inaccesible a la unión de un homólogo putativo de RPS5. Interesantemente, la proteína hipotética de *P. pinaster* contiene el motivo GDK y el motivo NNRAA expuesto en un bucle. Por otro lado, el *Ch. braunii* presenta GGETHV (potencialmente equivalente a los motivos GDK y SEMPH). Se desconoce la función de estos motivos, pero su mecanismo de acción podría ser diferente de PBS1, dada la disimilitud de secuencia. Un estudio en el que el motivo SEMPH de AthPBS1 fue reemplazado por el motivo NSRAA de *P. patens* da como resultado una pérdida de HR en *A. thaliana*, lo que demuestra que sus funciones son diferentes (Qi et al., 2014). Proteínas HvPBS1-1 y HvPBS1-2 de cebada son ortólogos de AthPBS1 y son escindidos por AvrPphB (Carter et al., 2019). Interesantemente, una proteína lejana relacionada con RPS5, PBR1, reconoce AvrPphB, apoyando la idea, en primer lugar, de que las proteínas que reconocen los efectores microbianos, en varios casos han experimentado evolución convergente; segundo, que diferentes proteínas del huésped pueden reconocer el mismo efector (y están bajo una fuerte selección de equilibrio en diferentes genotipos); y finalmente, el mecanismo de interacción de PBR1 con los homólogos de PBS1 difiere de RPS5 (Carter et al., 2019).

Las proteínas con repetidos ricos en leucina del dominio de unión a nucleótidos (NBS-LRR), como RPS5, son una categoría importante de receptores del sistema inmune (DeYoung y Innes, 2006). Un análisis genómico comparativo en plantas sugiere que los NBS-LRR están presentes en carófitos (Gao et al., 2018), plantas terrestres y 2 de 24 especies de clorofitas (Shao et al., 2019). El origen de las NBS-LRR se puede rastrear hasta un antepasado anterior de la divergencia de las algas verdes y las plantas terrestres (Gao et al., 2018; Shao et al., 2019).

Estos datos sugieren que PBS1 y RPS5 han evolucionado a partir de especies ancestrales, adquiriendo motivos útiles para la respuesta de defensa. En las proteínas NBS-LRR, incluida la RPS5, la ganancia y la pérdida de dominios, así

como la duplicación de genes y los eventos de translocación han ocurrido con frecuencia durante la evolución, lo que resulta en una alta variabilidad requerida para el reconocimiento de patógenos en la evolución. RPS5 es, por lo tanto, una proteína altamente polimórfica, lo que sugiere que equilibrar la selección tiene un papel importante en su evolución (Gao et al., 2018; Han, 2019; Meyers et al., 2003; Upson et al., 2018).

9. Conclusiones

La soya es tetraploide y recalcitrante a la transformación, siendo complicado y tardado obtener una planta modificada. Por lo tanto, esta investigación proporciona condiciones adecuadas para el método de transformación, así como la utilidad de la edición del genoma mediante el uso de CRISPR/Cas9. Empleando el gen *PBS1* como objetivo para su edición, se demostró mediante análisis moleculares la presencia del gen *cas9*, así como la modificación del gen *PBS1* en la generación T0 y sólo la presencia del gen *cas9* en la generación T1. Por lo tanto, se proporcionan datos alentadores para la generación de modificaciones genéticas en soya; sin embargo, no se sabe si esta modificación genera resistencia a *P. pachyrhizi*. De acuerdo con nuestros análisis bioinformáticos es muy probable que pueda ocurrir este evento.

Interesantemente, se identificaron conjuntos de proteínas ortólogas a *PBS1* en angiospermas, briófitas, gimnospermas, hepáticas, licofitas, pteridofitas, carófitas y clorofitas. Los análisis bioinformáticos en *PBS1* indican la presencia del motivo GDK en especies ancestrales, tales como briófitas, carófitas, hepáticas y licofitas, pero no en clorofitas y pteridofitas, lo que sugiere que *PBS1* podría haber surgido en la evolución de un grupo hermano de estas especies ancestrales, pero este *PBS1* ancestral posiblemente tenía una función no relacionada con la respuesta de defensa. Puede haber adquirido el motivo SEMPH posteriormente en dicotiledóneas, y haber dado origen a *PBS1* que está implicado en la respuesta a *P. syringae*. *RPS5* está dentro de un grupo de proteínas de resistencia que, como *PBS1*, han evolucionado a partir de especies ancestrales a través de la confrontación con varios patógenos para desencadenar ETI.

10. Bibliografía

- Ade J., DeYoung B. J., Golstein C., y Innes R. W. (2007). Indirect activation of a plant nucleotide binding site–leucine-rich repeat protein by a bacterial protease. *Proceedings of the National Academy of Sciences*, 104(7), 2531-2536.
- Afzal A. J., Wood A. J., y Lightfoot D. A. (2008). Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *Molecular Plant-Microbe Interactions*, 21(5), 507-517.
- Agricultura C. M. d., y Rural D. (2006). *Soya (glycine Max (l.) Merrill" Alternativa Para Los Sistemas de Produccion de la Orinoquia Colombiana: Corpoica*.
- Albers P., Üstün S., Witzel K., Kraner M., y Börnke F. (2019). A remorin from *Nicotiana benthamiana* interacts with the *Pseudomonas* type-III effector protein HopZ1a and is phosphorylated by the immune-related kinase PBS1. *Molecular Plant-Microbe Interactions*, 32(9), 1229-1242.
- Albert V., Barbazuk W., dePamphilis C., Der J., Leebens-Mack J., Ma H., . . . Schuster S. Amborella Genome Project (2013) The Amborella genome and the evolution of flowering plants. *Science*, 342, 1241089.
- Ashfield T., Egan A. N., Pfeil B. E., Chen N. W., Podicheti R., Ratnaparkhe M. B., . . . Doyle J. J. (2012). Evolution of a complex disease resistance gene cluster in diploid *Phaseolus* and tetraploid *Glycine*. *Plant physiology*, 159(1), 336-354.
- Bailey T. L., Johnson J., Grant C. E., y Noble W. S. (2015). The MEME suite. *Nucleic acids research*, 43(W1), W39-W49.
- Belhaj K., Chaparro-Garcia A., Kamoun S., Patron N. J., y Nekrasov V. (2015). Editing plant genomes with CRISPR/Cas9. *Current opinion in biotechnology*, 32, 76-84.
- Block A., Li G., Fu Z. Q., y Alfano J. R. (2008). Phytopathogen type III effector weaponry and their plant targets. *Current opinion in plant biology*, 11(4), 396-403.
- Boller T., y Felix G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual review of plant biology*, 60, 379-406.
- Bourges H. (1978). El papel de la soya en la alimentación humana. *Cuad Nutr*, 3, 365-371.
- Bromfield K. (1984). Soybean Rust Monogr. No. 11. American Phytopathological Society, St. Paul, MN. *Soybean Rust Monogr. No. 11. American Phytopathological Society, St. Paul, MN.*, -.
- Cai Y., Chen L., Liu X., Guo C., Sun S., Wu C., Hou W. (2018). CRISPR/Cas9-mediated targeted mutagenesis of GmFT2a delays flowering time in soya bean. *Plant biotechnology journal*, 16(1), 176-185.
- Caldwell K. S., y Michelmore R. W. (2009). *Arabidopsis thaliana* genes encoding defense signaling and recognition proteins exhibit contrasting evolutionary dynamics. *Genetics*, 181(2), 671-684.
- Carter M. E., Helm M., Chapman A. V., Wan E., Restrepo Sierra A. M., Innes R. W., . . . Wise R. P. (2019). Convergent evolution of effector protease recognition by *Arabidopsis* and barley. *Molecular Plant-Microbe Interactions*, 32(5), 550-565.

- Chisholm S. T., Coaker G., Day B., y Staskawicz B. J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*, 124(4), 803-814.
- Christie J. M. (2007). Phototropin blue-light receptors. *Annu. Rev. Plant Biol.*, 58, 21-45.
- Cui H., Tsuda K., y Parker J. E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. *Annual review of plant biology*, 66, 487-511.
- Delaux P.-M., Radhakrishnan G. V., Jayaraman D., Cheema J., Malbreil M., Volkening J. D., . . . Pokorny L. (2015). Algal ancestor of land plants was preadapted for symbiosis. *Proceedings of the National Academy of Sciences*, 112(43), 13390-13395.
- DeYoung B. J., y Innes R. W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature immunology*, 7(12), 1243-1249.
- DeYoung B. J., Qi D., Kim S. H., Burke T. P., y Innes R. W. (2012). Activation of a plant nucleotide binding-leucine rich repeat disease resistance protein by a modified self protein. *Cellular microbiology*, 14(7), 1071-1084.
- Do P. T., Nguyen C. X., Bui H. T., Tran L. T., Stacey G., Gillman J. D., . . . Stacey M. G. (2019). Demonstration of highly efficient dual gRNA CRISPR/Cas9 editing of the homeologous GmFAD2–1A and GmFAD2–1B genes to yield a high oleic, low linoleic and α -linolenic acid phenotype in soybean. *BMC plant biology*, 19(1), 1-14.
- Edgar R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5), 1792-1797.
- Fang Y., y Tyler B. M. (2016). Efficient disruption and replacement of an effector gene in the oomycete *Phytophthora sojae* using CRISPR/Cas9. *Molecular plant pathology*, 17(1), 127-139.
- Furukawa R., Toma W., Yamazaki K., y Akanuma S. (2020). Ancestral sequence reconstruction produces thermally stable enzymes with mesophilic enzyme-like catalytic properties. *Scientific reports*, 10(1), 1-13.
- Gao Y., Wang W., Zhang T., Gong Z., Zhao H., y Han G.-Z. (2018). Out of water: the origin and early diversification of plant R-genes. *Plant physiology*, 177(1), 82-89.
- Garrido-Bigotes A., Valenzuela-Riffo F., y Figueroa C. R. (2019). Evolutionary analysis of JAZ proteins in plants: an approach in search of the ancestral sequence. *International journal of molecular sciences*, 20(20), 5060.
- Gleditsch D., Pausch P., Müller-Esparza H., Özcan A., Guo X., Bange G., y Randau L. (2019). PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. *RNA biology*, 16(4), 504-517.
- Goellner K., Loehrer M., Langenbach C., Conrath U., Koch E., y Schaffrath U. (2010). *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. *Molecular plant pathology*, 11(2), 169-177.
- Gómez-Gómez L., y Boller T. (2000). FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular cell*, 5(6), 1003-1011.
- Gouy M., Guindon S., y Gascuel O. (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular biology and evolution*, 27(2), 221-224.
- Grassini P., Cafaro La Menza N., Rattalino Edreira J. I., Monzón J. P., Tenorio F. A., y Specht J. E. (2021). Chapter 8 - Soybean. In Sadras V. O. y Calderini D. F. (Eds.), *Crop Physiology Case Histories for Major Crops* (pp. 282-319): Academic Press.
- Han G. Z. (2019). Origin and evolution of the plant immune system. *New Phytol*, 222(1), 70-83. doi:10.1111/nph.15596

- Harms M. J., y Thornton J. W. (2013). Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nature Reviews Genetics*, 14(8), 559-571.
- Harrison M. M., Jenkins B. V., O'Connor-Giles K. M., y Wildonger J. (2014). A CRISPR view of development. *Genes & development*, 28(17), 1859-1872.
- Hochstrasser M. L., y Doudna J. A. (2015). Cutting it close: CRISPR-associated endoribonuclease structure and function. *Trends in biochemical sciences*, 40(1), 58-66.
- Hu N., Xian Z., Li N., Liu Y., Huang W., Yan F., . . . Li Z. (2019). Rapid and user-friendly open-source CRISPR/Cas9 system for single-or multi-site editing of tomato genome. *Horticulture research*, 6.
- Hymowitz T. (2008). The history of the soybean. In *Soybeans* (pp. 1-31): Elsevier.
- Intagri (Producer). (2015). Soya: Importancia Nacional e Internacional.
- Ishino Y., Krupovic M., y Forterre P. (2018). History of CRISPR-Cas from encounter with a mysterious repeated sequence to genome editing technology. *Journal of bacteriology*, 200(7), e00580-00517.
- Ito S., Song Y. H., Josephson-Day A. R., Miller R. J., Breton G., Olmstead R. G., y Imaizumi T. (2012). FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator CONSTANS in Arabidopsis. *Proceedings of the National Academy of Sciences*, 109(9), 3582-3587.
- Jaganathan D., Ramasamy K., Sellamuthu G., Jayabalan S., y Venkataraman G. (2018). CRISPR for crop improvement: an update review. *Frontiers in plant science*, 9, 985.
- Jiang F., y Doudna J. A. (2017). CRISPR-Cas9 structures and mechanisms. *Annual review of biophysics*, 46, 505-529.
- Jones J. D., y Dangl J. L. (2006). The plant immune system. *Nature*, 444(7117), 323-329. doi:10.1038/nature05286
- Jose J., Ghantasala S., y Roy Choudhury S. (2020). Arabidopsis transmembrane receptor-like kinases (RLKs): a bridge between extracellular signal and intracellular regulatory machinery. *International journal of molecular sciences*, 21(11), 4000.
- Joy J. B., Liang R. H., McCloskey R. M., Nguyen T., y Poon A. F. (2016). Ancestral reconstruction. *PLoS computational biology*, 12(7), e1004763.
- Kaku H., Nishizawa Y., Ishii-Minami N., Akimoto-Tomiya C., Dohmae N., Takio K., . . . Shibuya N. (2006). Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proceedings of the National Academy of Sciences*, 103(29), 11086-11091.
- Kim H. S., Park S.-C., Ji C. Y., Park S., Jeong J. C., Lee H.-S., y Kwak S.-S. (2016). Molecular characterization of biotic and abiotic stress-responsive MAP kinase genes, IbMPK3 and IbMPK6, in sweetpotato. *Plant Physiology and Biochemistry*, 108, 37-48.
- Kim S. H., Qi D., Ashfield T., Helm M., y Innes R. W. (2016). Using decoys to expand the recognition specificity of a plant disease resistance protein. *Science*, 351(6274), 684-687.
- Koch E., y Hoppe H. (1987). Effect of light on uredospore germination and germ tube growth of soybean rust (*Phakopsora pachyrhizi* Syd.). *Journal of Phytopathology*, 119(1), 64-74.
- Kourelis J., van der Hoorn R., y Sueldo D. J. (2016). Decoy engineering: the next step in resistance breeding. *Trends in plant science*, 21(5), 371-373.
- Larsson A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, 30(22), 3276-3278.
- Lehti-Shiu M. D., Zou C., Hanada K., y Shiu S.-H. (2009). Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant physiology*, 150(1), 12-26.

- Li J.-F., Norville J. E., Aach J., McCormack M., Zhang D., Bush J., . . . Sheen J. (2013). Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nature biotechnology*, 31(8), 688-691.
- Li S., Cong Y., Liu Y., Wang T., Shuai Q., Chen N., . . . Li Y. (2017). Optimization of Agrobacterium-mediated transformation in soybean. *Frontiers in plant science*, 8, 246.
- 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.
- Liu X., Wu S., Xu J., Sui C., y Wei J. (2017). Application of CRISPR/Cas9 in plant biology. *Acta pharmaceutica sinica B*, 7(3), 292-302.
- Lu D., Wu S., Gao X., Zhang Y., Shan L., y He P. (2010). A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. *Proceedings of the National Academy of Sciences*, 107(1), 496-501.
- Lule-Chávez A., Carballar-Lejarazú R., Cabrera-Ponce J., Lanz-Mendoza H., y Ibarra J. (2021). Genetic transformation of mosquitoes by microparticle bombardment. *Insect Molecular Biology*, 30(1), 30-41.
- Ma S., y Bohnert H. J. (2007). Integration of Arabidopsis thaliana stress-related transcript profiles, promoter structures, and cell-specific expression. *Genome biology*, 8(4), 1-22.
- Ma X., Zhu Q., Chen Y., y Liu Y.-G. (2016). CRISPR/Cas9 platforms for genome editing in plants: developments and applications. *Molecular plant*, 9(7), 961-974.
- Makarova K. S., Haft D. H., Barrangou R., Brouns S. J., Charpentier E., Horvath P., . . . Yakunin A. F. (2011). Evolution and classification of the CRISPR-Cas systems. *Nature Reviews Microbiology*, 9(6), 467-477.
- Makarova K. S., Wolf Y. I., Iranzo J., Shmakov S. A., Alkhnbashi O. S., Brouns S. J., . . . Horvath P. (2020). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nature Reviews Microbiology*, 18(2), 67-83.
- McGinn J., y Marraffini L. A. (2019). Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nature Reviews Microbiology*, 17(1), 7-12.
- Meurer C., Dinkins R., y Collins G. (1998). Factors affecting soybean cotyledonary node transformation. *Plant cell reports*, 18(3), 180-186.
- Meyers B. C., Kozik A., Griego A., Kuang H., y Michelmore R. W. (2003). Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *The Plant Cell*, 15(4), 809-834.
- Mojica J., y Almendros C. (2017). El descubrimiento del sistema CRISPR-Cas. *Investigación y Ciencia*(493), 20-29.
- Nakata A., Amemura M., y Makino K. (1989). Unusual nucleotide arrangement with repeated sequences in the Escherichia coli K-12 chromosome. *Journal of bacteriology*, 171(6), 3553-3556.
- Nidhi S., Anand U., Oleksak P., Tripathi P., Lal J. A., Thomas G., . . . Tripathi V. (2021). Novel CRISPR-Cas systems: An updated review of the current achievements, applications, and future research perspectives. *International journal of molecular sciences*, 22(7), 3327.
- Oerke E.-C. (2006). Crop losses to pests. *The Journal of Agricultural Science*, 144(1), 31-43.
- Olhoft P., y Somers D. (2001). L-Cysteine increases Agrobacterium-mediated T-DNA delivery into soybean cotyledonary-node cells. *Plant cell reports*, 20(8), 706-711.
- Paes de Melo B., Lourenço-Tessutti I. T., Morgante C. V., Santos N. C., Pinheiro L. B., de Jesus Lins C. B., . . . Grossi-de-Sa M. F. (2020). Soybean embryonic axis transformation: combining biolistic and Agrobacterium-mediated protocols to overcome typical complications of in vitro plant regeneration. *Frontiers in plant science*, 1228.

- Pareddy D., Chennareddy S., Anthony G., Sardesai N., Mall T., Minnicks T., . . . Sarria R. (2020). Improved soybean transformation for efficient and high throughput transgenic production. *Transgenic Research*, 29(3), 267-281. doi:10.1007/s11248-020-00198-8
- Paz M. M., Martinez J. C., Kalvig A. B., Fonger T. M., y Wang K. (2006). Improved cotyledonary node method using an alternative explant derived from mature seed for efficient Agrobacterium-mediated soybean transformation. *Plant cell reports*, 25(3), 206-213.
- Qi D., Dubiella U., Kim S. H., Sloss D. I., Dowen R. H., Dixon J. E., y Innes R. W. (2014). Recognition of the protein kinase AVRPPHB SUSCEPTIBLE1 by the disease resistance protein RESISTANCE TO PSEUDOMONAS SYRINGAE5 is dependent on s-acylation and an exposed loop in AVRPPHB SUSCEPTIBLE1. *Plant physiology*, 164(1), 340-351.
- Rech E. L., Vianna G. R., y Aragão F. J. L. (2008). High-efficiency transformation by biolistics of soybean, common bean and cotton transgenic plants. *Nature Protocols*, 3(3), 410-418. doi:10.1038/nprot.2008.9
- Rosas J. C., y Young R. (1989). El cultivo de la soya.
- SAGARPA. (2011). Servicio de información agroalimentaria y pesquera. *Secretaría de agricultura, ganadería, desarrollo rural, pesca y alimentación*.
- SAGARPA. (2022). Servicio de información agroalimentaria y pesquera. *Secretaría de agricultura, ganadería, desarrollo rural, pesca y alimentación*.
https://www.gob.mx/cms/uploads/attachment/file/392558/Soya_b.pdf
- Schmutz J., Cannon S. B., Schlueter J., Ma J., Mitros T., Nelson W., . . . Cheng J. (2010). Genome sequence of the palaeopolyploid soybean. *Nature*, 463(7278), 178-183.
- Schwessinger B., y Zipfel C. (2008). News from the frontline: recent insights into PAMP-triggered immunity in plants. *Current opinion in plant biology*, 11(4), 389-395.
- Shankar A., Agrawal N., Sharma M., Pandey A., y K Pandey G. (2015). Role of protein tyrosine phosphatases in plants. *Current genomics*, 16(4), 224-236.
- Shao Z.-Q., Xue J.-Y., Wang Q., Wang B., y Chen J.-Q. (2019). Revisiting the origin of plant NBS-LRR genes. *Trends in plant science*, 24(1), 9-12.
- Shapouri H., Duffield J., y Mcaloon A. (2001). United States Department of Agriculture (USDA). In: The.
- Shiu S.-H., Karlowski W. M., Pan R., Tzeng Y.-H., Mayer K. F., y Li W.-H. (2004). Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. *The Plant Cell*, 16(5), 1220-1234.
- Shiu S. H., y Bleecker A. B. (2001). Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. *Proceedings of the National Academy of Sciences of the United States of America*, 98(19), 10763-10768. doi:10.1073/pnas.181141598
- SIAP. (2021). Servicio de Información Agroalimentaria y Pesquera. Boletín mensual de avance de la producción de soya. <https://www.gob.mx/siap/documentos/boletin-mensual-de-avances-de-produccion-de-soya>
- Sun J., Huang G., Fan F., Wang S., Zhang Y., Han Y., . . . Lu D. (2017). Comparative study of Arabidopsis PBS1 and a wheat PBS1 homolog helps understand the mechanism of PBS1 functioning in innate immunity. *Scientific reports*, 7(1), 1-12.
- Tang D., Wang G., y Zhou J.-M. (2017). Receptor kinases in plant-pathogen interactions: more than pattern recognition. *The Plant Cell*, 29(4), 618-637.
- Thornton J. W. (2004). Resurrecting ancient genes: experimental analysis of extinct molecules. *Nature Reviews Genetics*, 5(5), 366-375.

- Thurtle-Schmidt D. M., y Lo T. W. (2018). Molecular biology at the cutting edge: a review on CRISPR/CAS9 gene editing for undergraduates. *Biochemistry and molecular biology education*, 46(2), 195-205.
- Tomes D., Ross M., y Songstad D. (1995). Direct DNA transfer into intact plant cells via microprojectile bombardment. In *Plant cell, tissue and organ culture* (pp. 197-213): Springer.
- Torres y Torres N., y Tovar-Palacio A. R. (2009). La historia del uso de la soya en México, su valor nutricional y su efecto en la salud. *Salud pública de México*, 51(3), 246-254.
- Tremblay A., Hosseini P., Alkharouf N. W., Li S., y Matthews B. F. (2010). Transcriptome analysis of a compatible response by *Glycine max* to *Phakopsora pachyrhizi* infection. *Plant Science*, 179(3), 183-193. doi:<https://doi.org/10.1016/j.plantsci.2010.04.011>
- Trifinopoulos J., Nguyen L.-T., von Haeseler A., y Minh B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic acids research*, 44(W1), W232-W235.
- Upton J. L., Zess E. K., Bialas A., Wu C.-h., y Kamoun S. (2018). The coming of age of EvoMPMI: evolutionary molecular plant–microbe interactions across multiple timescales. *Current opinion in plant biology*, 44, 108-116.
- Whitham S. A., Qi M., Innes R. W., Ma W., Lopes-Caitar V., y Hewezi T. (2016). Molecular soybean-pathogen interactions. *Annual review of phytopathology*, 54, 443-468.
- Wright D., y Lenssen A. (2013). Staging soybean development.
- Yamada T., Takagi K., y Ishimoto M. (2012). Recent advances in soybean transformation and their application to molecular breeding and genomic analysis. *Breeding Science*, 61(5), 480-494.
- Yorinori J., Paiva W., Frederick R., Costamilan L., Bertagnolli P., Hartman G., . . . Nunes Jr J. (2005). Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. *Plant Disease*, 89(6), 675-677.
- Zhang J., Li W., Xiang T., Liu Z., Laluk K., Ding X., . . . Chen S. (2010). Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a *Pseudomonas syringae* effector. *Cell host & microbe*, 7(4), 290-301.
- Zheng N., Li T., Dittman J. D., Su J., Li R., Gassmann W., . . . Yang B. (2020). CRISPR/Cas9-Based Gene Editing Using Egg Cell-Specific Promoters in Arabidopsis and Soybean. *Frontiers in plant science*, 11. doi:10.3389/fpls.2020.00800

11. Anexos

Artículo