



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

**ANÁLISIS DE LA FUNCIÓN DE PÉPTIDOS RALF Y
DE LA CINASA CITOPASMÁTICA RIPK EN FRIJOL
DURANTE LA SIMBIOSIS CON RHIZOBIA**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

M. en C. Jorge Esaú Solís Miranda

TUTOR PRINCIPAL

M en I.B.B. María del Carmen Quinto Hernández (Tutor)

Instituto de Biotecnología, UNAM

MIEMBROS DEL COMITÉ TUTORAL

Dr. José Luis Reyes Taboada

Instituto de Biotecnología, UNAM

Dr. Miguel Ángel Cevallos Gaos

Centro de Ciencias Genómicas, UNAM

Cuernavaca Morelos. Agosto, 2021



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

ESTE PROYECTO SE LLEVÓ A CABO EN EL INSTITUTO DE BIOTECNOLOGÍA DE LA UNAM, EN EL GRUPO DE LA M.I.B.B **MARÍA DEL CARMEN QUINTO HERNÁNDEZ**, EN EL DEPARTAMENTO DE BIOLOGÍA MOLECULAR DE PLANTAS.

AGRADEZCO AL **CONACYT** POR EL APOYO PARA LA REALIZACIÓN DEL PROYECTO: BECA **CONACYT**, NIVEL DE DOCTORADO NÚMERO **463767**.

AGRADEZCO AL **PROGRAMA DE APOYO A LOS ESTUDIOS DE POSGRADO (PAEP)** POR EL APOYO PARA ASISTIR A LOS CONGRESOS

AGRADECIMIENTOS

AGRADEZCO A LOS TECNICOS, NOREIDE NAVA NUÑEZ DEL INSTITUTO DE BIOTECNOLOGÍA Y AL TÉCNICO ALFONSO LEIJA DEL CENTRO DE CIENCIAS GEMÓMICAS; POR TODO SU APOYO TÉCNICO PARA LA REALIZACION DE ESTE PROYECTO.

A LOS MIEMBROS DEL COMITÉ TUTORAL, EL DOCTOR MIGUEL ANGEL CEBALLOS Y EL DOCTOR JOSE LUIS REYES POR TODO SU APOYO Y SUGERENCIAS PARA CONSEGUIR LOS OBJETIVOS DEL PROYECTO.

A LOS REVISORES DE LA TESIS Y MIEMBROS DEL JURADO, POR SUS ATENTAS OBSERVACIONES Y RECOMENDACIONES.

DEDICATORIA

ESTE TRABAJO LO DEDICO A MIS PADRES, JAVIER E IRINA, Y A MIS HERMANOS, JAPER Y LELI, QUE SIEMPRE ME HAN MOSTRADO SU AMOR Y SU APOYO, Y CON EL CUAL ME DAN LA FUERZA Y LA CONFIANZA PARA ALCANZAR MIS METAS. NO TENGO PALABRAS PARA EXPRESAR LA GRATITUD QUE SIENTO HACIA USTEDES.

A SAM, QUE COMPARTÍO CONMIGO TODOS LOS ÉXITOS, PERO SOBRE TODO LAS FRUSTRACIONES DURANTE ESTE PROCESO, SUS PALABRAS DE ALIENTO ME MOTIVARON A SEGUIR INTENTÁNDOLO Y MIRAR HACIA EL FUTURO ¡GRACIAS POR TODO!

A LA DOCTORA CARMEN QUINTO, QUE ME DIO SU APOYO DESDE QUE DECIDÍ ENTRAR AL POSGRADO, PERMITIENDO FORMAR PARTE DE SU GRUPO DE TRABAJO. EN USTED Y EN EL DOCTOR FEDERICO CONOCÍ A UNAS PERSONAS EXCEPCIONALES, QUE HAN MARCADO MI VIDA. GRACIAS A USTEDES A CRECIDO MI AMOR POR LA CIENCIA.

Y A LOS COMPAÑEROS DEL GRUPO DE TRABAJO, CON QUIENES COMPARTÍ ESTOS AÑOS DE MI VIDA, REALMENTE HICIERON QUE CADA DÍA DE TRABAJO FUERA MÁS DIVERTIDO QUE EL ANTERIOR.

INDICE

1.	RESUMEN	4
2.	ABSTRACT	5
3.	LISTA DE ABREVIATURAS.....	6
4.	INTRODUCCIÓN.....	7
4.1	El proceso de la simbiosis leguminosa-rhizobia.....	7
4.2	Funciones de los péptidos pequeños en plantas.....	9
4.3	Participación de péptidos pequeños en el proceso de simbiosis leguminosa-rhizobia	11
4.4	Los péptidos RALF.....	13
4.5	CrRLK1Ls, la subfamilia de receptores de los péptidos RALF	14
4.6	La cascada de señalización de RALF1-FER en el desarrollo de los pelos radicales.....	15
4.7	El papel de RALF, FER y RIPK en interacciones planta-microorganismos.....	19
5.	JUSTIFICACIÓN.....	21
6.	HIPÓTESIS.....	21
7.	OBJETIVOS	21
8.	METODOLOGÍA	22
8.1	Análisis bioinformático	22
8.2	Diseño de oligonucleótidos	22
8.3	Generación de las construcciones.....	23
8.4	Germinación de semillas	24
8.5	Generación de plantas de frijol con raíces transgénicas.....	24
8.6	Análisis de la actividad de promotor	25
8.7	Medición de los niveles de transcritos por RT-qPCR.....	25
8.8	Análisis de los fenotipos en la nodulación	26
8.8.1	Cuantificación del número y diámetro de los nódulos.....	27
8.8.2	Medición de la actividad de la nitrogenasa.....	27
8.9	Análisis del fenotipo sobre la inhibición de la nodulación por nitrógeno	27
9.	RESULTADOS	28
9.1	Análisis <i>in silico</i> de genes RALF y RIPK de frijol	28
9.1.1	Identificación y análisis <i>in silico</i> de los miembros de la familia RALF en frijol	28
9.1.2	Análisis de la expresión de los genes <i>RALF</i> de frijol	30
9.1.3	Determinación y análisis <i>in silico</i> de los genes <i>RIPK</i> de frijol	31

9.1.4	Análisis de la expresión de los genes <i>RIPK</i> de frijol	34
9.2	Acumulación de los transcritos <i>PvRALF1</i> , <i>PvRALF6</i> y <i>PvRIPK2</i> en raíces inoculadas con rhizobia y sin inocular.....	35
9.3	Artículo publicado: “Genome-wide identification of the CrRLK1L subfamily and comparative analysis of its role in the legume-rhizobia symbiosis”	36
9.4	Actividad de promotores <i>PvRALF1</i> , <i>PvRALF6</i> , <i>PvFER1</i> y <i>PvRIPK2</i> en raíces transgénicas de frijol	36
9.4.1	Análisis de la actividad del promotor <i>PvRALF1</i> en raíces transgénicas.....	36
9.4.2	Análisis de la actividad del promotor <i>PvRALF6</i> en raíces transgénicas.....	38
9.4.3	Análisis de la actividad del promotor <i>PvRIPK2</i> en raíces transgénicas	40
9.4.4	Análisis de la actividad del promotor <i>PvFER1</i> en raíces transgénicas.....	42
9.5	Análisis del fenotipo de raíces transgénicas de frijol silenciadas y sobreexpresantes <i>PvRALF1</i> , <i>PvRALF6</i> , <i>PvRIPK2</i> y <i>PvFER1</i> durante la simbiosis con rhizobia.....	44
9.5.1	Selección de clonas de <i>A. rhizogenes</i> portadoras de las construcciones para el silenciamiento y la sobreexpresión de <i>RALF1</i> y <i>RALF6</i>	44
9.5.2	Análisis del número de nódulos en las raíces silenciadas y raíces sobreexpresantes en <i>PvRALF1</i> y <i>PvRALF6</i>	45
9.5.3	Análisis del diámetro de los nódulos en las raíces silenciadas y raíces sobreexpresantes en <i>PvRALF1</i> y <i>PvRALF6</i>	47
9.5.4	Análisis de la actividad de la nitrogenasa en raíces transgénicas silenciadas en <i>RALF1</i> y <i>RALF6</i>	48
9.5.5	Análisis del número y diámetro de nódulos en las raíces silenciadas <i>PvRIPK2</i>	49
9.5.6	Análisis del número y diámetro de nódulos en las raíces silenciadas <i>PvFER1</i>	50
9.6	Análisis de la acumulación de transcritos <i>PvRALF1</i> , <i>PvRALF6</i> , <i>PvFER1</i> en raíces inoculadas con rhizobia en 0 y 5 mM de nitrógeno.	51
9.7	Análisis del fenotipo en la inhibición de la nodulación por nitrógeno	52
9.7.1	Efecto del silenciamiento de <i>PvRALF1</i> , <i>PvRALF6</i> y <i>PvFER1</i> en la inhibición de la nodulación por nitrógeno.....	53
9.7.2	Análisis del efecto de la sobreexpresión de <i>PvRALF1</i> , <i>PvRALF6</i> y <i>PvFER1</i> en la inhibición de la nodulación por nitrógeno	54
9.8	Efecto de los cambios en el nivel de expresión de <i>PvRALF1</i> , <i>PvRALF6</i> y <i>PvFER1</i> en genes relacionados con la organogénesis y la función del nódulo	55
9.9	Efecto del silenciamiento y la sobreexpresión de <i>PvRALF1</i> , <i>PvRALF6</i> y <i>PvFER1</i> en genes de frijol relacionados con las vías AON y NRN	57
10.	DISCUSIÓN	61
10.1	Características de los genes <i>RALF</i> y <i>CrRLK1L</i> de frijol	61
10.2	Características de la expresión de los genes <i>RALF</i> y <i>CrRLK1L</i> en frijol.....	62

10.3	Papel de PvRALF1, PvRALF6 y PvFER en la organogénesis del nódulo	63
10.4	Papel de los péptidos RALF y de FERONIA en las vías AON y NRN.....	64
11.	CONCLUSIONES	67
12.	PERSPECTIVAS	68
13.	REFERENCIAS	72
14.	MATERIAL SUPLEMENTARIO	81

1. RESUMEN

RALF y FERONIA forman un complejo ligando-receptor exclusivo de plantas que ha sido descrito en una gran variedad de procesos, desde desarrollo hasta respuestas a estrés biótico y abiótico. Debido a esto, han ganado mucha relevancia en la última década, siendo descritos una gran cantidad de interactores esenciales para que se desencadene el mecanismo de señalización mediado por RALF. Entre dichos interactores, la cinasa citoplasmática RIPK es de interés en este estudio debido a que el complejo FER-RIPK ha sido caracterizado como esencial en la regulación del desarrollo de los pelos radicales y la elongación de la raíz dependientes de RALF.

En el presente trabajo, se determinaron los diferentes genes *RALF* y *RIPK* de frijol, y se seleccionaron aquellos con función potencial en la nodulación mediante análisis *in silico*. Posteriormente se caracterizaron por genética reversa, con la finalidad de determinar su función en la nodulación. Se identificaron nueve RALFs y seis RIPKs en frijol, los cuales están altamente conservados. De estos, se seleccionaron los genes *PvRALF1*, *PvRALF6* y *PvRIPK2* para su caracterización basados en su expresión diferencial en raíces de frijol inoculadas respecto a no inoculadas. Los análisis de silenciamiento y sobreexpresión de estos tres genes, así como de *PvFER1*, en raíces transgénicas de frijol inoculadas con *Rhizobium tropici*, indican que *PvRALF1*, *PvRALF6* y *PvFER1*, pero no *PvRIPK2*, participan en la determinación del número de nódulos bajo condiciones promotoras de la simbiosis (CPS) y condiciones inhibitorias de la nodulación (CIN) por nitrógeno; la sobreexpresión de *PvRALF1* y *PvFER1*, provocan un incremento en el número de nódulos en CPS, mientras que raíces sobreexpresantes en cualquiera de los tres genes muestran una disminución en el número de nódulos en CIN. Por el contrario, el silenciamiento de *PvRALF6* y *PvFER1* causa una disminución en el número de nódulos durante CPS, mientras que el silenciamiento de cualquiera de los tres genes genera un incremento en el número de nódulos en CIN por nitrógeno. Mediante análisis de qPCR en dichas raíces transgénicas, demostramos una regulación positiva de genes participantes las vías de autorregulación de la nodulación (AON) y regulación de la nodulación por nitrógeno (NRN) en CPS, y regulación negativa de esos genes en CIN por nitrógeno.

Los datos presentados indican que *PvRALF1*, *PvRALF6*, y *PvFER1* participan en la regulación del número de nódulos en respuesta al contenido de nitrógeno en suelo, a través de la regulación de genes pertenecientes a las vías AON y NRN.

2. ABSTRACT

RALF and FERONIA form a ligand-receptor plant specific complex involved in a variety of processes, from plant development to biotic and abiotic stress responses. Due to that, this complex has acquired great relevance in the last decade, and many interactors that are essential for this RALF-triggered signaling has been described since then. Among them, the cytoplasm kinase RIPK is of interest in this work since the FER-RIPK complex has been proved as important for RALF-dependent root hair development and cell elongation in *A. thaliana* roots.

This work main objective was to determine common bean *RALF* and *RIPK* genes, and selection of those with a potential role in nodulation by several *in silico* analyzes. Then a functional characterization by reverse genetics to determine their function in such process was done. We determined that common bean harbors nine RALFs and six RIPKs highly conserved. From them, we select *PvRALF1*, *PvRALF6*, and *PvRIPK2* due to their differential expression in common bean inoculated roots comparing to non-inoculated ones for functional characterization, together with the previously characterized *PvFERONIA1*. Reverse genetic analyzes of *PvRALF1*, *PvRALF6*, and *PvRIPK2*, as well as *PvFER1*, in *P. vulgaris* transgenic roots inoculated with *R. tropici*, indicated that *PvRALF1*, *PvRALF6*, and *PvFER1*, but no *PvRIPK2*, participate in the regulation of the nodule number during symbiotic promoter (SPC) and nitrogen-mediated symbiotic inhibitory conditions (SIC). Overexpression of *PvRALF1* and *PvFER1*, and RNAi silencing of *PvRALF6* and *PvFER1* showed an increase and a reduction in the nodule number during SIC, respectively. In contrast, overexpression and silencing of any of the three genes showed a reduction and an increase in the nodule number in SPC, respectively. qPCR analysis of transgenic roots demonstrated a negative regulation of some autoregulation of nodulation and nitrate-mediated regulation of nodulation (AON and NRN) genes in SPC, and by positive regulating those genes under nitrogen-mediated SIC.

Data presented here indicates that *PvRALF1*, *PvRALF6*, and *PvFER1* participate in the regulation of the nodule number in response to soil nitrate concentration, through the regulation of genes associated with the AON and NRN regulatory pathways.

3. LISTA DE ABREVIATURAS

AON: Autorregulación de la nodulación (“Autoregulation of nodulation”)

CIN: condiciones inhibitorias de la nodulación

CLE: “CLAVATA3/endosperm-surrounding region”

CPS: condiciones promotoras de la simbiosis

CRP: péptidos ricos en cisteína

CrRLK1L: *Catharanthus roseous* receptor-like kinase 1L

FER: FERONIA

GS: glutamato sintetasa

Leg: Leghemoglobina

NARK: “Nodule autoregulation receptor kinase”

NCR; “nodule-specific cysteine rich peptide”

NIC: “Nitrate induced CLE”

NIN: “NODULE INCEPTION “

NRN: Regulación de la nodulación mediado por nitrógeno (“Nitrogen-mediated regulation of nodulation”)

PTM: péptidos pequeños modificados postraduccionalmente

PvGEA: atlas de expresión de frijol (“*P. vulgaris* gene expresión atlas”)

RIC: “Rhizobia induced CLE”

SS: sacarosa sintasa

TML: “TOO MUCH LOVE”

4. INTRODUCCIÓN

Las interacciones entre los microorganismos son esenciales para mantener los ecosistemas. Las plantas interactúan con una gran cantidad de microorganismos de la rizósfera, en parte, gracias a los exudados que liberan las raíces. Estos exudados están compuestos por aminoácidos, ácidos orgánicos, azúcares, compuestos fenólicos, mucilago, proteínas y otros metabolitos secundarios (Walker *et al.* 2003). Algunas de estas interacciones son benéficas, como en el caso de la asociación con bacterias gram-negativas del suelo o con hongos micorrízicos, mientras que otras son adversas, como en el caso de las interacciones con patógenos que causan enfermedades en las plantas.

4.1 El proceso de la simbiosis leguminosa-rhizobia

Una interacción benéfica conocida como simbiosis fijadora de nitrógeno ocurre entre las plantas de la familia leguminosa con bacterias del suelo del género rhizobia. Es un proceso importante desde el punto de vista agroecológico, que resulta en la fijación biológica de nitrógeno atmosférico en nitrógeno asimilable por la planta.

Para que esta simbiosis se establezca, es necesario que ocurra un intercambio de señales químicas entre la planta y la bacteria. Este proceso, conocido como diálogo molecular, inicia cuando las raíces de las plantas liberan a la rizósfera una serie de compuestos fenólicos de naturaleza flavonoide. Estas moléculas funcionan primeramente como quimioatrayentes, haciendo que el endosimbionte específico sea atraído hacia los pelos radicales (Ferguson *et al.* 2010). Posteriormente, los flavonoides se unen de manera específica a una proteína en la membrana de rhizobia, NodD, la cual es un regulador transcripcional de la familia LysR que, al unirse al flavonoide, sufre un cambio conformacional y promueve la transcripción de los genes bacterianos conocidos como genes *nod* (Ferguson *et al.* 2010). Los genes *nod* codifican genes involucrados en la síntesis de proteínas encargadas de la síntesis y secreción de los factores Nod. Estos compuestos de naturaleza lipoquito-oligosacarídica, son liberados por la bacteria a la rizosfera y percibidos de manera específica por los receptores de los factores Nod en la membrana de los pelos radicales de la raíz, induciéndose una serie de respuestas a nivel molecular, celular, y morfológico (Gage 2004).

Posterior al dialogo molecular, ocurre un redireccionamiento del crecimiento polar del pelo radical, generándose un enroscamiento en su ápice, donde la bacteria queda atrapada en una

cámara de infección. En este sitio la bacteria comienza a dividirse, y alrededor de 24 horas después se forma una estructura tubular en el pelo radical a partir de la cámara de infección, conocido como hilo de infección, producto de la invaginación de la pared y la membrana del pelo radical (Patriarca et al. 2004; Ferguson et al. 2010). De manera paralela, las células del córtex de la raíz sufren un proceso de desdiferenciación, tras lo cual reactivan el ciclo celular y proliferan; estas células dan origen al primordio del nódulo (Patriarca et al. 2004; Ferguson et al. 2010). El hilo de infección a su vez, se extiende a manera de túnel hasta alcanzar el primordio del nódulo, donde se ramifica. De cada uno de estas prolongaciones del hilo son liberadas bacterias, las cuales ingresan en algunas células del primordio del nódulo mediante un proceso parecido a la endocitosis, tras el cual las bacterias quedan rodeadas de una doble membrana de origen vegetal conocida como membrana peribacteroidal (Patriarca et al. 2004; Ferguson et al. 2010).

Dentro de este nuevo organelo la bacteria se diferencia en bacteroide, el cual es un proceso esencial para que ocurra la fijación de nitrógeno. Se han descrito tres tipos de bacteroides de acuerdo a sus características morfológicas: el primer tipo se observa en bacteroides de leguminosas pertenecientes al clado Dalbergoid, en el cuales los bacteroides poseen formas alargadas e incluso con forma de esferas (Bonaldi *et al.* 2011). El segundo tipo corresponde a bacteroides de leguminosas del grupo phaseoloid, en este la bacteria no sufre una diferenciación total, manteniendo la forma y el contenido de DNA similar a la bacteria de vida libre y manteniéndose viable (Mergaert *et al.* 2006). El tercer tipo corresponde a bacteroides de las leguminosas pertenecientes al clado denominado “carente de invertidos repetidos” (IRLC, “inverted repeat lacking clade”), los cuales sufren una diferenciación terminal perdiendo viabilidad, y presentan una morfología alargada y ramificada tras varios ciclos de replicación de DNA (Mergaert *et al.* 2006).

En general, han sido descritos dos tipos de nódulos, los de tipo determinado y los de tipo indeterminado. Los nódulos determinados, como los que se generan en frijol, se caracterizan por tener un meristemo determinado presentando una estructura redonda, por su parte los nódulos indeterminados, como los de *Medicago truncatula*, presentan un meristemo persistente, generando una estructura alargada.

Existen mecanismos a través de los cuales la planta controla la cantidad de nódulos que se generan, balanceando el gasto de energía para su formación y mantenimiento, con sus necesidades de nitrógeno. La autorregulación de la nodulación (AON) y la inhibición de la nodulación dependiente de nitrógeno (NRN) y por bajo fósforo (PRN) (Ferguson *et al.* 2010; Isidra-Arellano *et al.* 2020) son probablemente los mecanismos más importantes. La AON inicia en las raíces, pero es controlado en la parte aérea de la planta, y por tanto requiere de una molécula señal que llevará el mensaje de la raíz a los brotes. Por el contrario, la vía NRN actúa de manera local y casi inmediata (Ferguson *et al.* 2010). Estos procesos son regulados por péptidos, los cuales se discuten más adelante.

4.2 Funciones de los péptidos pequeños en plantas

Las plantas, al igual que los demás organismos, emplean una gran variedad de moléculas para llevar a cabo la comunicación entre células. Ejemplos de estas son las fitohormonas, como las auxinas importantes durante el desarrollo, (Vanneste and Friml 2009), factores de transcripción como SHORT ROOT que participa en el desarrollo del córtex y la endodermis, (Helariutta *et al.* 2000), moléculas de RNA como el micro RNA miR166 que participa en la determinación de la identidad abaxial de las hojas (Emery *et al.* 2003; Fahlgren *et al.* 2006), así como una gran cantidad de péptidos pequeños (van Norman *et al.* 2011). En la última década el análisis de la función de los péptidos en plantas ha adquirido relevancia debido a su participación en casi todos los procesos de desarrollo en las plantas (Matsubayashi and Sakagami 2006; Murphy *et al.* 2012). En el genoma de *Arabidopsis thaliana* se han descrito más de 1000 genes que presumiblemente codifican para la síntesis de distintos péptidos pequeños (Lease and Walker 2006).

A pesar de la enorme variedad y el bajo porcentaje de identidad entre ellos, los péptidos pequeños pueden clasificarse en dos categorías: péptidos de secreción y péptidos de no-secreción, basado en la presencia o ausencia de una señal de secreción. Entre los péptidos de no-secreción (Fig. 1) se encuentran las sisteminas y ENOD40. Los primeros, son péptidos en general de 18 aminoácidos, los cuales han sido caracterizados en respuesta a daño mecánico en solanáceas, induciendo la expresión de inhibidores de proteasas; por su parte, ENOD40 fue caracterizado como una proteína de respuesta temprana durante el proceso de simbiosis leguminosa-rhizobia (Charon *et al.* 1997, 1999); sin embargo, también ha sido descrita su

participación en el proceso de desarrollo tanto en leguminosas como en no leguminosas (Vlegheles *et al.* 2003). *ENOD40* está codificada por una secuencia de lectura abierta pequeña, de la cual se genera un RNA conservado y altamente estructurado. Se ha descrito que *ENOD40* posee un doble papel, tanto en forma de péptido como de RNA largo no codificante (Wan *et al.* 2007; Bardou *et al.* 2011).

Por su parte, los péptidos de secreción se subdividen en péptidos pequeños modificados postraduccionalmente (PTM) y péptidos ricos en cisteína (CRP) (Matsubayashi 2014; Murphy and De Smet 2014). Los PTM presentan tres características importantes: i) son péptidos de menos de 20 aminoácidos, ii) se generan por procesamiento proteolítico de precursores mayores con péptidos señal, y iii) sufren al menos una modificación postraduccional, que pueden ser hidroxilación en prolinas, sulfatación de tirosina o glicosilación de hidroxil-prolinas (Murphy *et al.* 2012; Matsubayashi 2014). La participación de diversos PTM ha sido reportada en distintos procesos celulares en plantas (Fig. 1), como es el caso de *Clavata 3 (CLV3)*, un PTM 13 aminoácidos, que participa en el mantenimiento y el balance del número de células troncales del meristemo apical de brotes (Fig. 1) (Fletcher 1999).

Los CRP también poseen tres características distintivas: i) son de tamaño mayor que los PTM pero menor a 160 aminoácidos, ii) poseen cuatro o más residuos de cisteína en la región carboxilo (C)-terminal, necesarios para la estabilidad del péptido maduro, y iii) presentan un péptido señal de secreción en la región amino (N)-terminal muy conservada entre toda esta clase de péptidos (Marshall *et al.* 2011; Murphy *et al.* 2012). Al igual que en el caso de las PTM, han sido descritos distintos CRP con patrones de expresión y funciones diferentes (Fig. 1). Ejemplos de CRPs son: *EPF (Epidermal factor) 1/2* y *STOMAGEN*, que funcionan como reguladores de la formación de estomas (Fig. 1) (Kondo *et al.* 2014) y la familia de factores de alcalinización rápida (RALFs, por sus siglas en inglés, Rapid, Alkalization Factors), cuyas funciones se discutirán más adelante.

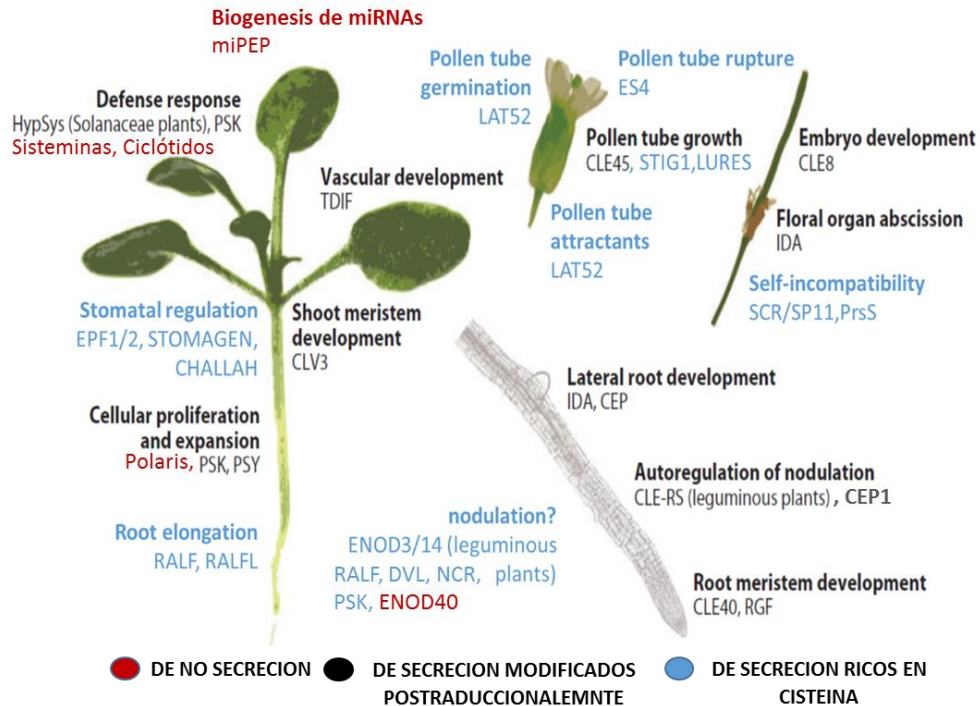


Figura 1. Ejemplos de péptidos pequeños de secreción. Funciones de algunos péptidos de no secreción (en rojo), de secreción modificados post-transcripcionalmente (en negro), y de péptidos ricos en cisteína (en azul) modificado de Marshall *et al.*, 2011 y Matsubayashi, 2014.

4.3 Participación de péptidos pequeños en el proceso de simbiosis leguminosa-rhizobia

La participación de algunos péptidos ha sido descrita durante diferentes etapas de la nodulación. Uno de los primeros péptidos descritos involucrados en este proceso es la nodulina temprana 40 (ENOD40), que se ha reportado que participa en la organogénesis de nódulo. Los genomas de *M. truncatula* y *L. japonicus* poseen dos copias de *ENOD40*, los cuales inducen una disminución en el número de primordios de nódulos tras su silenciamiento y un aumento debido a su sobreexpresión (Kereszt *et al.* 2018). La fitosulfoquina- α (PSK α) de *L. japonicus* es un PTM de cinco aminoácidos cuya sobreexpresión y su aplicación exógena, incrementa la nodulación probablemente mediante la inhibición de las respuestas de defensa (Wang *et al.* 2015). Otro péptido con funciones en la nodulación un péptido codificado en el microRNA miR172c (miPEP172c), el cual promueve la expresión de miR172c, un regulador positivo de la nodulación (Couzigou *et al.* 2015).

De manera exclusiva en las leguminosas IRLC, se ha determinado la presencia de CRP específicos de nódulos (NCR, “nodule-specific cysteine rich peptides”). Se han reportado la

presencia de 700 de estos péptidos NCR en el genoma de *M. truncatula*, de los cuales se ha detectado la expresión de 639 NCRs por RNAseq (Roux and Rodde 2014). Dichos péptidos tienen un papel primordial en la diferenciación del bacteroide: en mutantes *dnf1* de *M. truncatula*, en los que una vía secretoria está bloqueada, se observa que los bacteroides no se diferencian (Van de Velde *et al.* 2010); además, la expresión heteróloga de NCRs en *Lotus japonicus*, que pertenece al grupo phaseolid, resulta en la formación de bacteroides alargados (Van de Velde *et al.* 2010).

En *L. japonicus* y en *M. truncatula*, se ha descrito la participación de dos péptidos de la familia CLAVATA3/endosperm-surrounding región (CLE), LjCLE-R1/2 y MtCLE12/13, todos ellos ortólogos de los péptidos “Rhizobia-induced CLE 1 y 2” (RIC1 y RIC2) de soya (Reid *et al.* 2011; Mortier *et al.* 2012; Okamoto *et al.* 2013; Zhu *et al.* 2021). Estos CLE viajan de la raíz a la parte aérea donde son reconocidos por el receptor tipo cinasa GmNARK en soya, LjHAR1 en *L. japonicus*, y MtSUNN en *M. truncatula* (Ferguson *et al.* 2010, 2019; Mortier *et al.* 2010; Batut *et al.* 2011; Okamoto *et al.* 2013). La percepción de RIC1 y RIC2 por NARK provoca la inhibición de la expresión del microRNA 2111 (miR2111). En condiciones normales, miR2111 viaja a las raíces donde inhibe a una proteína que forma parte de un complejo E3 ubiquitina ligasa llamado “Too Much Love” (TML), el cual es un regulador negativo de la nodulación (Gautrat *et al.* 2020). Por tanto, estos CLEs inhiben la nodulación al promover la acumulación de TML en las raíces.

Además de su papel como señales que se transportan de las raíces al tallo, se han descrito CLEs que participan como reguladores negativos de la nodulación de manera local mediante la vía NRN. En soya, la presencia de nitrógeno promueve la expresión de “nitrate-induced CLE peptide 1” (GmNIC1), el cual es percibido en la raíz por GmNARK y se promueve la regulación local de factores aun desconocidos, que inhiben la nodulación (Ferguson 2019). En frijol, se determinó que los genes *PvRIC1* y *PvRIC2* se expresan en condiciones de bajo fósforo, y provocan la inhibición de la nodulación en dichas condiciones, proceso que depende del receptor NARK en el tejido aéreo (Isidra-Arellano *et al.* 2020). En *M. truncatula* se ha descrito otro PTM que participa en la regulación del número de nódulos, estos son llamados “C-TERMINALLY ENCODED PEPTIDES” (CEPs). Al igual que en el caso de los CLEs, estos son transportados al tejido aéreo y junto con su receptor “COMPACT ROOT

ARCHITECTURE 2” (CRA2), regulan el proceso de nodulación de manera positiva (Imin *et al.* 2013).

Un CRP cuya participación ha sido descrita en el desarrollo de nódulos, es el factor de alcalinización rápida (RALF, por sus siglas en inglés, Rapid, Alkalization Factor); estos forman una familia de péptidos interesantes debido a que están involucrados en muchos procesos de desarrollo en plantas, entre los que destaca su papel durante la simbiosis en *M. truncatula*, leguminosa que forma nódulos de tipo indeterminado (Combiér *et al.* 2008). Los autores encontraron que los niveles del transcrito de *MtRALF1* aumentan en respuesta a tratamiento con factores Nod, y que la sobreexpresión de *MtRALF1* provoca un aumento en el número de eventos de infección fallidas (Fig. 2A) y una disminución en el número de nódulos formados (Fig. 2B). Sin embargo, no se reportaron alteraciones en la formación o el desarrollo de las raíces primarias (Combiér *et al.* 2008), como fue descrito previamente en *A. thaliana* (Pearce *et al.* 2001), sugiriendo diferentes funciones para los diferentes péptidos de la familia RALF. A la fecha, no se conoce el mecanismo por el cual los péptidos RALF promueven la inhibición de la nodulación, ni se sabe si esta es una respuesta exclusiva de *M. truncatula* o si este mecanismo está conservado en procesos de formación de nódulos determinados (como en el caso de frijol) o es exclusiva de nódulos indeterminados, como en el caso de los péptidos NCR.

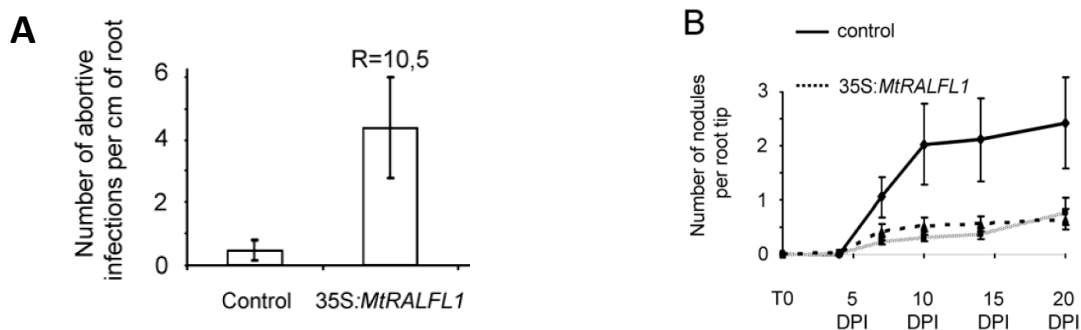


Figura 2. Efecto de la sobreexpresión de un péptido RALF en raíces de *M. truncatula* sobre la nodulación. La sobreexpresión de RALF en raíces de *M. truncatula* provoca una deficiencia en la infección por *Sinorhizobium melliloti*, generando un aumento en el número de infecciones fallidas (A) y una disminución en el número de nódulos (B) modificado de Combiér *et al.* 2008.

4.4 Los péptidos RALF

La familia de péptidos denominados RALF, son un grupo de péptidos de secreción CRP los cuales se encuentran conservados en todas las plantas terrestres. Estos se identificaron en un

ensayo en hojas de tabaco, diseñado para determinar péptidos que indujeran un aumento de pH en el medio de cultivo (Pearce *et al.* 2001). Los RALF se sintetizan como pre-péptidos que presentan un sitio RRXL el cual es reconocido por la enzima proteolítica de “sitio 1” (AtS1P), un sitio YISY, el cual es necesario para interactuar con su receptor, y cuatro cisteínas conservadas importantes para el plegamiento del péptido (Pearce *et al.* 2001). El número de genes de la familia RALF varía de especie a especie, y no depende del tamaño del genoma. Por ejemplo, en *A. thaliana*, que posee un genoma de 135 Mb, han sido identificados 34 genes *RALF*; *Selaginella moellendorffii*, con un genoma de aproximadamente 212 Mb, posee 2 *RALFs*; *Physcomitrella patens* tiene un genoma de 473 Mb y se han descrito 3 de estos genes; y en arroz, con un genoma de 389 Mb, existen 16 *RALFs* (Cao and Shi 2012).

La vía de señalización desencadenada por RALF1 ha sido analizada en la raíz principal en *A. thaliana* (Haruta *et al.* 2014); RALF1 se sintetiza como un pre-péptido de 115 aa, el cual es procesado proteolíticamente por AtS1P en Golgi, sugiriendo que los péptidos RALF1 siguen la vía de secreción para su maduración y secreción. Una vez en el apoplasto, RALF1 se une a su receptor FERONIA (FER) en la membrana de las células de las raíces. Una vez que FER se une a RALF, se inicia una cascada de fosforilación que inhibe la actividad de la ATPasa de protones en la membrana plasmática (AHA2), y de esta forma se induce el incremento en el pH del apoplasto, evitando la elongación celular de la raíz (Haruta *et al.* 2014).

Los genes *RALF* se expresan de manera diferencial en distintos tejidos, y su función está estrechamente relacionada con el perfil de expresión: en *A. thaliana*, RALF1 ha sido reportada en raíces, pelos radicales e hipocotilos, donde inhibe el crecimiento de cada uno de ellos (Pearce *et al.* 2001; Bergonci *et al.* 2014), mientras que AtRALF4 ha sido reportada en polen, donde inhibe la germinación (Morato do Canto *et al.* 2014). En *Solanum lycopersicum*, se han descrito dos péptidos RALF, SIRALF y SIPRALF, el primero se expresa en hoja y el segundo en polen; cuando se aplican *in vitro* SIRALF inhibe el desarrollo de raíces en semillas en germinación (Pearce *et al.* 2001), mientras que SIPRALF inhibe el crecimiento del tubo polínico (Covey *et al.* 2010).

4.5 CrRLK1Ls, la subfamilia de receptores de los péptidos RALF

Uno de los principales retos que existen en la caracterización de la señalización mediada por los péptidos, es la determinación de los receptores con los cuales interactúan. En el caso de

la identificación del receptor de RALF1, FERONIA (FER), esta se hizo mediante un análisis de fosfoproteómica acoplado con espectrometría de masas (Haruta *et al.* 2014). FER es un receptor de membrana tipo cinasa (RLK), perteneciente a la subfamilia CrRLK1L, llamada así ya que se describieron por primera vez en *Catharanthus roseus* (Schulze-Muth *et al.* 1996).

En *A. thaliana*, dicha subfamilia está constituida por 17 miembros, cuyos genes se expresan diferencialmente en todos los tejidos (Schulze-Muth *et al.* 1996). FER, el receptor más caracterizado de esta subfamilia, se expresa prácticamente en todos los tejidos, y participa en diferentes procesos, como son la respuesta a hormonas, en el desarrollo de tubos polínicos, y en el desarrollo de pelos radicales, entre otros (Tabla 1) (Escobar-Restrepo *et al.* 2007; Guo *et al.* 2009; Deslauriers and Larsen 2010; Duan *et al.* 2010; Yu *et al.* 2012). *ANXURI* (*ANX1*) y *ANX2* son dos CrRLK1Ls que se expresan exclusivamente en tubos polínicos y son necesarios para el crecimiento del mismo (Tabla 1) (Boisson-Dernier *et al.* 2009, 2011). La función de los 13 CrRLK1L restantes se resume en la Tabla 1.

Recientemente analizamos la subfamilia de CrRLK1Ls en 57 diferentes especies de plantas, encontrando que estas presentan una gran conservación en sus secuencias, principalmente en los dominios cinasa y malectina (Solis-Miranda *et al.* 2020). Observamos que se distribuyen filogenéticamente en 11 familias, de las cuales, la familia que denominamos “TINIA” se conforma exclusivamente de las CrRLK1Ls de tres musgos incluidos en el análisis, sugiriendo que dicha subfamilia de receptores surgió en un ancestro cercano a estos musgos, y posteriormente se diversificó en el resto de las plantas terrestres (Solis-Miranda *et al.* 2020).

4.6 La cascada de señalización de RALF1-FER en el desarrollo de los pelos radicales

A la fecha, se han descrito una gran cantidad de interactores de los CrRLK1Ls, los cuales son esenciales para que ocurra la señalización mediada por RALFs. Algunas proteínas posiblemente funcionan regulando el proceso de señalización a través de la modulación de la unión de péptido-receptor, así como la localización y la actividad del receptor. Recientemente, publicamos una revisión en la cual describimos los mecanismos de acción que ocurren en los procesos en los cuales están involucradas proteínas de la subfamilia

CrRLK1L, describiendo los interactores conocidos para estos CrRLK1Ls, requeridos en dichos procesos (Solis-miranda and Quinto 2021).

TABLA 1. CrRLK1Ls estudiados en diferentes especies (Solis-Miranda *et al.* 2020).

Gene Name	Plant Species	Mutant/RNAi	Phenotype	Reference
FERONIA	<i>A. thaliana</i>	<i>fer</i>	PT overgrowth, multiple PT reach one ovules	[15-19]
		<i>fer</i>	Collapsed, burst and short RH	[10]
		<i>fer</i>	Resistance to Powdery mildew infection, increased susceptibility to <i>Pseudomonas syringae</i> pv. <i>tomatum</i> DC3000	[11,20]
		<i>fer</i>	Ethylene hypersensitivity, brassinosteroid insensitivity, abscisic acid hypersensitivity, an increase of s-adenosyl methionine synthesis, inhibition of jasmonic acid responses	[5-7,9]
		FER RNAi	Dwarf phenotype	[4,21]
		<i>fer</i>	Larger seed size	[22]
		<i>fer</i>	Salt hypersensitivity	[23]
		<i>fer</i>	Increased starch accumulation in a sucrose medium. Hypersensitivity to high carbon/nitrogen ratio	[14,24]
		<i>fer</i>	Reduced induction of ErbB3-binding protein 1, alteration of ribosome synthesis	[13,25]
		<i>fer</i>	Delay in the flowering time under long day condition	[26]
<i>Oriza sativa</i> L.		<i>flr2</i>	Enhanced resistance to <i>Magnaporthe oryzae</i> infection	[28]
		<i>flr11</i>	Enhanced resistance to <i>M. oryzae</i> infection	[28]
		FaMRL47RNAi	Fruit ripening acceleration	[29]
		<i>bups1</i>	PT overgrowth	[30]
BUPS1/2	<i>A. thaliana</i>	<i>bups1 bups2</i>	Enhanced bups1 phenotype: PT overgrowth	
ANXUR1/2	<i>A. thaliana</i>	<i>anx1 anx2</i>	Reduced fertility, PT burst	[19,31,32]
HERKULES1	<i>A. thaliana</i>	<i>herk1 the</i>	Dwarf plants	
		<i>herk1</i>	Tolerance to cadmium, copper, nickel, and zinc	[19,27,33]
HERKULES2	<i>A. thaliana</i>	<i>herk1</i>	PT overgrowth	
		<i>herk2</i>	Tolerance to cadmium, copper, nickel, and lead	[27]
ANJEA	<i>A. thaliana</i>	<i>herk anj</i>	PT overgrowth	[33]
		<i>the prc</i>	Rescues of hypocotyl growth but without <i>prc</i> cellulose deficiency phenotype	[34]
THESEUS1	<i>A. thaliana</i>	<i>the</i>	Hypersensitivity to lead and zinc, tolerance to nickel	[27]
		<i>herk1 the</i>	Dwarf plants	[4]
CAP	<i>A. thaliana</i>	<i>cap</i>	Altered PT growth in low calcium	[35]
		<i>cap</i>	RH bursting and bulging	[36]
CURVY	<i>A. thaliana</i>	<i>cro</i>	Distortion of trichomes, altered pavement morphology	[37]
MEDOS1-4	<i>A. thaliana</i>	<i>med1,2,3,4</i>	Reduced growth in presence of metal ions	[38]

PT: pollen tubes, RH: root hairs, *prc*: *A. thaliana* mutant *procuste*.

Como se mencionó previamente, los pelos radicales funcionan generalmente como la puerta de entrada de la bacteria a las raíces de la leguminosa durante la simbiosis (Ferguson *et al.* 2010). Se sabe que RALF1 y FER son importantes en el desarrollo de los pelos radicales, ya que líneas de *Arabidopsis* silenciadas en RALF1, o mutantes de *fer* presenta defectos en el desarrollo de los pelos radicales, formándose muy pocos y más pequeños (Fig. 3 y 4A) (Duan *et al.* 2010; Haruta *et al.* 2014; Zhu *et al.* 2020). A la fecha se conoce parte del mecanismo por el cual RALF1-FER regulan el desarrollo de estas células epidérmicas, así como diferentes interactores de la vía de señalización (Fig. 4A).

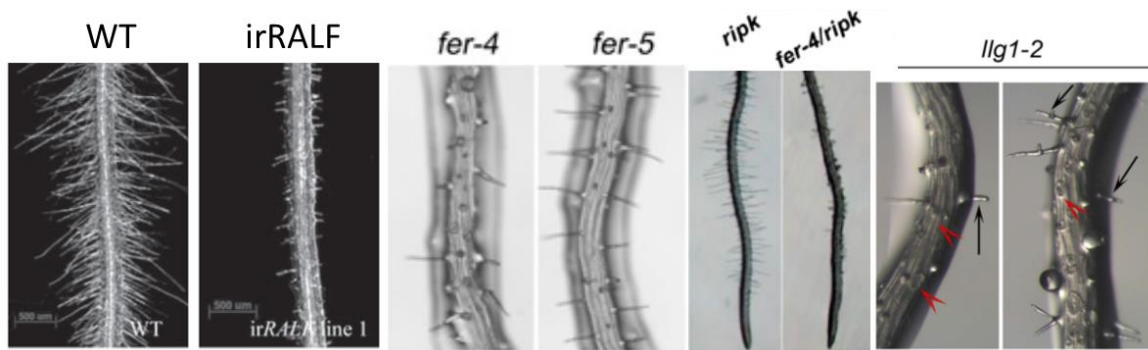


Figura 3. Fenotipo de mutantes de *fer*, *ripk* y *llg1*, y del silenciamiento de RALF sobre el desarrollo de los pelos radicales en *A. thaliana*. Las mutantes por pérdida de función de los genes *fer*, *ripk* y *llg*, así como el silenciamiento de RALF muestran un fenotipo similar, esto es, el crecimiento de los pelos radicales se detiene, y muchos de ellos estallan.

Una familia importante de interactores de CrRLK1Ls son las proteínas de membrana de unión a glicofosfatidil-inositol (GPI) LORELEI (LRL) y “LORELEI-LIKE GPI ANCHORING PROTEINS” (LLG), las cuales funcionan como co-receptores de FER, siendo requeridas para la localización de FER en la membrana plasmática y para su actividad (Li *et al.* 2015). Las mutantes *llg1* de *A. thaliana* muestran un fenotipo similar al observado en mutantes *fer*, así como las líneas silenciadas en RALF1 (Fig. 3). En la raíz principal de *A. thaliana*, FER interactúa con la cinasa citoplasmática RIPK (RPM1-induced protein kinase) (Du *et al.* 2016), la cual pertenece a la familia de cinasas citoplásmicas, involucradas en la transducción de señales procedentes de receptores tipo cinasas en la membrana (Zhang *et al.* 2010; Liu *et al.* 2011). Se ha demostrado que RIPK interactúa con FER en respuesta al tratamiento con RALF1, y la percepción del péptido promueve la fosforilación de ambas en un proceso interdependiente (Du *et al.* 2016). Las mutantes *ripk* de Arabidopsis muestran un fenotipo similar al descrito en mutantes *fer* y *llg*, (Fig. 3).

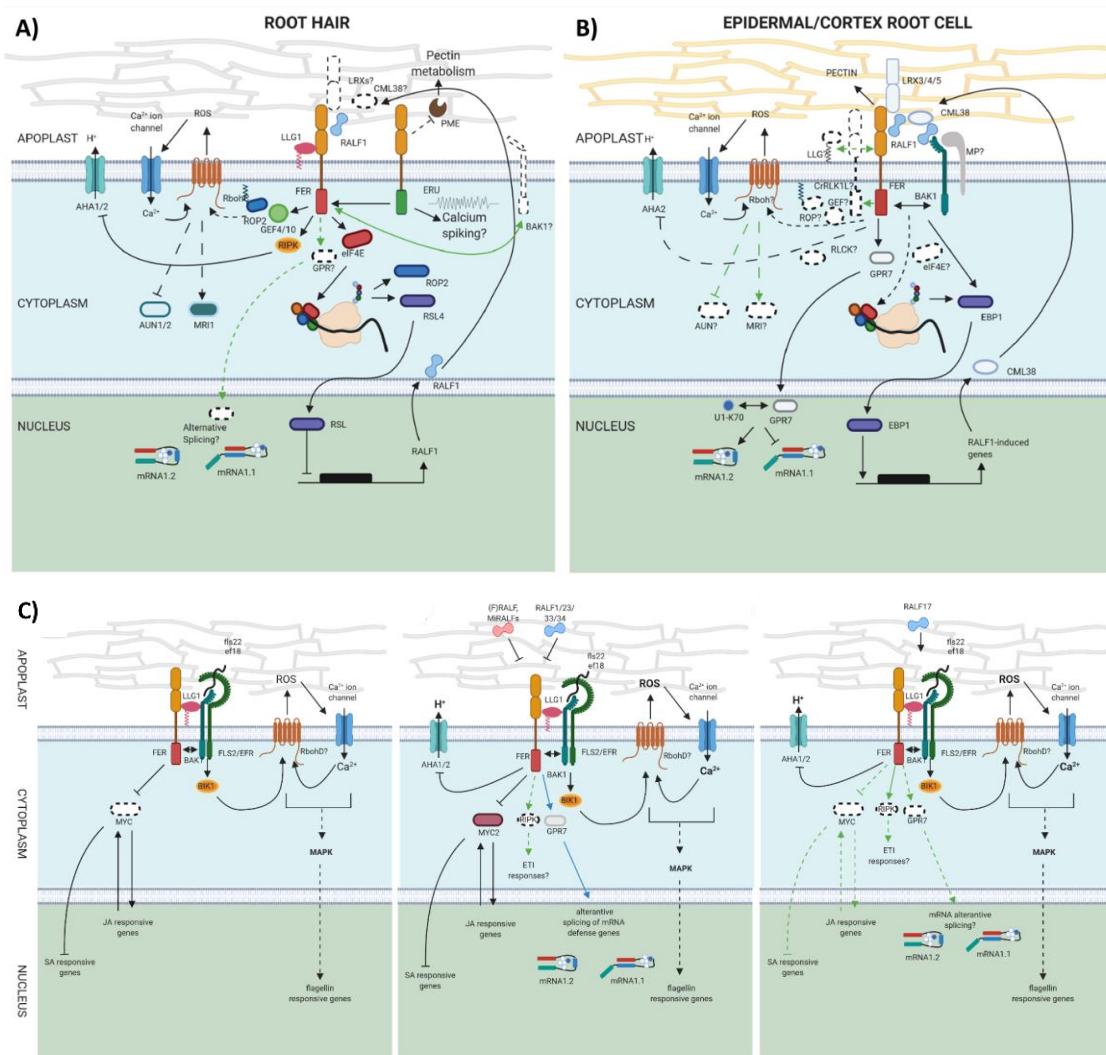


Figura 4. Mecanismo de acción de los receptores CrRLK1L en el desarrollo de los pelos radicales, la elongación de las células de la raíz y en estrés biótico. Esquema descriptivo del mecanismo desencadenado por FER durante el crecimiento del tubo polínico (A), la elongación de las células de la raíz (B) y en respuesta al estrés biótico (C). Los participantes conocidos se representan como figuras coloreadas, las líneas sólidas indican interacción directa mientras que las líneas punteadas indican interacción indirecta. Las puntas de flecha representan activación, los finales en línea representan represión, y la doble flecha indica una interacción inducida. Las figuras en blanco y líneas en verde indican participantes y vías sugeridas, respectivamente. CrRLK1L, *Catharanthus roseus* RECEPTOR-LIKE KINASE 1L; FER, FERONIA; ANX, ANXUR; ERU, ERULUS; HERK1, HERKULES 1; ANJ, ANJEA; BUPS1/2, BUDDHA'S PAPER SEAL 1/2; GEF, GUANINE EXCHANGE FACTOR; LRL, LORELEI; LLG, LORELEI-LIKE GPI ANCHORED PROTEIN; LRX, LEUCIN-RICH EXPANSIN; ROP, RAC OF PLANTS; AHA, H⁺-ATPase; BAK1, BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED RECEPTOR KINASE 1; CML38, CALMODULIN-LIKE 38; RLCK, RECEPTOR-LIKE CYTOPLASMIC KINASE; RIPK, RPM1-INDUCED PROTEIN KINASE; RALF, RAPID ALKALINIZATION FACTOR; AUN ANTUNIS; MRI, MARIS; GPR7, GLYCINE-RICH RNA BINDING PROTEIN 7; eIF4E1, EARLY TRANSLATION INITIATION FACTOR E1; EBP1, Erb3 BINDING PROTEIN 1; RSL4, ROOT HAIR DEFECTIVE 6-LIKE 4; PME; PECTIN METHYLESTERASE; ROS, reactive oxygen species (Solis-Miranda and Quinto 2021).

4.7 El papel de RALF, FER y RIPK en interacciones planta-microorganismos

Las plantas interactúan con una gran cantidad de microorganismos que se encuentran en la rizósfera. Por tanto, la planta debe ser capaz de diferenciar entre microorganismos benéficos y patógenos. Durante el proceso de simbiosis planta-rhizobia, el sistema inmune de la leguminosa excluye organismos patógenos de las raíces; esto ocurre a través de la precepción de moléculas en la superficie de la pared celular de las bacterias, mediante complejos cinasa en la superficie de la célula vegetal, y del reconocimiento de los efectores “inyectados” por la bacteria mediado por proteínas de resistencia (Cao *et al.* 2017). RIPK ha sido descrito como un componente esencial en la inmunidad de la planta. RIPK fosforila al modulador de la respuesta inmune “RPM1-interacting protein 4” (RIN4) al percibir efectores de la bacteria (Liu *et al.* 2011), el cual interactúa con bombas de protones de membrana plasmática, H⁺-ATPasas, regulando la apertura de estomas en las hojas y con el receptor de la respuesta inmune, desencadenando los mecanismos de respuesta en las plantas (Liu *et al.* 2009, 2011).

En línea con esto, se sabe que varios péptidos RALF y FER son moduladores de la respuesta inmune (Fig. 4C). Se ha descrito que mutantes *ralf1*, *fer* y *ripk* en *A. thaliana* poseen resistencia a la infección por *Powdery mildew* y *Pseudomonas syringae* (Kessler *et al.* 2010; Liu *et al.* 2011). Se sabe además que FER interactúa con el receptor “BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED RECEPTOR KINASE” (BAK1), promoviendo la formación de un complejo de respuesta inmune BAK1-“FLAGELLIN SENSING 2” (FLS2), el cual permite la percepción de la flagelina y la subsecuente respuesta inmune. La formación de este complejo es además inhibida por RALF1, 23, 33 y 34, e incrementado por RALF17 (Fig. 4C) (Shen *et al.* 2017; Stegmann *et al.* 2017). En el hongo fitopatógeno *Fusarium oxysporum* y en el nematodo fitopatógeno *Meloidogyne incognita* se ha descrito la presencia de genes que codifican péptidos funcionales con homología a RALF de las plantas (F-RALF y MiRALF) (Masachis *et al.* 2016; Thynne *et al.* 2017; Zhang *et al.* 2020). Estos péptidos F/Mi-RALF, promueven un incremento en la infección respecto a mutantes que carecen de éste (Masachis *et al.* 2016; Thynne *et al.* 2017; Zhang *et al.* 2020).

EL papel de un péptido RALF ha sido explorado en *M. truncatula* durante la nodulación (Combiér *et al.* 2008) (Fig. 2). Sin embargo, se desconoce el mecanismo de acción de dicho péptido, y si este se encuentra conservado en otras leguminosas, como frijol. Nosotros evaluamos el papel de *PvFER1* en raíces transgénicas de frijol durante la nodulación

mediante genética reversa (Tesis de maestría Jorge Esaú Solís Miranda (Solis-Miranda 2016)). Los resultados que obtuvimos indicaron que tanto el silenciamiento como la sobreexpresión de *AtFER* alteran el número de nódulos generados, observándose un aumento y una disminución en el número de los nódulos, respectivamente, a los 21 días post inoculación (Fig. 5A y B). Debido a que no se observaron diferencias en el número de nódulos a 7 ni 14 dpi (Fig. 5A y B), ni cambios importantes en los diámetros de dichos nódulos debido al silenciamiento o a la sobreexpresión (Fig. 5C y D), y solo se detectaron ligeros cambios en la actividad de la nitrogenasa en dichos nódulos (Fig. 5E y F) se puede descartar que esta diferencia en el número de los nódulos se deba a un retraso en el proceso. Estos datos sugieren que PvFER1 en frijol participa en regular el número de nódulos a tiempos tardíos, pero no en su desarrollo ni en su funcionamiento.

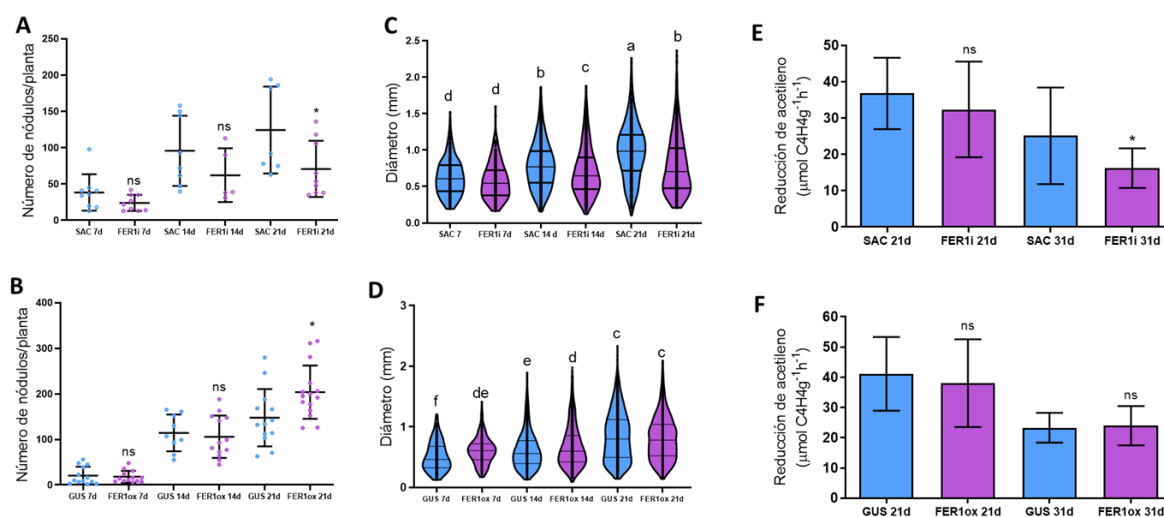


Figura 5. Fenotipo de raíces transgénicas silenciadas y sobreexpresantes *PvFER1* en la nodulación en frijol. Se observa la cantidad de nódulos (A y B), sus diámetros (C y D) y su tasa de reducción de acetileno (E y F) en las raíces transgénicas silenciadas y sobreexpresantes de *PvFER1* (puntos morados) y las raíces control (puntos azules) a 7, 14 y 21 dpi. Los puntos representan todos los datos, las líneas corresponden a la media \pm la desviación estándar ($n=12$); * $P < 0.05$, ** $P < 0.01$, ns=no significativo.

Los datos hasta aquí presentados indican que RALF-FER y RIPK forman parte de una misma vía que participan en diferentes procesos, algunos de los cuales están relacionados con la nodulación. Con base a todos estos antecedentes, proponemos analizar el papel de los genes que codifican RALF y RIPK que se expresan en raíces de frijol, durante la simbiosis con rizobias.

5. JUSTIFICACIÓN

El análisis del papel de *PvRALF*, *PvFER* y *PvRIPK* durante la simbiosis frijol-rhizobia generará conocimiento de frontera, y nos permitirá proponer un modelo de señalización mediada por *PvRALF* en esta interacción simbiótica.

6. HIPÓTESIS

PvRALF participa en el proceso de nodulación en frijol junto con *PvFER* y *PvRIPK*.

7. OBJETIVOS

a) General.

- Determinar el efecto de la sobreexpresión y el silenciamiento de los genes *PvRALF* y de *PvRIPK* durante la interacción simbiótica.

b) Específicos.

- Definir la familia génica de *PvRALF* y de *PvRIPK* en frijol.
- Evaluar la actividad de los promotores de los genes *PvRIPK* y *PvRALF* de interés, en raíces transgénicas de frijol.
- Evaluar la acumulación del transcrito de los genes de la familia *PvRALF* y de *PvRIPK* en raíces de frijol inoculadas y sin inocular con rhizobia.
- Evaluar el efecto del silenciamiento y la sobreexpresión de los genes *PvRALF* y *PvRIPK* de interés, en raíces transgénicas de frijol en el proceso de nodulación.
- Evaluar la expresión de genes marcadores de la nodulación en raíces transgénicas silenciadas y sobreexpresantes de los genes de interés.
- Proponer un posible modelo de señalización en raíces de frijol en simbiosis con rhizobia, que integre los genes evaluados.

8. METODOLOGÍA

8.1 Análisis bioinformático

Con la finalidad de identificar los péptidos de la familia RALF en frijol, se realizó una búsqueda informática mediante el algoritmo de BLASTP de la base de datos PHYTOZOME v12.1 (Goodstein *et al.* 2012), empleando la secuencia de RALF de *A. thaliana* como “query”. Para determinar las proteínas homólogas a RIPK en frijol, se empleó la secuencia de amino ácidos de RIPK de *A. thaliana* como “query” en una búsqueda por BLASTP en PHYTOZOME v12.1 (Goodstein *et al.* 2012). A continuación, se alinearon mediante el algoritmo de MUSCLE (Edgar 2004) las secuencias PvRALFs y PvRIPKs determinadas previamente y se analizaron mediante MEME (Bailey *et al.* 2009) para determinar los motivos conservados de dichas proteínas. Para la reconstrucción filogenética de RALF y de RIPK, primero se realizó un alineamiento múltiple mediante MUSCLE (Edgar 2004), incluyendo las secuencias de RALFs o RIPKs de *Phaseolus vulgaris*, *A. thaliana*, *M. truncatula* y *Glycine max*, seguido de una curación manual en el editor AliView (Larsson 2014). El árbol filogenético se reconstruyó por máxima verosimilitud (Guindon *et al.* 2010) mediante IQ-TREE 1.6.12 (Nguyen *et al.* 2015) con 1000 repeticiones de confiabilidad.

Para evaluar la organización, localización y estructura génica de los genes *RALF* y *RIPK* de frijol, se determinó la ubicación de los diferentes genes en el genoma de frijol mediante el recurso informático PhenoGram Plot (Wolfe *et al.* 2013), así como la estructura génica de los diferentes genes mediante el programa Gene Structure Display Server 2.0 (Sun *et al.* 2015). Por último, para seleccionar genes *RALF* y *RIPK* de frijol con función potencial en el proceso de nodulación, se analizaron los niveles de acumulación de transcritos de los diferentes genes de la familia génica *RALF* y de los genes *RIPK* en frijol, en el Atlas de Expresión Génica de *P. vulgaris* (O'Rourke *et al.* 2014).

8.2 Diseño de oligonucleótidos

Para la cuantificación de la acumulación de transcritos de los genes de la familia *RALF* y de los genes homólogos a *RIPK* en las raíces de frijol, se diseñaron oligonucleótidos específicos para cada secuencia. Para esto, se diseñaron los diferentes oligonucleótidos en regiones con baja identidad entre ellas. Además, se diseñaron oligonucleótidos para la cuantificación de transcritos de ocho genes de la familia CrRLK1L de frijol, que se expresan en raíz. Para la generación de las construcciones para el análisis de la actividad de promotor de los genes

RALF y *RIPK* seleccionados, así como de los dos genes *FERONIA*, se generaron oligonucleótidos para amplificar hasta 2000 pb corriente arriba del sitio de inicio de la traducción de cada uno de ellos.

En el caso de los oligonucleótidos para las construcciones de sobreexpresión de los genes *RALF* y *RIPK* seleccionados, estos se diseñaron para amplificar la secuencia codificante excluyendo el codón de paro, con la finalidad de realizar fusiones transcripcionales con la secuencia codificante de la proteína verde fluorescente (GFP). Para las construcciones de silenciamiento, se diseñaron oligonucleótidos que amplifican aproximadamente 120 pb de la región 5' o 3' no traducida de cada gen de interés, para que el silenciamiento sea específico. Debido a que la clonación de estos fragmentos se realizará inicialmente a través del vector donador pENTR Directional TOPO de Invitrogene (James L. Hartley, Gary F. Temple 2000), se adicionó al inicio de cada oligonucleótido directo la secuencia de 4 bases CACC, de acuerdo a las especificaciones del fabricante. La lista completa de oligonucleótidos puede consultarse en la tabla S1.

8.3 Generación de las construcciones

Con la finalidad de analizar la actividad de los promotores de los genes *RALF* y *RIPK* seleccionados, así como de los dos genes *FER* de frijol, se generaron distintas construcciones portadoras de la región promotora de cada uno de ellos. Para ello se amplificó por PCR, 2000 pb corriente arriba del sitio de inicio de la traducción de cada uno de los genes de interés. Para el análisis del fenotipo de la sobreexpresión y el silenciamiento de los genes *RALF* y *RIPK* seleccionados, se generaron las construcciones para cada condición. Para ello se amplificó por PCR la región codificante para el caso de la sobreexpresión, y un fragmento de 120 pb de las regiones 5' para el silenciamiento, de cada uno de los genes de interés. A continuación, los distintos fragmentos se clonaron en el vector donador pENTR Directional TOPO, y las construcciones se comprobaron por secuenciación. Los fragmentos se recombinaron en un vector “lanzadera”, basados en la tecnología Gateway de Invitrogene (James L. Hartley, Gary F. Temple 2000).

Para el análisis de la actividad de cada promotor de interés, se empleó el vector “lanzadera” pBGWSF7.0, el cual permite la expresión de la proteína reportera β -glucuronidasa (GUS) en fusión con la proteína GFP, bajo el control del promotor de interés (Fig. S1) (Karimi et al.

2002). Para la sobreexpresión, se empleó el vector pH7fwg2d (Karimi *et al.* 2002), el cual permite la transcripción de la secuencia clonada, en fusión transcripcional con la GFP, bajo el control del promotor 35S (Fig. S2). Por su parte, para el silenciamiento se utilizó el vector ptdT-RNAi (Valdés-López *et al.* 2008), el cual expresa bajo control del promotor 35S un duplicado de la región deseada separadas por una secuencia que permite la formación de una estructura tipo asa, necesaria para generar el RNA de interferencia para el gen de interés (Fig. S3). Este vector permite además la expresión de la proteína roja fluorescente (tdTomato) bajo promotor 35s en las raíces transformadas (Fig. S3). Tanto los vectores para el análisis de promotor, análisis de sobreexpresión y de silenciamiento presentan una secuencia para la expresión de proteínas fluorescentes (Fig. S1, S2 y S3), permitiendo la selección de raíces transformadas mediante fluorescencia.

8.4 Germinación de semillas

Con la finalidad de obtener material para el análisis de la acumulación de transcritos en raíces WT de frijol y para la generación de raíces transgénicas, se germinaron semillas de frijol negro Jamapa. Para esto, inicialmente se desinfectan las semillas mediante un lavado con etanol al 90 % y uno con cloro al 20 %. A continuación, las semillas desinfectadas se colocan sobre servitoallas humedecidas con medio nutritivo Fahraeus, y se colocan en cámaras de crecimiento a 28 °C durante 2 días (dpg). Para los análisis en raíces silvestres (WT) las plantas de 2dpg se transfirieron a macetas con sustrato vermiculita estéril, y se mantuvieron en riego con medio Fahraeus estéril sin nitrógeno, hasta el momento de su colecta.

8.5 Generación de plantas de frijol con raíces transgénicas

Con el fin de obtener raíces transgénicas portadoras de las diferentes construcciones de interés, se empleó el protocolo descrito por Estrada-Navarrete *et al.*, 2007 (Estrada-Navarrete *et al.* 2007). Primero se germinaron semillas de frijol con la metodología descrita en el apartado 5.4. A continuación, se seleccionaron plantas de 2 dpg, y se les realizó una herida en la raíz debajo del cotiledón; sobre la herida se colocó una azada de la clona de *Agrobacterium rhizogenes* K599 portadoras de las construcciones de interés. Las plantas se mantuvieron en medio Fahraeus estéril sin nitrógeno, hasta que se observó la emergencia de las raíces transgénicas en la zona de la herida (aproximadamente 10 a 12 días después de la herida). Las raíces peludas se analizaron bajo el microscopio de epifluorescencia y tanto las raíces WT como las raíces transgénicas que no mostraron fluorescencia fueron retiradas.

8.6 Análisis de la actividad de promotor

Con la finalidad de determinar la actividad espacio-temporal del promotor de los genes *RALF* y *RIPK* de interés, así como de los dos genes *FER*, en raíces transgénicas de frijol durante la nodulación, se generaron plantas con raíces transgénicas portadoras de las construcciones correspondientes (Fig. S1). Con este objetivo, se siguió el protocolo descrito en el apartado 7.5. Las raíces generadas se observaron bajo el microscopio de epifluorescencia para la selección de raíces transgénicas fluorescentes (GFP) portadoras de la construcción deseada para analizar la actividad del promotor. A continuación, las plantas se transfirieron a macetas con vermiculita, y se inocularon con la bacteria *R. tropici* CIAT 899. Las plantas se retiraron de la vermiculita a los 5, 7, 14 y 21 días post inoculación (dpi).

La actividad del promotor se analizó a través del gen reportero GUS mediante la tinción GUS (Lee and Schöffl 1997). Se analizaron las raíces de un total de 15 plantas por promotor analizado y raíces de 15 plantas transformadas con el vector vacío, provenientes de tres replicas biológicas. Las raíces seleccionadas se colocaron en cajas Petri, y se lavaron con las soluciones I (NaH_2PO_4 0.5 M, Na_2HPO_4 0.5 M, $\text{K}_3\text{Fe}(\text{CN})_6$ 0.1 M, $\text{K}_4\text{Fe}(\text{CN})_6$ 0.1 M) y II (solución I, EDTA 0.5 M, TRITON 10%) respectivamente. Posteriormente, las raíces se incubaron en la solución de tinción GUS (solución II, X-Gluc 50 mM) de 30 min a 2 hrs (Lee and Schöffl 1997), dependiendo de cada promotor analizado. Las raíces se lavaron con agua destilada estéril para detener la reacción y se aclararon con cloro al 1%. Finalmente, las raíces se almacenaron en glicerol (30%) + DMSO (1%) para su posterior observación en el microscopio de campo claro para el análisis de la actividad de GUS.

8.7 Medición de los niveles de transcritos por RT-qPCR

Con la finalidad de evaluar los niveles de acumulación de transcrito de los genes *RALF*, *RIPK* y *CrRLK1L* bajo estudio, así como de los genes marcadores de diferentes etapas de la nodulación en raíces de frijol, se extrajo RNA de dichos tejidos bajo los distintos tratamientos y se analizaron por PCR en tiempo real (RT-qPCR). Para los análisis en raíces de frijol WT se seleccionaron plántulas de 2 dpv por cada tiempo de tratamiento, y se trasplantaron en macetas con vermiculita. Las plántulas se inocularon con *R. tropici* CIAT 899 o únicamente con medio Fahraeus para el tratamiento sin inocular. A los tiempos seleccionados (5, 7, 14 y 21 dpi) se colectaron las raíces de las plantas y se congelaron con nitrógeno líquido hasta el momento de la extracción de RNA. Para los análisis en raíces transgénicas, estas se generaron

conforme a lo descrito en el apartado 7.5, y se seleccionaron por epifluorescencia. Las plantas se transfirieron a macetas con vermiculita y se inocularon con *R. tropici* CIAT 899. Las plantas se regaron con medio Fahraeus para los análisis sin nitrógeno, y con medio Fahraeus adicionado con nitrato para los análisis de 5 y 2 mM de nitrógeno (Tabla S2).

La extracción del RNA se realizó mediante el protocolo de extracción de RNA con TriZOL de Sigma, siguiendo la metodología del proveedor. La integridad y concentración del RNA se comprobó en un espectrofotómetro NanoDrop2000 (Thermo Fisher Scientific, Waltham, MA, USA). Para eliminar el DNA contaminante, las muestras de RNA se incubaron con DNasa libre de RNasas (1 U/ μ l; Roche, Basel, Switzerland); el DNA complementario (cDNA) se sintetizó utilizando la Transcriptasa Reversa de Thermo Scientific (RevertAid Reverse Transcriptase 200 U/ μ l), siguiendo las instrucciones del proveedor. La cuantificación de los niveles de transcrito se llevó a cabo mediante RT-qPCR usando el Maxima SYBR Green/ROX qPCR Master Mix (2X) (Thermo Scientific, USA), en el equipo de tiempo real QuantStudio 5 (Applied Biosystems, Waltham, MA, USA). Los niveles del transcrito del gen “*insulin degrading enzyme*” (IDE) y del *factor de elongación 1 α* (*EF1 α*), que se expresan constitutivamente (Borges *et al.* 2012; Montiel *et al.* 2012) se usaron como genes normalizadores. La lista completa de oligonucleótidos se encuentra en la tabla S3.

8.8 Análisis de los fenotipos en la nodulación

Para determinar el efecto del silenciamiento o la sobreexpresión de los genes *RALF* y *RIPK* de interés en las raíces transgénicas de frijol en el proceso de nodulación, se generaron plantas de frijol con raíces transgénicas portadoras de las construcciones correspondientes de acuerdo al protocolo descrito en el apartado 7.5. El aumento o la disminución de la acumulación del nivel de transcrito del gen correspondiente en las raíces transgénicas se determinó mediante qRT-PCR (apartado 7.7).

Las plantas con raíces sobreexpresantes o silenciadas se trasplantaron en vermiculita y fueron inoculadas con *R. tropici*-GUS (*Rhizobium* que expresa GUS bajo un promotor constitutivo, (Vinesa *et al.* 2003). Para analizar el fenotipo en la nodulación, se evaluó el número de nódulos formados, los diámetros de los mismos y la actividad de la nitrogenasa en plantas regadas con medio Fahraeus sin nitrógeno (0 mM) hasta el momento de su colecta.

8.8.1 Cuantificación del número y diámetro de los nódulos

Con la finalidad de determinar el efecto del silenciamiento o la sobreexpresión de los genes *RALF* y *RIPK* seleccionados en el proceso de nodulación, se evaluó el fenotipo analizando el número de nódulos formados y su tamaño. Para esto, se generaron plantas de frijol con raíces transgénicas portadoras de las construcciones de silenciamiento o de sobreexpresión (apartado 7.5) y se transfirieron a macetas con vermiculita. Las plantas se inocularon con *R. tropici*-GUS, y se regaron con medio Fahraeus a 0, 2 y 5 mM de nitrógeno. Las raíces se evaluaron a los 7, 14, 21 y 31 dpi en el caso de las plantas a 0 mM de nitrógeno, y a 21 y 31 dpi en plantas con 2 y 5 mM de nitrógeno.; para esto, las raíces se limpiaron de vermiculita y las raíces transgénicas fluorescentes se seleccionaron y se tiñeron con GUS (apartado 7.7); posteriormente se obtuvieron imágenes y se analizaron mediante el software imageJ.

8.8.2 Medición de la actividad de la nitrogenasa

Con el objetivo de evaluar el efecto del incremento o la disminución de los niveles del transcrito de los genes *RALF* y de los genes *RIPK* de interés, en la actividad de la nitrogenasa, se midió de manera indirecta la fijación de nitrógeno, a través de la evaluación de la reducción de acetileno (Hardy *et al.* 1968). Primero se generaron plantas de frijol con raíces transgénicas portadoras de las construcciones de silenciamiento o de sobreexpresión conforme a lo descrito en el apartado 7.5. Las plantas inoculadas se retiraron de la vermiculita a los 14, 21 y 30 dpi, y se realizó una segunda selección de raíces fluorescentes. Las raíces seleccionadas se colocaron en frascos de vidrio sellados completamente; a cada frasco se le retiró 2 mL de aire y se le adicionó 2 mL de acetileno. Los frascos se incubaron a temperatura ambiente un mínimo de 30 minutos, y la cantidad de acetileno reducido a etileno se midió mediante cromatografía de gases. Los valores obtenidos se procesaron y ajustaron en función del tiempo de incubación y del peso seco del total de los nódulos analizados de cada planta (Hardy *et al.* 1968)

8.9 Análisis del fenotipo sobre la inhibición de la nodulación por nitrógeno

Para determinar el efecto del silenciamiento o la sobreexpresión de los genes *RALF* y *RIPK* de interés, así como de *FER1* y *FER2* en raíces transgénicas de frijol en la inhibición de la nodulación por nitrógeno, se generaron plantas de frijol con raíces transgénicas portadoras de las construcciones correspondientes de acuerdo al apartado 7.5. Para analizar la inhibición de la nodulación por nitrógeno, las plantas se regaron con medio Fahraeus con nitrógeno, a

una concentración de 2 y 5 mM, y se evaluó el número de nódulos formados, así como el diámetro de los mismos, de acuerdo al protocolo descrito en el apartado 7.8.1.

9. RESULTADOS

9.1 Análisis *in silico* de genes RALF y RIPK de frijol

9.1.1 Identificación y análisis *in silico* de los miembros de la familia RALF en frijol

Los péptidos RALF conforman una familia de péptidos ricos en cisteína, exclusivo de las plantas terrestres (Campbell and Turner 2017). Mediante una búsqueda informática por BLASTP y el análisis de motivos mediante MEME, se determinó que en el genoma de frijol se encuentran codificados nueve RALFs, los cuales poseen una región amino terminal variable y una región carboxilo terminal conservada (Fig. S4). En la región carboxilo terminal, se localizan el motivo de procesamiento RRXL, el motivo YISI necesario para la actividad del péptido, y 4 cisteínas características de los péptidos tipo RALF (Fig. 6). Este resultado es consistente con datos publicados previamente (Campbell and Turner 2017).

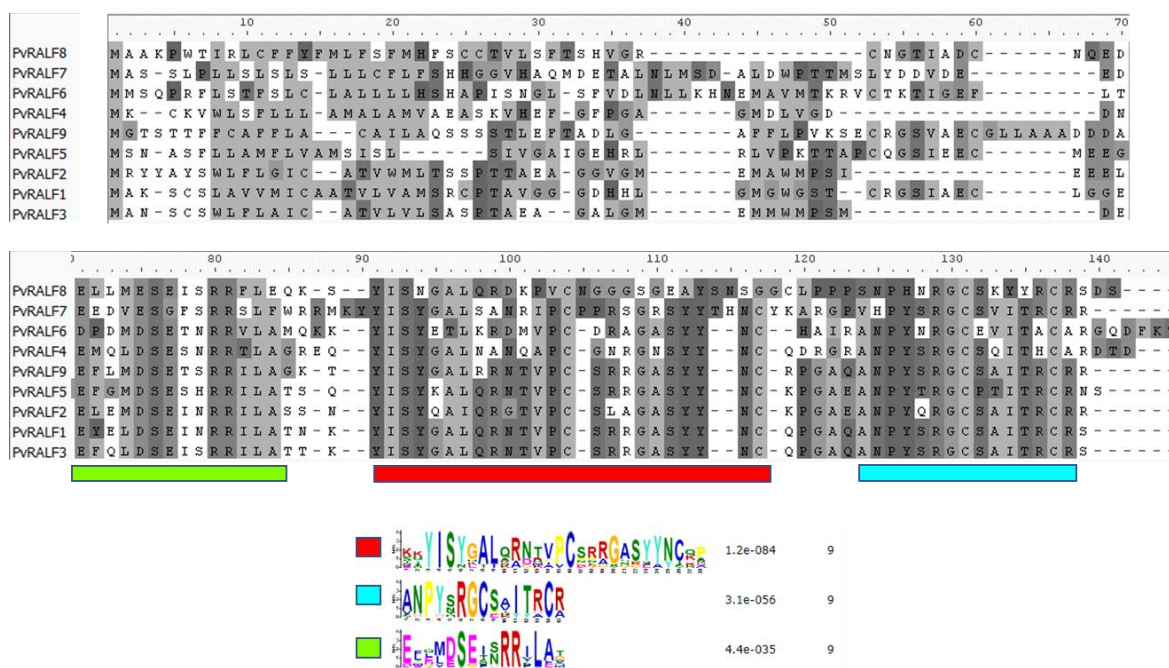


Figura 6. Motivos y alineamiento de la familia de péptidos RALF en frijol. Alineamiento de las nueve secuencias proteicas de frijol encontradas tras la búsqueda informática de RALF. Las barras de colores representan los motivos sobrerepresentados encontrados por MEME. Se muestran los logos correspondientes al código de colores.

A continuación, se compararon algunas características de los nueve genes *RALF* de frijol con los de la planta modelo *A. thaliana*, y la leguminosa *M. truncatula*, la cual genera nódulos de

tipo indeterminado. La localización cromosómica de los diferentes genes en sus respectivos genomas se llevó a cabo mediante el programa PhenoGram Plot y la estructura génica mediante Gene Structure Display Server 2.0. En la Fig. S5 se observa un esquema representativo de los cromosomas de *A. thaliana* (Fig. S5A), de *M. truncatula* (Fig. S5B) y de frijol (Fig. S5C). En dicho esquema, se puede apreciar que el número de genes *RALF* de Arabidopsis supera más de tres veces el número de genes presentes en ambas leguminosas, las cuales poseen un número similar de genes *RALF* entre ellas (nueve en frijol y doce en *M. truncatula*). En el caso de Arabidopsis, puede verse una gran cantidad de genes *RALF* formando clústeres, lo cual sugiere una rápida expansión de la familia debido a divisiones múltiples (Campbell and Turner, 2017) (Fig. S6A). Al analizar la localización de los genes *RALF* en frijol y de *M. truncatula*, se deduce que dicha expansión no ocurrió en estas leguminosas, y por tanto no se generaron clústeres de genes (Fig. S6B y C).

Con el objetivo de hacer un análisis filogenético de los péptidos RALF de frijol, 92 secuencias RALF procedentes de cinco especies: *A. thaliana*, *M. truncatula*, *G. max* y *L. japonicus* y frijol, fueron utilizadas. Los resultados obtenidos se muestran en la Fig. 7, en donde se puede observar que las 92 secuencias se agruparon en 4 clados principales, los cuales fueron numerados en relación a un análisis descrito previamente (Campbell and Turner 2017). En el clado 1, se ubica la proteína de *A. thaliana* RALF1 (AT1G02900), la cual está cercana filogenéticamente a las proteínas codificadas en los genes *Phvul.007G197000*, *Phvul.008G122900* y *Phvul.006G089900* en frijol. En el clado 3A se ubica la proteína MtRALF1 (*Medtr3g084350*) de *M. truncatula*, que se ha descrito previamente por poseer un papel en la nodulación en esta planta: esta proteína se encuentra cercana filogenéticamente a la proteína codificada por el gen *Phvul.001G266400* de frijol (Fig. 7).

Estos resultados indican que los nueve genes *PvRALF* presentan las firmas características de la familia RALF, presentan regiones codificantes con tamaños similares y sin intrones, sugiriendo un alto grado de conservación.

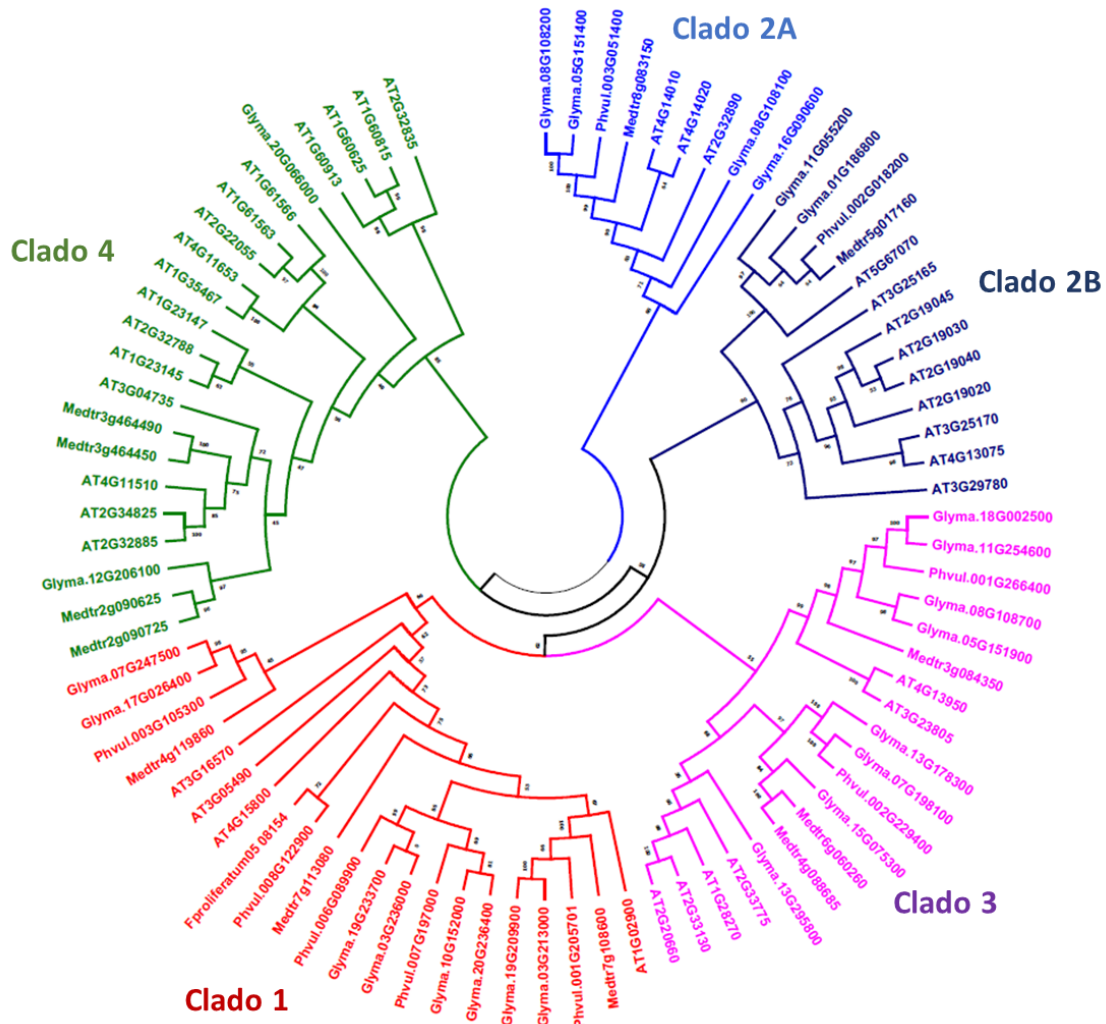


Figura 7. Árbol filogenético de la familia de péptidos RALF. Cladograma que muestra la reconstrucción filogenética de los péptidos RALF para *A. thaliana* y 4 fabáceas (*P. vulgaris*, *M. truncatula*, *G. max* y *L. japonicus*). A la derecha se marcan clados en los cuales se pueden agrupar los miembros, basado en el reporte de Campbell y Turner de 2017. En el clado 1 se ubica RALF1 de Arabidopsis, y en el clado 3A se localiza MtRALF1.

9.1.2 Análisis de la expresión de los genes *RALF* de frijol

Con la finalidad de seleccionar genes de la familia RALF frijol que tienen potencialmente una función en la nodulación, se realizó un análisis *in silico* de sus perfiles de expresión en diferentes tejidos de frijol, de acuerdo a lo reportado en el PvGEA (O'Rourke *et al.* 2014). Se observa que los genes *RALF2*, *RALF4*, *RALF5*, *RALF6* y *RALF8* presentan los niveles de expresión baja de los nueve genes o nula, los genes *PvRALF7* y *PvRALF9* muestran niveles de expresión intermedia, y los genes *PvRALF1* y *PvRALF3* presentan la expresión más alta de todos ellos (Fig. 8A); *PvRALF1* tiene su máxima expresión en raíces con 5 días de inoculación (R5) mientras que *PvRALF3* se observa más elevado en raíz completa (RT) (Fig.

8A). Al analizar los niveles de expresión de *PvRALF6* de manera individual, se observa que, aunque los valores son bajos, el nivel máximo de expresión de este gen es en nódulos de 5 dpi (Fig. 8B).

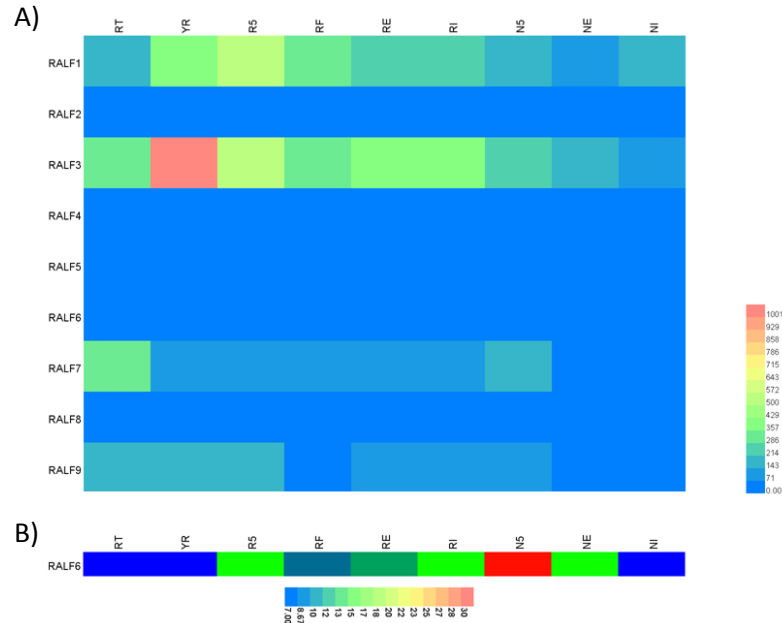


Figura 8. Mapa de calor representativo de la expresión de los nueve genes *RALF* de frijol. Mapa de calor de la expresión de los nueve genes *RALF* en frijol construido a partir de datos PvGEA (O'Rourke *et al.*, 2014). (A) Los niveles de expresión de los nueve miembros de la familia *RALF* de frijol en diferentes tejidos son representados de acuerdo a la leyenda de colores. (B) Se muestra de manera individual los valores de expresión del gen *RALF6* de frijol en diferentes tejidos de acuerdo a la leyenda de colores. RT- ápices de raíz colectada de plantas fertilizadas en el estado de desarrollo de aparición del segundo trifolio; YR- raíces completas colectadas de plantas en el estado de desarrollo de aparición del segundo trifolio; R5-raíces completas separadas de nódulos de 5 dpi; RF-raíces completas de plantas fertilizadas colectadas al mismo tiempo que RE y RI; RE-raíces completas separadas de nódulos fijadores de 21 dpi; RI- raíces completas separadas de nódulos no fijadores de 21 dpi; N5-nodulos de 5 dpi; NE-nódulos fijadores de 21 dpi; NI-nódulos no fijadores de 21 dpi.

Para su posterior análisis funcional, se seleccionaron los péptidos codificados en los genes de frijol *PvRALF1* (Phvul.007G197000) y *PvRALF6* (Phvul.001G266400), basados en la cercanía filogenética con *RALF1* de *Arabidopsis* y *MtRALF1* de *M. truncatula*, respectivamente (Fig. 7), en su expresión en raíces de frijol inoculadas y en nódulos (Fig. 8), y en la presencia de los dominios necesarios para la proteólisis (RRXL) y su función (YISY, y los residuos cinasa) (Fig. 6).

9.1.3 Determinación y análisis *in silico* de los genes *RIPK* de frijol

RIPK es una cinasa citoplásmica, perteneciente a la familia RLK citoplásmicos-VII (CRLK-VII), la cual ha sido caracterizada como un componente de la respuesta inmune de la planta (Liu *et al.* 2011). Mediante una búsqueda informática por BLASTP empleando la secuencia

RIPK de *Arabidopsis* como “query” y el análisis de sus motivos mediante MEME, se determinó que en frijol existen seis PvRIPKs (Fig. 9), los cuales presentan ocho motivos sobrerrepresentados (Fig. S7). Seis de los ocho motivos encontrados, se localizan en el dominio cinasa, abarcándolo completamente (Fig. 9), lo cual sugiere una alta conservación.

Posteriormente, se analizó la localización y la estructura génica de los seis genes *PvRIPK* de frijol, y se compararon con los genes *RIPK* en *A. thaliana* y *M. truncatula* (Fig. S8). Comparando los genomas de dichas especies, se observa que hay dos genes más en el genoma de ambas leguminosas (seis) que en el genoma de *A. thaliana* (cuatro), presentando una distribución aislada en los cromosomas en las tres especies (Fig. S8). Se encontraron en su estructura génica intrones de tamaños y posiciones variables en las tres especies (Fig. S9); Los genes *RIPK* de *A. thaliana* presentan entre tres y cuatro intrones (Fig. S9A), mientras que en el caso de los genes en las dos leguminosas se observa que cinco de los seis genes poseen cuatro intrones, y uno presenta un quinto intrón, localizado en la región 5' no traducida (UTR) (Fig. S9B y C). Además, se observa una gran variación en los tamaños de los transcritos primarios entre las tres especies; en *A. thaliana* van de 2000 a 2500 pb, en *M. truncatula* de entre 2200 a 3700 pb y en frijol desde 2200 hasta 200 pb (Fig. S9).

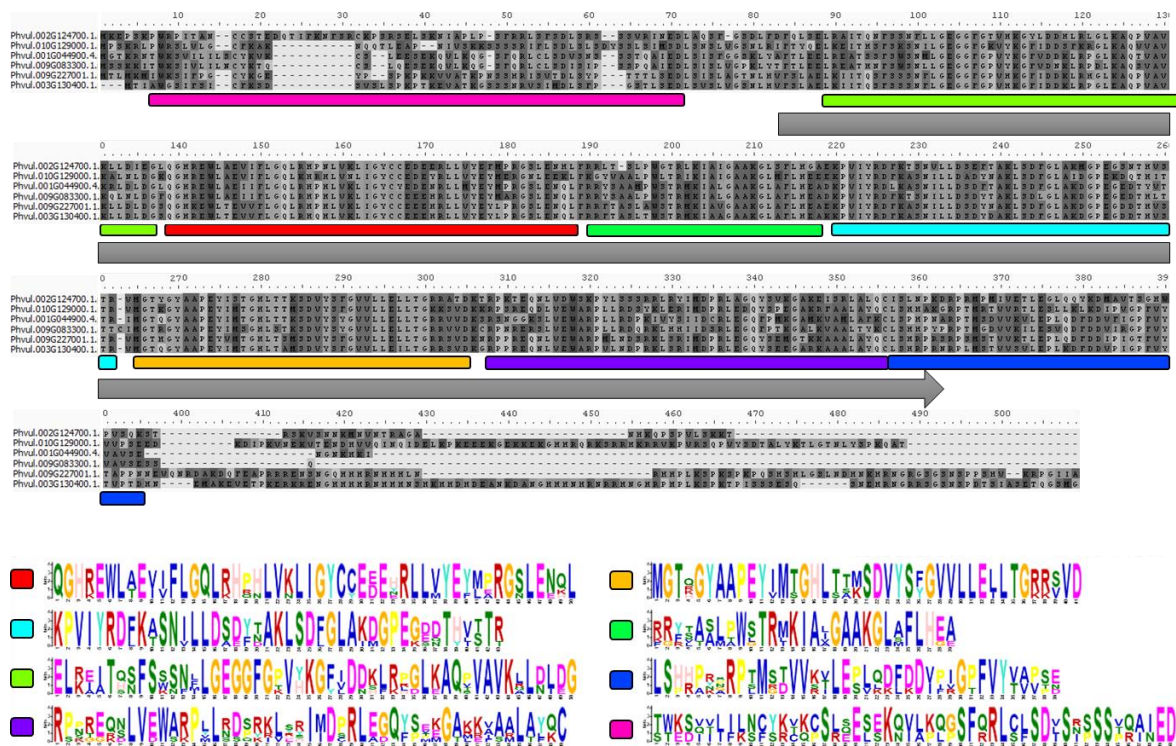


Figura 9. Motivos y alineamiento de la familia de péptidos RIPK en frijol. Alineamiento de las 6 seis secuencias proteicas de frijol encontradas tras la búsqueda informática de PvRIPK. Las barras de colores representan los motivos sobrerrepresentados encontrados por MEME, mientras que la línea gris representa la localización del dominio cinasa. Se muestran los logos correspondientes al código de colores.

Para identificar el posible gen homólogo de *RIPK* de *A. thaliana* en frijol, se realizó un análisis filogenético incluyendo las proteínas RIPK de frijol, *A. thaliana*, *M. truncatula* y *G. max*. Con este objetivo, se realizó la reconstrucción filogenética empleando las 29 secuencias proteicas procedentes de estas cuatro especies de plantas, en la cual se agruparon en tres clados distintos (Fig. 10). En el clado 1, el cual incluye a RIPK de *A. thaliana* se agruparon dos proteínas de frijol, denominadas PvRIPK1 y PvRIPK2. En el clado dos se agruparon tres de las proteínas de frijol analizadas, PvRIPK3/4/5. En el caso de PvRIPK6, ésta se agrupó en el clado tres. Estos resultados indican un alto grado de conservación del dominio cinasa en las seis cinasas PvRIPK, y la región codificante de estos genes *PvRIPK* presenta longitudes variables debido a los intrones.

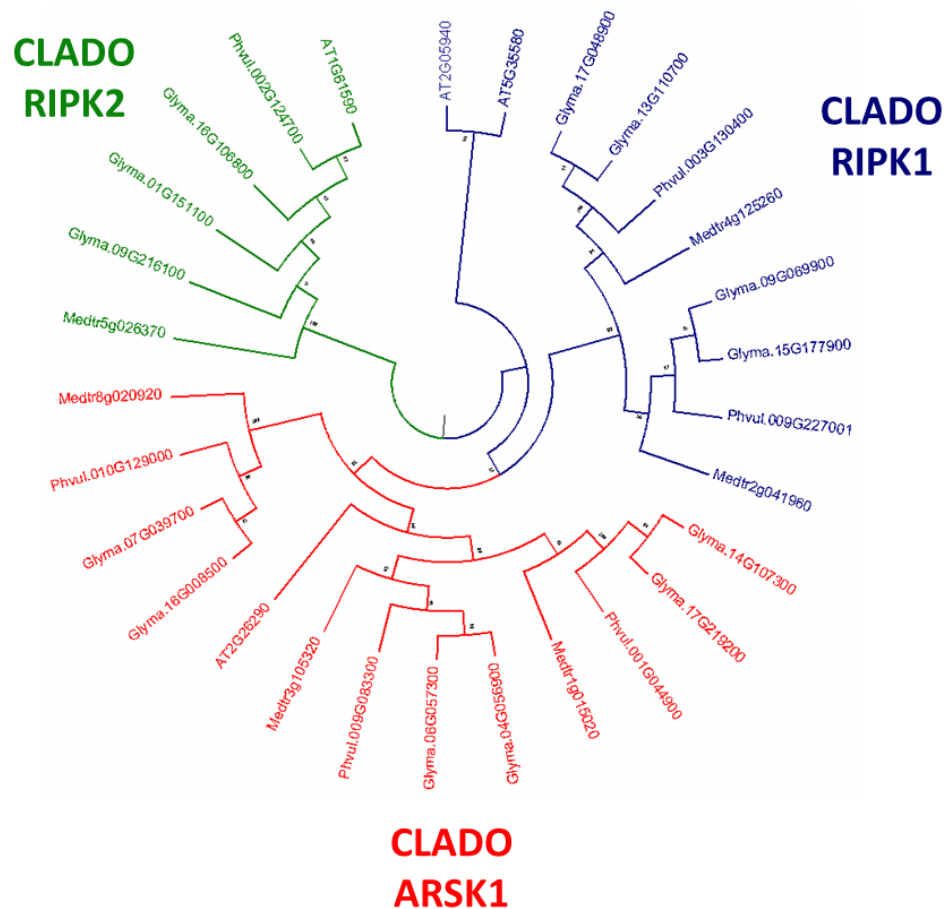


Figura 10. Árbol filogenético de RIPK. En este filograma se muestra la reconstrucción filogenética de las proteínas emparentadas con RIPK: 4 proteínas de *A. thaliana*, 6 de *P. vulgaris*, 6 de *M. truncatula* y 13 de *G. max*. A la derecha se marcan 3 clados en los cuales se pueden agrupar. El clado RIPK1 agrupa dos proteínas de *A. thaliana*, a RIPK (AT5G05940) y a PBL13 (AT5G35580), una cinasa descrita como regulador negativo de la respuesta inmune (Lin *et al.*, 2015). El clado ARSK1 incluye la proteína de *A. thaliana* ARSK1 (AT2G26290), una cinasa específica de raíz que se induce por estrés abiótico (Hwaan y Goodman, 1995). El clado RIPK2 incluye una proteína de *A. thaliana* (AT1G61590).

9.1.4 Análisis de la expresión de los genes *RIPK* de frijol

Con el objetivo de identificar los genes que potencialmente tienen un papel en la nodulación, se procedió a analizar la expresión de los genes *RIPK* de frijol seleccionados. Para esto se generó un mapa de calor representativo de los datos de expresión de dichos genes provenientes del PvGEA (Fig. 11). En el mapa de calor, se observa que la expresión de los genes *PvRIPK4*, 5 y 6 no fue detectada en los tejidos analizados, mientras que *PvRIPK1* y 3 presentan baja expresión en raíces y nódulos (Fig. 11). *PvRIPK2* presentan la mayor expresión de todos los genes en los tejidos evaluados, principalmente en nódulos (Fig. 11). Con base en todos estos resultados, se seleccionó al gen *PvRIPK2* para su posterior análisis en la nodulación en frijol.

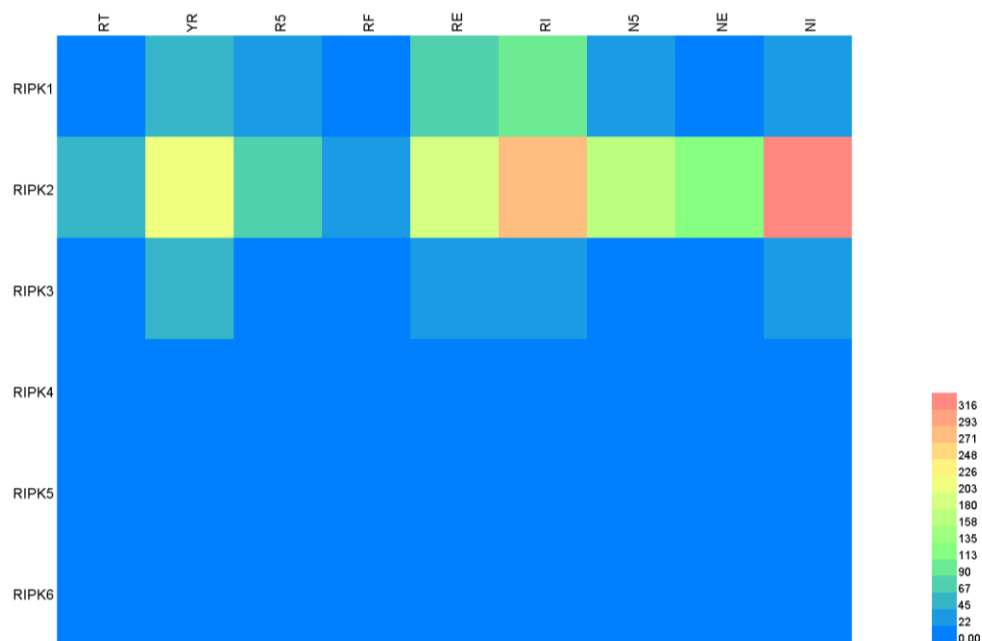


Figura 11. Mapa de calor representativo de la expresión de los seis genes *RIPK* de frijol de acuerdo al atlas de frijol (PvGEA). Mapa de calor representativo de la expresión de los seis genes de la familia *RIPK* en frijol, construido a partir de datos del PvGEA (O'Rourke *et al.* 2014). Los niveles de expresión son representados de acuerdo a la leyenda de colores. RT- ápices de raíz colectada de plantas fertilizadas en el estado de desarrollo de aparición del segundo trifolio; YR- raíces completas colectadas de plantas en el estado de desarrollo de aparición del segundo trifolio; R5-raíces completas separadas de nódulos de 5 dpi; RF-raíces completas de plantas fertilizadas colectadas al mismo tiempo que RE y RI; RE-

raíces completas separadas de nódulos fijadores de 21 dpi; RI- raíces completas separadas de nódulos no fijadores de 21 dpi; N5-nódulos de 5 dpi; NE-nódulos fijadores de 21 dpi; NI-nódulos no fijadores de 21 dpi.

9.2 Acumulación de los transcritos *PvRALF1*, *PvRALF6* y *PvRIPK2* en raíces inoculadas con rhizobia y sin inocular

Con el objetivo de analizar los niveles de acumulación de los transcritos de los genes *PvRALF1*, *PvRALF6* y *PvRIPK2* durante la nodulación, se generaron plantas de frijol inoculadas con rhizobia y sin inocular. Se colectaron las raíces y se extrajo RNA de acuerdo con el protocolo descrito en el apartado 7.7. Para el análisis por RT-qPCR de cada uno de los genes a evaluar se utilizaron oligonucleótidos específicos, lo cual nos permite detectar únicamente los transcritos del gen de interés.

Al comparar los niveles del transcrito *PvRALF1* en raíces inoculadas respecto a raíces sin inocular (Fig. 12A), se observa una diferencia significativa entre ellas a los distintos tiempos evaluados (3, 5, 7, 14 días); tanto a los 3 como a los 7 dpi se observan mayores niveles de transcrito en raíces inoculadas respecto a no inoculadas, mientras que, a los 5 y 14 dpi, los niveles de transcrito son menores en raíces inoculadas que en raíces sin inocular. En el caso de los niveles de transcrito del gen *PvRALF6* se puede observar que a los 3, 5 y 7 dpi, los niveles de acumulación de transcrito son similares en las raíces sin inocular e inoculadas, sin embargo, se observa una tendencia a un incremento progresivo en la acumulación de los transcritos en las raíces inoculadas (Fig. 12B). A los 14 dpi los niveles de transcrito en las raíces inoculadas son mayores que en raíces sin inocular de 14 dpi (Fig. 12B). Con respecto a los niveles de acumulación del transcrito *PvRIPK2* son muy bajos respecto al gen de referencia; no se observan diferencias significativas entre raíces inoculadas y no inoculadas, posiblemente debido a la variabilidad y los niveles tan bajos observados (Fig. 12C).

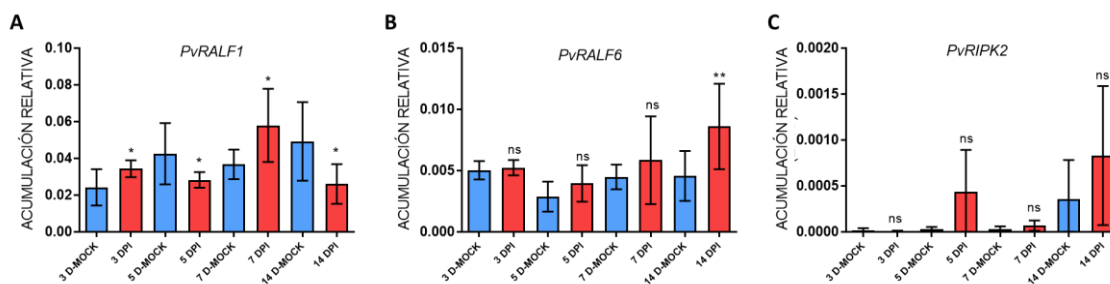


Figura 12. Niveles de acumulación de transcrito *PvRALF1*, *PvRALF6* y *PvRIPK2*, en raíces de frijol inoculadas con rhizobia y en raíces sin inocular. Cinética de la acumulación de transcritos de *PvRALF1* (A), *PvRALF6* (B) y *PvRIPK2* (C) en raíces de frijol inoculadas con rhizobia (barras azules) y raíces no inoculadas (barras rojas) de 3, 5, 7 y 14 días. Se observan

diferencias en el nivel de transcrito de *PvRALF1* a lo largo de los tiempos evaluados (A). Los niveles de *PvRALF6* no parecen cambiar a lo largo del tiempo evaluado en raíces inoculadas respecto a las raíces sin inocular (B). No se observan diferencias en los niveles de transcrito de *PvRIPK2* en los tiempos evaluados (C). Los datos representan la media \pm desviación estándar (n=9); *P <0.05, ns=no significativo (Student test). Acumulación relativa al gen IDE

9.3 Artículo publicado: “Genome-wide identification of the CrRLK1L subfamily and comparative analysis of its role in the legume-rhizobia symbiosis”

Debido a la relación estrecha entre los péptidos RALF y los receptores de membrana de la subfamilia CrRLK1L, se realizó un estudio bioinformático profundo de estos receptores que inicialmente incluyó 63 especies de plantas, y posteriormente el estudio se enfocó en un análisis comparativo entre cuatro leguminosas, *A. thaliana*, sorgo y el musgo *Physcomitrella patens* con la finalidad de identificar genes con funciones potenciales en la simbiosis (ver artículo anexo, incluido en la sección de material suplementario).

9.4 Actividad de promotores *PvRALF1*, *PvRALF6*, *PvFER1* y *PvRIPK2* en raíces transgénicas de frijol

Con el fin de analizar espacio-temporalmente la actividad de los promotores de los genes *PvRALF1*, *PvRALF6* y *PvRIPK2* seleccionados para su análisis, primero se procedió a la construcción de los vectores que llevan la región promotora de cada gen, conforme al protocolo descrito en la sección 5.3. Además, se analizó la actividad de promotor de *PvFER1* de frijol.

Para analizar la actividad de los promotores de los genes de interés en frijol, se obtuvieron raíces transgénicas inducidas con las clonas de *A. rhizogenes* portadoras de las distintas construcciones generadas para este fin. Una vez obtenidas las raíces transgénicas con las diferentes construcciones, las plantas se inocularon con *R. tropici* CIAT899, y se analizó la actividad de los promotores a través de la actividad de la glucoronidasa en las raíces inoculadas a los 7, 14 y 21 dpi. Para demostrar que la actividad de la glucoronidasa es debida a la enzima expresada en las raíces transgénicas bajo control del promotor estudiado, se analizaron raíces portadoras del vector vacío como control negativo (Fig. S10).

9.4.1 Análisis de la actividad del promotor *PvRALF1* en raíces transgénicas

Las raíces transgénicas de frijol generadas con la construcción para el análisis del promotor *PvRALF1* inoculadas con rhizobia y sin inocular, se colectaron y se realizó la tinción de GUS. La aparición del precipitado azul, producto de la actividad de la glucoronidasa, se observó aproximadamente 25 min. después de que las raíces se colocaron en la solución de tinción.

Los resultados obtenidos muestran actividad de GUS en los ápices de las raíces, que corresponden a la zona meristemática, dicha actividad disminuye al aproximarse a la región de elongación, localizándose principalmente en el cilindro central (Fig. 13A y B). En las zonas más alejadas del ápice de la raíz, no se observa actividad de promotor (Fig. 13C). Se detectó actividad del promotor en los pelos radicales (Fig. 13D). La actividad también se evaluó en primordios de raíces laterales, encontrándose inicialmente en el córtex interno, en las células que darán origen al primordio de raíz lateral (Fig. 13E). En las etapas posteriores del desarrollo, se encontró actividad en la base de los primordios de las raíces laterales previo (Fig. 13F y G) y posterior a la emergencia de la raíz (Fig. 13H e I).

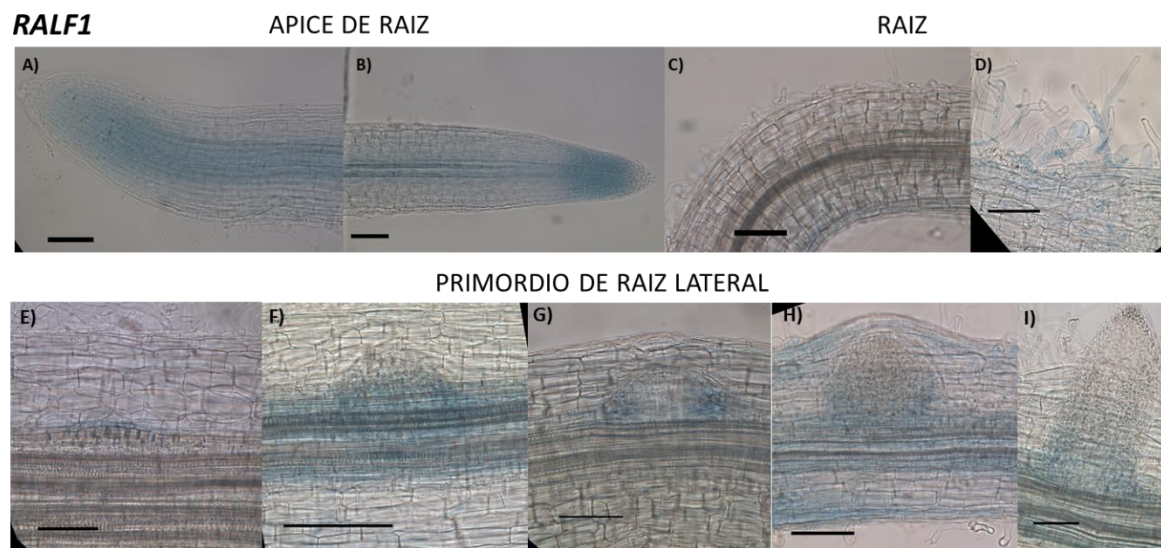


Figura 13. Actividad del promotor *PvRALF1* en raíces y primordios de raíz lateral determinada mediante tinción GUS. La actividad del promotor *PvRALF1* se observó en los ápices de las raíces transgénicas de frijol, la cual está localizada en el cilindro vascular central, al analizar la zona de elongación (A y B); esto no fue así en zonas más viejas de las raíces, en donde no se observó actividad del promotor en ninguna sección de la raíz (Fig. C). Se detectó actividad de GUS en pequeñas secciones aleatorias de las raíces, principalmente en los pelos radicales y la epidermis (D). La actividad del promotor *PvRALF1* se analizó en primordios de raíces laterales (E-I), observándose en las primeras células en el córtex interno de la raíz (E), y durante su emergencia (I). Barras de escala = 100 μ m. n=15.

La actividad del promotor *PvRALF1* se evaluó durante varias etapas del desarrollo de los nódulos (Fig. 14). La actividad fue observada en el córtex externo durante las primeras etapas de división celular y en las células en donde el primordio del nódulo comienza a formarse (3 dpi, Fig. 14A). La actividad se mantiene en las células del primordio del nódulo, desde su emergencia hasta las etapas de infección de las células del primordio del nódulo (5-14 dpi,

Fig. 14B-D). En los nódulos maduros, la actividad del promotor *PvRALF1* se localiza principalmente asociada a la vasculatura del mismo (14-21 dpi, Fig. 14E).

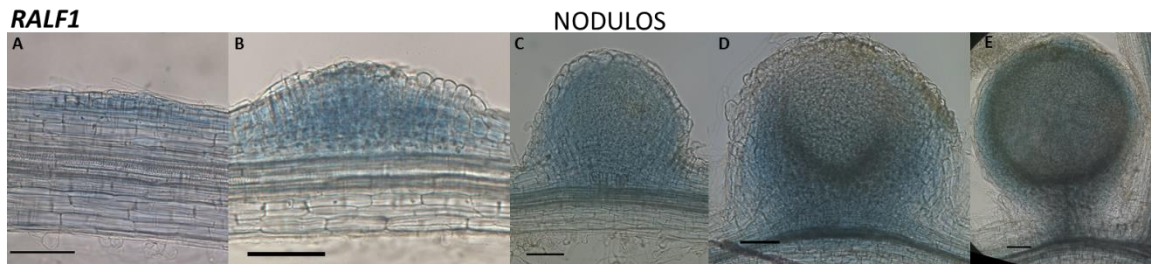


Figura 14. Actividad del promotor *PvRALF1* durante el desarrollo de nódulos. La actividad del promotor *PvRALF1* en las distintas etapas del desarrollo del nódulo se detectó mediante tinción de GUS. Durante las primeras etapas, se aprecia la actividad de promotor en el córtex externo en las células en división que darán origen al primordio del nódulo (A), durante el crecimiento se observa actividad en las células del primordio (B-D) y en nódulos maduros se aprecia mayor actividad asociada a la vasculatura de los nódulos (D y E). Barras de escala = 100 μ m. n=15.

Los resultados del análisis de la actividad del promotor *PvRALF1* indican que este gen es activo en zonas con gran actividad meristemática: en los ápices de las raíces en los primordios de nódulos y en primordios de raíces laterales. Además, que el promotor sea activo en la zona de elongación de las raíces y en los primordios de nódulos de etapas más tardías podría reflejar una función de *PvRALF1* en la elongación celular en estos tejidos (la actividad de *RALF1* en *Arabidopsis* y otros modelos se ha asociado a elongación celular). Así mismo, que el promotor sea activo en nódulos maduros, sugiere alguna(s) función(es) en su mantenimiento.

9.4.2 Análisis de la actividad del promotor *PvRALF6* en raíces transgénicas

Las raíces transgénicas de frijol transformadas con la construcción del promotor *PvRALF6* se analizaron en distintas secciones de la raíz para determinar su expresión espacio-temporal. La actividad GUS se detectó aproximadamente 40 min. después de que las raíces se colocaron en la solución de tinción. En la Fig. 15, se observa la actividad del promotor del gen en el ápice de la raíz, correspondiente a la zona meristemática (Fig. 15A y B). En las zonas de elongación y diferenciación, la actividad se localiza en el cilindro central de las raíces, y se mantiene ese patrón en las zonas de la raíz más alejadas del ápice (Fig. 15C y D). La actividad del promotor se analizó también en función del tiempo, en distintas etapas de desarrollo de las raíces laterales. El promotor mostró actividad en las células del córtex interno, en donde se origina el primordio de raíz lateral (Fig. 15E). Esta actividad se mantiene en todo el primordio de la raíz durante su desarrollo (Fig. 15F-H) y hasta su emergencia (Fig. 15I y J).

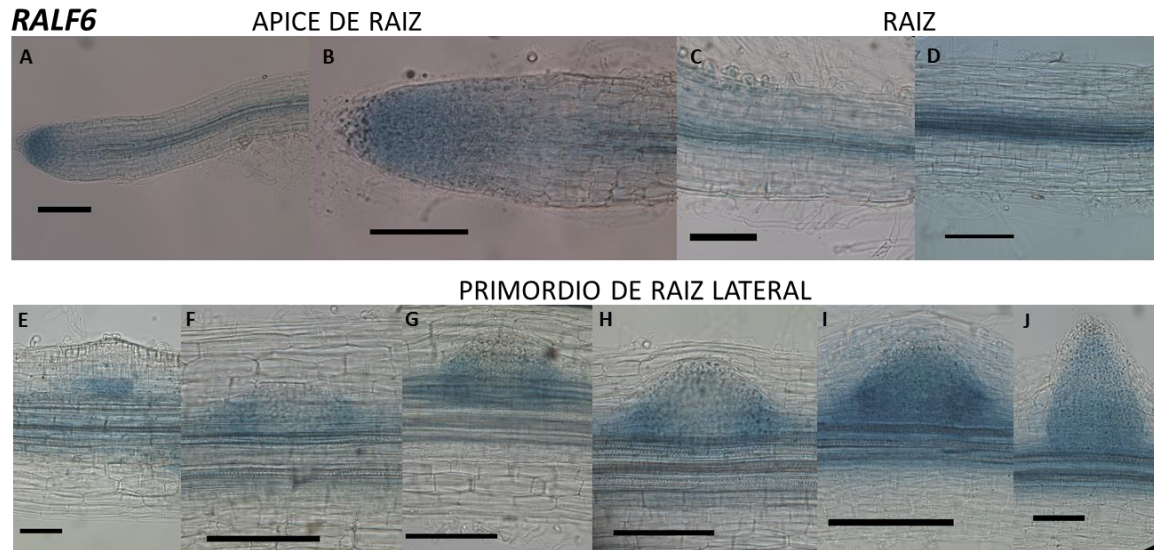


Figura 15. Actividad del promotor *PvRALF6* en raíces y primordios de raíz lateral. La actividad del promotor *PvRALF6* es evidente en la región meristemática en los ápices de las raíces (A), así como en el cilindro vascular a partir de la zona de elongación y por toda la raíz (B y D). Así mismo, la actividad del promotor *PvRALF6* fue analizada en primordios de raíces laterales (E-J), la actividad se observa desde las primeras divisiones celulares en el córtex interno de la raíz (E), y se mantiene en las células del primordio de raíz (F-H) hasta su emergencia (I y J). Barras de escala = 100 μ m. n=15.

La actividad del promotor *PvRALF6* también se analizó durante el desarrollo de los nódulos. Al igual que en el caso del primordio de raíz, la actividad del promotor se observa en las células que entran en división y que dan origen al primordio del nódulo, que, en el caso de los primordios de nódulo, esto ocurre en el córtex externo (3 dpi, Fig. 16A). A medida que el primordio del nódulo crece, la actividad del promotor *PvRALF6* se observa asociada al primordio del nódulo hasta su emergencia (de 5 a 14 dpi, Fig. 16B-D). En nódulos maduros, la actividad de GUS se mantiene, y se observa principalmente en la vasculatura del nódulo (14 a 21 dpi, Fig. 16E).

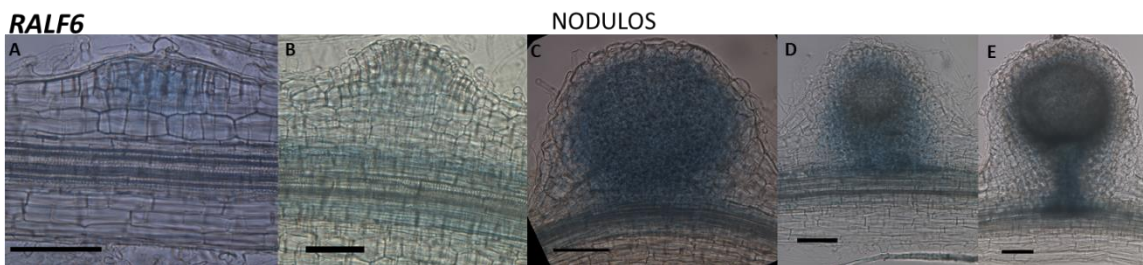


Figura 16. Actividad del promotor *PvRALF6* durante el desarrollo de nódulos. La actividad del promotor del gen *PvRALF6* en raíces transgénicas de frijol, se observa durante el proceso de nodulación. Durante la infección, se observa actividad

del promotor *PvRALF6* en las células del córtex externo, debajo del pelo radical enroscado (A). Durante el desarrollo del primordio del nódulo, la actividad se detecta en las células del primordio del nódulo en crecimiento (B), y dicha actividad se mantiene en primordios de nódulo más desarrollados (C, D). En los nódulos maduros, la actividad del promotor se observó asociada principalmente al tejido vascular del nódulo (E). Barras de escala = 100 μ m. n=15.

Los resultados de la actividad del promotor *PvRALF6* indican que este es activo en zonas con gran actividad meristemática (como fue el caso del promotor *PvRALF1*), los ápices de las raíces y los primordios de raíces laterales y de nódulos. Así mismo, el hecho de que el promotor esté activo en la zona de elongación de las raíces y en los primordios de nódulos de etapas más tardías, sugiere una función de *PvRALF6* en la elongación celular en estos tejidos. Finalmente, que el promotor sea activo en nódulos maduros y en raíces más viejas, es indicativo de posibles funciones en el mantenimiento en ambos órganos.

9.4.3 Análisis de la actividad del promotor *PvRIPK2* en raíces transgénicas

Para determinar la actividad espacio temporal del promotor *PvRIPK2*, seguimos la misma estrategia usada previamente, para lo cual, se generaron raíces compuestas de frijol con la construcción correspondiente. En este caso, la formación del precipitado azul fue observable a partir de 40 min. de incubación. Los resultados obtenidos indican que el promotor es activo en las células meristemáticas del ápice de las raíces, y su actividad disminuye a medida que se observa en zonas arriba del ápice (Fig. 17A y B); en las zonas de la raíz más alejadas del ápice, esta actividad deja de observarse (Fig. 17C). Al igual que lo observado tras el análisis de la actividad del promotor *PvRALF1*, se determinó la actividad del promotor *PvRIPK2* en los pelos radicales (Fig. 17D). Además, el promotor es activo en las células del córtex interno a partir de las cuales se origina el primordio de raíz lateral (Fig. 17E), manteniéndose dicha actividad en el primordio de la raíz lateral durante su desarrollo (Fig. 17F-G) y su emergencia (Fig. 17H e I).

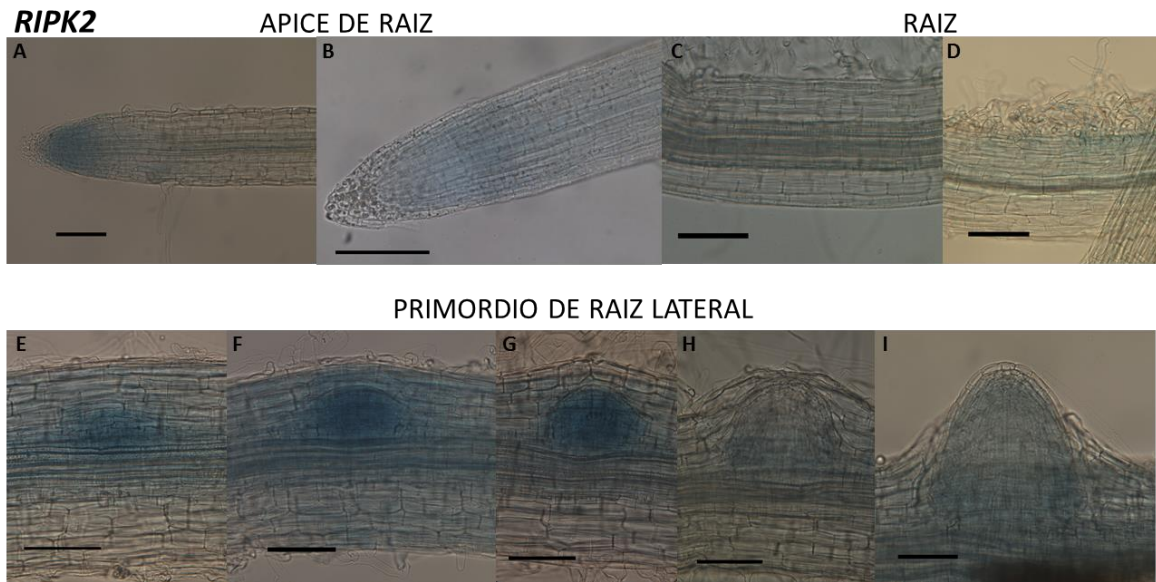


Figura 17. Actividad del promotor *PvRIPK2* en raíces y primordios de raíz lateral. La actividad de promotor (revelado por actividad de GUS) del gen *PvRIPK2* se observa en los ápices de raíces transgénicas de frijol, y desaparece en las zonas de elongación y diferenciación (A y B). En las zonas de la raíz más alejadas del ápice no se observa actividad del promotor (C) salvo en pequeñas regiones en donde se aprecia actividad GUS asociada a los pelos radicales (D). Durante el desarrollo de raíces laterales, el promotor *PvRIPK2* es activo en la región meristemática, desde las primeras divisiones celulares (E) hasta el crecimiento (F y G) y emergencia de las raíces laterales (H e I). Barras de escala = 100 μ m. n=15.

El análisis de la actividad GUS durante la nodulación mostró que el promotor está activo durante las diferentes etapas del desarrollo del nódulo. La actividad de *PvRIPK2* se observó en las células que se dividen tras reactivar su ciclo celular (3 dpi, Fig. 18A), conforme el primordio del nódulo crece, la actividad del promotor *PvRIPK2* se detecta en todo el primordio del nódulo, hasta su emergencia (5-7 dpi, Fig. 18B y C). Esta actividad se mantiene después de la emergencia del nódulo (7-14 dpi, Fig. 18D). En los nódulos maduros, sin embargo, la actividad se detectó principalmente en la base del nódulo y en la vasculatura del nódulo (14-21 dpi, Fig. 18E).

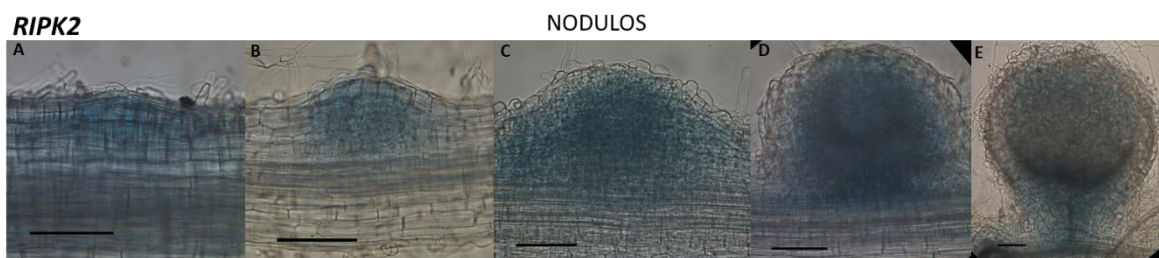


Figura 18. Actividad del promotor *PvRIPK2* durante el desarrollo de nódulos. El promotor *PvRIPK2* es activo durante el desarrollo del nódulo en frijol. En la zona de división celular, se observa actividad GUS en el córtex externo, por debajo del

pelo radical enroscado (A). La actividad del promotor *PvRIPK2* se mantiene en las células del primordio durante su crecimiento (B) y hasta su emergencia (C y D). En nódulos maduros, se observa una mayor actividad asociada a su vasculatura (E). Barras de escala = 100 μm . n=15.

Los resultados obtenidos de la actividad del promotor *PvRIPK2* confirman inicialmente que el gen es activo tanto en raíces de frijol como en nódulos. Además, las observaciones de que el promotor es activo en ápices de raíz, primordios de nódulo y primordios de raíz lateral, zonas con gran actividad meristemática, indican un posible papel de *PvRIPK2* en el desarrollo de raíces y de nódulos. Así mismo, que el gen sea activo en primordios de nódulo más avanzados en su desarrollo sugiere un papel en el crecimiento. Por último, la actividad del promotor *PvRIPK2* en nódulos maduros, sugiere una función en el mantenimiento de estos.

9.4.4 Análisis de la actividad del promotor *PvFER1* en raíces transgénicas

Previamente, se realizó un análisis por genética reversa del papel del receptor *PvFER1* durante la nodulación en frijol (Tesis de maestría Jorge Esaú Solís Miranda (Solís-Miranda 2016)). Estas proteínas han sido descritas como los receptores de los péptidos RALF, y posiblemente podrían estar actuando en conjunto con *PvRALF1*, *PvRALF6* y *PvRIPK2*. Por ello, decidimos analizar la actividad de promotor de *PvFER1* en raíces de frijol sin inocular y durante el proceso de nodulación.

Los resultados obtenidos muestran que el promotor *PvFER1* es activo en los ápices de las raíces, correspondiente a la zona meristemática (Fig. 19A y B), dicha actividad se mantiene a lo largo de la raíz principal, localizándose principalmente en el cilindro central (Fig. 19C y D). La actividad del promotor *PvFER1* se observa también en las células que dan origen al primordio de raíz lateral en el córtex interno (Fig. 19E), en el primordio de las raíces laterales, y en el ápice de éstas, previo (Fig. 19F y G) y posterior a su emergencia (Fig. 19H e I).

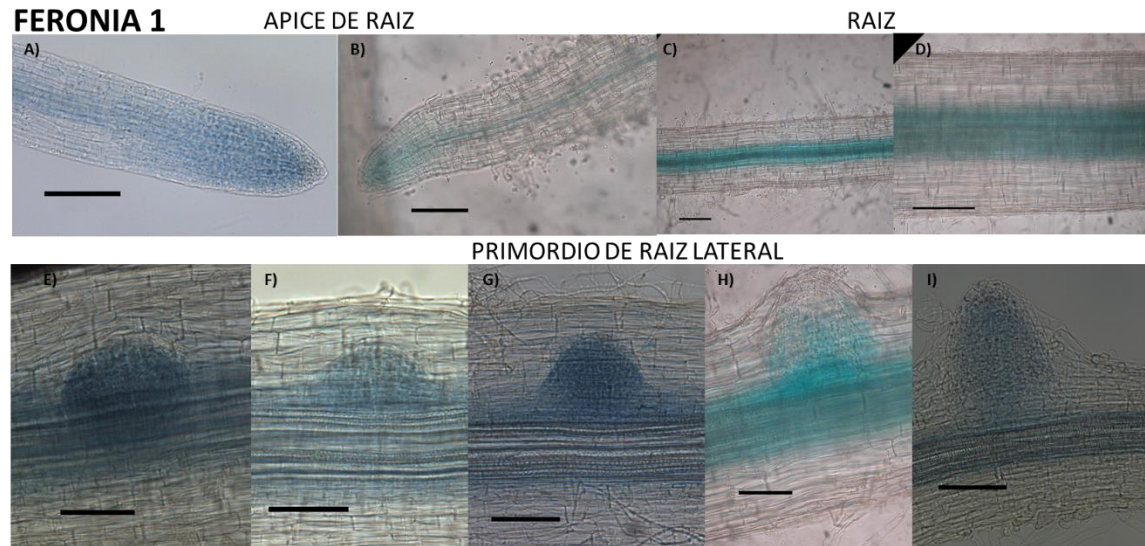


Figura 19. Actividad del promotor *PvFER1* en raíz principal y en primordios de raíces laterales. La actividad de promotor (revelado por actividad GUS) del gen *PvFER1* se observa en los ápices de raíces transgénicas de frijol, en las zonas de elongación y de diferenciación, asociada a la vasculatura (A y B), la cual se mantiene en zonas de la raíz más alejados del ápice (C y D). Durante el desarrollo de raíces laterales, el promotor *PvFER1* es activo en la región meristemática, desde las primeras divisiones celulares (E) hasta el crecimiento (F y G) y emergencia de las raíces laterales (H e I). Barras de escala = 100 μm . n=15.

Durante el desarrollo del nódulo se observó actividad del promotor *PvFER1*, a partir de las primeras etapas, en las células que entran en división en el córtex externo, debajo del pelo enroscado (3 dpi, Fig. 20A). El promotor se mantiene activo durante el desarrollo del primordio del nódulo, hasta su emergencia (7-14 dpi, Fig. 20B y C) así como en nódulos maduros (14-21 dpi, Fig. 20D y E).

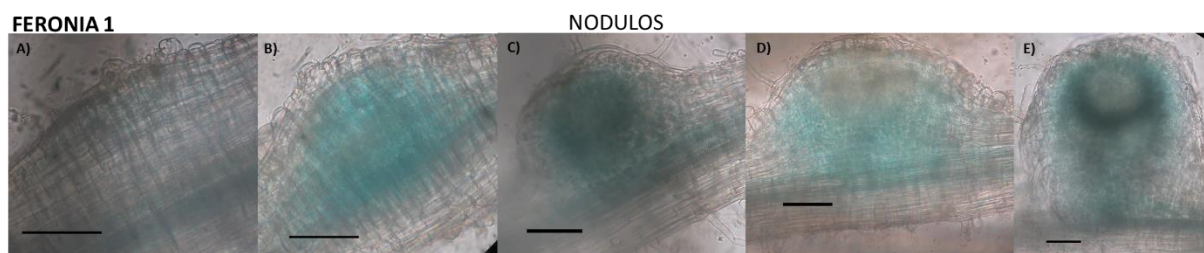


Figura 20. Actividad del promotor *PvFER1* durante el desarrollo de nódulos. Durante la infección por rhizobia en raíces transgénicas de frijol, se observa actividad del promotor *PvFER1* en las células del córtex externo, por debajo del pelo radical enroscado (A); la actividad se mantiene durante el desarrollo del primordio del nódulo, en las células del córtex interno debajo del sitio de infección, y en las células del primordio previo a su emergencia (B), y después de la emergencia (C), durante la formación de la vasculatura, y en el nódulo maduro, la actividad del promotor *PvFER1* se observa asociado a estas células (D y E). Barras de escala = 100 μm . n=15.

9.5 Análisis del fenotipo de raíces transgénicas de frijol silenciadas y sobreexpresantes *PvRALF1*, *PvRALF6*, *PvRIPK2* y *PvFER1* durante la simbiosis con rhizobia

Con la finalidad de analizar la función de los diferentes genes de interés (*PvRALF1*, *PvRALF6* y *PvRIPK2*) durante la simbiosis, se generaron construcciones de silenciamiento o sobreexpresión de cada gen de interés. Dichos vectores se construyeron de acuerdo al protocolo descrito en el apartado 7.3. Se obtuvieron construcciones para el silenciamiento de cada uno de los genes de interés (ptdT-RALF1i, ptdT-RALF6i y ptdT-RIPK2i), así como para la sobreexpresión de los mismos (pH7-RALF1ox, pH7-RALF6ox y pH7-RIPK2ox).

Previo al análisis del efecto del silenciamiento y la sobreexpresión de los genes *PvRALF1*, *PvRALF6* y *PvRIPK2* de frijol, se generaron raíces transgénicas utilizando 10 clonas de *A. rhizogenes* portadoras de alguna de las distintas construcciones previamente generadas y se seleccionaron clonas en las que se observaron los cambios esperados en los niveles del transcrito correspondiente. Para los análisis de silenciamiento y sobreexpresión de *PvFER1*, se emplearon las construcciones generadas previamente (ptdT-FER1i y pH7-FER1ox, Tesis de maestría Jorge Esaú Solís Miranda (Solis-Miranda 2016).

Es pertinente mencionar que no fue posible obtener raíces transgénicas fluorescentes con las distintas clonas de *A. rhizogenes* generadas para la construcción de sobreexpresión de *PvRIPK2*. Debido a esto, no se realizó el análisis del efecto de la sobreexpresión de *PvRIPK2* en la nodulación en frijol.

9.5.1 Selección de clonas de *A. rhizogenes* portadoras de las construcciones para el silenciamiento y la sobreexpresión de *RALF1* y *RALF6*

Una vez seleccionadas las clonas de *A. rhizogenes* y previo al análisis fenotípico, se evaluaron por RT-qPCR los niveles de acumulación de los transcritos *PvRALF1*, *PvRALF6* y *PvRIPK2* de las raíces transgénicas seleccionadas, con el objetivo de comprobar que la clona de *A. rhizogenes* seleccionada genera los cambios esperados en los niveles del transcrito correspondiente.

En las raíces silenciadas *PvRALF1* (barra roja) se encontró una disminución de hasta un 70% en el nivel del transcrito en comparación al nivel en las raíces transgénicas control (SAC, barra azul) (Fig. 21A). Los niveles del transcrito *PvRALF6* (barra verde) presentan una disminución de hasta un 80% en las raíces silenciadas, comparado con las raíces control (Fig.

21B). Por otro lado, los niveles de acumulación de transcrito en las raíces transgénicas portadoras de la construcción de sobreexpresión *PvRALF1* (barra roja) muestran un aumento de hasta 35 veces, comparado con las raíces control (pH7-GUS, barra azul) (Fig. 21C). Los niveles de acumulación del transcrito *PvRALF6* (barra verde) en las raíces sobreexpresantes presentan un incremento de más de 150 veces respecto al control (Fig. 21D). Debido a los niveles de expresión muy bajos del transcrito *PvRIPK2* en las raíces transgénicas, no fue posible evaluar dicho transcrito en las plantas silenciadas.

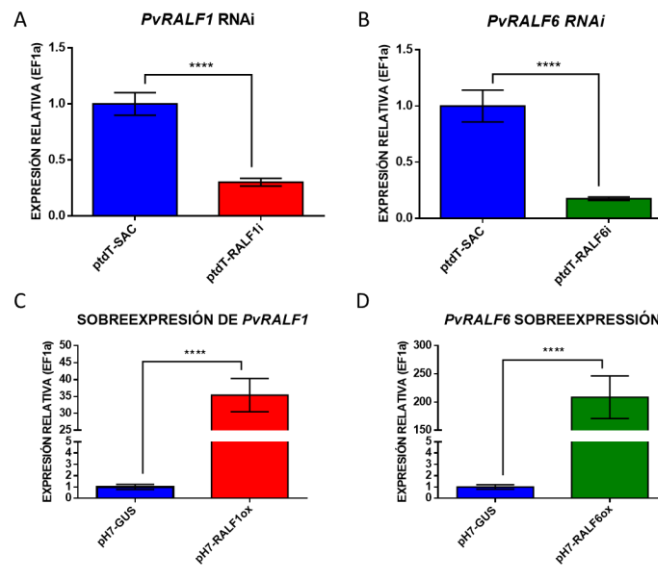


Figura 21. Análisis de la acumulación de transcritos en raíces transgénicas de frijol. En A y B) Se observa una disminución en los niveles de acumulación de transcrito *PvRALF1* y *PvRALF6* en las raíces transgénicas silenciadas para cada gen (barra roja y barra verde) al comparar con el control (barras azules). En C) y D) se aprecia un incremento en los niveles de transcrito *PvRALF1* y *PvRALF6* en las raíces sobreexpresantes de cada gen (barra roja y barra verde) respecto al control (barras azules). El nivel del transcrito IDE se usó como referencia. Los datos representan la media \pm error estándar (n=9); *** $P < 0.001$ (Student test).

9.5.2 Análisis del número de nódulos en las raíces silenciadas y raíces sobreexpresantes en *PvRALF1* y *PvRALF6*

Para determinar el efecto del silenciamiento y la sobreexpresión en la generación de nódulos, estos se cuantificaron a cuatro tiempos de la nodulación (7, 14, 21 y 31 dpi), de acuerdo a lo descrito en el apartado 7.8.1. En las raíces silenciadas *PvRALF1* se observó un ligero incremento en el número de nódulos generados a los 7 días (puntos rojos) respecto a las raíces control (puntos azules, Fig. 22A), sin embargo, a los 14, 21 y 31 dpi, la cantidad de nódulos entre las raíces silenciadas *PvRALF1* no difiere significativamente de las raíces control (Fig.

22A). En el caso de las raíces silenciadas *PvRALF6*, se puede ver que a los 21 y 31 días hay una disminución en el número de nódulos en las raíces silenciadas (puntos verdes), con respecto al control (Fig. 22B), mientras que a los 7 y 14 dpi no se observan diferencias significativas (Fig. 22B).

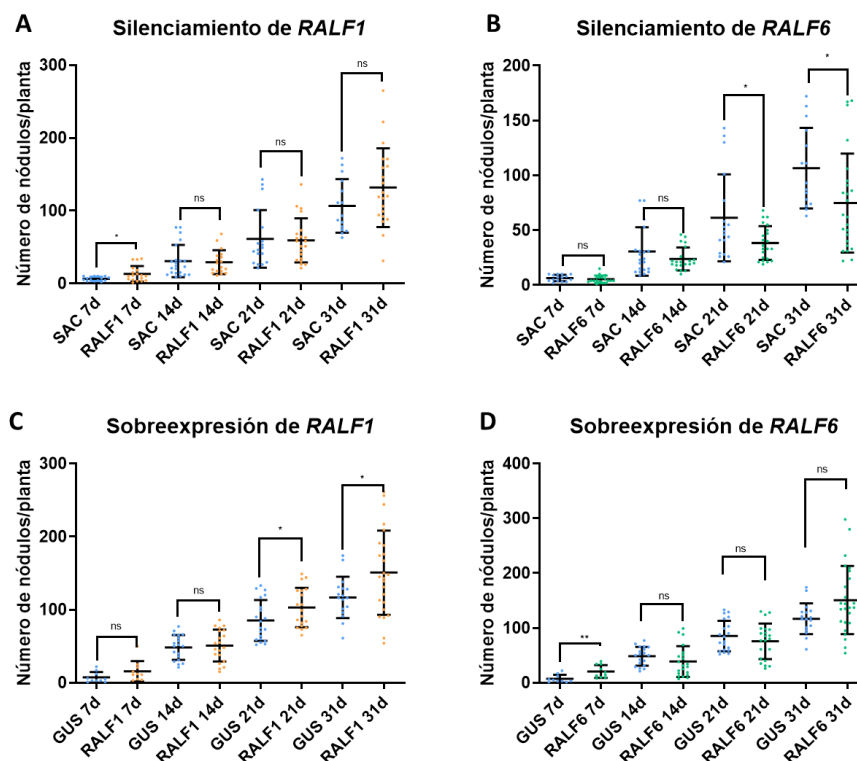


Figura 22. Número de nódulos en raíces transgénicas de frijol silenciadas y sobreexpresantes de *PvRALF1* y *PvRALF6*. A) Número de nódulos generados en las raíces transgénicas silenciadas *PvRALF1* (puntos naranjas) y las raíces control (puntos azules). B) Número de nódulos generados en raíces silenciadas *PvRALF6* (puntos verdes) a 7, 14, 21 y 31 dpi. C) Número de nódulos en las raíces transgénicas sobreexpresantes *PvRALF1* (puntos naranjas), y en raíces control. D) Cuantificación del número de nódulos generados en raíces silenciadas *PvRALF6* (puntos verdes). Los puntos representan todos los datos, las líneas corresponden a la media \pm la desviación estándar ($n=12$); * $P < 0.05$, ** $P < 0.01$, ns=no significativo.

En los análisis de sobreexpresión, de manera interesante, se observa un patrón opuesto; la sobreexpresión de *PvRALF1* provoca un incremento en el número de nódulos a los 21 y 31 dpi al comparar con las raíces control, pero no a 7 y 14 dpi (Fig. 22C). En cambio, en las raíces sobreexpresantes *PvRALF6* se encontró un ligero incremento en el número de nódulos a los 7 dpi al comparar con el control (Fig. 22D), no obstante, no hay diferencia significativa a los 14, 21 y 31 dpi (Fig. 22D).

9.5.3 Análisis del diámetro de los nódulos en las raíces silenciadas y raíces sobreexpresantes en *PvRALF1* y *PvRALF6*

Dado que la participación de RALF durante la expansión celular en raíces de *A. thaliana* ha sido reportada (Canto *et al.*, 2013), se decidió explorar si los nódulos generados en las raíces silenciadas o sobreexpresantes de *PvRALF1* y *PvRALF6* muestran cambios en su expansión celular. Para esto, se procedió a medir el diámetro de los nódulos generados en las raíces transgénicas silenciadas y sobreexpresantes de *PvRALF1* y *PvRALF6*, a los 7, 14, 21 y 31 dpi, utilizando la metodología descrita en el apartado 7.8.2.

Los resultados del análisis de los diámetros de los nódulos de raíces control (barras azules), silenciadas y sobreexpresantes en *PvRALF1* (barras naranjas) y en *PvRALF6* (barras verdes) muestran un incremento progresivo en los diámetros de la población total de nódulos (la cual incluye primordios de nódulos y nódulos maduros) durante los tiempos evaluados para las condiciones control como silenciadas y sobreexpresantes (Fig. 23). Este comportamiento es el esperado ya que a los 7 dpi la población de nódulos estaría compuesta por primordios de nódulos; a los 14 dpi la población estaría constituida principalmente por nódulos cercanos a la maduración y por nuevos primordios de nódulo; a los 21 dpi, gran parte de los primordios de han completado su desarrollo y nuevos primordios aparecen; y a los 31 dpi, la mayoría de los nódulos han alcanzado su madurez.

Al comparar los diámetros de raíces silenciadas en *PvRALF1* y *PvRALF6* con los nódulos de las raíces control, se observa una distribución similar de las medias y de la varianza de los diámetros de los nódulos de las raíces silenciadas respecto a los controles a los 7, 14 y 21, y a 31 dpi se observan diámetros ligeramente menores en las raíces silenciadas (Fig. 23A y B). De la misma forma, en raíces sobreexpresantes de *PvRALF1* y de *PvRALF6*, los nódulos de 7 y 14 dpi presentan una distribución y una media similares a los de los nódulos de las raíces control. A los 21 y 31 dpi, las distribuciones y las medias de los diámetros son ligeramente menores en las raíces silenciadas respecto a los nódulos de las raíces control (Fig. 23C y D).

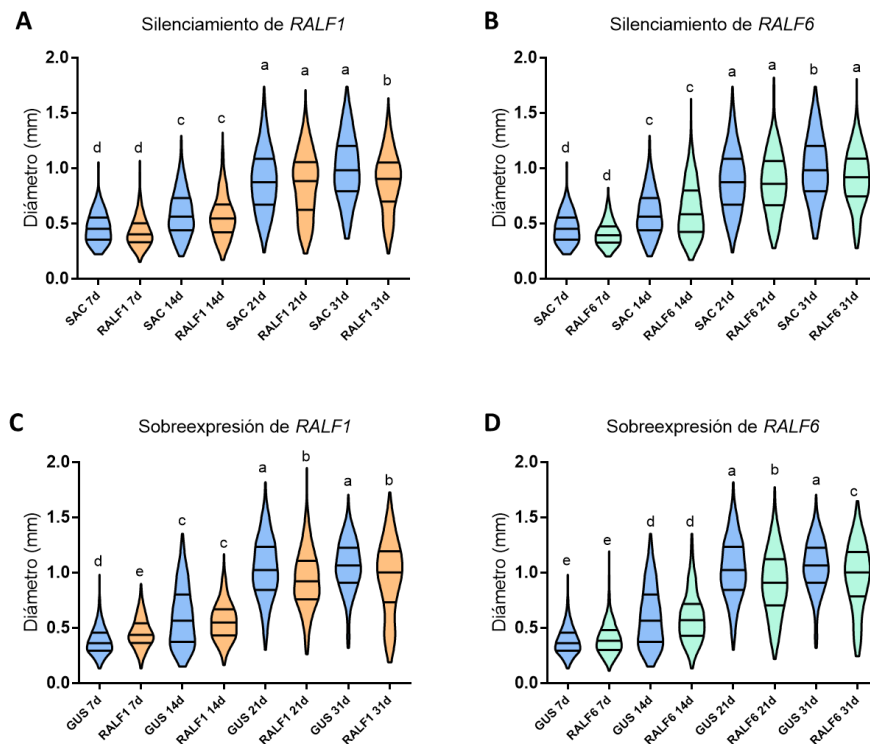


Figura 23. Diámetros de nódulos en raíces transgénicas de frijol silenciadas y sobreexpresantes de *PvRALF1* y *PvRALF6*. Distribución de los diámetros de los nódulos generados en raíces silenciadas *PvRALF1* (barras naranjas, A), raíces silenciadas *PvRALF6* (barras verdes, B), raíces sobreexpresantes *PvRALF1* (barras naranjas, C) y raíces sobreexpresantes *PvRALF6* (barras verdes, D) comparados con los de nódulos de raíces control (barras azules) a los 7, 14, 21 y 31 dpi nódulos de raíces control. Los datos representan los resultados de tres experimentos independientes. La línea central representa la mediana de los datos y los límites de la grafica representan la distribución de la población de datos. Los datos representan más de 1000 mediciones diferentes analizados por ANOVA de una vía.

9.5.4 Análisis de la actividad de la nitrogenasa en raíces transgénicas silenciadas en *RALF1* y *RALF6*

Debido a que el silenciamiento de *PvRALF6* y la sobreexpresión de *PvRALF1* indujeron un cambio en el número de los nódulos a 21 y 31 dpi, se decidió analizar también la actividad de la nitrogenasa en esos nódulos. Para esto, se generaron raíces transgénicas portadoras de las construcciones correspondientes y se procesaron conforme al apartado 7.8.3.

En la Fig. 24, se muestran los resultados obtenidos a los 14, 21 y 31 dpi en las raíces silenciadas y sobreexpresantes de *PvRALF1* y *PvRALF6* (a los 7 dpi los nódulos no han madurado, siendo incapaces de fijar nitrógeno). En los nódulos de raíces silenciadas *PvRALF1* y *PvRALF6*, se observa que a los 14 dpi la una reducción de acetileno es ligeramente menor (barras naranjas y verdes) respecto a nódulos de raíces control (barras azules, Fig. 24A y B). A los 21 y 31 dpi, los niveles de reducción de acetileno de los nódulos

en las raíces silenciadas en ambos genes son similares a los niveles en el control (Fig. 24A y B).

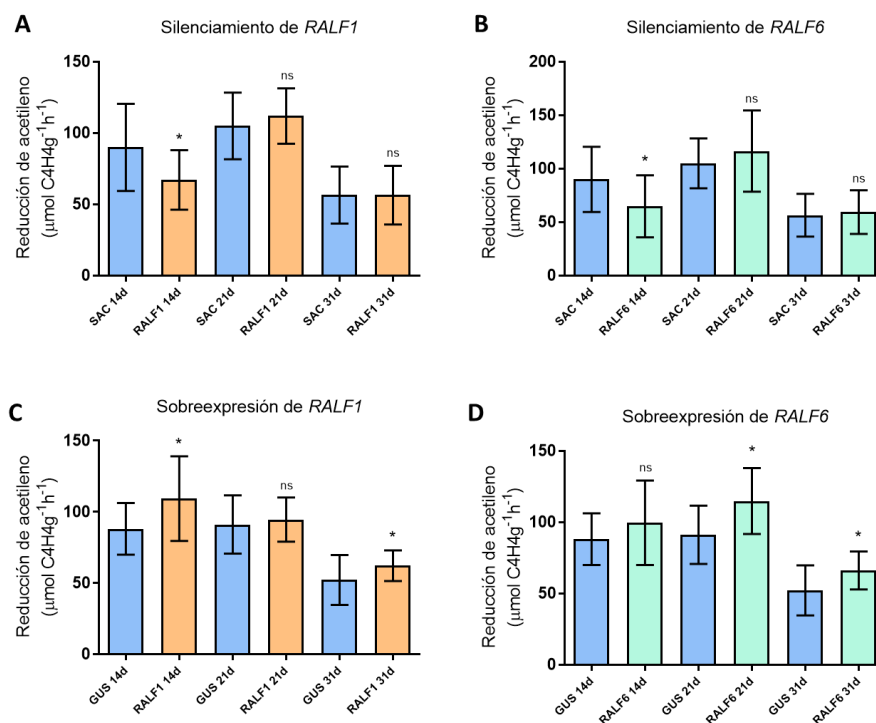


Figura 24. Reducción de acetileno y número de nódulos en raíces transgénicas silenciadas y sobreexpresantes en *RALF1* y *RALF6*. A) Gráficas representativas de la tasa de reducción de acetileno correspondiente a los nódulos generados en las raíces transgénicas correspondientes al silenciamiento de *RALF1* (barras rojas) y *RALF6* (barras verdes), comparados con su respectivo control (barras azules). Se observa una menor tasa de reducción de acetileno a los 14 dpi; esa diferencia desaparece a los 21 y 31 dpi. B) Actividad de la nitrogenasa evaluada en nódulos de raíces sobreexpresantes de *RALF1* (barras rojas) y de *RALF6* (barras verdes), y de raíces control (barras azules). La sobreexpresión de *RALF1* o de *RALF6* provoca un ligero incremento en la reducción de acetileno. Los datos representan la media \pm desviación estándar ($n=20$); * $P < 0.05$, ns=no significativo (Student test).

Por otro lado, los nódulos de raíces sobreexpresantes *PvRALF1* presentan un ligero incremento en la actividad de la nitrogenasa a los 14 y 31 dpi (barras naranjas) respecto a los nódulos de las raíces control (barras azules, Fig. 24C). Los nódulos de las raíces sobreexpresantes *PvRALF6* (barras verdes), muestran un incremento pequeño en la tasa de reducción de acetileno a los 21 y 31 dpi al comparar con los nódulos de las raíces control (Fig. 24D).

9.5.5 Análisis del número y diámetro de nódulos en las raíces silenciadas *PvRIPK2*
 En *Arabidopsis*, *RIPK1* participa corriente debajo de RALF-FER en la elongación celular de la raíz principal, desarrollo de pelos radicales, y es un modulador positivo de la respuesta inmune (Liu et al. 2011; Lee et al. 2015; Du et al. 2016). Para determinar si el silenciamiento

de *PvRIPK2* induce un fenotipo similar al observado en el silenciamiento de *PvRALF1* y *PvRALF6* durante la nodulación, se evaluó el número y el diámetro de los nódulos generados en las raíces transgénicas de acuerdo al protocolo descrito en la sección 7.8.2.

Los resultados obtenidos indican que el silenciamiento de *PvRIPK2* (puntos rosas) no induce cambios significativos en el número de nódulos al comparar con los de las raíces control (puntos azules) en ninguno de los tiempos evaluados (Fig. 25A). Además, dichos nódulos no muestran cambios importantes en los diámetros en las raíces silenciadas al comparar con raíces control; los diámetros de las raíces silenciadas *PvRIPK2* (barras rosas) no presentan diferencia significativa al comparar con los diámetros de raíces control (barras azules) a los 7 y 14 dpi, mientras que a los 21 y 31 dpi los nódulos de esas raíces silenciadas presentan diámetros inferiores a los de los nódulos control (Fig. 25B).

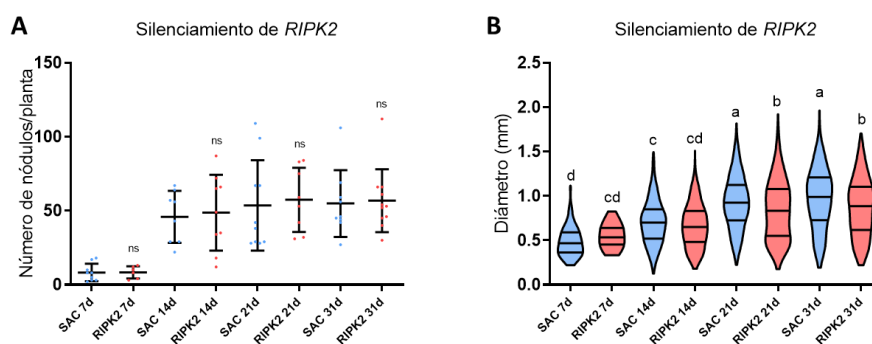


Figura 25. Número y diámetro de los nódulos en raíces transgénicas de frijol silenciadas de *RIPK2*. A) Número de nódulos generado en raíces transgénicas silenciadas en *RIPK2*. Se observa la cantidad de nódulos generados en las raíces transgénicas silenciadas en *RIPK2* (barras grises) y la raíces control (barras azules). No existe diferencia significativa en el número de nódulos en las raíces silenciadas comparados con el control a ninguno de los tiempos evaluados. Los datos representan la media \pm error estándar ($n=12$); * $P < 0.05$, ** $P < 0.01$ (Student's test). B) Diámetro de los nódulos generados en raíces transgénicas silenciadas en *RIPK2* (barras grises) y de nódulos de las raíces control (barras azules) a los 7, 14, 21 y 31 dpi. A) Los datos representan la media \pm desviación estándar ($n=20$); * $P < 0.05$, ns=no significativo (Student test). B) La línea central representa la mediana de los datos y los límites de las grafica representan la distribución de la población de datos. Los datos representan más de 1000 mediciones diferentes analizados por ANOVA de una vía.

9.5.6 Análisis del número y diámetro de nódulos en las raíces silenciadas *PvFER1*

Previamente evaluamos el fenotipo del silenciamiento y la sobreexpresión de *PvFER1* en la nodulación a 7, 14 y 21 dpi, en donde se encontró únicamente un efecto en el número de nódulos a 21 dpi (Tesis de maestría (Solis-Miranda 2016)). Debido a que los resultados observados en el silenciamiento de *PvRALF6* y la sobreexpresión de *PvRALF1* son a 21 y 31 dpi, decidimos evaluar el fenotipo del silenciamiento y la sobreexpresión de *PvFER1* a 31 dpi (Fig. 26). El número de nódulos a 21 y 31 dpi, disminuyó debido al silenciamiento de

PvFER1 (Fig. 26A), mientras que su sobreexpresión provocó un incremento en el número de nódulos en estos mismos tiempos evaluados (Fig. 26B). Al analizar los diámetros de dichos nódulos a los tiempos evaluados, se observó que estos son similares tanto en raíces silenciadas y sobreexpresantes *PvFER1* (Fig. 26C y D).

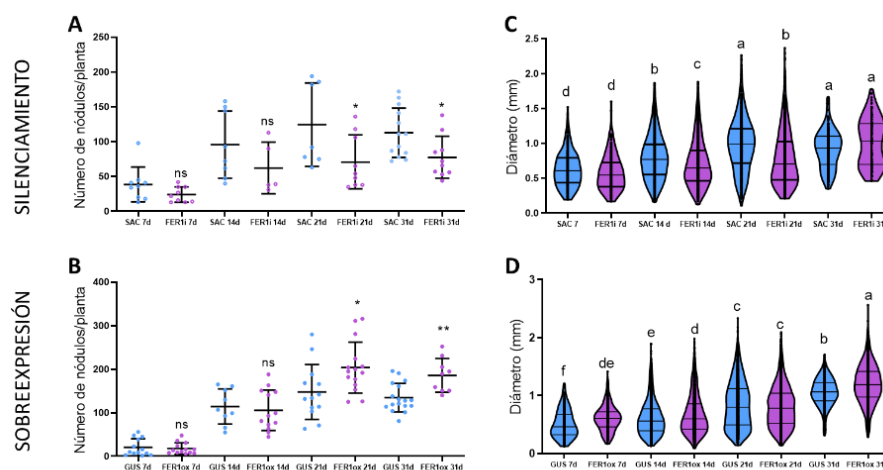


Figura 26. Número y diámetro de los nódulos en raíces transgénicas silenciadas y sobreexpresantes *PvFER1* a 21 y 31 dpi. Se muestra la distribución de los diámetros de los nódulos generados en raíces silenciadas (A) y sobreexpresantes *PvFER1* (B). Los diámetros de los nódulos generados en raíces silenciadas (C) y sobreexpresantes *PvFER1* (D). En (A) y en (B) los datos representan la media \pm desviación estándar ($n=9$); * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, ns=no significativo (Student test). En (C) y en (D) la línea central representa la mediana de los datos y los límites de las gráficas representan la distribución de la población de datos. Los datos representan más de 1000 mediciones diferentes analizados por ANOVA de una vía.

9.6 Análisis de la acumulación de transcritos *PvRALF1*, *PvRALF6*, *PvFER1* en raíces inoculadas con rhizobia en 0 y 5 mM de nitrógeno.

La nodulación es un proceso costoso energéticamente, cuya principal función es la de permitir a la bacteria la fijación de nitrógeno atmosférico, para que este sea utilizado por la planta. Está bien documentado que la presencia de compuestos de nitrógeno en el suelo inhibe el proceso de nodulación (Reid 2011, Ferguson 2019). En *A. thaliana* se ha descrito que RALF1 y FER son importantes en la respuesta de la planta cuando hay cambios en la relación C/N, lo cual ocurre a través de la regulación de la acumulación de proteínas 14.3.3 (Xu *et al.* 2019). Con la finalidad de determinar un posible papel de *PvRALF1*, *PvRALF6* y *PvFER1* de frijol durante la inhibición de la nodulación por nitrógeno, decidimos analizar la respuesta de dichos genes en raíces de frijol inoculadas con rhizobia y fertilizadas con medio Fahraeus

con 5 mM de nitrógeno, concentración a la cual se sabe que se inhibe la nodulación sin afectar el desarrollo de la planta.

En el caso de los genes *PvRALF1* y *PvRALF6*, se encontró que estos genes responden a la presencia de nitrógeno, incrementando la acumulación de su transcrito (Fig. 27A y B). El transcrito *PvRALF1* presenta valores similares en raíces a los 3 dpi tanto en 0 y 5 mM de nitrógeno, mientras que, *PvRALF6* muestra valores ligeramente menores en raíces inoculadas con nitrógeno (barras verdes) respecto a raíces sin nitrógeno (barras rojas). Ambos genes presentan mayor acumulación en raíces de 5, 7 y 14 dpi en condiciones con 5 mM de nitrógeno (barras verdes) que sin nitrógeno (barras rojas) (Fig. 27A y B). En el caso del *PvFER1*, se encontró un decremento en el nivel de transcrito a los 3, 5 y 14 dpi en condiciones con nitrógeno (barras verdes) comparado a condiciones sin nitrógeno (Fig. 27C). A los 7 dpi el transcrito es similar en ambas condiciones (Fig. 27C). Estos datos indican que los tres genes responden de manera diferente en condiciones inoculadas con y sin nitrógeno, sugiriendo un papel potencial de estos genes en la inhibición de la nodulación por nitrógeno.

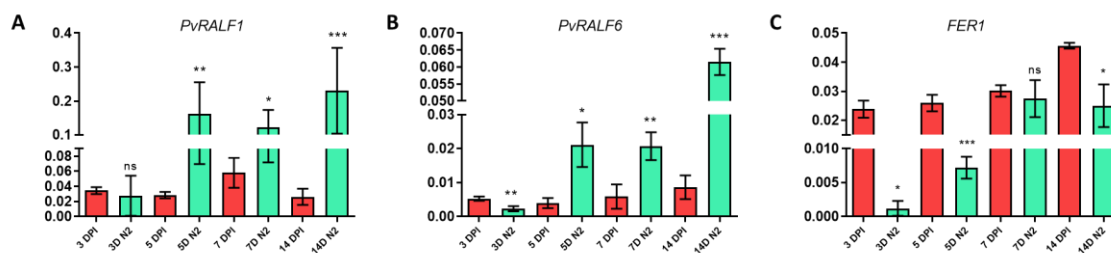


Figura 27. Niveles de acumulación de los transcritos *PvRALF1*, *PvRALF6* y *PvFER1* en raíces de frijol inoculadas con rhizobia a 0 y 5mM de nitrógeno. Nivel de acumulación del transcrito *PvRALF1* (A), *PvRALF6* (B) y *PvFER1* (C) en raíces de frijol inoculadas con rhizobia, sin nitrógeno (barras rojas) y con 5 mM de nitrógeno (barras verdes) a 3, 5, 7 y 14 días. Los datos representan la media \pm desviación estándar (n=9); *P < 0.05, **P < 0.01, ***P < 0.001, ns=no significativo (Student test).

9.7 Análisis del fenotipo en la inhibición de la nodulación por nitrógeno

Se evaluó el efecto que tiene el silenciamiento y la sobreexpresión de *PvRALF1*, *PvRALF6* y *PvFER1* en la inhibición de la nodulación por nitrógeno. Inicialmente, se generaron plantas compuestas de acuerdo a lo descrito en el apartado 7.5, las cuales se transplantaron a macetas con vermiculita como sustrato y se inocularon con *R. tropici* CIAT 899. Las plantas fueron regadas con medio Fahraeus con 2 y 5 mM de nitrógeno, y se procesaron para su análisis a los 21 y 30 dpi.

9.7.1 Efecto del silenciamiento de *PvRALF1*, *PvRALF6* y *PvFER1* en la inhibición de la nodulación por nitrógeno

Para determinar el posible papel de *PvRALF1*, *PvRALF6* y *PvFER1* en la inhibición de la nodulación por nitrógeno, se evaluó el número y los diámetros de los nódulos generados en plantas silenciadas en cada uno de estos genes, en condiciones de nitrógeno inhibitorias (2 y 5 mM).

Al evaluar el fenotipo en la nodulación en plantas silenciadas *PvRALF1*, *PvRALF6* y *PvFER1* en 5 mM de nitrógeno, se encontró que hay una mayor cantidad de nódulos en las raíces silenciadas tanto a 21 como 31 dpi comparado con el control (Fig. 28A-C); en las raíces silenciadas *PvRALF1*, *PvRALF6* y *PvFER1* se encontró un incremento de 60, 50 y 50%, respectivamente, en el número de nódulos, al comparar con las raíces control a 21dpi (Fig. 28A-C); a 31 dpi hubo un aumento de 50, 70 y 40% del número de nódulos en raíces silenciadas *PvRALF1*, *PvRALF6* y *PvFER1*, respectivamente (Fig. 28A-C). Al evaluar el efecto en 2 mM de nitrógeno, se observó un efecto menor del nitrógeno sobre la nodulación. Esto es, una diferencia significativa en el número de nódulos en las plantas silenciadas en los tres genes a 21 dpi comparado con el control, pero no a 31 dpi (Fig. S11A-C). No detectamos cambios significativos en los diámetros de los nódulos tanto en 5 (Fig. 28A-C) como 2 mM de nitrógeno (Fig. S11D a F).

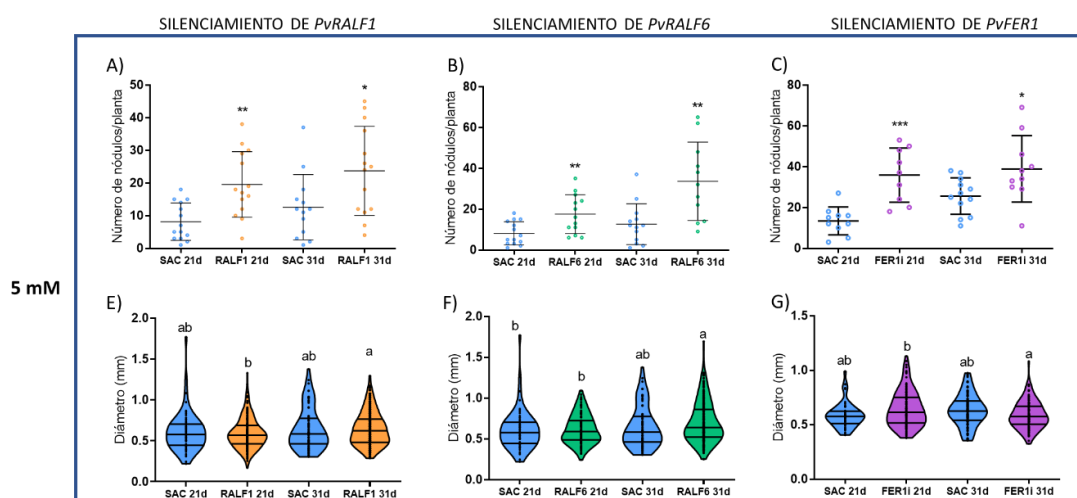


Figura 28. Número y diámetro de los nódulos en las raíces silenciadas en *PvRALF1*, *PvRALF6* y *PvFER1* tratadas con nitrógeno. Numero de nódulos generados a los 21 y 31 dpi, en las raíces transgénicas silenciadas en *PvRALF1* (puntos naranjas, A), *PvRALF6* (puntos verdes, B), *PvFER1* (puntos morados, C) y raíces control (puntos azules) tratadas con 5 mM de nitrógeno. Cuantificación del diámetro de nódulos generados a los 21 y 31 dpi, en raíces transgénicas silenciadas en

PvRALF1 (graficas naranjas, D), *PvRALF6* (graficas verdes, E), *PvFER1* (graficas moradas, F) y raíces control (graficas azules) tratadas con 5 mM de nitrógeno. Los datos representan la media \pm desviación estándar (n=10); *P <0.05, **P <0.01, ***P <0.001 ns=no significativo (Student test).

9.7.2 Análisis del efecto de la sobreexpresión de *PvRALF1*, *PvRALF6* y *PvFER1* en la inhibición de la nodulación por nitrógeno

El fenotipo de las raíces sobreexpresantes *PvRALF1*, *PvRALF6* y *PvFER1* fue evaluada en la inhibición de la nodulación por nitrógeno. En las raíces sobreexpresantes *PvRALF1* y *PvRALF6* hay una disminución en el número de nódulos en las raíces tratadas con 5 mM de nitrógeno tanto a 21 dpi como 31 dpi (Fig. 29A y B); en las raíces sobreexpresantes *PvRALF1* se encontró una disminución de 40 y 42% en el número de nódulos, respectivamente (Fig. 29A) y se observó una disminución del 40 y 36% en el número de nódulos en raíces sobreexpresantes *PvRALF6* (Fig. 29B), a 21 y 31 dpi, respectivamente, comparado con el control (Fig. 29A y B). De forma interesante, no se observaron cambios significativos en el número de nódulos en raíces con sobreexpresión de *PvFER1* (Fig. 29C).

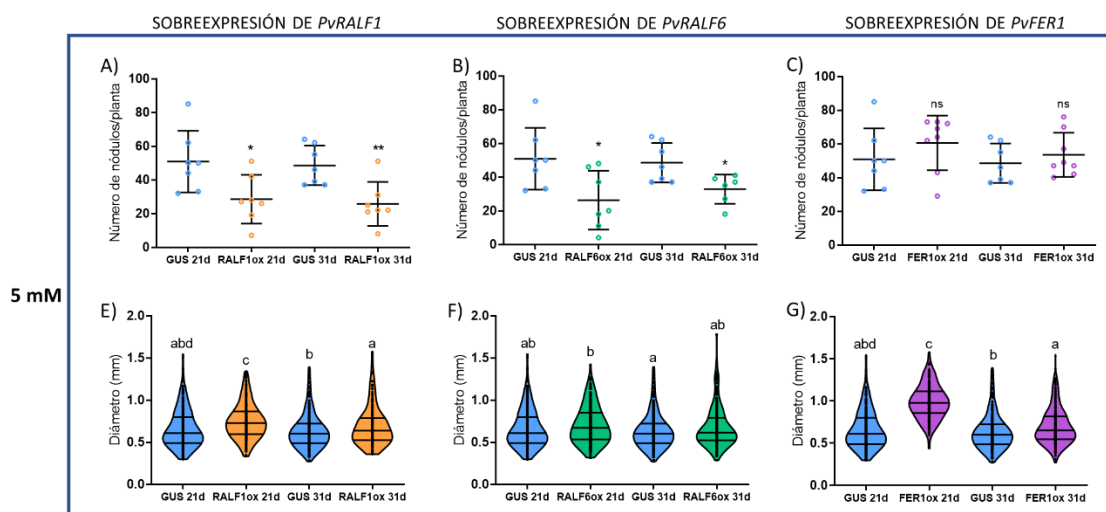


Figura 29. Número de nódulos en raíces transgénicas sobreexpresantes *PvRALF1*, *PvRALF6* Y *PvFER1* tratadas con nitrógeno. Numero de nódulos generados a los 21 y 31 dpi, en raíces transgénicas sobreexpresantes de *PvRALF1* (puntos naranjas, A), *PvRALF6* (puntos verdes, B), *PvFER1* (puntos morados, C) y raíces control (puntos azules) tratadas con 5 mM de nitrógeno. Cuantificación del número de nódulos generados a los 21 y 31 dpi, en raíces transgénicas sobreexpresantes de *PvRALF1* (puntos naranjas, A), *PvRALF6* (puntos verdes, B), *PvFER1* (puntos morados, C) y raíces control (puntos azules) tratadas con 2 mM de nitrógeno. Los datos representan la media \pm desviación estándar (n=8); *P <0.05, **P <0.01, ***P <0.001 ns=no significativo (Student test).

Al evaluar el fenotipo en 2 mM de nitrógeno, se encontró que la inhibición de la nodulación es menor que con 5 mM, tanto en las raíces sobreexpresantes como en las raíces control. Las raíces sobreexpresantes *PvRALF1* muestran una disminución del número de nódulos a los 21 dpi comparado con el control, pero no a los 31 dpi (Fig. S12A); la sobreexpresión de *PvRALF6* provoca una disminución en el número de nódulos a 21 y 31 dpi en 2 mM de nitrógeno comparado con plantas control (Fig. S12B). La sobreexpresión de *PvFER1* presenta un ligero incremento y disminución en el número de nódulos a 21 y 31 dpi, respectivamente (Fig. S12C). Al igual que en las raíces silenciadas, no se detectaron cambios significativos en los diámetros de los nódulos en ninguna de las condiciones analizadas (Fig. 29D a F y Fig. S12D a F).

9.8 Efecto de los cambios en el nivel de expresión de *PvRALF1*, *PvRALF6* y *PvFER1* en genes relacionados con la organogénesis y la función del nódulo

Con el objetivo de determinar el proceso en el cual el silenciamiento y la sobreexpresión de los genes de interés afectan la nodulación, se realizó el análisis de la acumulación de transcrito (RT-qPCR) de genes relacionados con diferentes etapas de la nodulación en dichas raíces transgénicas a 0 y 5 mM de nitrógeno. Los genes “*Nodule Inception*” (*NIN*) y *Leghemoglobina* (*Leg*) se usaron como marcadores de la etapa inicial de la organogénesis del nódulo y de nódulos maduros funcionales, respectivamente (Roy *et al.* 2020). También, usamos como marcadores los genes que codifican las enzimas sacarosa sintasa (*SS*) y glutamato sintetasa (*GS*), las cuales son esenciales en el metabolismo nitrogenado en los nódulos; el primero para proveer a los bacteroides de esqueletos carbonados y el segundo para incorporar del nitrógeno fijado por el bacteroide en compuestos nitrogenados (Carvalho *et al.* 2000; Cordoba *et al.* 2003; Baier *et al.* 2007; Aleman *et al.* 2010; Roy *et al.* 2020).

Al cuantificar los niveles de acumulación de los transcritos de dichos genes en las raíces inoculadas sin nitrógeno, no se observaron diferencias significativas en los transcritos *PvNIN* y *PvSS*, ni en condiciones silenciadas ni en sobreexpresantes de los genes evaluados (Fig. 30A y B y Fig. 31A y B). De manera interesante, al evaluar los niveles de los transcritos *PvLEG*, y *PvGS*, se observó un incremento y una disminución de los transcritos en las raíces silenciadas y sobreexpresantes *PvFER1*, respectivamente (Fig. 30E y F y Fig. 31E y F, barras rosas), pero no de *PvRALF1* ni *PvRALF6*. Estos datos concuerdan con la ausencia de fenotipo observada a tiempos cortos y en los diámetros de los nódulos, indicando que *PvRALF1*,

PvRALF6 y *PvFER1* no participan en el proceso de organogénesis del nódulo y sugiere un papel menor de estos genes en el metabolismo de los nódulos.

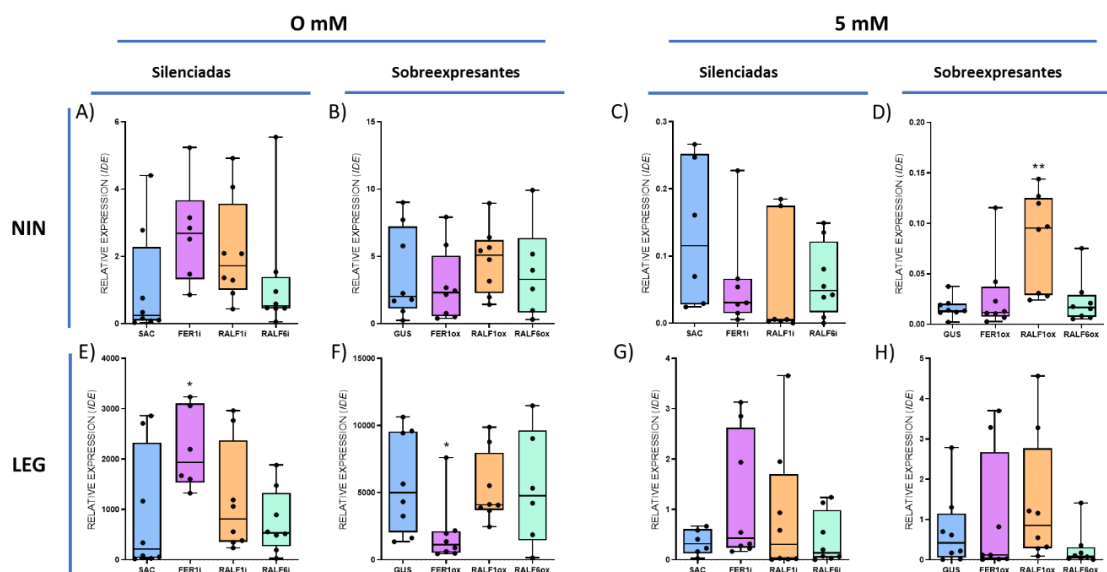


Figura 30. Acumulación de transcritos de los genes *PvNIN* y *PvLEG* en raíces transgénicas de frijol. Niveles de acumulación de los transcritos *PvNIN* (A-D) y *PvLEG* (E-H) en las distintas raíces silenciadas y sobreexpresantes de *PvFER1* (barras moradas) *PvRALF1* (barras anaranjadas), *PvRALF6* (barras verdes) y raíces control (barras azules) inoculadas con 0 y 5 mM de nitrógeno. Los puntos negros representan los datos para cada condición, los bigotes corresponden a la desviación estándar (n=12); *P < 0.05, **P < 0.01, ns=no significativo.

En el caso de las raíces inoculadas adicionales con nitrógeno (5 mM), no se observan cambios en los niveles de acumulación del transcrito *PvNIN* debido al silenciamiento de los genes evaluados (Fig. 30C), tampoco cambios en el transcrito de *PvNIN* en la sobreexpresión de *PvFER1* o *PvRALF6* (Fig. 30D). Únicamente se encontró un incremento de *PvNIN* en raíces sobreexpresantes *PvRALF1* comparado con raíces control (Fig. 30D, barra naranja). Los niveles del transcrito *PvLEG*, por su parte, no se alteran en ninguna de las raíces transgénicas analizadas, al comparar con los controles (Fig. 30G y H).

De manera interesante, en raíces silenciadas *PvFER1*, y no *PvRALF1* ni *PvRALF6*, se observa una disminución de la expresión de *PvSS* y *PvGS* (Fig. 31C y G), mientras que únicamente la sobreexpresión de *PvRALF1* provoca un incremento de estos transcritos (Fig. 31D y H). Estos datos sugieren nuevamente que estos genes no participan en el proceso de organogénesis del nódulo y una participación menor en la regulación del metabolismo de los nódulos.

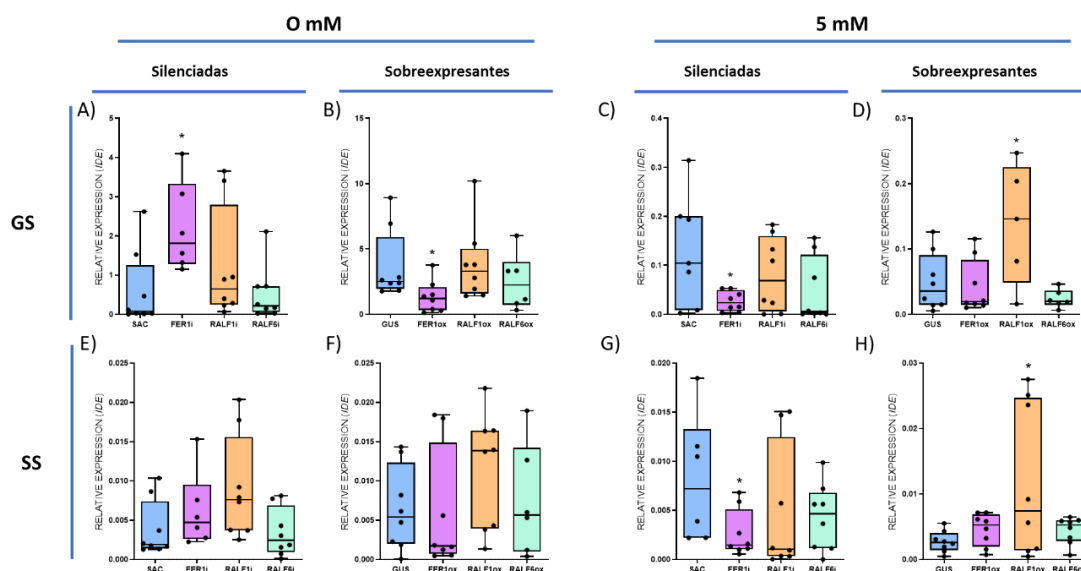


Figura 31. Acumulación de transcritos de los genes *PvGS* y *PvSS* en raíces transgénicas de frijol. Niveles de acumulación de transcritos *PvGS* (A-D) y *PvSS* (E-H) en las distintas raíces silenciadas y sobreexpresantes *PvFER1* (barras moradas) *PvRALF1* (barras anaranjadas), *PvRALF6* (barras verdes) y raíces control (barras azules) inoculadas con 0 y 5 mM de nitrógeno. Los puntos negros representan los datos para cada condición, los bigotes corresponden a la desviación estándar (n=12); *P < 0.05, **P < 0.01, ns=no significativo.

9.9 Efecto del silenciamiento y la sobreexpresión de *PvRALF1*, *PvRALF6* y *PvFER1* en genes de frijol relacionados con las vías AON y NRN

La vía de AON ha sido descrita como un mecanismo de control del número de nódulos por parte de la planta, con la finalidad de regular el gasto energético (Ferguson *et al.* 2019). En el caso de la vía NRN, ésta regula el número de nódulos en función del contenido de nitrógeno en el suelo (Ferguson *et al.* 2019). Debido a los efectos observados en el número de nódulos debido al silenciamiento y sobreexpresión de los genes de interés, se decidió evaluar en raíces de frijol la expresión de los genes “*Too Much Love*” (*TML*) y *NARK* en dichas raíces transgénicas, los cuales son importantes para la regulación del número de nódulos por la vía AON y NRN en otras leguminosas (Gautrat *et al.* 2020; Roy *et al.* 2020).

En condiciones sin nitrógeno, no se observaron diferencias en la expresión de *PvTML* en las raíces silenciadas ni en las sobreexpresantes al comparar con el control (Fig. 32A y B). Sin embargo, en 5 mM de nitrógeno se observa que las raíces silenciadas *PvFER1* y *PvRALF6*, la acumulación del transcrito de este gen es más bajo comparado con el control, y solo se observa una tendencia a menor expresión en las raíces silenciadas *PvRALF1* (Fig. 32C). Por el contrario, en las raíces sobreexpresantes *PvRALF1*, se encontró un incremento en la

acumulación del transcrito *PvTML* en 5 mM de nitrógeno al comparar con el control (Fig. 32E); no se detectaron diferencias en las raíces sobreexpresantes *PvFER1* ni *PvRALF6* (Fig. 32D). Estos datos sugieren que *PvFER1*, *PvRALF1* y *PvRALF6* podrían estar participando en la regulación del número de nódulos en presencia de nitrógeno a través del inhibidor de la nodulación TML.

Por su parte, al analizar los niveles de transcrito del gen *PvNARK* en raíces silenciadas en los tres genes de interés, se encontró un incremento en raíces inoculadas en 0 mM de nitrógeno, y un aumento en estas mismas raíces en 5 mM de nitrógeno (Fig. 32E y G). *PvNARK* y sus homólogos en otras leguminosas son inhibidores de la nodulación en diferentes leguminosas (Ferguson *et al.* 2019), los cambios en los niveles de expresión de *PvNARK* correlacionan con la disminución y el incremento de nódulos en las raíces silenciadas en 0 y 5 mM de nitrógeno, respectivamente. De manera interesante, no se observan cambios en las raíces sobreexpresantes de ninguno de los genes, tanto a 0 como 5 mM de nitrógeno (Fig. 32F y H).

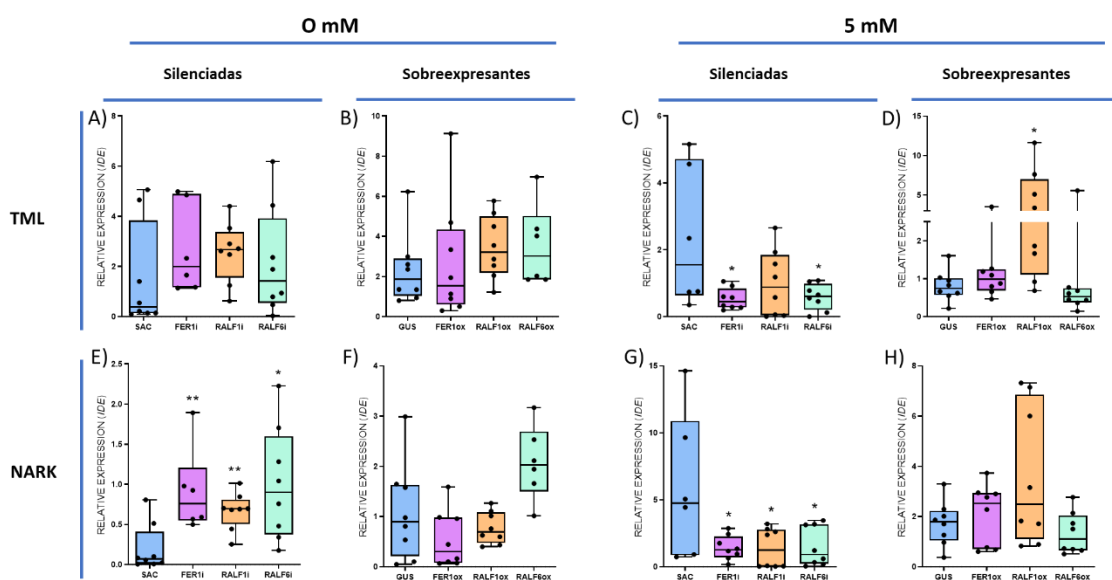


Figura 32. Acumulación de transcritos de los genes *PvTML* y *PvNARK* en raíces transgénicas de frijol. Niveles de acumulación de transcritos *PvTML* (A-D) y *PvNARK* (E-H) en las distintas raíces silenciadas y sobreexpresantes de *PvFER1* (barras moradas) *PvRALF1* (barras anaranjadas), *PvRALF6* (barras verdes) y raíces control (barras azules) inoculadas con 0 y 5 mM de nitrógeno. Los puntos negros representan los datos para cada condición, los bigotes corresponden a la desviación estándar (n=12); *P < 0.05, **P < 0.01, ns=no significativo.

Además de *NARK* y *TML*, las vías AON y NRN dependen de la función de diferentes péptidos CLEs. En un reporte previo, se encontró que el gen *PvCLE21* de frijol es el

homólogo de *NICI* de soya, inhibición de la nodulación por nitrógeno en esta leguminosa (Hastwell *et al.* 2015). Así mismo, *PvCLE38* se describió como el homólogo de *RIC2* de soya, el cual promueve la inhibición de la nodulación a través de la vía AON en esta leguminosa (Hastwell *et al.* 2015). Se reportó en frijol que los genes *PvCLE19* y *PvCLE38* se expresan en condiciones de bajo fosfato en frijol, y que participan en la inhibición de la nodulación a través de la interacción de su producto peptídico con el receptor NARK en estas condiciones (Isidra-Arellano *et al.* 2020). A partir de estos antecedentes, y tras la evaluación del patrón de expresión de los 40 genes que codifican péptidos CLEs en frijol (Fig. S13) en el PvGEA (O'Rourke *et al.* 2014); se seleccionaron cuatro genes para su análisis en raíces transgénicas de frijol, *PvCLE4* que se expresa principalmente en raíces inoculadas y fertilizadas, *PvCLE18* en raíces inoculadas a 5 dpi, *PvCLE21* (*PvNIC1*) en nódulos maduros y *PvCLE38* (*PvRIC2*) en nódulos de 5dpi (Fig. S13).

Al evaluar la expresión por RT-qPCR de los cuatro genes *PvCLE* en raíces silenciadas sin nitrógeno, se observó un incremento en el nivel de acumulación de los transcritos de los genes *PvCLE4*, *PvCLE18* y *PvNIC1* en las raíces silenciadas *PvFER1*, *PvRALF1* y *PvRALF6*, al comparar con las raíces control (Fig. 33A Fig. 34A y E); además, encontramos un incremento del transcrito *PvRIC2* en las raíces silenciadas *PvFER1* y *PvRALF1*, y no en *PvRALF6* (Fig. 33E). Contrario a esto, se observó una disminución en la expresión de *PvNIC1* y *PvRIC2* en raíces sobreexpresantes *PvFER1* y *PvRALF1* (Fig. 33B y F). No se encontró ningún cambio en los niveles de los transcritos *PvCLE4* y *PvCLE18* comparando con el control (Fig. 34B y F). En el caso de las raíces sobreexpresantes *PvRALF6*, hubo un incremento en la acumulación del transcrito *PvRIC2* y una disminución del de *PvCLE4* (Fig. 33B y Fig. 34B). Estos resultados sugieren que la disminución del número de nódulos en raíces silenciadas *PvFER1* y *PvRALF6* y el incremento de nódulos en raíces sobreexpresantes de *PvFER1* y *PvRALF1* en condiciones sin nitrógeno se debe en parte a la desregulación en la expresión de *PvCLE4*, *PvCLE18*, *PvNIC1* y *PvRIC2*.

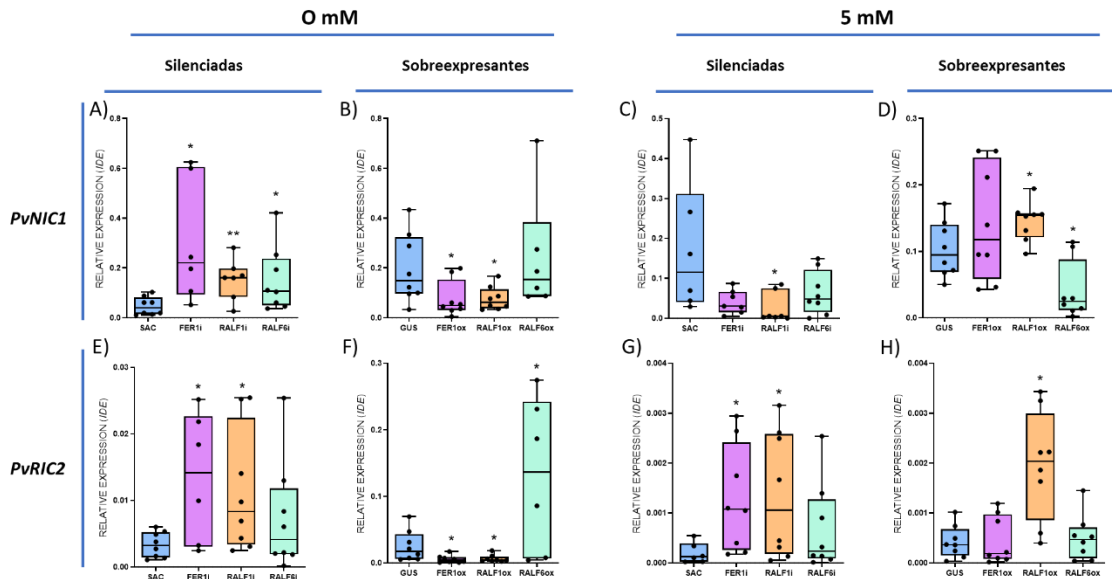


Figura 33. Acumulación de transcritos de los genes *PvNIC1* y *PvRIC2* en raíces transgénicas de frijol. Niveles de acumulación de transcritos *PvNIC1* (A-D) y *PvRIC2* (E-H), en las distintas raíces silenciadas y sobreexpresantes de *PvFER1* (barras moradas) *PvRALF1* (barras anaranjadas), *PvRALF6* (barras verdes) y raíces control (barras azules) inoculadas con 0 y 5 mM de nitrógeno. Los puntos negros representan los datos para cada condición, los bigotes corresponden a la desviación estándar (n=12); *P < 0.05, **P < 0.01, ns=no significativo.

Posteriormente, se evaluó la acumulación de los transcritos de estos cuatro genes *PvCLE* en las raíces transgénicas inoculadas en 5 mM de nitrógeno. Los resultados obtenidos muestran que en las raíces silenciadas *PvFER1* hay un incremento en el nivel del transcrito *PvCLE4* y en las silenciadas *PvRALF1K* existe una disminución de la expresión de *PvNIC1* y un incremento de *PvCLE4*, en tanto que los demás genes que codifican dichos péptidos no presentan cambios en sus niveles de transcrito. Así mismo, el silenciamiento de *PvRALF6* no afecta los niveles de expresión de *PvCLEs*. Al comparar los niveles de acumulación de transcrito en las raíces sobreexpresantes comparados con el control, se observó un aumento en la acumulación del transcrito *PvCLE18* en las raíces sobreexpresantes *PvFER1*, un incremento en la expresión de *PvCLE18*, *PvCLE21* y *PvRIC2* en las raíces sobreexpresantes *PvRALF1*, y una disminución de los niveles del transcrito *PvNIC1* debido a la sobreexpresión de *PvRALF6*.

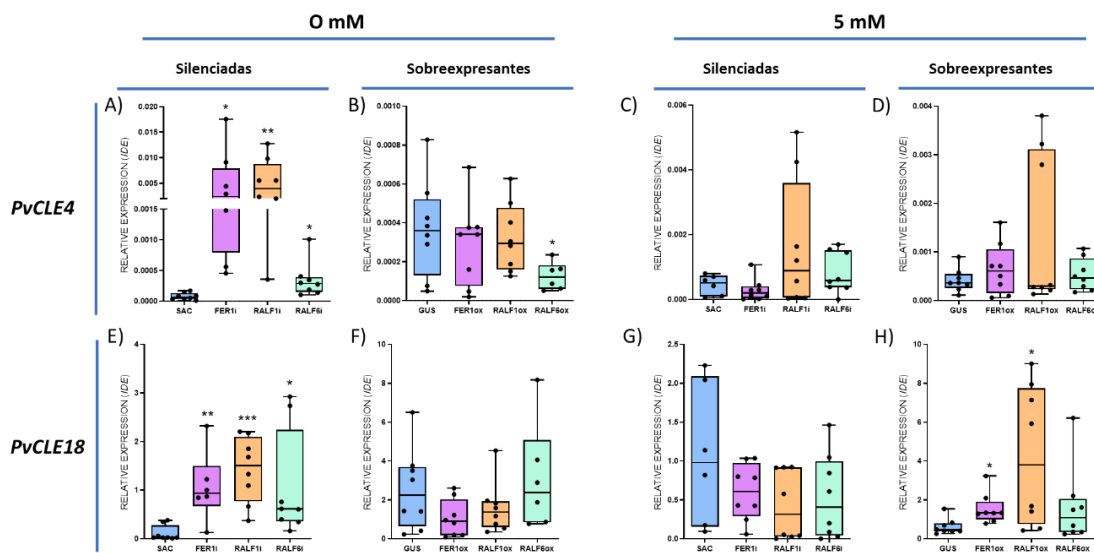


Figura 34. Acumulación de transcritos de los genes *PvCLE4* y *PvCLE18* en raíces transgénicas de frijol. Niveles de acumulación de transcritos *PvCLE4* (A-D) y *PvCLE8* (E-H), en las distintas raíces silenciadas y sobreexpresantes de *PvFER1* (barras moradas) *PvRALF1* (barras anaranjadas), *PvRALF6* (barras verdes) y raíces control (barras azules) inoculadas con 0 y 5 mM de nitrógeno. Los puntos negros representan los datos para cada condición, los bigotes corresponden a la desviación estándar (n=12); *P <0.05, **P <0.01, ns=no significativo.

Estos datos sugieren que al modificarse la señalización mediada por *PvFER1*, *PvRALF1* y *PvRALF6*, se desregulan las vías AON, NRN en frijol, y de esta forma el número de nódulos se ve afectado.

10. DISCUSIÓN

La familia RALF, está constituida por un grupo de péptidos ricos en cisteína, cuya función se ha asociado con su capacidad de promover la alcalinización del apoplasto. En *A. thaliana*, la percepción de estos péptidos por receptores tipo cinasa de la subfamilia CrRLK1L en la membrana celular ha sido ampliamente documentada en diversos de procesos, así como una gran cantidad interactores involucrados en esta vía de señalización (Tabla 1). La participación tanto de RALF, FER, y RIPK en procesos de desarrollo de los pelos radicales, expansión celular y respuesta a patógenos nos ha permitido proponerlos como potenciales participantes del proceso de nodulación.

10.1 Características de los genes *RALF* y *CrRLK1L* de frijol

Las plantas terrestres surgieron hace aproximadamente 500 millones de años (Niklas and Kutschera 2010; Morris et al. 2018). Durante su evolución, han sufrido grandes expansiones génicas debido principalmente a duplicaciones dispersas, en tándem y de genoma completo (Lynch and Conery 2000; Bowers et al. 2003; Coghlan et al. 2005; Freeling 2009; Wang et

al. 2011). Con base en los análisis *in silico* de los miembros de la familia RALF y la subfamilia CrRLK1L de frijol, se pueden hacer varias observaciones. De acuerdo a Campbell y Turner, en las Brassicaceas, como *A. thaliana*, la familia de genes *RALF* sufrieron una duplicación génica por duplicaciones en tándem, formándose grandes clústeres de estos genes en diferentes cromosomas (Campbell and Turner 2017). En el caso de los *RALFs* de frijol, se deduce que esto no ocurrió ya que hay cuatro veces menos péptidos RALF que en *A. thaliana*, y se encuentran distribuidos de manera aislada en el genoma de frijol. Por el contrario, existen 33 genes *CrRLK1L* en frijol, los cuales están altamente conservados y sufrieron una expansión génica, posiblemente por duplicaciones en tándem, presentándose clústeres de estos genes en dicho genoma (Solis-Miranda *et al.* 2020).

Se sabe que la presencia de intrones determina la función del gen, afectando al estabilidad genómica y el procesamiento del mensajero primario (Mattick and Gagen 2001; Schwartz *et al.* 2009). Así mismo, la expansión de familias de genes permite una mayor divergencia génica que cuando su cantidad se mantiene restringida, existiendo además una relación estrecha entre los tiempos de divergencia entre especies con los patrones de los intrones (Zhu *et al.* 2009). Nuestras observaciones sugieren que los patrones en el número de intrones en ambos grupos de genes se asocian con las duplicaciones génicas y la evolución de cada especie. Los péptidos RALF de frijol, al igual que en *A. thaliana*, no presentan intrones, mientras que la subfamilia de receptores CrRLK1L solo una proporción menor (6 en frijol y 3 en *A. thaliana*) presenta intrones.

Los datos aquí presentados sugieren que hay una alta conservación de los péptidos RALF y los receptores CrRLK1L en frijol, posiblemente debido a que los péptidos RALF son los únicos ligandos de los receptores CrRLK1L reportados a la fecha.

10.2 Características de la expresión de los genes *RALF* y *CrRLK1L* en frijol

En *A. thaliana*, se ha descrito que los péptidos RALF presentan expresión diferencial en toda la planta. Al analizar la expresión de los diferentes genes *PvRALF* de frijol, aun cuando en frijol son menos genes, se observa que estos también se expresan en todos los tejidos de la planta (Fig. 8). Se ha descrito que los péptidos RALF de diferentes especies de plantas presentan la misma actividad, sugiriendo que su función está asociada al tejido donde se expresa y no depende de cada péptido (Morato do Canto *et al.* 2014; Murphy and De Smet

2014). En el caso de los genes *CrRLK1L*, encontramos que la expresión de estos genes en seis diferentes especies de plantas, presentan una expresión similar a lo observado en *A. thaliana*, con genes con expresión restringida a ciertos tejidos (Solis-Miranda *et al.* 2020). Debido a la estrecha relación entre ellos, es posible que la función de los péptidos RALF se encuentre también muy conservada al igual que los receptores CrRLK1L.

El hecho de que la actividad de los promotores *PvRALF1*, *PvRALF6* y *PvFER1* de frijol se encontró en los ápices de las raíces y en los primordios de raíces laterales y de nódulos, sugiere que estos genes son relevantes en la división celular o en el mantenimiento en estas zonas meristemáticas, funciones en las que no ha sido descrita la participación de ninguno de estos genes. Que el promotor se mantenga activo en zonas arriba de los ápices, podría estar relacionado con sus funciones en la elongación de la raíz, así como en procesos de regulación del crecimiento, en la integridad de la pared celular, en el estado metabólico, y en respuesta al estrés, procesos en los cuales ha sido descrita la participación de genes homólogos a los aquí estudiados (Tabla 1).

PvRALF1, *PvRALF6* y *PvFER1* se expresan tanto en raíces inoculadas como no inoculadas; por su parte, los análisis de la actividad de promotor indican que el promotor de estos genes es activo durante diferentes fases de la nodulación, lo que sugiere que estos genes son necesarios durante las distintas etapas de dicho proceso. Por otro lado, los niveles de acumulación de los transcrito *PvRALF1*, *PvRALF6* y *PvFER1* presentaron cambios en respuesta a la presencia de nitrógeno en medio, tras la inoculación con rhizobia (Fig. 27), sugiriendo que estos genes podrían participar en la inhibición de la nodulación por nitrógeno.

10.3 Papel de *PvRALF1*, *PvRALF6* y *PvFER* en la organogénesis del nódulo

La nodulación es un proceso complejo, que requiere la integración y regulación de una serie de mecanismos por parte de la planta y la bacteria, con la finalidad de establecer un nuevo órgano en las raíces de la planta, el nódulo (Ferguson *et al.* 2019; Roy *et al.* 2020). La participación de muchos genes en la regulación de la organogénesis del nódulo y la etapa en la cual está involucrados, ha sido determinada usando diferentes enfoques. Por ejemplo, plantas mutantes de *L. japonicus* en los receptores LjNFR1 y LjNFR5 (*NOD FACTOR PERCEPTION*) son incapaces de responder a la presencia de dichos factores nod y por tanto no nodulan (Radutoiu *et al.* 2003). Líneas vegetales con un transposón insertado en la

secuencia del regulador maestro de la nodulación *NIN* son incapaces de formar nódulos e iniciar la infección, a pesar de que se genera un incremento de pelos radicales enroscados (Schauser *et al.* 1999). Mutantes en genes relacionados al tráfico vesicular, como son *VAMP721d* y *VAMP721e* (*VESICLE-ASSOCIATED MEMBRAN PROTEIN*), son incapaces de liberar las bacterias en los nódulos, generando nódulos vacíos (Ivanov *et al.* 2012). Plantas mutantes obtenidas por CRISPR-Cas9 en tres genes de *leghemoglobina* generan nódulos incapaces de fijar nitrógeno debido a la inhibición de la nitrogenasa (Wang *et al.* 2019). La reducción en los niveles de las enzimas *sacarosa sintasa* en *M. truncatula* y *glutamato sintasa* en alfalfa provoca la generación de nódulos con baja capacidad de fijación y disminución en la síntesis de aminoácidos en los nódulo (Cordoba *et al.* 2003; Baier *et al.* 2007). A pesar de que se conoce la función de una gran cantidad de genes en el proceso de nodulación (Roy *et al.* 2020), aún hay mucho que se explore en relación a los mecanismos celulares y moleculares que ocurren en este proceso.

En condiciones de promoción e inhibición de la nodulación (0 y 5 mM de nitrógeno, respectivamente), nosotros encontramos que *PvRALF1*, *PvRALF6* y *PvFER1* no parecen participar directamente en el desarrollo de los nódulos, y posiblemente tengan un rol no esencial en su metabolismo nitrogenado (Fig. 22 a 24 y 26). Esto correlaciona con la ausencia de cambios en los niveles de expresión de genes marcadores de la organogénesis del nódulo (*PvNIN* y *PvLEG*, Fig. 30) y con algunos cambios en genes marcadores del metabolismo del nódulo (*PvGS* y *PvSS*, Fig. 31) tanto en raíces silenciadas como sobreexpresantes en estos genes.

10.4 Papel de los péptidos RALF y de FERONIA en las vías AON y NRN

En diferentes modelos de leguminosas, se ha descrito una vía de regulación del número de nódulos que involucra un proceso de comunicación sistémica entre las raíces de la planta y el tejido aéreo, conocido como autorregulación de la nodulación (AON), así como un proceso de regulación local mediado por la presencia de nitrógeno asimilable en el suelo, llamado regulación de la nodulación por nitrógeno (NRN) (Ferguson *et al.* 2019; Roy *et al.* 2020). Análisis de líneas mutantes en diferentes genes relacionados con estas vías, muestran un fenotipo relacionado al número de nódulos sin afectar el desarrollo de estos. La adición exógena a raíces de varias leguminosas de los péptidos de la familia CLE participantes de estas vías, reducen la formación de nódulos (Okamoto *et al.* 2009, 2013; Mortier *et al.* 2010,

2012; Reid et al. 2011). Mutantes en los receptores de estos péptidos (*NARK* en *Glycine max*, *SUNN* en *M. truncatula* y *HARI* en soya) generan un número excesivo de nódulos (Krusell et al. 2002; Searle et al. 2003; Schnabel et al. 2005). Por otro lado, el silenciamiento del gen *TML* por RNA interferente induce un incremento en el número de nódulos (Gautrat et al. 2019, 2020). En un trabajo previo en frijol, se reportó la inducción de dos genes *CLE* (*PvRIC1* y *PvRIC2*) en raíces en condiciones limitantes de fósforo, y que dichos péptidos provocan una inhibición de la nodulación a través del receptor *NARK* en tejido aéreo (Isidra-Arellano et al. 2020), y una sobreposición entre estas vías de inhibición de la nodulación.

Los resultados obtenidos en este trabajo indican que *PvRALF1*, *PvRALF6* y *PvFER1* juegan un papel en determinar del número de nódulos a tiempos largos (21 y 31 dpi) en condiciones con y sin nitrógeno (Fig. 22, 26, 28 y 29). Esto correlaciona con datos obtenidos en relación a niveles de expresión de los genes involucrados en la AON y NRN. Esto es, la disminución en los niveles de transcritos *PvRALF1*, *PvRALF6* y *PvFER1* induce un incremento en la acumulación de los transcritos *PvNARK* y *PvCLEs* en condiciones promotoras de la nodulación (0 mM de nitrógeno) (Fig. 32 a 39), y la disminución de los niveles de transcritos *PvNARK*, *PvTML*, y *PvCLEs* en condiciones inhibitoras de la nodulación (5 mM de nitrógeno) (Fig. 32 a 39), que se corresponde con una disminución y un incremento en el número de nódulos observado, respectivamente (Fig. 22, 26 y 28). Por su parte, el aumento en el nivel de transcritos *PvRALF1*, *PvRALF6* y *PvFER1* provoca el efecto opuesto (Fig. 32 a 39).

Estos datos indican que los fenotipos observados en raíces transgénicas pueden ser explicados, al menos en parte, por la desregulación de la expresión de *PvNARK*, *PvTML*, *PvCLE4*, *PvCLE18*, *PvRIC2* y *PvNIC1*, y posiblemente de otros *CLEs* no evaluados aquí, sugiriendo un papel de *PvRALF1*, *PvRALF6* y *PvFER1* como reguladores negativos de las vías AON y NRN en condiciones sin nitrógeno, y como reguladores positivos de ambas vías en condiciones inhibitoras de nitrógeno (Fig. 35). El modelo que proponemos y que integra los resultados obtenidos en este proyecto, se muestra en la Fig. 35.

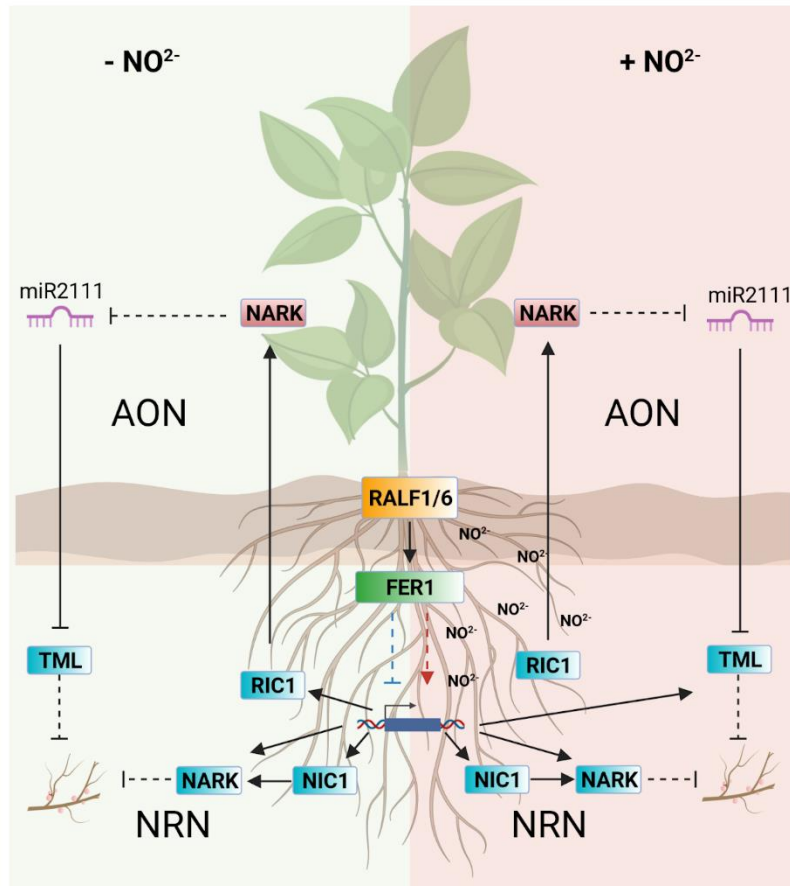


Figura 35. Modelo propuesto de la participación de PvRALF1, PvRALF6 y PvFER1 de frijol en raíces de frijol durante la nodulación. En condiciones promotoras de la nodulación (panel izquierdo), PvRALF1, PvRALF6 y PvFER1 funcionan como reguladores negativos de la expresión de genes reguladores del número de nódulos de las vías AON y NRN. En condiciones inhibitoras de la nodulación (panel derecho), PvRALF1, PvRALF6 y PvFER1 funcionan como reguladores positivos de la expresión de genes reguladores del número de nódulos de las vías AON y NRN.

11. CONCLUSIONES

- La familia de péptidos PvRALF en frijol está formada por nueve miembros
- La familia de cinasas citoplasmáticas PvRIPK en frijol está constituida por seis miembros
- Las secuencias de aminoácidos en la región carboxilo terminal correspondiente al péptido maduro en PvRALF de frijol y al dominio cinasa de PvRIPK en frijol están altamente conservadas
- Los genes *PvRALF1*, *PvRALF6*, *PvRIPK2* y *PvFER1* se expresan en nódulos de frijol durante diferentes etapas, y sus promotores son activos a lo largo de la nodulación
- *PvRALF1*, *PvRALF6* y *PvFER1* participan en la determinación del número de nódulos a 0 mM de nitrógeno a través de la regulación negativa de genes de las vías AON y NRN
- *PvRALF1*, *PvRALF6* y *PvFER1* participan en determinar el número de nódulos a 5 mM de nitrógeno, a través de la regulación positiva de genes que participan en las vías AON y NRN
- *PvRALF1*, *PvRALF6* y *PvFER1* no parecen tener funciones en el desarrollo de los nódulos en frijol

12. PERSPECTIVAS

Este trabajo se determinó a dos péptidos RALF y al receptor FER de frijol con un rol negativo en la determinación del número óptimo de nódulos en condiciones promotoras de simbiosis, y un rol positivo en condiciones inhibitorias por nitrógeno. Reportamos que el proceso de señalización empleado por dichos péptidos parece ser a través de FER, el cual ha sido descrito por primera vez involucrado en este proceso simbiótico, y que dicho mecanismo de control parece ser a través de la regulación de la vía AON y NRN. Sin embargo, varias cuestiones surgen alrededor de este proyecto, abriendo ventanas hacia diferentes proyectos para investigación a futuro. Entre ellas surge la necesidad de demostrar la posible interacción entre PvRALF1 y PvRALF6 con PvFER1, surge la duda de si el mecanismo de regulación del número de nódulos aquí descrito está conservado en otras leguminosas, y queda pendiente por determinarse el mecanismo de acción a través del cual PvRALF1, y PvFER1 regulan de manera diferencial el número de nódulos en respuesta a la presencia de nitrógeno y otros nutrientes esenciales para la planta.

La respuesta a estas preguntas permitirá entender de manera más profunda el mecanismo por el cual los péptidos RALF y los receptores FER pueden regular tanto positiva como negativamente la cantidad óptima de nódulos en respuesta a las condiciones del suelo, en este caso por nitrógeno. Algunos estudios que permitirían entender de manera más profunda el proceso aquí descrito se enumeran a continuación.

1. Análisis de la posible interacción de PvRALF1 y PvRALF6 con PvFER1. En numerosos trabajos en *Arabidopsis* ha sido reportada la interacción entre RALFs y receptores de la subfamilia CrRLK1L, principalmente con FER (Tabla 1), y en un estudio reciente se reportó que al menos la mitad de los 37 RALFs de *Arabidopsis* interactúan con FER, y que esta interacción es necesaria para activar su función (Abarca et al. 2021). Determinar la interacción de PvRALF1 y PvRALF6 con PvFER1 en frijol nos permitiría comprobar que esta interacción entre RALFs y FER ocurre fuera de la planta modelo *Arabidopsis*, y permitiría comprobar el modelo propuesto en el presente trabajo. Experimentos de complementación bimolecular de

la fluorescencia (BiFC), co-inmunoprecipitación (CoIP) y de doble híbrido de levadura (Y2H) permitirían responder esta cuestión.

2. Análisis morfológico de nódulos de las raíces silenciadas y sobreexpresantes. De acuerdo a los datos obtenidos aquí, *PvRALF1*, *PvRALF6* y *PvFER1* no parecen tener una función en la organogénesis o la función de los nódulos, sin embargo, el análisis de la morfología de los nódulos generados en dichas raíces transgénicas permitirá evaluar a mayor profundidad si existen alteraciones en el desarrollo de los nódulos o durante la diferenciación del bacteroide debido al silenciamiento o la sobreexpresión de *PvRALF1*, *PvRALF6* y *PvFER1*. Esta cuestión se puede abordar fácilmente mediante el análisis de cortes de nódulos procedentes de raíces silenciadas y sobreexpresantes por microscopía.
3. Determinar el efecto de la adición de péptidos sintéticos durante la nodulación, y confrontar con el fenotipo observado en las raíces silenciadas y sobreexpresantes de *PvRALF1*, *PvRALF6* y *PvFER1*. La gran mayoría de los estudios de los péptidos RALF, han sido empleados péptidos sintéticos para determinar su función (Blackburn et al. 2020; Abarca et al. 2021). El análisis del efecto de la adición de péptidos sintéticos RALF en el proceso de nodulación en raíces de frijol, nos permitiría evaluar si el fenotipo observado en las raíces transgénicas silenciadas de *PvRALF1* y *PvRALF6* se recupera en presencia del péptido sintético, y si el fenotipo observado en raíces transgénicas sobreexpresantes de *PvRALF1* y *PvRALF6* sucede en raíces WT tratadas con el péptido.
4. Análisis de la conservación de este mecanismo regulatorio para el control del número óptimo de nódulos en otras leguminosas. Se ha descrito la participación de MtRALF1 en el proceso de nodulación en *M. truncatula*, sin embargo, en esta leguminosa se reportaron efectos en la morfología del nódulo además del número de nódulos. Sería interesante determinar si en otras leguminosas, como *Lotus japonicus* y *Glycine max* tanto los péptidos RALF como receptores FER presentan funciones en la nodulación, mediante el estudio de mutantes en genes FER y RALF en estas leguminosas bajo condiciones simbióticas.

5. Análisis de la participación de otros interactores conocidos de FER conocidos. En el presente trabajo, determinamos que *PvRALF1*, *PvRALF6* y *PvFER1* de frijol tienen un papel en la determinación del número de nódulos tanto en condiciones promotoras como inhibitorias de la simbiosis por nitrógeno. Sin embargo, este proceso parece ser independiente de la cinasa RIPK, uno de los interactores descritos que participan corriente debajo de FER. Analizar la participación de otros interactores de FER conocidos permitiría determinar la vía a través de la cual ocurre este proceso. Para ello se puede realizar el estudio por genética reversa en frijol de manera similar al realizado en este trabajo, y complementarlo con estudios de mutantes en otras leguminosas.
6. Análisis del perfil de expresión por RNAseq de las raíces transgénicas silenciadas y sobreexpresantes de *PvRALF1*, *PvRALF6* y *PvFER1*. Se ha descrito en *A. thaliana* que RALF1 a través de FER, regula la expresión de genes asociados a respuesta inmune, desarrollo de la raíz, entre otros (Tabla 1). En el presente trabajo, determinamos cambios en la expresión de algunos genes de las vías AON y NRN mediante análisis por RT-qPCR, sin embargo, un análisis por RNAseq de las raíces transgénicas silenciadas y sobreexpresantes de *PvRALF1*, *PvRALF6* y *PvFER1* permitirá determinar cambios globales en la expresión de todo el genoma, y poder determinar con mayor profundidad las vías que están siendo afectadas en dichas raíces transgénicas.
7. Analizar el papel de *PvRALF1*, *PvRALF6* y *PvFER1* en el proceso de nodulación en condiciones inhibitorias por limitación de fosfato inorgánico (Pi). Se sabe que en leguminosas la limitación de Pi impacta fuertemente el proceso de nodulación (Hernández et al. 2009; Nasr Esfahani et al. 2017; Isidra-Arellano et al. 2020). En un trabajo reciente en frijol, se reportó la inducción de dos genes *CLE* (*PvRIC1* y *PvRIC2*) en raíces en condiciones limitantes de fósforo, y que dichos péptidos participan en la inhibición de la nodulación a través del receptor NARK en tejido aéreo (Isidra-Arellano et al. 2020), indicando que la vía de AON es activada durante la inhibición de la nodulación por bajo Pi. El análisis del fenotipo en la nodulación del silenciamiento y la sobreexpresantes de *PvRALF1*, *PvRALF6* y *PvFER1* bajo

condiciones limitantes de Pi permitiría determinar si existe una regulación de los genes de las vías de AON por parte de RALF y FER en estas condiciones limitantes.

13. REFERENCIAS

- Abarca A, Franck CM, Zipfel C (2021) Family-wide evaluation of RAPID ALKALINIZATION FACTOR peptides. *Plant Physiol* 1–15. <https://doi.org/10.1093/plphys/kiab308>
- Aleman L, Ortega JL, Martinez-Grimes M, et al (2010) Nodule-enhanced expression of a sucrose phosphate synthase gene member (MsSPSA) has a role in carbon and nitrogen metabolism in the nodules of alfalfa (*Medicago sativa* L.). *Planta* 231:233–244. <https://doi.org/10.1007/s00425-009-1043-y>
- Baier MC, Barsch A, Küster H, Hohnjec N (2007) Antisense repression of the *Medicago truncatula* nodule-enhanced sucrose synthase leads to a handicapped nitrogen fixation mirrored by specific alterations in the symbiotic transcriptome and metabolome. *Plant Physiol* 145:1600–1618. <https://doi.org/10.1104/pp.107.106955>
- Bailey TL, Boden M, Buske FA, et al (2009) MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res* 37:202–208. <https://doi.org/10.1093/nar/gkp335>
- Bardou F, Merchan F, Ariel F, Crespi M (2011) Dual RNAs in plants. *Biochimie* 93:1950–1954. <https://doi.org/10.1016/j.biochi.2011.07.028>
- Batut J, Mergaert P, Masson-Boivin C (2011) Peptide signalling in the rhizobium-legume symbiosis. *Curr Opin Microbiol* 14:181–187. <https://doi.org/10.1016/j.mib.2010.12.010>
- Bergonci T, Ribeiro B, Ceciliato PHO, et al (2014) Arabidopsis thaliana RALF1 opposes brassinosteroid effects on root cell elongation and lateral root formation. *J Exp Bot* 65:2219–2230. <https://doi.org/10.1093/jxb/eru099>
- Blackburn MR, Haruta M, Moura DS (2020) Twenty years of progress in physiological and biochemical investigation of RALF peptides. *Plant Physiol* 182:1657–1666. <https://doi.org/10.1104/PP.19.01310>
- Boisson-Dernier A, Kessler SA, Grossniklaus U (2011) The walls have ears: The role of plant CrRLK1Ls in sensing and transducing extracellular signals. *J Exp Bot* 62:1581–1591. <https://doi.org/10.1093/jxb/erq445>
- Boisson-Dernier A, Roy S, Kritsas K, et al (2009) Disruption of the pollen-expressed FERONIA homologs ANXUR1 and ANXUR2 triggers pollen tube discharge. *Development* 136:3279–3288. <https://doi.org/10.1242/dev.040071>
- Bonaldi K, Gargani D, Prin Y, et al (2011) Sp. Strain ORS285: The Nod-Dependent Versus the Nod-Independent Symbiotic Interaction. *Mol Plant-Microbe Interact* 24:1359–1371
- Borges A, Tsai SM, Caldas DGG (2012) Validation of reference genes for RT-qPCR normalization in common bean during biotic and abiotic stresses. *Plant Cell Rep* 31:827–838. <https://doi.org/10.1007/s00299-011-1204-x>
- Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422:433–438. <https://doi.org/10.1038/nature01521>
- Campbell L, Turner SR (2017) A comprehensive analysis of RALF proteins in green plants suggests there are two distinct functional groups. *Front Plant Sci* 8:1–14. <https://doi.org/10.3389/fpls.2017.00037>

- Cao J, Shi F (2012) Evolution of the RALF gene family in plants: Gene duplication and selection patterns. *Evol Bioinforma* 2012:271–292. <https://doi.org/10.4137/EBO.S9652>
- Cao Y, Halane MK, Gassmann W, Stacey G (2017) The Role of Plant Innate Immunity in the Legume-Rhizobium Symbiosis. *Annu Rev Plant Biol* 68:535–561. <https://doi.org/10.1146/annurev-arplant-042916-041030>
- Carvalho H, Lima L, Lescure N, et al (2000) Differential expression of the two cytosolic glutamine synthetase genes in various organs of *Medicago truncatula*. *Plant Sci* 159:301–312. [https://doi.org/10.1016/S0168-9452\(00\)00360-5](https://doi.org/10.1016/S0168-9452(00)00360-5)
- Charon C, Johansson C, Kondorosi E, et al (1997) Enod40 Induces Dedifferentiation and Division of Root Cortical Cells in Legumes. *Proc Natl Acad Sci U S A* 94:8901–8906. <https://doi.org/10.1073/pnas.94.16.8901>
- Charon C, Sousa C, Crespi M, Kondorosi A (1999) Alteration of enod40 expression modifies *Medicago truncatula* root nodule development induced by *Sinorhizobium meliloti*. *Plant Cell* 11:1953–1965. <https://doi.org/10.1105/tpc.11.10.1953>
- Coghlan A, Eichler EE, Oliver SG, et al (2005) Chromosome evolution in eukaryotes: A multi-kingdom perspective. *Trends Genet* 21:673–682. <https://doi.org/10.1016/j.tig.2005.09.009>
- Combiér JP, Küster H, Journet EP, et al (2008) Evidence for the involvement in nodulation of the two small putative regulatory peptide-encoding genes MtRALFL1 and MtDVL1. *Mol Plant-Microbe Interact* 21:1118–1127. <https://doi.org/10.1094/MPMI-21-8-1118>
- Cordoba E, Shishkova S, Vance CP, Hernández G (2003) Antisense inhibition of NADH glutamate synthase impairs carbon/nitrogen assimilation in nodules of alfalfa (*Medicago sativa* L.). *Plant J* 33:1037–1049. <https://doi.org/10.1046/j.1365-313X.2003.01686.x>
- Couzigou J, Laressergues D, Combiér J (2015) of miRNAs in plant biology. 12:1178–1180
- Covey PA, Subbaiah CC, Parsons RL, et al (2010) A pollen-specific RALF from tomato that regulates pollen tube elongation. *Plant Physiol* 153:703–715. <https://doi.org/10.1104/pp.110.155457>
- Deslauriers SD, Larsen PB (2010) FERONIA is a key modulator of brassinosteroid and ethylene responsiveness in *Arabidopsis hypocotyls*. *Mol Plant* 3:626–640. <https://doi.org/10.1093/mp/ssq015>
- Du C, Li X, Chen J, et al (2016) Receptor kinase complex transmits RALF peptide signal to inhibit root growth in *Arabidopsis*. *Proc Natl Acad Sci U S A* 113:E8326–E8334. <https://doi.org/10.1073/pnas.1609626113>
- Duan Q, Kita D, Li C, et al (2010) FERONIA receptor-like kinase regulates RHO GTPase signaling of root hair development. *Proc Natl Acad Sci U S A* 107:17821–17826. <https://doi.org/10.1073/pnas.1005366107>
- Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Emery JF, Floyd SK, Alvarez J, et al (2003) Radial Patterning of *Arabidopsis* Shoots by Class III HD-ZIP and KANADI Genes. *Curr Biol* 13:1768–1774. <https://doi.org/10.1016/j.cub.2003.09.035>
- Escobar-Restrepo JM, Huck N, Kessler S, et al (2007) The Feronia receptor-like kinase mediates

- male-female interactions during pollen tube reception. *Science* (80-) 317:656–660. <https://doi.org/10.1126/science.1143562>
- Estrada-Navarrete G, Alvarado-Affantranger X, Olivares JE, et al (2007) Fast, efficient and reproducible genetic transformation of *Phaseolus* spp. by *Agrobacterium rhizogenes*. *Nat Protoc* 2:1819–1824. <https://doi.org/10.1038/nprot.2007.259>
- Fahlgren N, Montgomery TA, Howell MD, et al (2006) Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA Affects Developmental Timing and Patterning in Arabidopsis. *Curr Biol* 16:939–944. <https://doi.org/10.1016/j.cub.2006.03.065>
- Ferguson BJ, Indrasumunar A, Hayashi S, et al (2010) Molecular analysis of legume nodule development and autoregulation. *J Integr Plant Biol* 52:61–76. <https://doi.org/10.1111/j.1744-7909.2010.00899.x>
- Ferguson BJ, Mens C, Hastwell AH, et al (2019) Legume nodulation: The host controls the party. *Plant Cell Environ* 42:41–51. <https://doi.org/10.1111/pce.13348>
- Fletcher JC (1999) Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. *Science* (80-) 283:1911–1914. <https://doi.org/10.1126/science.283.5409.1911>
- Freeling M (2009) Bias in Plant Gene Content Following Different Sorts of Duplication: Tandem, Whole-Genome, Segmental, or by Transposition. *Annu Rev Plant Biol* 60:433–453. <https://doi.org/10.1146/annurev.arplant.043008.092122>
- Gage DJ (2004) Infection and Invasion of Roots by Symbiotic, Nitrogen-Fixing Rhizobia during Nodulation of Temperate Legumes. *Microbiol Mol Biol Rev* 68:280–300. <https://doi.org/10.1128/MMBR.68.2.280-300.2004>
- Gautrat P, Laffont C, Frugier F (2020) Compact Root Architecture 2 Promotes Root Competence for Nodulation through the miR2111 Systemic Effector. *Curr Biol* 30:1339–1345.e3. <https://doi.org/10.1016/j.cub.2020.01.084>
- Gautrat P, Mortier V, Laffont C, et al (2019) Unraveling new molecular players involved in the autoregulation of nodulation in *Medicago truncatula*. *J Exp Bot* 70:1407–1417. <https://doi.org/10.1093/jxb/ery465>
- Goodstein DM, Shu S, Howson R, et al (2012) Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res* 40:1178–1186. <https://doi.org/10.1093/nar/gkr944>
- Guindon S, Dufayard JF, Lefort V, et al (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321. <https://doi.org/10.1093/sysbio/syq010>
- Guo H, Li L, Ye H, et al (2009) Three related receptor-like kinases are required for optimal cell elongation in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 106:7648–7653. <https://doi.org/10.1073/pnas.0812346106>
- Hardy RWF, Holsten RD, Jackson EK, Burns RC (1968) The Acetylene - Ethylene Assay for N₂ Fixation : Laboratory and Field Evaluation '. 43:1185–1207
- Haruta M, Sabat G, Stecker K, et al (2014) A Peptide Hormone and Its Receptor. *Science* 343:408–412. <https://doi.org/10.1126/science.1244454>

- Hastwell AH, Gresshoff PM, Ferguson BJ (2015) Genome-wide annotation and characterization of CLAVATA/ESR (CLE) peptide hormones of soybean (*Glycine max*) and common bean (*Phaseolus vulgaris*), and their orthologues of *Arabidopsis thaliana*. *J Exp Bot* 66:5271–5287. <https://doi.org/10.1093/jxb/erv351>
- Helariutta Y, Fukaki H, Wysocka-Diller J, et al (2000) The SHORT-ROOT gene controls radial patterning of the *Arabidopsis* root through radial signaling. *Cell* 101:555–567. [https://doi.org/10.1016/S0092-8674\(00\)80865-X](https://doi.org/10.1016/S0092-8674(00)80865-X)
- Hernández G, Valdés-López O, Ramírez M, et al (2009) Global changes in the transcript and metabolic profiles during symbiotic nitrogen fixation in phosphorus-stressed common bean plants. *Plant Physiol* 151:1221–1238. <https://doi.org/10.1104/pp.109.143842>
- Imin N, Mohd-Radzman NA, Ogilvie HA, Djordjevic MA (2013) The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in *Medicago truncatula*. *J Exp Bot* 64:5395–5409. <https://doi.org/10.1093/jxb/ert369>
- Isidra-Arellano MC, Pozas-Rodríguez EA, del Rocío Reyero-Saavedra M, et al (2020) Inhibition of legume nodulation by Pi deficiency is dependent on the autoregulation of nodulation (AON) pathway. *Plant J* 103:1125–1139. <https://doi.org/10.1111/tpj.14789>
- Ivanov S, Fedorova EE, Limpens E, et al (2012) Rhizobium-legume symbiosis shares an exocytotic pathway required for arbuscule formation. *Proc Natl Acad Sci U S A* 109:8316–8321. <https://doi.org/10.1073/pnas.1200407109>
- James L. Hartley, Gary F. Temple and MAB (2000) DNA Cloning Using In Vitro Site-Specific Recombination. 1788–1795. <https://doi.org/10.1101/gr.143000.that>
- Karimi M, Inzé D, Depicker A (2002) GATEWAY vectors for *Agrobacterium*-mediated plant.pdf. *Trends Plant Sci* 7:193–195
- Kereszt A, Mergaert P, Montiel J, et al (2018) Impact of plant peptides on symbiotic nodule development and functioning. *Front Plant Sci* 9:1–16. <https://doi.org/10.3389/fpls.2018.01026>
- Kessler SA, Shimosato-Asano H, Keinath NF, et al (2010) Conserved molecular components for pollen tube reception and fungal invasion. *Science* (80-) 330:968–971. <https://doi.org/10.1126/science.1195211>
- Kondo Y, Hirakawa Y, Fukuda H (2014) Peptide Ligands in Plants, 1st edn. Elsevier Inc.
- Krusell L, Madsen LH, Sato S, et al (2002) Shoot control of root development and nodulation is mediated by a receptor-like kinase. *Nature* 420:422–426. <https://doi.org/10.1038/nature01207>
- Larsson A (2014) AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30:3276–3278. <https://doi.org/10.1093/bioinformatics/btu531>
- Lease KA, Walker JC (2006) The *Arabidopsis* unannotated secreted peptide database, a resource for plant peptidomics. *Plant Physiol* 142:831–838. <https://doi.org/10.1104/pp.106.086041>
- Lee DH, Bourdais G, Yu G, et al (2015) Phosphorylation of the plant immune regulator RPM1-INTERACTING PROTEIN4 enhances plant plasma membrane h⁺-ATPase activity and inhibits flagellin-triggered immune responses in *Arabidopsis*. *Plant Cell* 27:2042–2056.

<https://doi.org/10.1105/tpc.114.132308>

- Lee JH, Schöffl F (1997) GUS activity staining — a powerful tool in plant molecular biology. *Plant Mol Biol Man* 23–32. https://doi.org/10.1007/978-94-011-5400-0_2
- Li C, Yeh FL, Cheung AY, et al (2015) Glycosylphosphatidylinositol-anchored proteins as chaperones and co-receptors for FERONIA receptor kinase signaling in Arabidopsis. *Elife* 4:1–21. <https://doi.org/10.7554/eLife.06587>
- Liu J, Elmore JM, Coaker G (2009) Investigating the functions of the RIN4 protein complex during plant innate immune responses. *Plant Signal Behav* 4:1107–1110. <https://doi.org/10.4161/psb.4.12.9944>
- Liu J, Elmore JM, Lin ZJD, Coaker G (2011) A receptor-like cytoplasmic kinase phosphorylates the host target RIN4, leading to the activation of a plant innate immune receptor. *Cell Host Microbe* 9:137–146. <https://doi.org/10.1016/j.chom.2011.01.010>
- Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* (80-) 290:1151–1155. <https://doi.org/10.1126/science.290.5494.1151>
- Marshall E, Costa LM, Gutierrez-Marcos J (2011) Cysteine-Rich Peptides (CRPs) mediate diverse aspects of cell-cell communication in plant reproduction and development. *J Exp Bot* 62:1677–1686. <https://doi.org/10.1093/jxb/err002>
- Masachis S, Segorbe D, Turrà D, et al (2016) A fungal pathogen secretes plant alkalinizing peptides to increase infection. *Nat Microbiol* 1:. <https://doi.org/10.1038/nmicrobiol.2016.43>
- Matsubayashi Y (2014) Posttranslationally Modified Small-Peptide Signals in Plants. *Annu Rev Plant Biol* 65:385–413. <https://doi.org/10.1146/annurev-arplant-050312-120122>
- Matsubayashi Y, Sakagami Y (2006) Peptide Hormones in Plants. *Annu Rev Plant Biol* 57:649–674. <https://doi.org/10.1146/annurev.arplant.56.032604.144204>
- Mattick JS, Gagen MJ (2001) The evolution of controlled multitasked gene networks: The role of introns and other noncoding RNAs in the development of complex organisms. *Mol Biol Evol* 18:1611–1630. <https://doi.org/10.1093/oxfordjournals.molbev.a003951>
- Mergaert P, Uchiumi T, Alunni B, et al (2006) Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis. *Proc Natl Acad Sci U S A* 103:5230–5235. <https://doi.org/10.1073/pnas.0600912103>
- Montiel J, Nava N, Cárdenas L, et al (2012) A Phaseolus vulgaris NADPH Oxidase Gene is Required for Root Infection by Rhizobia. *Plant Cell Physiol* 53:1751–1767. <https://doi.org/10.1093/pcp/pcs120>
- Morato do Canto A, Ceciliato PHO, Ribeiro B, et al (2014) Biological activity of nine recombinant AtRALF peptides: Implications for their perception and function in Arabidopsis. *Plant Physiol Biochem* 75:45–54. <https://doi.org/10.1016/j.plaphy.2013.12.005>
- Morris JL, Puttick MN, Clark JW, et al (2018) The timescale of early land plant evolution. *Proc Natl Acad Sci U S A* 115:E2274–E2283. <https://doi.org/10.1073/pnas.1719588115>
- Mortier V, De Wever E, Vuylsteke M, et al (2012) Nodule numbers are governed by interaction between CLE peptides and cytokinin signaling. *Plant J* 70:367–376.

<https://doi.org/10.1111/j.1365-313X.2011.04881.x>

- Mortier V, den Herder G, Whitford R, et al (2010) CLE peptides control medicago truncatula nodulation locally and systemically. *Plant Physiol* 153:222–237.
<https://doi.org/10.1104/pp.110.153718>
- Murphy E, De Smet I (2014) Understanding the RALF family: A tale of many species. *Trends Plant Sci* 19:664–671. <https://doi.org/10.1016/j.tplants.2014.06.005>
- Murphy E, Smith S, De Smet I (2012) Small signaling peptides in Arabidopsis development: How cells communicate over a short distance. *Plant Cell* 24:3198–3217.
<https://doi.org/10.1105/tpc.112.099010>
- Nasr Esfahani M, Inoue K, Chu HD, et al (2017) Comparative transcriptome analysis of nodules of two Mesorhizobium–chickpea associations with differential symbiotic efficiency under phosphate deficiency. *Plant J* 91:911–926. <https://doi.org/10.1111/tpj.13616>
- Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274.
<https://doi.org/10.1093/molbev/msu300>
- Niklas KJ, Kutschera U (2010) The evolution of the land plant life cycle. *New Phytol* 185:27–41.
<https://doi.org/10.1111/j.1469-8137.2009.03054.x>
- O’Rourke JA, Iniguez LP, Fu F, et al (2014) An RNA-Seq based gene expression atlas of the common bean. *BMC Genomics* 15:. <https://doi.org/10.1186/1471-2164-15-866>
- Okamoto S, Ohnishi E, Sato S, et al (2009) Nod factor/nitrate-induced CLE genes that drive HAR1-mediated systemic regulation of nodulation. *Plant Cell Physiol* 50:67–77.
<https://doi.org/10.1093/pcp/pcn194>
- Okamoto S, Shinohara H, Mori T, et al (2013) Root-derived CLE glycopeptides control nodulation by direct binding to HAR1 receptor kinase. *Nat Commun* 4:1–7.
<https://doi.org/10.1038/ncomms3191>
- Patriarca EJ, Tatè R, Ferraioli S, Iaccarino M (2004) Organogenesis of Legume Root Nodules. *Int Rev Cytol* 234:201–262. [https://doi.org/10.1016/S0074-7696\(04\)34005-2](https://doi.org/10.1016/S0074-7696(04)34005-2)
- Pearce G, Moura DS, Stratmann J, Ryan CA (2001) RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. *Proc Natl Acad Sci* 98:12843–12847.
<https://doi.org/10.1073/pnas.201416998>
- Radutoiu S, Madsen LH, Madsen EB, et al (2003) Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 425:585–592. <https://doi.org/10.1038/nature02039>
- Reid DE, Ferguson BJ, Gresshoff PM (2011) Inoculation- and nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation. *Mol Plant-Microbe Interact* 24:606–618. <https://doi.org/10.1094/MPMI-09-10-0207>
- Roux B, Rodde N (2014) An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. 817–837.
<https://doi.org/10.1111/tpj.12442>
- Roy S, Liu W, Nandety RS, et al (2020) Celebrating 20 Years of Genetic Discoveries in Legume

- Nodulation and Symbiotic Nitrogen Fixation. *Plant Cell* 32:15–41.
<https://doi.org/10.1105/tpc.19.00279>
- Schauser L, Roussis A, Stiller J, Stougaard J (1999) A plant regulator controlling development of symbiotic root nodules. *Nature* 402:191–195. <https://doi.org/10.1038/46058>
- Schnabel E, Journet EP, De Carvalho-Niebel F, et al (2005) The *Medicago truncatula* SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. *Plant Mol Biol* 58:809–822. <https://doi.org/10.1007/s11103-005-8102-y>
- Schulze-Muth P, Irmeler S, Schröder G, Schröder J (1996) Novel type of receptor-like protein kinase from a higher plant (*Catharanthus roseus*): cDNA, gene, intramolecular autophosphorylation, and identification of a threonine important for auto- and substrate phosphorylation. *J Biol Chem* 271:26684–26689. <https://doi.org/10.1074/jbc.271.43.26684>
- Schwartz S, Meshorer E, Ast G (2009) Chromatin organization marks exon-intron structure. *Nat Struct Mol Biol* 16:990–995. <https://doi.org/10.1038/nsmb.1659>
- Searle IR, Men AE, Laniya TS, et al (2003) Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. *Science* (80-) 299:109–112.
<https://doi.org/10.1126/science.1077937>
- Shen Q, Bourdais G, Pan H, et al (2017) Arabidopsis glycosylphosphatidylinositol-anchored protein LLG1 associates with and modulates FLS2 to regulate innate immunity. *Proc Natl Acad Sci U S A* 114:5749–5754. <https://doi.org/10.1073/pnas.1614468114>
- Solis-miranda J, Quinto C (2021) Plant Physiology and Biochemistry The CrRLK1L subfamily : One of the keys to versatility in plants. *Plant Physiol Biochem* 166:88–102.
<https://doi.org/10.1016/j.plaphy.2021.05.028>
- Solis-Miranda J (2016) Análisis de la participación de feronia, un receptor tipo cinasa, en el desarrollo de los pelos radicales de frijol y en su asociación con rhizobia
- Solis-Miranda J, Fonseca-garcía C, Nava N, et al (2020) Genome-wide identification of the crrlk1l subfamily and comparative analysis of its role in the legume-rhizobia symbiosis. *Genes (Basel)* 11:1–29. <https://doi.org/10.3390/genes11070793>
- Stegmann M, Monaghan J, Smakowska-Luzan E, et al (2017) The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. *Science* (80-) 355:287–289.
<https://doi.org/10.1126/science.aal2541>
- Sun X, Hu Z, Chen R, et al (2015) Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Sci Rep* 5:1–10. <https://doi.org/10.1038/srep10342>
- Thynne E, Saur IML, Simbaqueba J, et al (2017) Fungal phytopathogens encode functional homologues of plant rapid alkalization factor (RALF) peptides. *Mol Plant Pathol* 18:811–824. <https://doi.org/10.1111/mpp.12444>
- Valdés-López O, Arenas-Huertero C, Ramírez M, et al (2008) Essential role of MYB transcription factor: PvPHR1 and microRNA: PvmiR399 in phosphorus-deficiency signalling in common bean roots. *Plant, Cell Environ* 31:1834–1843. <https://doi.org/10.1111/j.1365-3040.2008.01883.x>
- Van de Velde W, Zehirov G, Szatmari A, et al (2010) Plant Peptides Govern Terminal Differentiation

- of Bacteria in Symbiosis. *Science* (80-) 327:1122–1126.
<https://doi.org/10.1126/science.1184057>
- van Norman JM, Breakfield NW, Benfey PN (2011) Intercellular communication during plant development. *Plant Cell* 23:855–864. <https://doi.org/10.1105/tpc.111.082982>
- Vanneste S, Friml J (2009) Auxin: A Trigger for Change in Plant Development. *Cell* 136:1005–1016.
<https://doi.org/10.1016/j.cell.2009.03.001>
- Vinuesa P, Neumann-Silkow F, Pacios-Bras C, et al (2003) Genetic analysis of a pH-regulated operon from *Rhizobium tropici* CIAT899 involved in acid tolerance and nodulation competitiveness. *Mol Plant-Microbe Interact* 16:159–168.
<https://doi.org/10.1094/MPMI.2003.16.2.159>
- Vleghels I, Hontelez J, Ribeiro A, et al (2003) Expression of ENOD40 during tomato plant development. *Planta* 218:42–49. <https://doi.org/10.1007/s00425-003-1081-9>
- Walker TS, Bais HP, Grotewold E, Vivanco JM (2003) Update on Root Exudation and Rhizosphere Biology Root Exudation and Rhizosphere Biology 1. 132:44–51.
<https://doi.org/10.1104/pp.102.019661.Although>
- Wan X, Hontelez J, Lillo A, et al (2007) *Medicago truncatula* ENOD40-1 and ENOD40-2 are both involved in nodule initiation and bacteroid development. *J Exp Bot* 58:2033–2041.
<https://doi.org/10.1093/jxb/erm072>
- Wang C, Yu H, Zhang Z, et al (2015) Phytosulfokine is involved in positive regulation of *Lotus japonicus* nodulation. *Mol Plant-Microbe Interact* 28:847–855.
<https://doi.org/10.1094/MPMI-02-15-0032-R>
- Wang L, Rubio MC, Xin X, et al (2019) CRISPR/Cas9 knockout of leghemoglobin genes in *Lotus japonicus* uncovers their synergistic roles in symbiotic nitrogen fixation. *New Phytol* 224:818–832. <https://doi.org/10.1111/nph.16077>
- Wang Y, Wang X, Tang H, et al (2011) Modes of gene duplication contribute differently to genetic novelty and redundancy, but show parallels across divergent angiosperms. *PLoS One* 6:.
<https://doi.org/10.1371/journal.pone.0028150>
- Wolfe D, Dudek S, Ritchie MD, Pendergrass SA (2013) Visualizing genomic information across chromosomes with PhenoGram. *BioData Min* 6:1. <https://doi.org/10.1186/1756-0381-6-18>
- Xu G, Chen W, Song L, et al (2019) FERONIA phosphorylates E3 ubiquitin ligase ATL6 to modulate the stability of 14-3-3 proteins in response to the carbon/nitrogen ratio. *J Exp Bot* 70:6375–6388. <https://doi.org/10.1093/jxb/erz378>
- Yu F, Qian L, Nibau C, et al (2012) FERONIA receptor kinase pathway suppresses abscisic acid signaling in *Arabidopsis* by activating ABI2 phosphatase. *Proc Natl Acad Sci U S A* 109:14693–14698. <https://doi.org/10.1073/pnas.1212547109>
- Zhang J, Li W, Xiang T, et al (2010) Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a *Pseudomonas syringae* effector. *Cell Host Microbe* 7:290–301. <https://doi.org/10.1016/j.chom.2010.03.007>
- Zhang X, Peng H, Zhu S, et al (2020) Nematode-encoded RALF Peptide Mimics Facilitate Parasitism of Plants through the FERONIA Receptor Kinase. *Mol Plant*.

<https://doi.org/10.1016/j.molp.2020.08.014>

Zhu F, Ye Q, Chen H, et al (2021) Multigene editing reveals that MtCEP1 / 2 / 12 redundantly control lateral root and nodule number in *Medicago truncatula*. *J Exp Bot* 72:3661–3676. <https://doi.org/10.1093/jxb/erab093>

Zhu L, Zhang Y, Zhang W, et al (2009) Patterns of exon-intron architecture variation of genes in eukaryotic genomes. *BMC Genomics* 10:1–12. <https://doi.org/10.1186/1471-2164-10-47>

Zhu S, Estévez JM, Liao H, et al (2020) The RALF1–FERONIA complex phosphorylates eIF4E1 to promote protein synthesis and polar root hair growth. *Mol Plant*. <https://doi.org/10.1016/j.molp.2019.12.014>

14. MATERIAL SUPLEMENTARIO

PROTOCOLO PARA TINCION DE GUS (Lee and Schöffl 1997).

1. Colocar la raíz en una caja y sumergir en la solución i durante 5 min a temperatura ambiente (el volumen depende de la cantidad de raíz).
2. Retirar la solución i.
3. Sumergir las raíces en la solución ii durante 5 min a temperatura ambiente.
4. Retirar la solución ii.
5. Revelar la actividad de GUS sumergiendo en la solución iii. El tiempo varía dependiendo del experimento.
6. Lavar las raíces con agua estéril y desgasificar en el sistema de trampa de agua.
7. Aclarar las raíces en una solución de hipoclorito de sodio al 0.6% hasta retirar el color amarillo
8. Lavar las raíces 3 veces con agua estéril.

PROTOCOLO CON TRIZOL PARA EXTRACCIÓN DE RNA EN UN SOLO PASO, MODIFICADO DE CHOMCZYNSKI AND SACCHI, 1987.

1. Macerar el tejido con ayuda de nitrógeno líquido, evitando que este se caliente.
2. Adicionar 1 mL de solución de TRIZOL al tejido macerado.
3. Mezclar por vortex durante 15 seg y dejar 5 min a temperatura ambiente.
4. Adicionar 200 μ L de cloroformo:isoamílico (24:1), mezclar por vortex e incubar 3 min a temperatura ambiente.
5. Centrifugar durante 15 min a 11800 rpm a 4 °C.
6. Recuperar la fase superior en un nuevo tubo
7. Adicionar 1 volumen de isopropanol total e incubar 10 min a temperatura ambiente
8. Centrifugar durante 10 min a 5900 rpm a 4 °C.
9. Descartar el sobrenadante y lavar el pellet en 1 volumen LiCl 4M.
10. Centrifugar durante 20 min a 5900 rpm a 4 °C.
11. Descartar todo el sobrenadante.
12. Resuspender el pellet en un volumen de buffer TE 1x a pH de 8.
13. Adicionar 1 volumen de cloroformo:isoamílico (24:1) y mezclar por vortex.
14. Centrifugar durante 10 min a 5900 rpm a 4 °C.
15. Recuperar la fase superior en un tubo nuevo, agregar un volumen de isopropanol total y 66 μ L de acetato de sodio 3M a pH DE 5.2.
16. Mezclar por vortex y centrifugar durante 10 min a 11800 rpm a 4 °C.
17. Descartar el sobrenadante y secar completamente
18. Resuspender la pastilla en 50 μ L de agua DEPC.

Table S1: Soluciones de lavado (Tincion GUS)

SOLUCIÓN	FÓRMULA QUÍMICA	CONCENTRACIÓN (mM)
i	NaH ₂ PO ₄ (0.5 M)	50 mM
	Na ₂ HPO ₄ (0.5 M)	50 mM
	K ₃ Fe(CN) ₆ (0.1 M)	0.5 mM
	K ₄ Fe(CN) ₆ (0.1 M)	0.5 mM
ii	SOLUCIÓN i	
	EDTA (0.5 M)	2 mM
	TRITON 10%	0.1%
iii	SOLUCIÓN ii	
	X-GLUCORO-CHA	10 µL/mL de solución

Table S2: Medio Fahraeus

FÓRMULA QUÍMICA	CONCENTRACIÓN FINAL (µM)
CaCl ₂	.680
K ₂ HPO ₄	.734
Fe-citrato	.0185
MgSO ₄	.468
Na ₂ PO ₄	1.042
MnSO ₄	.012
H ₃ BO ₃	.046
ZnSO ₄	.765
CuSO ₄	.320
H ₂ MoO ₄	.038
*KNO ₃	2/5

*Para condiciones inhibitorias de la nodulación por nitrógeno

Table S3: Oligonucleótidos utilizados

Nombre	Sequence	Amplicon (pb)
Análisis por qPCR		
qPCR_RALF1_FW	CACCATGGCGAAGTCGTGTTTCCTT	123
qPCR_RALF1_RV	GGAACCCCATCCCATTCCCAG	
qPCR_RALF6_FW	CACCATGATGTCCCAACCCAGA	127
qPCR_RALF6_RV	ATTGTGTTTCAACAGATTGAGGTcc	
qPCR_RIPK2_FW	ATCCCTATCTATCTTCTTCAA	198
qPCR_RIPK2_RV	TTACCCTGTGTTTTTATCTGG	
qPCR_FER1_FW	TTTGCTGCGTCCCATCGT	137
qPCR_FER1_RV	GCATAACTTCTGTGTGTTGGTC	
qPCR_NIN_FW	GGGGATTCAGAGATTTGCAG	101
qPCR_NIN_RV	AACCCACTCTTGAGCATCGT	
qPCR_LEG_FW	TGGCTGATGCTGCACTTGG	166
qPCR_LEG_RV	GCTGCCAATTCATCGTAGG	
qPCR_GS_FW	CAGGAGGGGTAACAATATCCTGG	189
qPCR_GS_RV	GATGTTCAAGTGGCCTGTTGGATG	
qPCR_SS_FW	CGTCTCTCTTTGTGTTTGGC	154
qPCR_SS_RV	GGATGATGTGCCATCTTCAG	
qPCR_NARK_FW	TTGGTCATCCACATTCACCTC	104
qPCR_NARK_RV	CAGATTTGTTTTGCTGCTCCA	
qPCR_TML_FW	AGGAGAGTGAAAATGGCAAGG	89
qPCR_TML_RV	GGGAAAAGATAGCACTTGAGGTT	
qPCR_RIC2_FW	CCACTAGCCATCTCAAGCTT	111
qPCR_RIC2_RV	CAAGCATGGAGAGTTGGGTT	
qPCR_NIC1_FW	TTGGGGACAATGGCCAAC	126
qPCR_NIC1_RV	ACCTTCGAGTTGGGTCTTC	
EF1 α _LW	GCACCCAGGCATACTTGAATGACC	146
EF1 α _UP	GGTCATTGGTCATGTCGACTCTGG	
IDE_LW	GCAACCAACCTTTCATCAGC	156
IDE_UP	AGAAATGCCTCAACCCTTG	
Análisis de actividad de promotor		
pRALF1_FW	CACCGTAAAAACATATCCTTGACAC	2402
pRALF1_RV	TGATCAGAGGCTTCTCTTTT	
pRALF6_FW	CACCAGGGTTCAAATGAAGATGG	2020
pRALF6_RV	TGTGTGTCGAGAGATGTTTTG	

pRIPK2_FW	CACCACATGGGTAAATGAATTAG	2345
pRIPK2_RV	GTCAGTTACCCTTGTGTTTTT	
pFER1_FW	CACCTATTTGCTACTAGAAATTGAGT	2004
pFER1_RV	CACGACACAGAATTCCAAACC	
Construcciones de silenciamiento/sobreexpresión		
RALF1i_FW	CACCATGATGAAACGAGAGATTTG	274
RALF1i_RV	TGATCAGAGGCTTCTCTTTT	
RALF6i_FW	CACCTGACAGTAATGATGGCAC	168
RALF6i_RV	TGTGTGTCGAGAGATGTTTTG	
RIPK2i_FW	CACCAAGCCGAGGAAAATCAAC	202
RIPK2i_RV	TGAGTCTTATGCTGCAAAAGT	
FER1i_FW	CACCATCCTCACCTTTATTAGA	123
FER1i_RV	AATTACAATCACGCACACCCA	
RALF1ox_FW	CACCATGGCGAAGTCGTGTTTCCTT	348
RALF1ox_RV	GCTCCTGCACCTCGTAATGG	
RALF6ox_FW	CACCATGATGTCCCAACCCAGA	388
RALF6ox_RV	AGTCTTGAAGTCTTGACCTCT	
RIPK2ox_FW	CACCATGACAATCGCATGGGGAT	1441
RIPK2ox_RV	CACCGGTTGAGCCTCGAG	
FER1ox_FW	CACCATGTGGAGCATGAATCG	2701
FER1ox_RV	ACGCCCTTTTGGATTCATGATC	

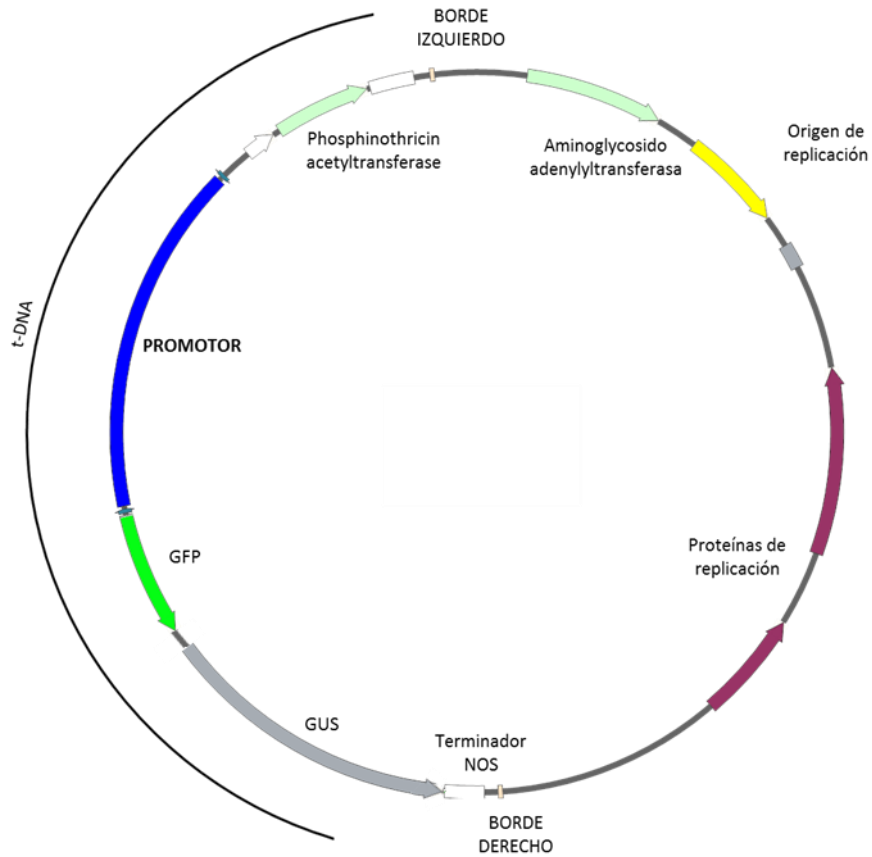


Figura S1. Esquema del plásmido pBGWSF7.0. Mapa representativo del plásmido pBGWSF7.0 empleado para el análisis de la actividad de promotor del gen de interés. Entre los bordes izquierdo y derecho (cintas amarillo pálido) se representa la secuencia correspondiente al “t-DNA” el cual es transferido por *Agrobacterium rhizogenes* al genoma de las células de frijol durante la infección. La cinta azul representa la región promotora del gen bajo estudio que controla la expresión del gen que codifican la proteína reportera GFP (Cinta verde) en fusión transcripcional con el gen que codifica la enzima glucoronidasa (GUS, Cinta gris). Las cintas moradas representan los genes que codifican las proteínas necesarias para la replicación del plásmido en *E. coli* y *A. rhizogenes*, la cinta amarilla representa el sitio de origen de replicación del plásmido, las cintas verdes claro representan los genes que codifican para las enzimas de resistencia a herbicidas a las raíces transformada (Phosphinitricin acetyltransferasa) y al antibiótico espectinomycin a *E. coli* y a *A. rhizogenes* (Aminoglycosido adenyltransferasa) para la selección de clones positivas.

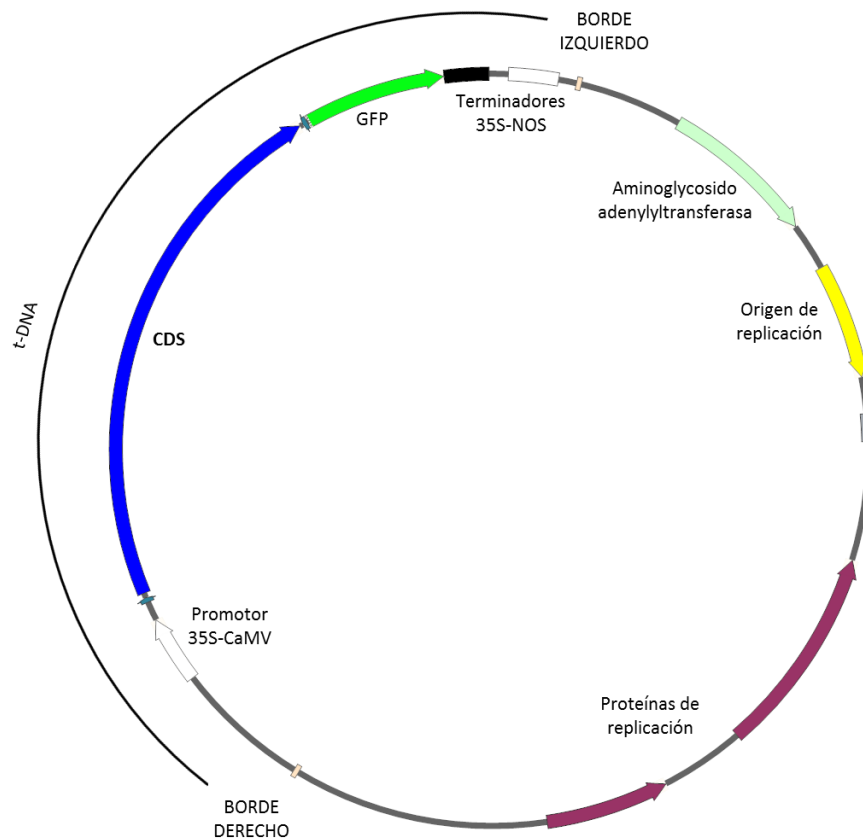


Figura S2. Esquema del plásmido pH7fwg2d. Esquema representativo del plásmido pH7fwg2d empleado para el análisis de la sobreexpresión del gen de interés. Entre los bordes izquierdo y derecho (cintas amarillo claro) se representa la secuencia correspondiente al “tDNA” el cual es transferido por *Agrobacterium rhizogenes* a las células de frijol durante la infección, donde se inserta aleatoriamente en el genoma de la planta. La cinta azul representa la región codificante del gen bajo estudio en fusión transcripcional con el gen que codifican la proteína reportera GFP (Cinta verde), cuya expresión es controlada por el promotor 35S del virus del mosaico de la calabaza (Cinta blanca). Las cintas moradas representan los genes que codifican las proteínas necesarias para la replicación del plásmido en *E. coli* y *A. rhizogenes*, la cinta amarilla representa el sitio de origen de replicación del plásmido, las cintas verdes claro representan el gen que codifican para la enzima que le otorga resistencia al antibiótico espectinomycin a *E. coli* y a *A. rhizogenes* (Aminoglycosido adenyltransferasa) para la selección de clones positivas (Karimi et al. 2002).

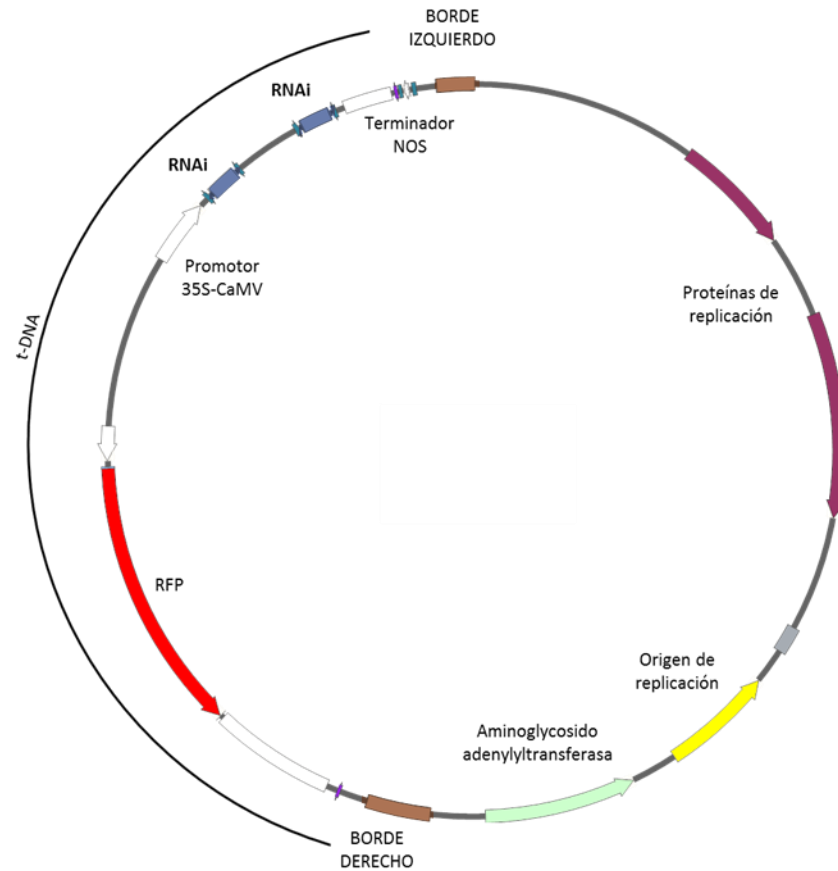


Figura S3. Esquema del plásmido ptdT-RNAi. Esquema representativo del plásmido ptdT-RNAi empleado para el análisis del silenciamiento del gen de interés. Entre los bordes izquierdo y derecho (cintas de color café) se representa la secuencia correspondiente al “tDNA” el cual es transferido por *Agrobacterium rhizogenes* a las células de frijol durante la infección, donde se inserta aleatoriamente en el genoma de la planta. Las cintas azules representan las secuencias específicas del gen bajo estudio, bajo expresión es controlada por el promotor 35S del virus del mosaico de la calabaza (Cinta blanca). La cinta roja representa el gen que codifica la proteína roja fluorescente (RFP). Las cintas moradas representan los genes que codifican las proteínas necesarias para la replicación del plásmido en *E. coli* y *A. rhizogenes*, la cinta amarilla representa el sitio de origen de replicación del plásmido, y la cintas verde clara representan el gen que codifican para la enzima que le otorga a *E. coli* y *A. rhizogenes* resistencia al antibiótico espectinomomicina (Aminoglicosido adenyltransferasa) para la selección de clones positivos (Valdés-López *et al.* 2008).

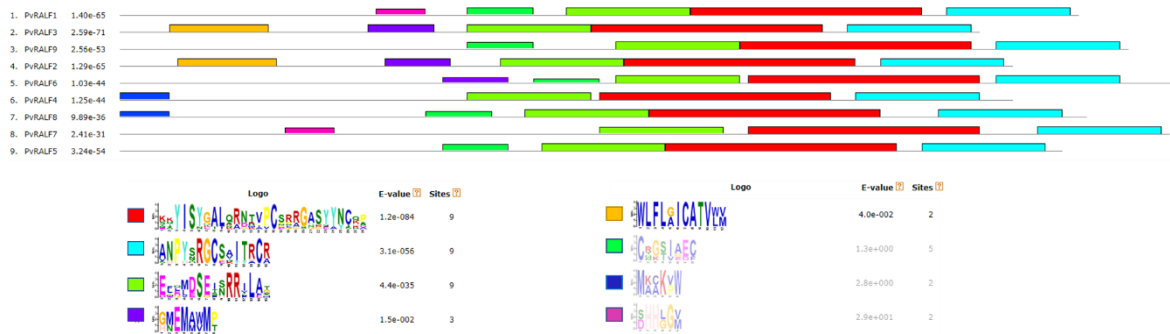


Figura S4. Motivos de la familia de péptidos RALF en frijol. Identificación de motivos en las secuencias proteicas de péptidos RALF de frijol. Mediante MEME, se identificaron motivos en los nueve RALFs de frijol. Los motivos más representativos se representan en barras de colores ubicadas en su posición correspondiente. El código de colores corresponde a los logos representativos.

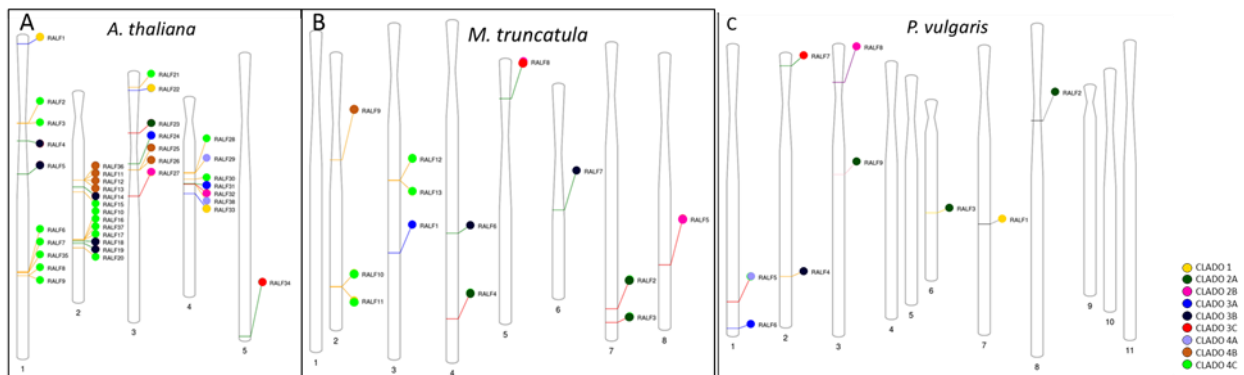


Figura S5. Ubicación de los genes RALF de *A. thaliana*, *M. truncatula* y frijol en su respectivo genoma. Representación de la ubicación de los genes RALF en los cromosomas de *A. thaliana*, *M. truncatula* y *P. vulgaris*. Se puede observar la gran cantidad de genes y los clústeres formados en el genoma de Arabidopsis, que no sucede en el caso de frijol ni de *M. truncatula*, lo cual refleja la rápida expansión que refieren Campbell y Turner en 2017. Los círculos de colores indican el clado al que pertenece cada gen de acuerdo a dichos autores.

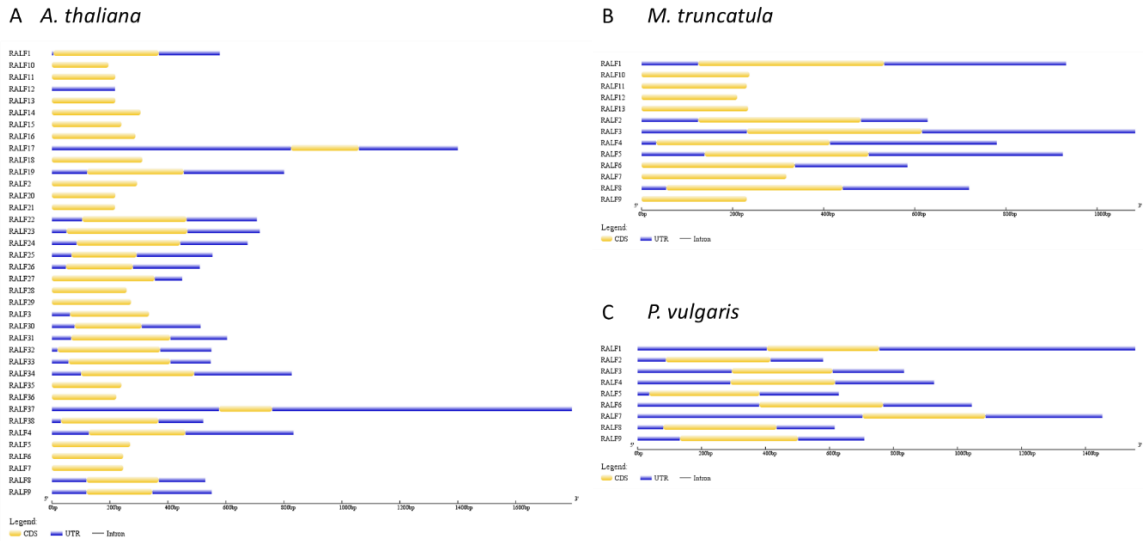


Figura S6. Estructura génica del transcrito primario de los genes *RALF* en *A. thaliana*, *M. truncatula* y frijol. Representación gráfica de la estructura génica de los transcritos primarios de los genes *RIPK* de *A. thaliana*, *M. truncatula* y *P. vulgaris*. Ningún gen *RALF* presenta intrones en su estructura de transcrito primario en ninguna de las especies analizadas, además, se observa variación en los tamaños de las regiones 3' o 5' UTR entre los diferentes genes, sin embargo, el tamaño de las regiones codificantes está conservada en todos los genes.

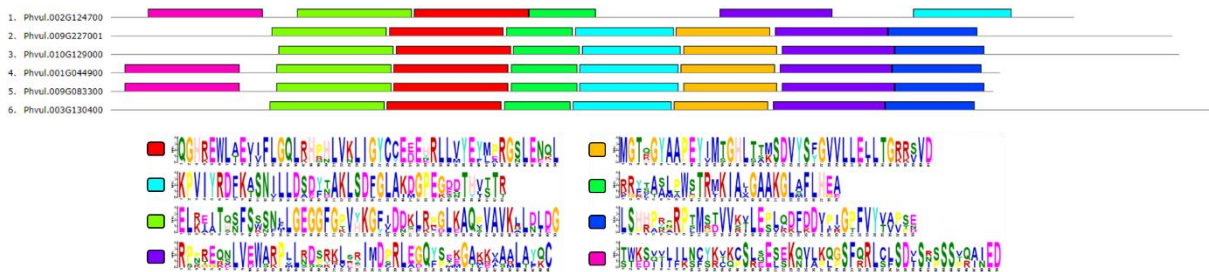


Figura S7. Motivos y alineamiento de la familia de péptidos RIPK en frijol. A) Identificación de motivos en las secuencias proteicas de RIPK de frijol. Mediante MEME, se identificaron motivos en los seis PvRIPKs. Los motivos más representativos se representan en barras de colores ubicadas en su posición correspondiente. B) Logos representativos de los motivos encontrados por MEME. Se observan los ocho motivos más representativos encontrados por MEME, con su respectivo logo, el cual permite apreciar la conservación de los residuos. El código de colores corresponde a las barras representadas en A) y C). C) Se muestra el alineamiento de las seis secuencias proteicas de frijol encontradas tras la búsqueda informática de PvRIPK. Las barras de colores representan los motivos sobrerrepresentados encontrados por MEME mientras que la línea en verde representa la localización del dominio cinasa. Se observa que seis de los ocho motivos sobrerrepresentados se localizan dentro de dominio cinasa, lo que habla de su alta conservación.

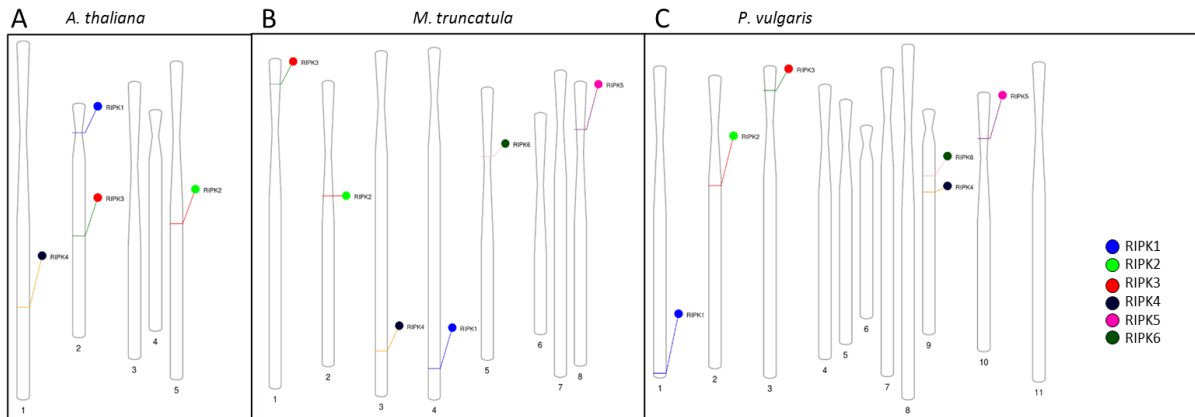


Figura S8. Ubicación de los genes *RIPK* de *A. thaliana*, *M. truncatula* y frijol en su respectivo genoma. Representación de la ubicación de los genes *RIPK* en los cromosomas de *A. thaliana* (A), *M. truncatula* (B) y *P. vulgaris* (C).

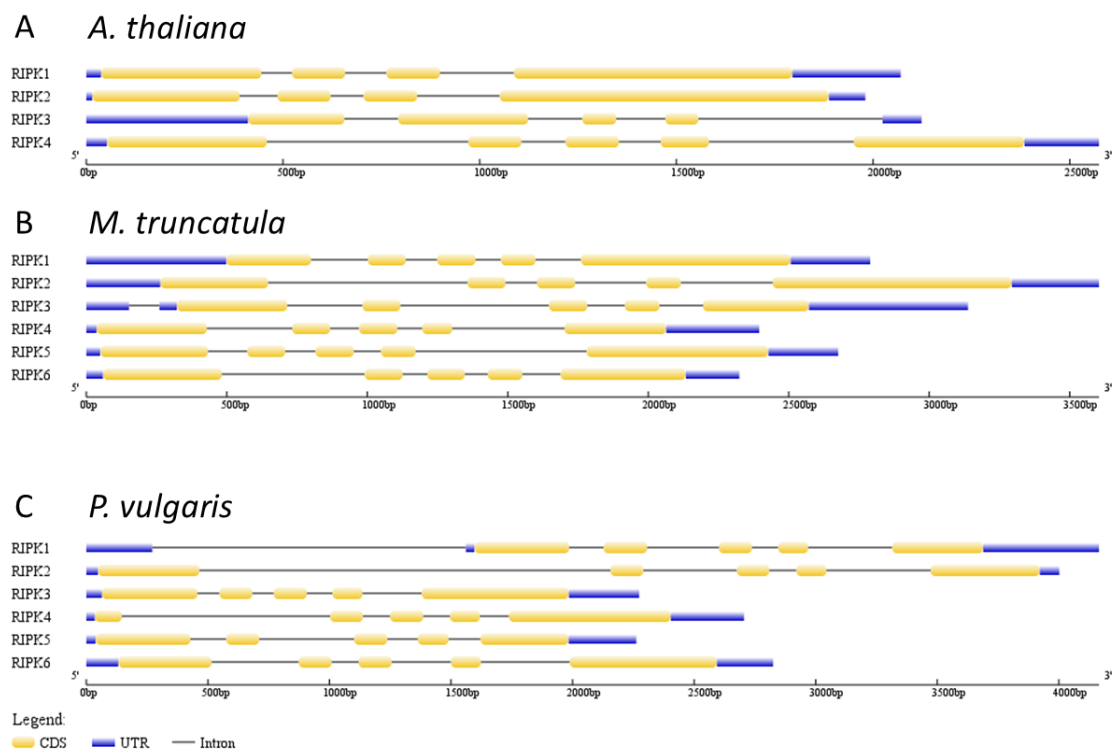


Figura S9. Estructura génica de los transcritos *RIPK* de *A. thaliana*, *M. truncatula* y frijol. Estructura génica de los transcritos primarios de los genes *RIPK* de *A. thaliana* (A), *M. truncatula* (B) y *P. vulgaris* (C). Se observa la presencia de tres y cuatro intrones en el genoma de Arabidopsis, y cuatro y cinco en los genomas de *M. truncatula* y de frijol.

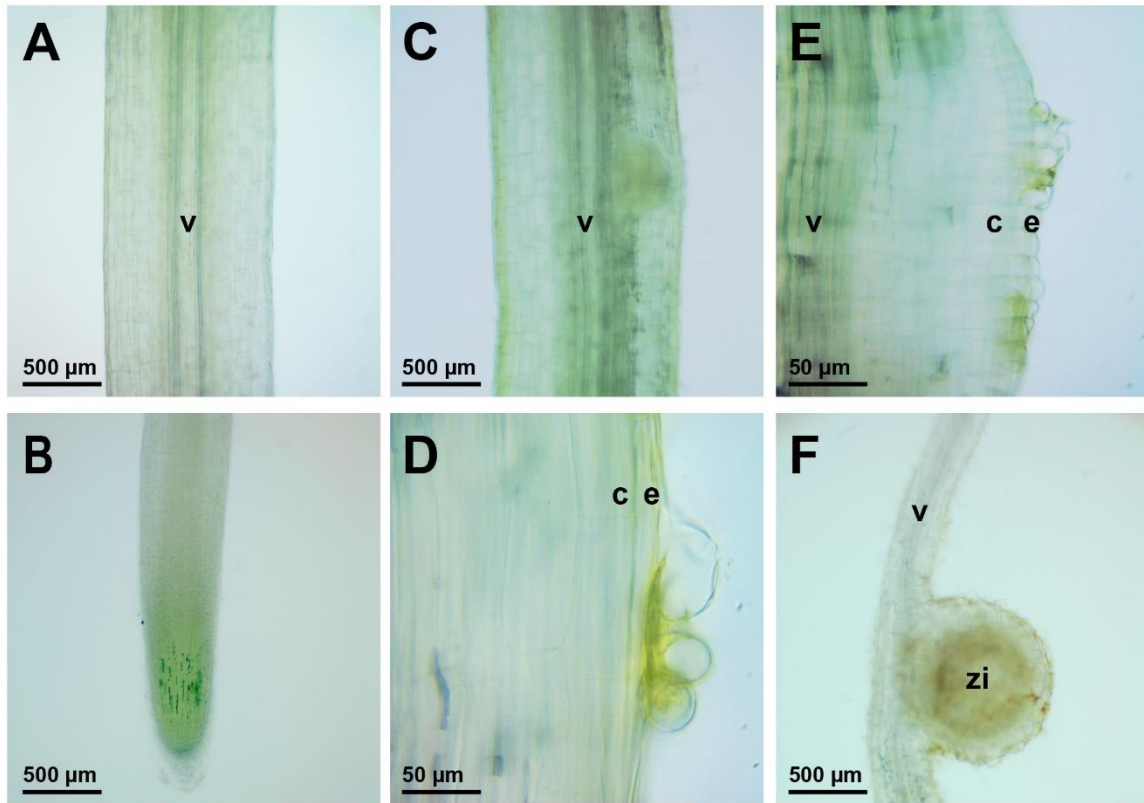


Figura S10. Tinciones de GUS en raíces transgénicas portadoras del vector pBGWFS7.0 vacío . La actividad del promotor se evaluó en raíces transgénicas portadoras de la construcción del vector vacío. Se muestra la ausencia de señal GUS en la vasculatura (A) en el ápice de la raíz principal (B), en los primordios de raíces laterales (C), y durante las etapas iniciales (D y E) y finales de la nodulación (F). v, vasculatura; c, cortex; e, epidermis; zi, zona de infección. n=15.

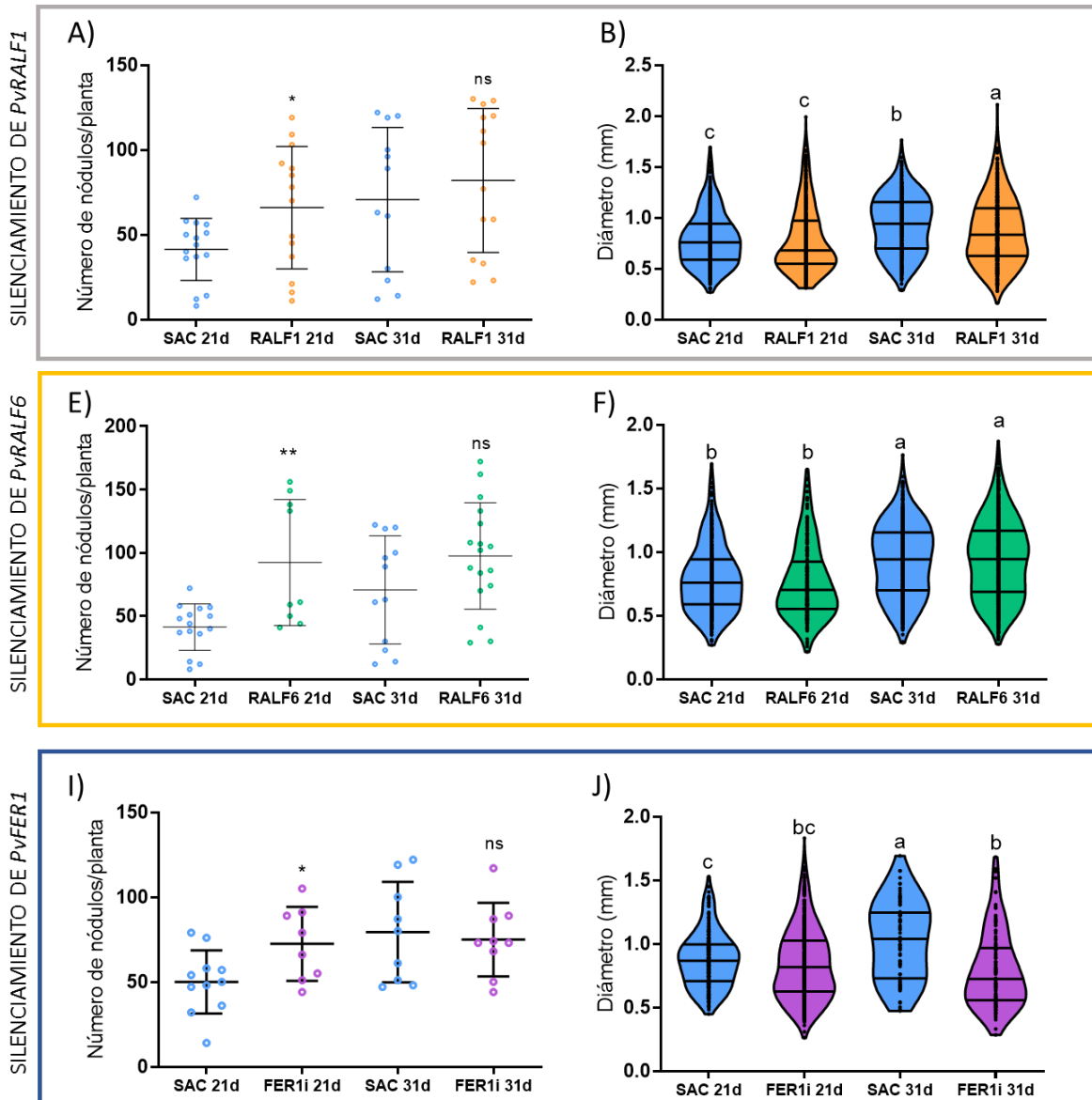


Figura S11. Número y diámetro de nódulos en raíces silenciadas *PvRALF1*, *PvRALF6* y *PvFER1* tratadas con 2 mM de nitrógeno. Número de nódulos generados a los 21 y 31 dpi, en raíces transgénicas silenciadas en *PvRALF1* (puntos naranjas, A), *PvRALF6* (puntos verdes, B), *PvFER1* (puntos morados, C) y raíces control (puntos azules) tratadas con 2 mM de nitrógeno. Cuantificación del diámetro de los nódulos generados a los 21 y 31 dpi, en raíces transgénicas silenciadas en *PvRALF1* (graficas naranjas, D), *PvRALF6* (graficas verdes, E), *PvFER1* (graficas moradas, F) y raíces control (graficas azules) tratadas con 2 mM de nitrógeno. Los datos representan la media \pm desviación estándar (n=10); *P <0.05, **P <0.01, ***P <0.001 ns=no significativo (Student test).

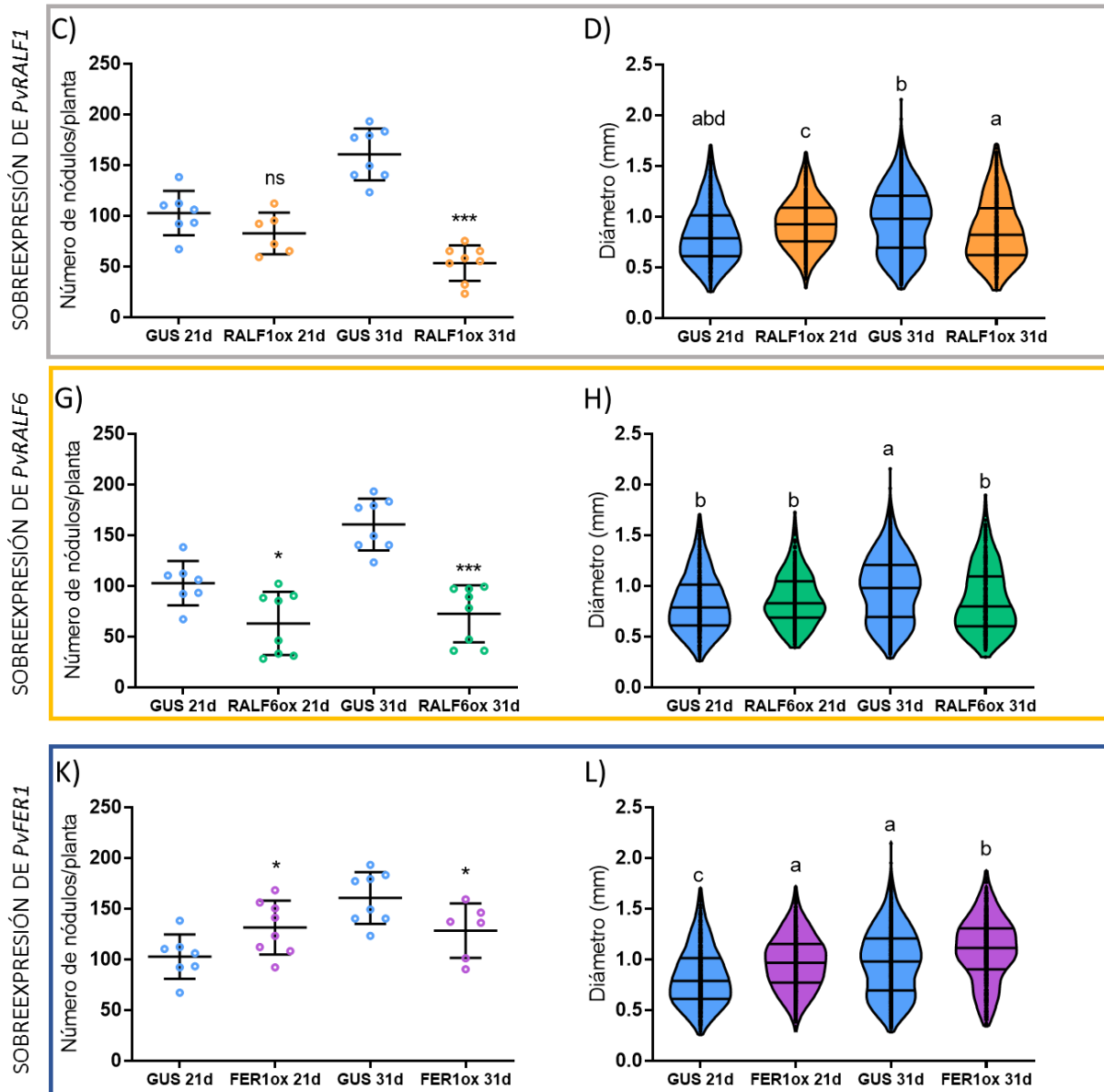


Figura S12. Número y diámetro de nódulos en raíces sobreexpresantes *PvRALF1*, *PvRALF6* y *PvFER1* tratadas con 2 mM de nitrógeno. Número de nódulos generados a los 21 y 31 dpi, en raíces transgénicas sobreexpresantes en *PvRALF1* (puntos naranjas, A), *PvRALF6* (puntos verdes, B), *PvFER1* (puntos morados, C) y raíces control (puntos azules) tratadas con 2 mM de nitrógeno. Cuantificación del diámetro de los nódulos generados a los 21 y 31 dpi, en raíces transgénicas sobreexpresantes en *PvRALF1* (graficas naranjas, D), *PvRALF6* (graficas verdes, E), *PvFER1* (graficas moradas, F) y raíces control (graficas azules) tratadas con 2 mM de nitrógeno. Los datos representan la media \pm desviación estándar (n=10); *P < 0.05, **P < 0.01, ***P < 0.001 ns=no significativo (Student test).

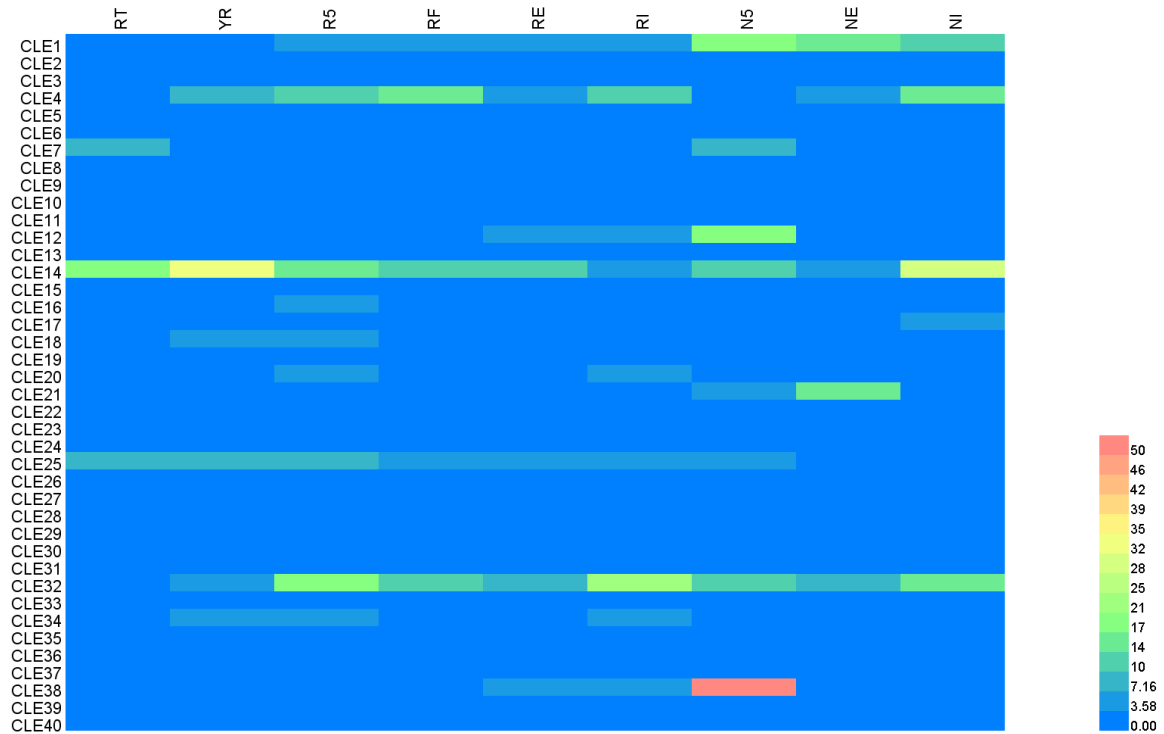


Figura S13. Mapa de calor representativo de la expresión de los 40 genes *CLE* de frijol de acuerdo al PvGEA. Mapa de calor representativo de la expresión de los genes *CLE* en frijol, construido a partir datos del atlas de expresión de frijol PvGEA (O'Rourke *et al.* 2014). Los niveles de expresión son representados de acuerdo a la leyenda de colores. RT- ápices de raíz colectada de plantas fertilizadas en el estado de desarrollo de aparición del segundo trifolio; YR- raíces completas colectadas de plantas en el estado de desarrollo de aparición del segundo trifolio; R5-raíces completas separadas de nódulos de 5 dpi; RF-raíces completas de plantas fertilizadas colectadas al mismo tiempo que RE y RI; RE-raíces completas separadas de nódulos fijadores de 21 dpi; RI- raíces completas separadas de nódulos no fijadores de 21 dpi; N5-nodulos de 5 dpi; NE-nódulos fijadores de 21 dpi; NI-nódulos no fijadores de 21 dpi.

Artículo de revisión publicado

Debido a las restricciones de la revista, solo la primera página del artículo de revisión es mostrada

Plant Physiology and Biochemistry 166 (2021) 88–102



Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Review

The CrRLK1L subfamily: One of the keys to versatility in plants

Jorge Solis-Miranda, Carmen Quinto*

Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Avenida Universidad 2001, Colonia Chamilpa, Cuernavaca, Morelos, 62210, Mexico



ARTICLE INFO

Keywords:

CrRLK1L
FERONIA
RALF
Cell wall
Biotic and abiotic stress

ABSTRACT

Catharanthus roseus kinase 1L receptors (CrRLK1Ls) are a subfamily of membrane receptors unique to plant cells that perceive internal and external signals, integrate metabolic, physiological, and molecular processes, and regulate plant development. Recent genomic studies have suggested that this receptor subfamily arose during the emergence of terrestrial plants and has since diversified, preserving its essential functions. Participation of some of these CrRLK1Ls in different processes is presented and discussed herein, as well as the increasing number of interactors necessary for their function. At least five different responses have been detected after activating these receptors, such as physiological changes, formation or disassembly of protein complexes, metabolic responses, modification of gene expression, and modulation of phytohormone activity. To date, a common response mechanism for all processes involving CrRLK1Ls has not been described. In this review, the information available on the different functions of CrRLK1Ls was compiled. Additionally, the physiological and/or molecular mechanisms involved in the signaling processes triggered by these receptors are also discussed. In this review, we propose a possible common signaling mechanism for all processes regulated by CrRLK1Ls and pose questions to be answered in the future.

1. Introduction

Plants are sessile organisms that sense external cues to accurately respond to environmental conditions and adjust their growth and development under stressful stimuli. Plant cells have many receptors on the plasma membrane, allowing the perception of exogenous signals on the cell surface. Among these receptors, the receptor-like kinase (RLK) family is one of the largest receptor families. For example, in *Arabidopsis thaliana*, approximately 600 RLKs have been reported and are distributed in several subfamilies (Shiu and Bleeker, 2001). RLKs typically contain a single extracellular domain, a transmembrane domain, and a cytoplasmic kinase domain. The first domain perceives the signal and leads to a conformational change and activation of the cytoplasmic domain, promoting intracellular phosphorylation events and activating signaling cascades (Couto and Zipfel, 2016; Hohmann et al., 2017).

Catharanthus roseus RLK-1Ls (CrRLK1Ls) are a subfamily of RLKs exclusive to plant cells (Schulze-Muth et al., 1996). They are characterized by a malectin-like domain in the extracellular region, a single transmembrane domain, and a cytoplasmic kinase domain (Kanaoka and Torii, 2010). Several *in silico* analyses have revealed high conservation of both malectin and kinase domains among different species (Niu

et al., 2016; Zhang et al. 2016, 2020a, 2020a; Kou et al., 2017; Solis-Miranda et al., 2020). Two independent analyses of the crystallographic structure of the extracellular domain of two CrRLK1Ls revealed that they are similar to each other (Du et al., 2018; Moussu et al., 2018). However, a domain exchange analysis between various members of this subfamily in *Arabidopsis*, revealed some specificity of function in the malectin domain (Kessler et al., 2015), which implies certain specialization of the members of this subfamily. CrRLK1Ls have been widely studied in recent years because of their versatility, participating in processes such as responses to biotic and abiotic stress, and plant growth regulation. This subfamily of receptors probably arose during the emergence of terrestrial plants, and subsequently, there was a genetic expansion of the subfamily along with the increase in the size of the genomes (Solis-Miranda et al., 2020).

Previously, we conducted a comparative genetic analysis of 57 species of terrestrial plants, in which we determined that the members of this subfamily of receptors are distributed in 11 clades: 10 composed of at least one gene of *A. thaliana*, and one clade was exclusively composed of CrRLK1Ls of the four mosses studied (Solis-Miranda et al., 2020). Interestingly, in 150 different proteins from five species, high conservation of genomic and proteomic sequences was observed. Furthermore,

* Corresponding author.

E-mail addresses: jorge.solis@ibt.unam.mx (J. Solis-Miranda), carmen.quinto@ibt.unam.mx (C. Quinto).<https://doi.org/10.1016/j.plaphy.2021.05.028>

Received 29 March 2021; Accepted 18 May 2021


Available online 1 June 2021

0981-9428/© 2021 Elsevier Masson SAS. All rights reserved.

Artículo original publicado

Article

Genome-Wide Identification of the CrRLK1L Subfamily and Comparative Analysis of Its Role in the Legume-Rhizobia Symbiosis

Jorge Solis-Miranda , Citlali Fonseca-García , Noreide Nava, Ronal Pacheco  and Carmen Quinto *

Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Avenida Universidad 2001, Colonia Chamilpa, Cuernavaca, Morelos 62210, Mexico; jsolis@ibt.unam.mx (J.S.-M.); fonsecac@ibt.unam.mx (C.F.-G.); noreide@ibt.unam.mx (N.N.); ronal.pacheco@mail.ibt.unam.mx (R.P.)

* Correspondence: quinto@ibt.unam.mx; Tel.: +52-7773291642

Received: 11 June 2020; Accepted: 9 July 2020; Published: 14 July 2020



Abstract: The plant receptor-like-kinase subfamily CrRLK1L has been widely studied, and CrRLK1Ls have been described as crucial regulators in many processes in *Arabidopsis thaliana* (L.), Heynh. Little is known, however, about the functions of these proteins in other plant species, including potential roles in symbiotic nodulation. We performed a phylogenetic analysis of CrRLK1L subfamily receptors of 57 different plant species and identified 1050 CrRLK1L proteins, clustered into 11 clades. This analysis revealed that the CrRLK1L subfamily probably arose in plants during the transition from chlorophytes to embryophytes and has undergone several duplication events during its evolution. Among the CrRLK1Ls of legumes and *A. thaliana*, protein structure, gene structure, and expression patterns were highly conserved. Some legume CrRLK1L genes were active in nodules. A detailed analysis of eight nodule-expressed genes in *Phaseolus vulgaris* L. showed that these genes were differentially expressed in roots at different stages of the symbiotic process. These data suggest that CrRLK1Ls are both conserved and underwent diversification in a wide group of plants, and shed light on the roles of these genes in legume–rhizobia symbiosis.

Keywords: CrRLK1L; expression profile; legumes; nodule symbiosis; phylogenetic analysis

1. Introduction

Plants are continually exposed to many environmental conditions that they must contend with to survive. These conditions are perceived by plant cells as physical or chemical signals that are sensed by plasma membrane receptors. The receptor-like kinase (RLK) family is one of the largest receptor families and is represented in all organisms. RLKs are involved in many processes, including the perception of pathogens and symbiotic partners. Defense-associated RLKs are activated by pathogen-derived molecules (such as flagellin or fungal chitin) and initiate defense responses. Other specific RLKs bind to signal molecules from mycorrhizal fungi or rhizobia, triggering symbiosis.

In the model plant *Arabidopsis thaliana* (L.), Heynh, about 600 RLKs have been described, and they have been classified into several subfamilies [1]. The *Catharanthus roseus* (L.), D. Don RLK-1L (CrRLK1L) subfamily is unique to plants [2] and has been widely studied in *A. thaliana*. CrRLK1L receptors are characterized by a carbohydrate-binding domain known as the malectin-like domain for its similarity to

the animal protein malectin [3]. The *A. thaliana* genome harbors 17 CrRLK1Ls [2], of which FERONIA (FER) has been the most studied. FER was initially characterized as a regulator of female fertility; later, it was described as an important regulator in some phytohormone signaling pathways [4–9] and was shown to be essential for polar growth in root hair cells (Table S1) [10]. More recently, FER has been reported to be a negative regulator of the immune response in plants [11,12], an activator of protein synthesis [13], and a regulator of growth in response to metabolic status (the C/N ratio) (Table 1) [14].

Table 1. CrRLK1L genes studied in different plant species.

Gene Name	Plant Species	Mutant/RNAi	Phenotype	Reference		
FERONIA	<i>A. thaliana</i>	<i>fer</i>	PT overgrowth, multiple PT reach one ovules	[15–19]		
		<i>fer</i>	Collapsed, burst and short RH	[10]		
		<i>fer</i>	Resistance to Powdery mildew infection, increased susceptibility to <i>Pseudomonas syringae</i> pv. <i>tomaeum</i> DC3000	[11,20]		
		<i>fer</i>	Ethylene hypersensitivity, brassinosteroid insensitivity, abscisic acid hypersensitivity, an increase of s-adenosyl methionine synthesis, inhibition of jasmonic acid responses	[5–7,9]		
		FER RNAi	Dwarf phenotype	[4,21]		
		<i>fer</i>	Larger seed size	[22]		
		<i>fer</i>	Salt hypersensitivity	[23]		
		<i>fer</i>	Increased starch accumulation in a sucrose medium.	[14,24]		
		<i>fer</i>	Hypersensitivity to high carbon/nitrogen ratio	[13,25]		
		<i>fer</i>	Reduced induction of ErbB3-binding protein 1, alteration of ribosome synthesis	[26]		
		<i>fer</i>	Delay in the flowering time under long day condition	[27]		
		<i>fer</i>	Hypersensitivity to nickel, tolerance to cadmium, copper, zinc, and lead	[28]		
			<i>Oriza sativa</i> L.	<i>flr2</i>	Enhanced resistance to <i>Magnaporthe oryzae</i> infection	[28]
				<i>flr11</i>	Enhanced resistance to <i>M. oryzae</i> infection	[28]
	<i>Fragaria x ananassa</i>	FaMRL47RNAi	Fruit ripening acceleration	[29]		
BUPS1/2	<i>A. thaliana</i>	<i>bups1</i>	PT overgrowth	[30]		
ANXUR1/2	<i>A. thaliana</i>	<i>bups1 bups2</i>	Enhanced <i>bups1</i> phenotype: PT overgrowth	[19,31,32]		
HERKULES1	<i>A. thaliana</i>	<i>anx1 anx2</i>	Reduced fertility, PT burst	[19,31,32]		
		<i>herk1 the</i>	Dwarf plants	[19,27,33]		
HERKULES2	<i>A. thaliana</i>	<i>herk1</i>	Tolerance to cadmium, copper, nickel, and zinc	[27]		
ANJEA	<i>A. thaliana</i>	<i>herk1</i>	PT overgrowth	[33]		
		<i>herk2</i>	Tolerance to cadmium, copper, nickel, and lead	[34]		
THESEUS1	<i>A. thaliana</i>	<i>herk2</i>	Tolerance to cadmium, copper, nickel, and lead	[27]		
		<i>the prc</i>	Rescues of hypocotyl growth but without <i>prc</i> cellulose deficiency phenotype	[4]		
		<i>the</i>	Hypersensitivity to lead and zinc, tolerance to nickel	[35]		
CAP	<i>A. thaliana</i>	<i>herk1 the</i>	Dwarf plants	[36]		
		<i>cap</i>	Altered PT growth in low calcium	[37]		
CURVY	<i>A. thaliana</i>	<i>cap</i>	RH bursting and bulging	[38]		
MEDOS1-4	<i>A. thaliana</i>	<i>cro</i>	Distortion of trichomes, altered pavement morphology	[37]		
		<i>med1,2,3,4</i>	Reduced growth in presence of metal ions	[38]		

PT: pollen tubes, RH: root hairs, *prc*: *A. thaliana* mutant *procuste*.

In association with FER, other CrRLK1Ls, such as HERKULES1 (HERK1), HERK2, and THESEUS1 (THE1), are involved in cell wall maintenance and cytoplasmic membrane homeostasis (Table 1) [4]. During fertilization, two CrRLK1Ls, HERK1 and ANJEA, together with FER, mediate male–female gametophyte interaction at the synergid cells (Table 1) [33]. Four other CrRLK1Ls, ANXUR1 (ANX1), ANX2, BUDDHAs PAPER SEAL1 (BUPS1), and BUPS2 are essential for preserving the integrity of the pollen tube during growth (Table 1) [30,31]. The CrRLK1L CAP regulates calcium-dependent pollen tube growth, and is also implicated in maintaining cell wall composition in root hairs during tip growth (Table 1) [35,36]. CURVY1, another CrRLK1L, is important in trichome and tapetal cell morphogenesis, the vegetative-to-reproductive state transition, and seed production (Table 1) [37]. Four other CrRLK1Ls, MEDOS1-4, are associated with the regulation of plant development in response to the presence of metal ions (Table 1) [38].

RALF (Rapid alkalization factor) peptides have been described as ligands of some of the CrRLK1L receptors [8,20,39,40]. These peptides are widely distributed in all land plants, and their activity is associated with pH modulation and the production of reactive oxygen species (ROS) [39,41,42]. *A. thaliana* has 34 RALF peptides, which are differentially expressed in different plant tissues [43,44], and a total of 795 RALFs have been identified in 51 different plant species (monocots, eudicots, and early-diverging lineages) [44]. The cysteine-rich peptide RALF1 was the first peptide described as a FER ligand [39]. The RALF1-FER complex is important for fine-tuning the plant response to non-peptide hormones, root elongation, and polar root hair growth in *A. thaliana* [8]. RALF34, RALF4, and RALF19 interact with the CrRLK1L complex BUPS1/2-ANX1/2 during fertilization [40]. RALF34 also binds to THE1 in roots, a signaling step required for division of the pericycle during lateral root initiation [40]. RALF23 acts as a negative regulator of immunity through its interaction with FER [20]. In symbiotic interactions, it has been reported that the *Medicago truncatula* Gaertn. homolog of RALF1 (MtRALF1) functions as a negative regulator of nodule formation during the development of nitrogen-fixing symbioses; however, the receptor that recognizes MtRALF1 and triggers this inhibition of nodule formation is unknown [45].

The formation of nitrogen-fixing root nodules is a complex process and occurs almost exclusively in legumes, a large family of plants [46]. In this process, the plant roots interact with the Gram-negative soil bacteria, known as rhizobia, which through a molecular dialogue between these two partners, induce the formation of a new structure, the nodule, where the rhizobia gain the ability to fix atmospheric nitrogen. Symbiotic development includes changes in gene expression, suppression of defense mechanisms, induction of root cell division, and formation of nitrogen-fixing nodules. Because this is an expensive process for the plant, the establishment of symbiosis is highly regulated. Inhibition of nitrogen fixation, inhibition of symbiosis when interacting with incompatible or non-fixing bacteria, and control of the number of nodules are three of the essential regulatory mechanisms that match the degree of nodulation to the needs of the plant [47].

Although most of the CrRLK1Ls have been studied in *A. thaliana*, little is known about these proteins in other plant models, including legumes. Therefore, our knowledge is very limited about the function of CrRLK1Ls during the legume-rhizobia symbiosis. To address this gap, we first performed a robust phylogenetic analysis of the CrRLK1L subfamily members of more than 60 plant species, including four species of legumes. We compared the gene features and expression profiles of CrRLK1Ls between different organs in four legumes and *A. thaliana* and demonstrated that some *CrRLK1L* genes are expressed in legume nodules. Among these are eight genes that were differentially expressed over the course of nodule development in *P. vulgaris* roots inoculated with rhizobia. This study provides a robust and comprehensive phylogenetic analysis of the CrRLK1L subfamily and unveils, for the first time, relevant information about the presumed role of this receptor subfamily in legume–rhizobia symbiosis.

2. Materials and Methods

2.1. Identification of CrRLK1 Subfamily Proteins in 62 Plant Species

To identify all CrRLK1L proteins in 61 plant genomes available in the Phytozome v12.1 database (<https://phytozome.jgi.doe.gov>) [48] and in the *Lotus japonicus* L. genome (<https://lotus.au.dk/>) [49], BLASTP searches using the *A. thaliana* CrRLK1L protein sequences as query were performed. In both databases, default settings for e-values (e^{-1} value) and the number of hit sequences (100 hits) were used. To confirm that the sequences were part of the CrRLK1L family, they were analyzed with Pfam 32.0 (<http://pfam.xfam.org>) [50] and filtered by the presence of the characteristic malectin-like and kinase domains in this subfamily. A total of 1050 proteins sequences were confirmed as CrRLK1L proteins and downloaded from both databases.

2.2. Phylogenetic Analysis of the CrRLK1L Subfamily

All 1050 protein sequences were aligned using the MUSCLE algorithm [51] followed by a manual optimization of the misaligned sequences in the AliView editor [52]. An approximately maximum-likelihood phylogenetic tree [53] was created for edited sequence alignment with IQ-TREE 1.6.12 [54], using the JTT+F+R10 substitution model with 1000 bootstraps and default parameters. The Pearson correlation coefficient was calculated to determine the relation between the number of CrRLK1L genes and the genome size or the number of CrRLK1L genes between the total number of genes, for the analyzed species.

To explore the possibility that CrRLK1Ls participate in legume–rhizobia symbiosis, a phylogenetic analysis of the CrRLK1L protein sequences of *P. vulgaris*, *L. japonicus*, *Glycine max* (L.), Merr. and *M. truncatula* was performed. To compare legumes with other model plants, we selected *A. thaliana* as the model plant in which the CrRLK1Ls genes have been more studied, and *Physcomitrella patens* (Hedw.) Bruch & Schimp as a representative moss species. All of these plant species have complete accurate genome and proteome annotations in the Phytozome and Lotus Base databases, as well as available expression profile data. Alignment of CrRLK1L protein sequences from these six species was also done using the MUSCLE algorithm [51] within the AliView alignment editor [52], and a manual optimization of the misaligned regions. Then, an approximately maximum-likelihood phylogenetic unrooted tree [53] was established for full-length aligned protein sequences with IQ-TREE 1.6.12 [54] with a JTT+F+R7 substitution model and 1000 bootstraps for reliability, using the default parameters. The clades and subclades of both phylogenetic trees were analyzed using MEGA7 [55].

2.3. Analysis of CrRLK1L Protein Motif Conservation in Legumes and *A. thaliana*

Protein motif conservation of the 150 CrRLK1Ls present in *A. thaliana* and in the four legumes analyzed were determined using the conserved sequence motif analyzer MEME (<http://meme-suite.org>) [56]. The analysis was done using the full-length amino acid sequences, setting the maximum number to 15 motifs, the number of expected motifs to any number of repetitions, and the length of the motif to 10–200 amino acids. The other parameters were kept as default. To calculate the theoretical molecular weight and isoelectric point, the 150 proteins sequences were submitted to the ExpASY web server (https://web.expasy.org/compute_pi/) [57].

2.4. Gene Structure, Chromosomal Localization, and Synteny Analysis of the CrRLK1L Gene Subfamily of Legumes, *A. thaliana*, and *Sorghum bicolor* (L.), Moench

The gene structure and chromosomal localization data of the 33 *P. vulgaris*, 18 *L. japonicus*, 46 *G. max*, 36 *M. truncatula*, 17 *A. thaliana*, and 14 *S. bicolor* CrRLK1L genes were retrieved from the Phytozome v12.1 [48] database and Lotus Base [49]. *S. bicolor* was used to evaluate the differences between eudicot and monocot CrRLK1L genes features, since it is a monocot model with complete genome sequence and gene expression information. The gene structure map for each species was represented using the free resource Gene Structure Display Server 2.0 (http://gsds.gao-lab.org/Gsds_about.php). For the chromosome distribution, data were uploaded into the free resource PhenoGram Plot (<http://visualization.ritchielab.org/phenograms/plot>) [58]. For synteny analysis, the protein sequences and annotation files of the full genomes of *P. vulgaris*, *L. japonicus*, *G. max*, and *L. japonicus* were downloaded from the previously mentioned databases [48,49]. For each case, an m8 format BLASTP file and a simplified gff file were used as inputs to the collinearity scanner toolkit MCScanx (<http://chibba.pgml.uga.edu/mcscan2/>) [59] to determine synteny between CrRLK1L genes in the legume species and to compare it with that in *A. thaliana*.

2.5. In Silico Analysis of the CrRLK1L Gene Family Expression in Legumes, *A. thaliana*, and *P. patens*

Expression profiles of the 33 members of the *P. vulgaris* CrRLK1L gene subfamily were retrieved from the Common bean Gene Expression Atlas, PvGEA (<https://plantgrn.noble.org/PvGEA/>) [60]. *L. japonicus* expression profile data were downloaded from the *L. japonicus* reference genome transcript explorer in Lotus Base [49]. *M. truncatula* expression profile data were downloaded from the *M. truncatula* Gene Expression Atlas (MtGEA) [61,62] by BLASTN. The expression profiles of the 46 *G. max*, 17 *A. thaliana*, and 6 *P. patens* CrRLK1L genes were obtained from the Bio-Analytic Resource for Plant Biology (BAR) [61,63–65]. The distribution and abundance of the expression profile of the genes were presented in heatmaps with the function `heatmap.2` of the `gplot` package [66] using R. To identify the shared genes expressed in nodules of *P. vulgaris*, *L. japonicus*, and *G. max*, a Venn diagram was drawn using the Venn diagram drawing tool (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

2.6. Plant Growth Conditions and RT-qPCR Assays

Common bean (*P. vulgaris* cv. Negro Jamapa) seeds were surface-sterilized and germinated for 2 days (dpg) at 28 °C in darkness. For RT-qPCR accumulation profile analysis during nodulation, 2 dpg seedlings were transplanted into pots with vermiculite and inoculated with *Rhizobium tropici* CIAT899 at an OD₆₀₀ of 0.05, or only with Fahraeus media as mock. Roots were harvested at 5, 7, 14, and 21 days post-inoculation (dpi). The tissues selected for RT-qPCR were immediately frozen in liquid nitrogen and stored at −70 °C until RNA extraction. RNA was isolated from the frozen tissues using Trizol reagent (Sigma-Aldrich, St. Louis, MO, USA), following the manufacturer's instructions. RNA integrity was verified by electrophoresis and the concentration was assessed using a NanoDrop2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). To eliminate DNA contamination, the RNA samples were incubated with RNase-free DNase (1 U/μL; Roche, Basel, Switzerland).

Complementary DNA (cDNA) was synthesized using Thermo Scientific RevertAid Reverse Transcriptase (200 U/μL, Thermo Scientific, Waltham, MA, USA), with 200 ng of DNA-free RNA as template and following the manufacturer's instructions. RT-qPCR assays were performed using Maxima SYBR Green/ROX qPCR Master Mix (2X) (Thermo Scientific, Waltham, MA, USA), in a real time PCR system (QuantStudio 5; Applied Biosystems, Waltham, MA, USA) with the following thermal cycle: 95 °C for 10 min, 30 cycles of 95 °C for 15 s, and 60 °C for 60 s. Experiments were normalized with the reference gene *Elongation Factor 1α* (*EF1 α*) [67]. Relative expression values were calculated using the formula $2^{-\Delta Ct}$, where the cycle threshold value ΔCt is equal to the Ct of the gene of interest minus the Ct of the reference gene. Three biological replicates with three technical repeats were performed for each dataset. The gene-specific oligonucleotides used in this study are listed in Supplementary Materials Table S1.

3. Results

3.1. Identification and Phylogenetic Analysis of CrRLK1L Proteins in Diverse Plant Species

Previous reports identifying CrRLK1L subfamily receptors have focused on a few model species, such as *A. thaliana* and *Oryza sativa* L. To study the potential function of this receptor subfamily in legume-rhizobia symbiosis, we expanded our analysis of these proteins to other plant species by searching for CrRLK1Ls in 61 plant genomes deposited in Phytozome v12.1 and also in the *L. japonicus* genome. We searched for CrRLK1L homologs in the 62 species, followed by a domain analysis of the identified proteins to confirm the presence of the characteristic malectin-like and kinase domains of the CrRLK1L subfamily. We identified a total of 1050 CrRLK1L proteins in 57 of the 62 species analyzed. None of the five chlorophyte genomes we searched had a significant hit related to the CrRLK1L subfamily. By contrast,

at least one CrRLK1L was encoded in every land plant genome in our analysis (Table S2). The complete data are summarized in Table S2, and the IDs of the 1050 CrRLK1L proteins are listed in Table S3.

Based on a phylogenetic analysis of the amino acid sequences of these 1050 proteins, we established that they are distributed into 11 clades (Figure 1A, Figure S1). One of these clades consisted exclusively of CrRLK1Ls of the most ancient plant species included in this analysis: three bryophytes (*Marchantia polymorpha* L., *P. patens*, and *Sphagnum fallax* (H.Klinggr) H.Klinggr) and a clubmoss (*Selaginella moellendorffii* Hieron). This clade was named TINIA (after the first of the Etruscan gods and father of Herkules), following the mythological nomenclature used for other CrRLK1L clades (Figure 1A, Figure S1). Of the remaining ten clades, nine were named according to the nomenclature employed for the *A. thaliana* protein belonging to each clade (Figure 1A, Figure S1). A clade carrying the two uncharacterized *A. thaliana* CrRLK1Ls was named CADMUS (after the Etruscan king founder of Thebes) (Figure 1A, Figure S1). Eudicots had an average of 22 CrRLK1L proteins, whereas monocots had fewer, with an average of 13 (Figure 1A, Figure S1, Table S1). This difference in the number of CrRLK1Ls could be associated with the greater size of eudicot genomes compared to those of monocots, since there is a moderate correlation between the number of CrRLK1Ls and the genome size and the number of total genes in both groups (eudicots, $r = 0.55$ and $r = 0.47$, respectively; monocots, $r = 0.51$ and $r = 0.62$, respectively) (Figure 1B,C). These data suggest that the CrRLK1L subfamily probably appeared during the transition from chlorophytes to embryophytes, and that the number of members increased along with the size of the genome and the total number of genes during evolution.

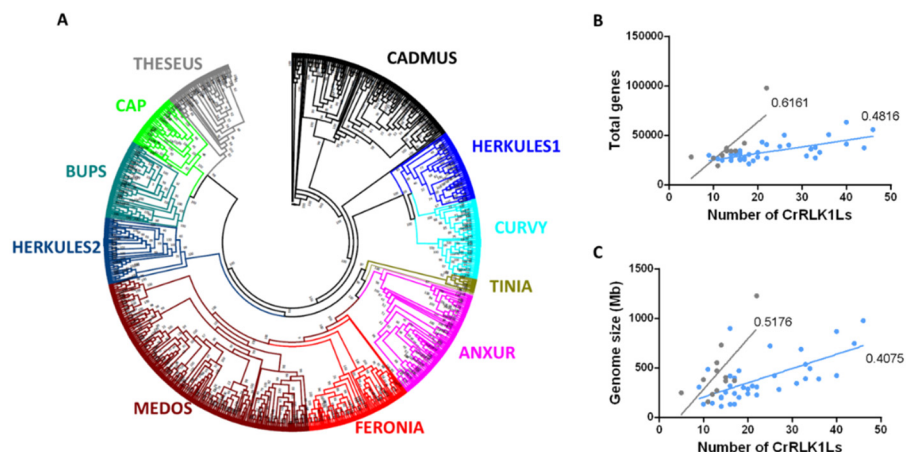


Figure 1. Phylogenetic relationship among 1050 CrRLK1L proteins and the relationship between *CrRLK1L* number versus gene size and gene number. **(A)** Unrooted approximately maximum-likelihood phylogenetic tree inferred from the 1050 CrRLK1L proteins present in 57 plant species. The clades, indicated in different colors, are named based on the *A. thaliana* CrRLK1L names. The CADMUS clade contains uncharacterized CrRLK1Ls from *A. thaliana*, and the TINIA clade corresponds to a clade formed only with the CrRLK1L proteins of *S. moellendorffii*, *S. fallax*, *P. patens*, and *M. polymorpha*. The phylogenetic tree was constructed using IQ-TREE software with the JTT+F+R10 substitution model with 1000 bootstrap iterations. **(B)** Relationship between total number of genes within the genome and number of *CrRLK1L* genes for monocotyledons (gray) and eudicots (blue). **(C)** Relationship between genome size and number of *CrRLK1L* genes for monocotyledons (gray) and eudicots (blue).

3.2. Phylogenetic Analysis of the *CrRLK1L* Subfamily in Legumes, *A. thaliana*, and *P. patens*

Legumes have the ability to establish a symbiotic relationship with rhizobia and form nitrogen-fixing nodules. To explore the possibility that CrRLK1Ls participate in legume–rhizobia symbiosis, we constructed

a phylogenetic tree that included all the CrRLK1Ls of four model legumes (*L. japonicus*, *M. truncatula*, *G. max*, and *P. vulgaris*), *A. thaliana*, and the moss *P. patens*.

As expected, all five *P. patens* proteins were placed in the basal TINIA clade, separated from the proteins of the four legumes and *A. thaliana* (Figure S2). The remaining CrRLK1L proteins were distributed among ten clades, each containing at least one CrRLK1L from *A. thaliana* and one to several proteins from the legumes. Although no clade was confined exclusively to legumes, some clades had more members in the legumes than are present in *A. thaliana*. In this context, the MEDOS clade was particularly interesting because *P. vulgaris*, *M. truncatula*, and *G. max* each have a relatively large number of these proteins (16, 18, and 24, respectively), whereas *A. thaliana* has only four (Table 2). These observations indicate that although there is no group of proteins exclusively associated with legumes, there is at least a four-fold increase in the number of CrRLK1L proteins in this plant family compared to *A. thaliana*, particularly in the MEDOS clade.

Table 2. Transcript lengths and protein properties of the CrRLK1L subfamily members in *P. vulgaris*, *G. max*, *A. thaliana*, *L. japonicus*, and *M. truncatula*.

Gene ID *	Gene Name	CDS Length, bp	Protein Length, aa	iP	Molecular Weight, kDa
<i>P. vulgaris</i>					
PhvuL006G102700	ANX1	2589	862	5.71	95.82
PhvuL007G188300	ANX2	2589	862	5.6	96
PhvuL011G210400	BUPS	2670	889	5.41	97.19
PhvuL003G188000	CAP	2469	822	5.91	92.06
PhvuL004G109500	CRV1	2505	834	5.49	92.38
PhvuL007G074000	CRV2	2535	844	5.7	93.64
PhvuL008G081000	FER1	2700	899	5.99	98.37
PhvuL008G082400	FER2	2697	898	6.67	98.1
PhvuL005G139800	HERK1A	2514	837	5.64	92.51
PhvuL008G000200	HERK1B	2472	823	5.38	92.18
PhvuL011G069600	HERK1C	2514	837	5.79	92.94
PhvuL006G127900	HERK2	2547	848	7.3	93.01
PhvuL004G038800	MEDOS1A	2676	891	7.29	99.43
PhvuL004G039200	MEDOS1B	2235	744	7.05	83.82
PhvuL004G039600	MEDOS1C	2598	865	5.78	96.98
PhvuL004G039700	MEDOS1D	2406	801	5.79	90.35
PhvuL004G039800	MEDOS1E	2442	813	5.99	92.04
PhvuL004G039900	MEDOS1F	2490	829	6.51	93.64
PhvuL004G040000	MEDOS1G	2460	819	7.31	92.32
PhvuL004G040300	MEDOS1H	1824	607	6.26	68.13
PhvuL004G040900	MEDOS1I	1353	450	8.46	50.59
PhvuL004G039400	MEDOS2A	2463	820	5.99	92.31
PhvuL008G030200	MEDOS3A	2595	864	6.15	96.81
PhvuL008G030400	MEDOS3B	2517	838	6.24	93.86
PhvuL008G030700	MEDOS3C	2577	858	6.28	96.37
PhvuL008G030800	MEDOS3D	2601	866	4.83	97.08
PhvuL003G038700	MEDOS4A	2385	794	8.67	89.03
PhvuL003G038800	MEDOS4B	2523	840	5.95	93.94
PhvuL005G085600	THE1	2523	840	5.93	92.73
PhvuL011G148700	THE2	2538	845	5.87	93.05
PhvuL003G239300	CAD1	2499	832	6.63	93.06
PhvuL003G239400	CAD2	2589	862	8.33	96.18
PhvuL003G239500	CAD3	2586	861	5.86	96.32
<i>G. max</i>					
Glyma.03G247800	ANXUR1	2610	869	5.24	96.22
Glyma.10G163200	ANXUR2	2589	862	5.67	95.95
Glyma.19G245800	ANXUR3	2601	866	5.31	95.71
Glyma.20G225800	ANXUR4	2532	843	5.8	93.73
Glyma.12G235900	BUPS1	2637	878	5.77	96.33
Glyma.13G201400	BUPS2	2610	869	5.85	95.29
Glyma.17G102600	CAP1	2586	861	6.55	95.69
Glyma.09G273300	FERONIA1	2691	896	5.64	98.07

Table 2. Cont.

Gene ID *	Gene Name	CDS Length, bp	Protein Length, aa	iP	Molecular Weight, kDa
Glyma.18G215800	FERONIA2	2685	894	5.66	97.76
Glyma.12G074600	HERKULES1A	2514	837	5.86	92.74
Glyma.15G042900	HERKULES1B	2226	741	7.93	81.93
Glyma.U033500	HERKULES1C	2436	811	6.5	89.91
Glyma.09G024700	HERKULES2	2559	852	5.59	93.59
Glyma.02G121900	MEDOS1A	2463	820	8.23	92.15
Glyma.02G122000	MEDOS1B	1944	647	5.81	72.73
Glyma.02G196000	MEDOS1C	2481	826	5.83	93.46
Glyma.08G248900	MEDOS2A	2529	842	6.25	92.88
Glyma.08G249200	MEDOS2B	2616	871	6.24	96.62
Glyma.08G249400	MEDOS2C	2373	790	6.02	88.62
Glyma.13G054400	MEDOS3A	2691	896	6.11	99.53
Glyma.13G053800	MEDOS3B	2109	702	6.44	77.97
Glyma.13G053700	MEDOS3C	2460	819	5.63	91.47
Glyma.13G053600	MEDOS3D	2685	894	5.9	99.35
Glyma.13G054200	MEDOS3E	2364	787	8.48	88.74
Glyma.13G054300	MEDOS3F	2535	844	5.98	94.104
Glyma.18G269900	MEDOS4A	2610	869	6.26	97.24
Glyma.18G270100	MEDOS4B	2607	868	6.09	97.39
Glyma.18G270600	MEDOS4C	3372	1123	5.77	124.54
Glyma.18G270700	MEDOS4D	2574	857	6.25	95.77
Glyma.18G270900	MEDOS4E	2628	875	5.82	97.45
Glyma.18G271000	MEDOS4F	2592	863	5.9	96.86
Glyma.18G271100	MEDOS4G	2652	883	6.02	98.05
Glyma.18G270800	MEDOS4I	2730	909	5.98	102.67
Glyma.18G271200	MEDOS4J	2550	849	5.83	95.14
Glyma.19G033100	MEDOS5A	3561	1186	6.49	133.95
Glyma.U027000	MEDOS5B	2364	787	8.48	88.76
Glyma.U027100	MEDOS5C	2460	819	5.75	91.75
Glyma.12G148200	THESEUS1	2541	846	5.68	93.19
Glyma.12G220400	THESEUS2	2070	689	6.44	75.72
Glyma.05G099900	CAD1	2382	793	6.3	88.42
Glyma.05G100000	CAD2	2517	838	5.58	94.02
Glyma.09G133000	CAD3	2457	818	7.04	99.92
Glyma.10G231500	CAD4	2481	826	7.94	91.86
Glyma.16G179600	CAD5	2322	773	8.81	86.06
Glyma.17G166200	CAD6	2523	840	5.8	93.93
Glyma.20G162300	CAD7	2523	840	8.17	93.19
<i>A. thaliana</i>					
AT3G04690	ANX1	2837	895	6.47	98.16
AT5G28680	ANX2	2577	850	6.54	94.06
AT4G39110	BUPS1	2637	858	5.76	94.31
AT2G21480	BUPS2	2616	873	5.66	97.18
AT5G61350	CAP	2529	880	5.92	97.96
AT2G39360	CRV	2683	815	6.13	91.33
AT3G51550	FER	3298	830	5.82	91.48
AT5G59700	HERK/ANJ	3041	849	5.76	93.96
AT3G46290	HERK1	3158	855	5.91	93.31
AT1G30570	HERK2	2550	842	6.16	92.7
AT5G39000	MEDO2	2622	878	5.75	96.52
AT5G38990	MEDOS1	2785	871	5.51	95.95
AT5G39020	MEDOS3	2442	829	6.5	91.97
AT5G39030	MEDOS4	2421	824	5.65	91.84
AT5G54380	THE1	2789	834	5.7	93.39
AT5G24010	CAD1	2821	813	7.6	90.64
AT2G23200	CAD2	2633	806	5.97	90.68
<i>L. japonicus</i>					
Lj1g3v4996200	ANXUR	2592	863	5.46	95.4
Lj0g3v0115159	BUPS	2643	880	5.93	96.29
Lj3g3v3639930	CURVY1	2472	823	6.08	90.91
Lj3g3v3639940	CURVY2	2121	706	5.98	77.77
Lj1g3v2533770	FERONIA	2094	697	5.96	75.96

Table 2. Cont.

Gene ID *	Gene Name	CDS Length, bp	Protein Length, aa	iP	Molecular Weight, kDa
Lj0g3v0249939	HERKULES1A	2514	837	5.41	91.74
Lj3g3v3132890	HERKULES1B	1848	615	6.61	67.66
Lj6g3v1641160	HERKULES2A	2532	843	5.77	92.66
Lj6g3v1641170	HERKULES2B	2532	843	5.77	92.66
Lj2g3v1102970	MEDOS1	2676	891	6.75	97.76
Lj2g3v1226730	MEDOS2	1542	513	7.59	58.1
Lj2g3v1226740	MEDOS3	2466	821	6.93	92.49
Lj2g3v1226750	MEDOS4	2277	758	5.49	84.89
Lj0g3v0346559	THESEUS	2535	844	5.54	92.59
Lj0g3v0151929	CAD1	1554	517	8.87	57.31
Lj2g3v0322770	CAD2	2493	830	6.46	92.25
Lj2g3v1902230	CAD3	2169	722	8.75	80.67
Lj5g3v1988700	CAD4	2535	844	6.98	94.3
<i>M. truncatula</i>					
Medtr1g080740	ANX1	2607	868	6.15	96.86
Medtr7g115300	ANX2	2619	872	5.32	97.04
Medtr8g037700	BUPS1	2406	801	5.94	96.47
Medtr4g109010	CAP1	3459	1152	6.5	129.96
Medtr4g111925	FER1	2106	701	6.5	76.58
Medtr7g073660	FER2	2700	899	5.89	97.98
Medtr4g061930	HERK1A	2523	840	5.82	93.13
Medtr2g096160	HERK1B	2544	847	5.57	92.93
Medtr4g061833	HERK1C	2523	840	5.82	93.13
Medtr2g030310	HERK2	2628	875	5.89	96.34
Medtr6g015805	MEDOS1A	2703	900	6.8	100.32
Medtr5g047120	MEDOS2A	2430	809	7.09	92.07
Medtr5g047070	MEDOS2B	1707	568	6.23	64.53
Medtr7g015390	MEDOS3A	2670	889	6.25	101.42
Medtr7g015550	MEDOS3B	2667	888	5.92	100.81
Medtr7g015670	MEDOS3C	2679	892	5.75	101.65
Medtr7g015510	MEDOS3D	2667	888	5.84	100.38
Medtr7g015240	MEDOS3E	2529	842	6.16	96.25
Medtr7g015280	MEDOS3F	2577	858	6.44	97.98
Medtr7g015420	MEDOS3G	3417	1138	6.02	130.78
Medtr4g052290	MEDOS3H	2658	885	7.8	101.29
Medtr7g015250	MEDOS3I	2733	910	6.16	103.04
Medtr7g015310	MEDOS3J	2622	873	6.08	98.98
Medtr7g015230	MEDOS3K	2652	883	6.91	100.7
Medtr7g015320	MEDOS3L	2637	878	7.58	99.19
Medtr7g015620	MEDOS3M	2064	687	8.81	78.9
Medtr5g047060	MEDOS4A	2502	833	5.21	94.17
Medtr5g047110	MEDOS4B	2454	817	5.5	92.57
Medtr6g048090	OG1	2385	794	6.84	88.69
Medtr1g100110	OG2	2445	814	6.57	91.66
Medtr2g080220	THE1	2532	843	5.61	92.73
Medtr4g095042	CAD1	2556	851	5.85	94.87
Medtr4g095012	CAD2	2460	819	6.12	91.27
Medtr4g095032	CAD3	2361	786	5.82	87.7
Medtr1g040073	CAD4	2256	751	6.23	83.97
Medtr8g467150	CAD5	2277	758	7.33	85.16

* Phytozome ID. bp: base pairs. CDS: coding sequence. aa: amino acids. iP: isoelectric point. kDa: kiloDalton.

3.3. Features of the CrRLK1L Subfamily Proteins in Legumes and *A. thaliana*

Since we identified no CrRLK1L clade that was exclusive to legumes, we wondered whether some of these proteins, which have important functions in other plant processes, could have been recruited to function in legume-rhizobia symbiosis. To assess this possibility, we analyzed the molecular characteristics of the CrRLK1L proteins of the four legumes previously examined and *A. thaliana*. There are 33 CrRLK1L proteins encoded in the *P. vulgaris* genome, whereas in *A. thaliana* there are 17, in both cases distributed

among ten different clades. *M. truncatula*, *G. max*, and *L. japonicus* have 36, 46, and 18 CrRLK1Ls, respectively, scattered among nine clades (Figure S2).

The CrRLK1L proteins are defined by the presence of a malectin-like domain in the amino-terminal region and a kinase domain in the carboxy-terminal region. To characterize these conserved motifs, the CrRLK1L sequences of the four legumes under study and *A. thaliana* were examined using MEME software (Figure 2). In the 150 sequences analyzed, ten different motifs were identified; seven of these were located in the kinase domain, and only three in the malectin domain, two of them duplicated. The motifs located in the kinase domain were longer and more conserved than those in the malectin-like domain (Figure 2). No additional features were observed that could be associated with a given species or phylogenetic clade, beyond the particularities of individual proteins, such as shorter or longer amino acid sequences.

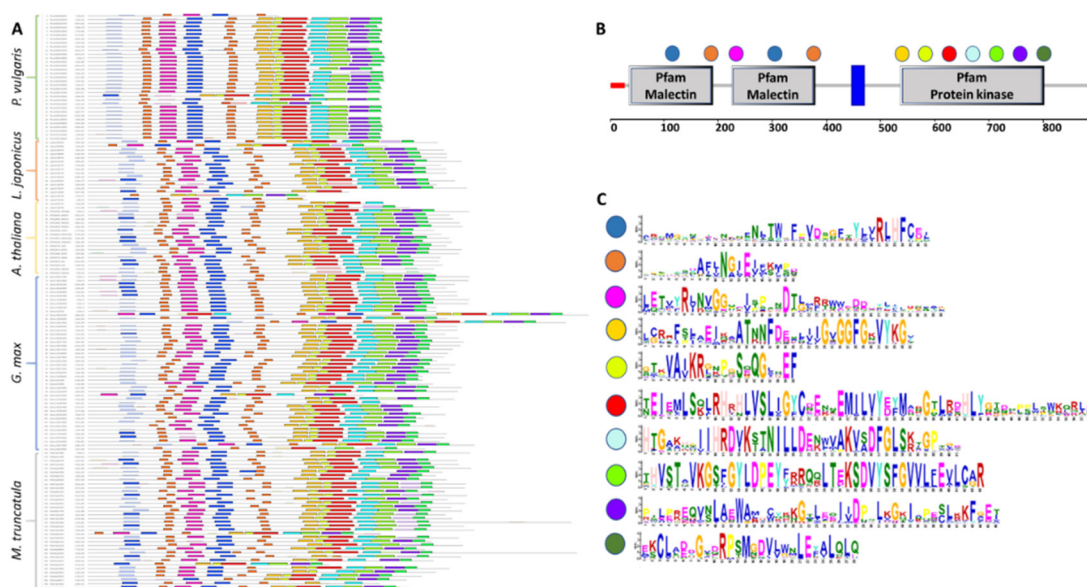


Figure 2. CrRLK1L protein sequence conservation and characteristic motifs in legumes and *A. thaliana*. MEME was used to identify motifs in the 150 CrRLK1Ls from four legumes and *A. thaliana*. (A) Diagram of the motifs in the CrRLK1L protein sequences from *P. vulgaris*, *L. japonicus*, *G. max*, *M. truncatula*, and *A. thaliana*. Significant overrepresented motifs are graphically depicted by bars corresponding to their predicted position. (B) Localization of overrepresented motifs identified using MEME in the CrRLK1L protein domains. (C) Logo of the overrepresented motifs identified with MEME; the color code corresponds with that used in (A,B).

The 33 *P. vulgaris* CrRLK1L proteins ranged from 450 to 899 amino acids (aa) in length and 50.59 to 99.43 kDa in molecular weight (MW) (Table 2). The theoretical isoelectric point (iP) of most of the common bean proteins is slightly acidic (4.83 to 6.67), though seven proteins are slightly alkaline (7.05 to 8.67) (Table 2). The *L. japonicus* CrRLK1Ls showed similar features to those of common bean, with MWs of 57.31 to 97.76 kDa and lengths of 513 to 891 aa. Furthermore, most of the *L. japonicus* proteins have acidic iPs (5.93 to 6.98), whereas only three of them have alkaline iPs (7.59 to 8.87) (Table 2). The CrRLK1L proteins have a broader MW range in *M. truncatula* and *G. max* (64.53 to 130.78 and 72.73 to 133.95 kDa, respectively) and are longer (568 to 1152 aa and 647 to 1186 aa, respectively) than to those of *P. vulgaris* (Table 2). In *M. truncatula*, only five proteins have alkaline iPs (7.05 to 8.81), whereas the remaining 31 have acidic iPs (5.21 to 6.91) (Table 2). In *G. max*, eight of the proteins were alkaline (7.04 to 8.81) and 38

are acidic (5.24 to 6.55) (Table 2). In comparison to the legume proteins, the *A. thaliana* CrRLK1Ls have narrower ranges of MW (90.68 to 98.16 kDa) and length (806 to 895 aa). Only one *A. thaliana* protein has an alkaline iP (7.6), whereas 16 have a slightly acidic iP (5.51 to 6.54) (Table 2).

Despite high conservation of the CrRLK1L domains, there were some physicochemical differences between the legume proteins we studied and those of *A. thaliana*. This variation could be associated with the higher number of proteins observed in *G. max*, *P. vulgaris*, and *M. truncatula* compared to *A. thaliana*, which could have allowed more divergence of the proteins over time.

3.4. Chromosomal Localization and Synteny of CrRLK1L Genes in Legumes and *A. thaliana*

To compare the genome distributions of CrRLK1L genes in *A. thaliana* and in the four legumes under study, we used PhenoGram Plot to map the chromosome locations of the CrRLK1L genes in each plant species. The *P. vulgaris* CrRLK1L genes are distributed among seven of the 11 chromosomes, mainly on chromosomes four and eight (Figure S3A). The *P. vulgaris* MEDOS genes were mapped to chromosomes three (two genes), four (ten genes), and eight (four genes). *M. truncatula* and *G. max* have a similar gene distribution, the 46 *G. max* CrRLK1L genes are distributed among 14 of the 20 chromosomes, with two groups of MEDOS clade genes, one on chromosome 13 (six genes) and the other on chromosome 18 (eight genes) (Figure S3C). *M. truncatula* has 36 CrRLK1L genes, located on seven chromosomes, and two clusters of MEDOS clade genes on chromosomes five (four genes) and seven (12 genes) (Figure S3D). In the *L. japonicus* and *A. thaliana* genomes, CrRLK1L genes are distributed on five chromosomes, with only one small cluster, consisting of MEDOS genes, in each species, on chromosome two in *L. japonicus* (Figure S3B) and on chromosome five in *A. thaliana* (Figure S3E). Previously, we found that the MEDOS clade is absent in most monocots; therefore, we decided to also examine the distribution of the CrRLK1L genes in *S. bicolor*. In this species, all CrRLK1L genes were located on six chromosomes, and no clustering was observed (Figure S3F). These data suggest that there has been an expansion of the MEDOS genes in eudicots, and that some species have undergone greater expansion than others.

To further explore the evolutionary trajectories of the CrRLK1L genes, we evaluated the local synteny among the CrRLK1L genes in *P. vulgaris*, *G. max*, *L. japonicus*, *M. truncatula*, and *A. thaliana*. Chromosomal synteny was evaluated in these five species individually and also between species using MCScanx software. This analysis showed that four pairs of genes are syntenic in *P. vulgaris*, corresponding to 25% of the CrRLK1L genes. This was close to the percentage of gene synteny in the *P. vulgaris* genome overall (28.68%) (Figure S4, Table S4). In *L. japonicus*, only one pair of syntenic genes was identified, corresponding to 11% of CrRLK1L genes. Despite being lower than the CrRLK1L gene synteny in *P. vulgaris*, this percentage is above the median for the *L. japonicum* genome overall (4.74%) (Figure S4, Table S4). In *G. max*, 21 pairs of the CrRLK1L genes are syntenic, corresponding to 25 different genes (54.35% of the CrRLK1L genes). This gene synteny is slightly low given that 68.13% of all the genes in *G. max* genome are syntenic (Figure S4, Table S4). In *A. thaliana*, two pairs of genes have synteny, which corresponds to 23.53% of the CrRLK1Ls, similar to the 27.1% synteny of the *A. thaliana* genome overall (Figure S4, Table S4).

Collinearity of genes was also examined between pairs of legume species that develop determinate nodules, namely *P. vulgaris*, *G. max*, and *L. japonicus*. The results indicate that 35 CrRLK1L genes have collinearity between *P. vulgaris* and *G. max*, 12 genes between *P. vulgaris* and *L. japonicus*, and 20 genes between *L. japonicus* and *G. max* (Figure 3A–C, Table S4). Collinearity between the CrRLK1Ls of these three legumes versus *A. thaliana* was also explored. This analysis revealed 24 legume genes that are syntenic to those of *A. thaliana*: 9 from *P. vulgaris*, 1 from *L. japonicus*, and 14 from *G. max* (Figure 3A–C, Table S4). We noted that in most of the syntenic gene pairs identified, the genes belong to the same clade. Thus, some FER genes in *P. vulgaris* are syntenic to FER genes in *L. japonicus* and *G. max*, and the same was true for

some of the *MEDOS* and *HERK* genes (Figure 3A–C, Table S4). Moreover, some of the genes maintain this collinearity between legumes and *A. thaliana*. These data strongly suggest that these genes are orthologous.

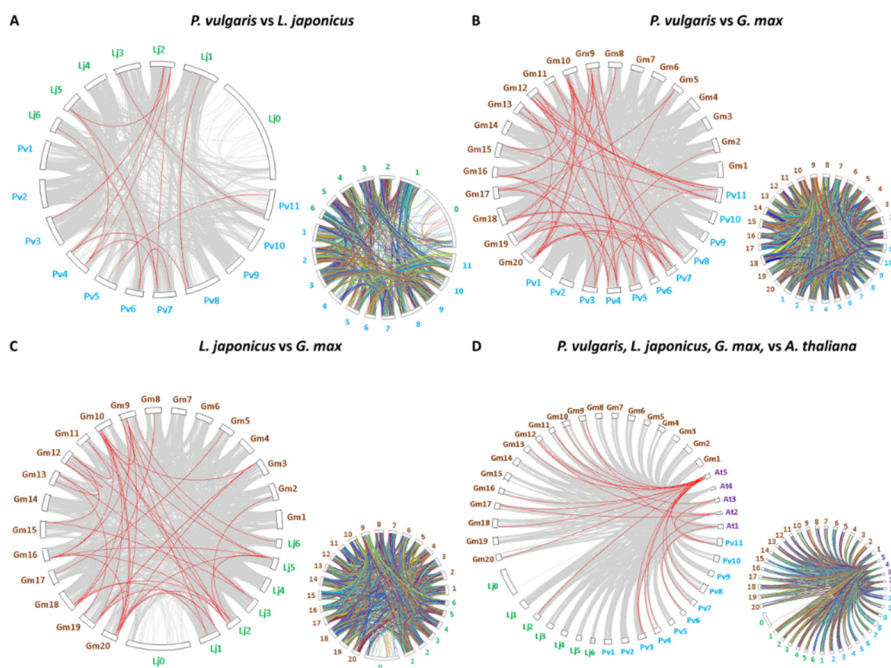


Figure 3. Synteny of the *CrRLK1L* genes in *P. vulgaris*, *L. japonicus*, *G. max*, and *A. thaliana*. MCScan software was used to analyze the syntenic correlation between species. (A) Synteny map of *CrRLK1L*s between *L. japonicus* and *P. vulgaris*. (B) Synteny map of *CrRLK1L*s between *G. max* and *P. vulgaris*. (C) Synteny map of *CrRLK1L*s between *L. japonicus* and *G. max*. (D) Synteny map of *CrRLK1L*s between *A. thaliana* and the three legumes examined. Chromosome numbers are indicated outside each figure as follows: in turquoise, *P. vulgaris*; in green, *L. japonicus*; in brown, *G. max*; and in purple, *A. thaliana*. The smaller synteny maps to the right of each image represent the synteny of all the genes in the genomes compared here.

3.5. Exon–Intron Structure of *CrRLK1L* Genes in Legumes and *A. thaliana*

To analyze the structural organization and evolution of the *CrRLK1L* genes following their duplication, we analyzed the exon–intron distribution in these genes in *P. vulgaris*, *M. truncatula*, *G. max*, *L. japonicus*, and *A. thaliana*, using Gene Structure Display Server 2.0 software for better visualization (Figure 4). Most of the *CrRLK1L* genes in *P. vulgaris* have no introns (26 genes), though four of them have one intron (*PvCRV2*, *PvFER2*, *PvHERK1C*, and *PvTHE1*) and three have two introns (*PvCRV1*, *PvHERK1B*, and *PvTHE2*) (Figure 4A). In *L. japonicus*, half of the *CrRLK1L* genes have no introns, two have one intron (*LjCRV2* and *LjMEDOS3*), three have two introns (*LjFER*, *LjHERK2B*, and *LjMEDOS4*), and one has seven introns (*LjMEDOS1*) (Figure 4B). More than half of the *G. max* *CrRLK1L* genes have no introns, and 17 genes have one to seven introns (*GmHERK1A*, *GmHERK1B*, *GmHERK1C*, *GmHERK2*, *GmMEDOS1B*, *GmMEDOS2C*, *GmMEDOS3A*, *GmMEDOS3B*, *GmMEDOS3D*, *GmMEDOS4C*, *GmMEDOS4I*, *GmMEDOS4J*, *GmMEDOS5A*, *GmTHE1*, *GmTHE2*, *GmCAD1*, and *GmCAD3*) (Figure 4C). In *M. truncatula*, more than half of the genes (19 genes) have no introns, and 16 of them possess one to three introns (*MtANX1*, *MtCAP*, *MtFER1*, *MtHERK1A*, *MtHERK1B*, *MtMEDOS1A*, *MtMEDOS3C*, *MtMEDOS3D*, *MtMEDOS3G*, *MtMEDOS3I*, *MtMEDOS3L*, *MtMEDOS4B*, *MtTHE1*, *MtTHE2*, *MtCAD1*, *MtCAD2*, and *MtCAD3*) (Figure 4D). Most of the 17 *A. thaliana* genes have no introns, with only four of them having one intron (*AtANX1*, *AtANX2*, *AtFER*, and *AtHERK1*) (Figure 4E). Although some intron conservation was detected among the five plants

analyzed, in *FER*, *HERK*, and *MEDOS* genes, many of the legume *CrRLK1L* genes have more introns than the corresponding *A. thaliana* genes.

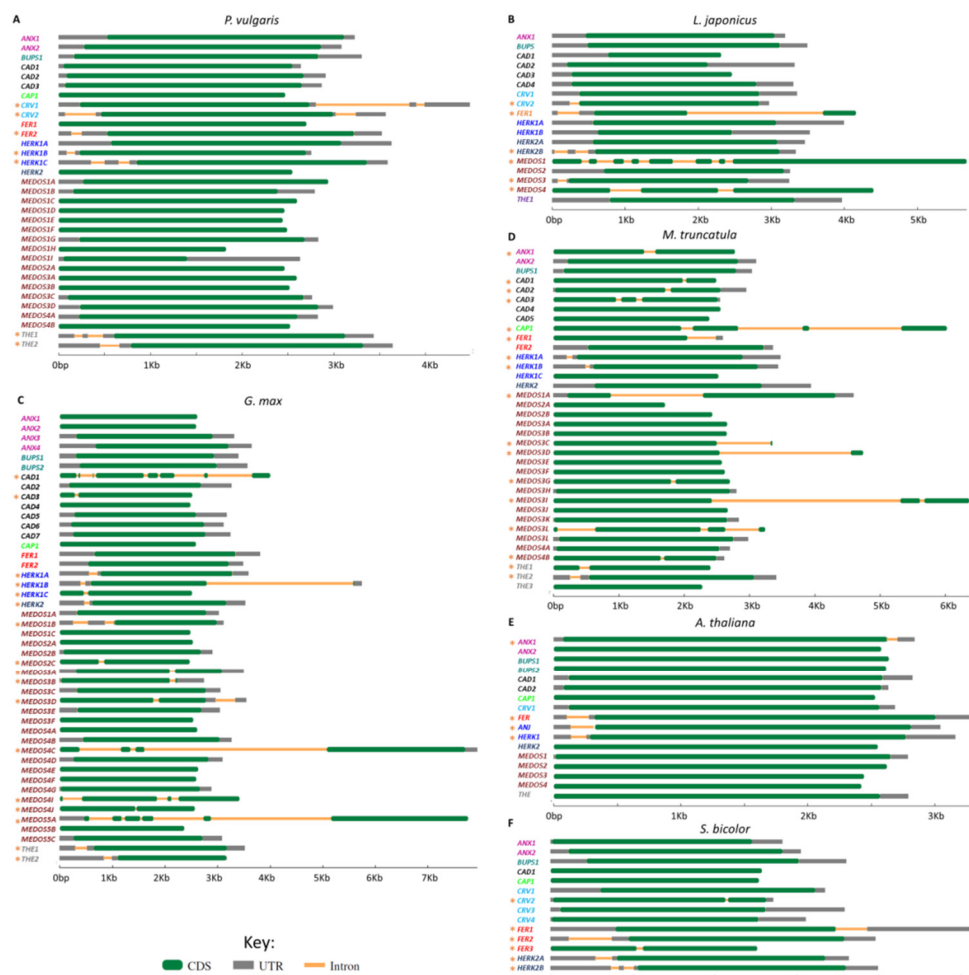


Figure 4. Gene structure of the *CrRLK1L* subfamily genes in four legumes, *A. thaliana*, and *S. bicolor*. The exon–intron structures of all *CrRLK1L* genes from (A) *P. vulgaris*, (B) *L. japonicus*, (C) *G. max*, (D) *M. truncatula*, (E) *A. thaliana*, and (F) *S. bicolor* were analyzed using the Gene Structure Display Server database. Exons (CDS), introns, and untranslated regions (UTRs) are represented according to the key. Gene names are highlighted in colors as follows: ANX in purple, BUPS in green, CAD in black, CAP in lime, CRV in blue, FER in red, HERK1 (ANJ) in blue, HERK2 in dark blue, MEDOS in brown, and THE in gray. Orange asterisks to the left of the names indicate genes with introns.

3.6. Analysis of the Expression Patterns of the *CrRLK1L* Genes in Legumes, *A. thaliana*, and *P. patens*

To evaluate the expression patterns of the *CrRLK1L*s genes between different organs in four legumes and compare them with those in *A. thaliana* and *P. patens*, we retrieved and compared the expression data for *P. vulgaris* *CrRLK1L*s from the common bean gene expression atlas (PvGEA) [60], for *L. japonicus* from the Lotus Base [49], for *M. truncatula* from MtGEA [61,62], and for *G. max*, *A. thaliana*, and *P. patens* from the BAR resource [61,63–65]. Expression data are represented as heat maps for each species (Figure 5).

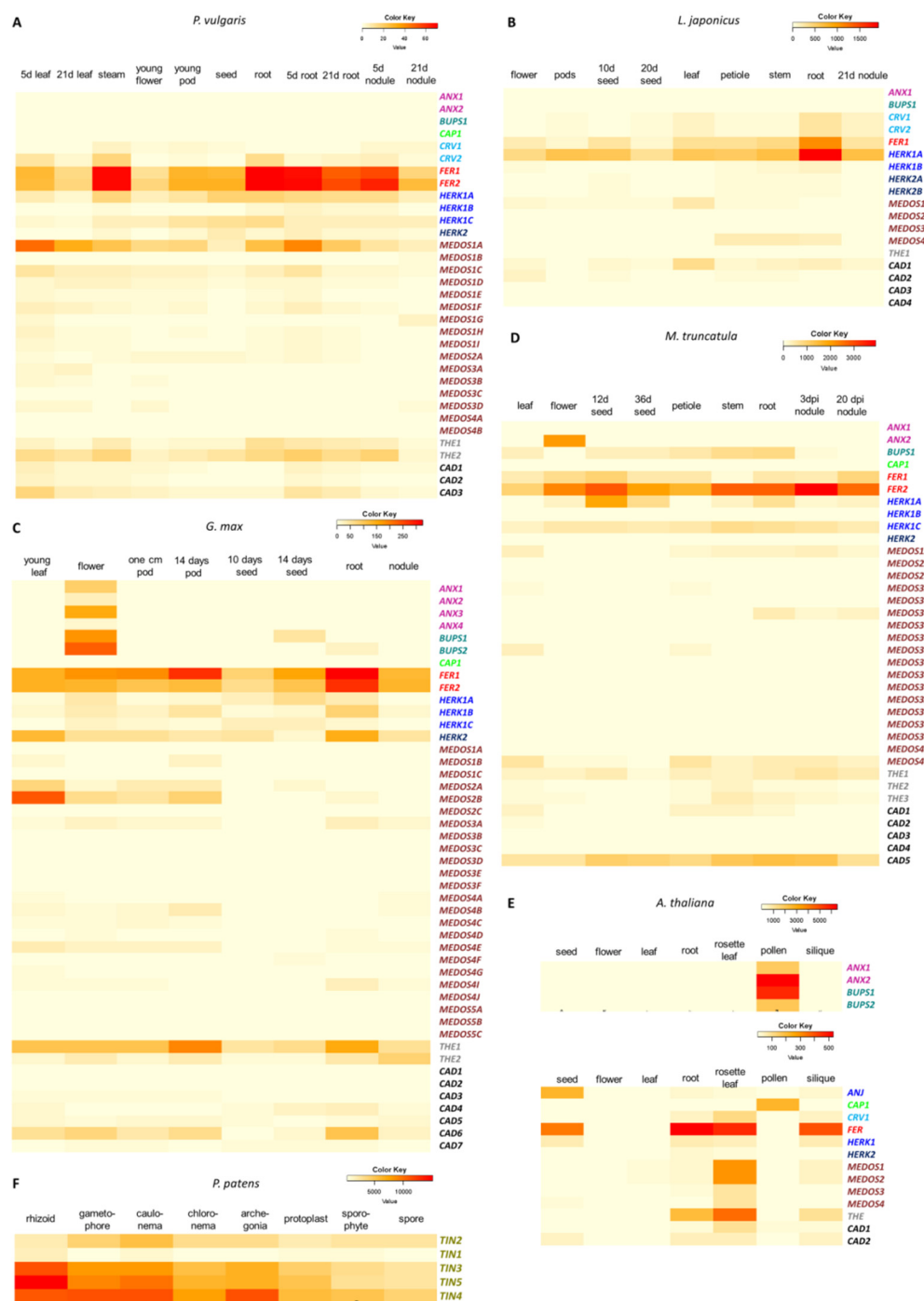


Figure 5. Gene expression profiles of *CrRLK1L*s in four legumes, *A. thaliana*, and *P. patens*. Heat map of expression profiles of *CrRLK1L* in (A) *P. vulgaris*, (B) *L. japonicus*, (C) *G. max*, (D) *M. truncatula*, (E) *A. thaliana*, and (F) *P. patens*. Transcriptome data were extracted from the PvGEA, LotusBASE, and BAR databases. RPKM values are represented as color key codes above each heat map. Gene names are indicated as in the following color key; ANX in purple, BUPS in green, CAP in lime, CRV in blue, FER in red, HERK1 (ANJ) in blue, HERK2 in dark blue, MEDOS in brown, THE in gray, and CAD in black.

As described in the following sections, it was observed that most of the genes showed similar expression patterns in the four legumes examined (*P. vulgaris*, *L. japonicus*, *G. max*, and *M. truncatula*) and in *A. thaliana* (*FER*, *ANX*, *BUPS*, *CAP*, *HERK*, *THE*, *MEDOS*, and *CRV*). Moreover, two genes differed in their expression patterns two to ten-fold in some tissues of the four legumes compared to their expression in *A. thaliana* (*CAD* and *MEDOS*). In addition, some *CrRLK1L* genes are expressed in legume nodules. Every *P. patens CrRLK1L* gene is expressed in all tissues analyzed; however, the levels of accumulation varied among the different tissues.

3.6.1. *FER* Genes Are Broadly Expressed in All Tissues in Four Different Plant Species

FER is the most studied gene of the *CrRLK1L* subfamily in *A. thaliana*, and it has key roles in diverse plant processes [3–11,15–29,68–70]. *A. thaliana* has only one *FER* gene, which is expressed in almost every tissue, with transcript levels being especially high in roots and rosette leaves (Figure 5E). The *FER* genes in *P. vulgaris*, *L. japonicus*, and *G. max* are expressed at high levels in almost every tissue, similar to the expression pattern of *AtFER* in *A. thaliana*. The two *FER* genes identified in *P. vulgaris* (*PvFER1* and *PvFER2*) are mainly expressed in roots and stems (Figure 5A). The single *FER* gene in *L. japonicus* (*LjFER*) shows expression in all tissues analyzed, with the highest levels in roots and stems (Figure 5B). *G. max* has two *FER* genes (*GmFER1* and *GmFER2*), which are both expressed at high levels, mainly in roots and pods (Figure 5C). In *M. truncatula*, there are two *FER* genes, both widely expressed in the tissues analyzed. *MtFER1* is mainly expressed in nodules, and seeds, while *MtFER2* shows the highest expression levels in in nodule, root, and stem (Figure 5D). These expression patterns suggest a presumed conservation of *FER* gene function between *A. thaliana* and the legumes.

3.6.2. *ANX*, *BUPS*, and *CAP* Genes Are Expressed Only in *A. thaliana* Pollen Tubes and *G. max* Flowers

Five *CrRLK1L* genes in *A. thaliana* have been reported to be essential for pollen tube growth, *AtANX1*, *AtANX2*, *AtBUPS1*, *AtBUPS2*, and *AtCAP* [30–32,36]. The expression of these five genes in *A. thaliana* is limited to pollen tubes and shows the highest accumulation levels of all *CrRLK1L* genes (Figure 5E). Four of these genes (*PvANX1*, *PvANX2*, *PvBUPS*, and *PvCAP*) are present in *P. vulgaris*, but no expression was detected according to PvGEA (Figure 5A). In *L. japonicus*, there are only two of these genes, *LjANX* and *LjCAP*, and no expression was detected in any of the tissues evaluated (Figure 5B). By contrast, in *G. max*, there are four *GmANX* and two *GmBUPS* genes, all of which exhibit expression in flower (Figure 5C); *GmCAP* show no expression in any of the tissues analyzed (Figure 5C). *M. truncatula* have two *ANX*, one *CAP*, and one *BUPS* gene. *MtANX1* and *MtCAP1* show no expression in any tissue examined, *MtANX2* are expressed exclusively in flowers, and *MtBUPS1* is expressed in several tissues (Figure 5D). Since the PvGEA and Lotus Base databases do not include expression data for pollen tubes, the expression pattern observed for *ANX*, *BUPS*, and *CAP* genes in *P. vulgaris* and *L. japonicum* probably resembles that of *A. thaliana*, while in *G. max* and *M. truncatula* *ANX* and *BUPS* genes probably expanded their expression to other tissues beyond pollen.

3.6.3. *HERK* and *THE* Genes Are Expressed in Roots, Leaves, and Pods/Siliques

HERK and *THE* genes have been reported to regulate cell wall homeostasis in *A. thaliana* root and leaf, and *HERK* genes have also been reported to be essential for fertilization [4,33]. *AtHERK1*, *AtHERK2*, and *AtTHE1* show highest levels of expression in rosette leaves, roots, and siliques (Figure 5E). In common bean, *PvHERK1A* and *PvHERK1C* are expressed in almost every tissue analyzed, mainly in roots and pods (Figure 5A), *PvHERK1B* and *PvHERK2* show low expression, and *PvTHE1* and *PvTHE2* have the highest levels of expression in leaves, roots, and stems (Figure 5A). In *L. japonicus*, the four *HERK* genes show their highest expression in roots and nodules, followed by stem, leaves, and pods; *LjTHE* is not expressed

(Figure 5B). In *G. max*, *GmHERK1A* and *GmHERK1C* are most strongly expressed in seeds, *GmHERK1B* in roots, and *GmHERK2* shows the maximum expression in roots and leaves, and the two *GmTHE* genes are primarily expressed in pods and roots (Figure 5C). In *M. truncatula*, meanwhile *MtHERK1A*, *MtHERK1C*, and *MtTHE1* are expressed in almost every tissues analyzed, mainly in roots and seeds, *MtHERK1B* and *MtHERK2* are not expressed at all (Figure 5D). These data indicate that *HERK* and *THE* genes, which are expressed mostly in roots, leaves, and siliques in *A. thaliana*, have similar expression patterns in legumes, although the latter have many more copies of these genes.

3.6.4. MEDOS Genes Are Mostly Expressed in Leaves

It has recently been reported that *MEDOS* genes are important for regulating growth in the presence of metal ions in *A. thaliana* [38]. The four *AtMEDOS* genes are expressed mainly in rosette leaves (Figure 5E). In common bean, we detected 15 *PvMEDOS* genes. *PvMEDOS1A* is the third most expressed gene of all the *CrRLK1L* genes in this legume. *PvMEDOS1B*, *PvMEDOS3C*, *PvMEDOS4A*, and *PvMEDOS4B* show no expression in the analyzed tissues, and the other 10 *PvMEDOS* genes are expressed at low levels (Figure 5A). Despite these differences in transcript levels, *PvMEDOS* genes are mainly expressed in leaves and roots (Figure 5A). *L. japonicus* has four *LjMEDOS* genes. *LjMEDOS1* is mainly expressed in leaves, *LjMEDOS4* is mainly expressed in stem and petiole, and *LjMEDOS2-3* genes show low expression in all tissues tested (Figure 5B). In *G. max*, there are 24 *GmMEDOS* genes, most of which show low or no expression in the tissues evaluated (14 of 24 genes). Nine *GmMEDOS* genes are mostly expressed in leaves and pods, whereas *GmMEDOS4I* is most strongly expressed in roots (Figure 5C). Eighteen *MEDOS* genes were founded in *M. truncatula*, of them only five (*MtMEDOS1A*, *MtMEDOS3A*, *MtMEDOS3C*, *MtMEDOS3F*, and *MtMEDOS4B*) show low expression levels, mainly in leaves, petioles, and roots (Figure 5D). These data suggest some conservation in the expression of *MEDOS* genes in leaves of legumes and *A. thaliana*. Nonetheless, some of these genes probably have additional functions, in legumes, which could be related to their expression in other tissues (Figure 5).

3.6.5. CRV Gene Expression Is Observed in Roots and Leaves, but Is Absent in *G. max*

CURVY (*CRV*) is a *CrRLK1L* receptor that is important for the development of leaves and seeds, as well as for the transition from vegetative to reproductive growth [37]. In *A. thaliana*, *AtCRV* is mainly expressed in rosette leaves, roots, and siliques (Figure 5E). In *P. vulgaris* there are two *CRV* genes, *PvCRV1* and *PvCRV2*, both of which are expressed at very low levels, mostly in stems, leaves, and roots (Figure 5A). *L. japonicus* also has two *LjCRV* genes, both with low expression, mainly in roots, leaves, and nodules (Figure 5B). No *CRV* genes were identified in the *G. max* and *M. truncatula* genomes, suggesting a possible loss of these genes during their evolution (Figure 5C,D). These observations indicate that expression of *CRV* genes in roots and leaves is conserved; however, the decrease in *CRV* expression in legumes and the loss of this gene in *G. max* and *M. truncatula* suggest a gradual loss of function.

3.6.6. The Tissue Specificity of CAD Gene Expression Is Broader in Legumes Than in *A. thaliana*

The *A. thaliana* *AtCAD1* and *AtCAD2* genes, which have not yet been characterized, are mainly expressed in rosette leaves and roots (Figure 5E). The three *PvCAD* genes identified in common bean show low expression, mostly in leaves and inoculated roots (Figure 5A). *L. japonicus* has four *LjCAD* genes; *LjCAD1* exhibits high expression in roots, leaves, and nodules, whereas the other three show low expression (Figure 5B). Seven *GmCAD* genes were detected in *G. max*: *GmCAD1* and *GmCAD2* show low expression; *GmCAD3* and *GmCAD5* are mainly expressed in pods, leaves, and flowers; *GmCAD4* and *GmCAD7* are predominantly expressed in seeds and roots; and *GmCAD6* is expressed in roots, flowers, pods, and leaves (Figure 5C). Three of the five *CAD* genes in *M. truncatula* show no expression, *MtCAD1*

is expressed at low levels in leaves, petiole and stems, and *MtCAD5* is expressed in most of the tissues, but mainly in stem, roots and nodules (Figure 5D). Thus, *CAD* genes are expressed in a wider range of tissues in legumes than in *A. thaliana*.

3.6.7. *P. patens* *CrRLK1L* Genes Are Widely but Differentially Expressed in All Tissues Tested

Our phylogenetic analysis showed that the five *CrRLK1L* genes in *P. patens* (*PpTIN1-5*) are clustered in a single distinct clade (Figure 1, Figure S1). All five genes are abundantly expressed in every tissue analyzed (Figure 5F). *PpTIN1*, *PpTIN2*, and *PpTIN4* are mainly expressed in the rhizoid (an organ functionally related to the roots of land plants) and the caulonema (an organ necessary for colonization and nutrient acquisition). *PpTIN3* is mostly expressed in the archegonia (the female reproductive organs in the moss) and in the caulonema. Maximum levels of *PpTIN5* accumulation are observed in the caulonema and the gametophore (the tissue carrying the sex organs in moss) (Figure 5F). These variations in the expression patterns of the *P. patens* *CrRLK1L* genes suggest a certain amount of functional specialization of these genes in this moss, since the five genes probably originated from duplication of a single *CrRLK1L* gene.

3.6.8. Certain *CrRLK1L* Genes Are Differentially Expressed during Nodulation

We observed that almost none of the legume *CrRLK1L* genes are expressed specifically in symbiotic organs; however, some of them are highly expressed in these symbiotic organs (Figure 5). In *P. vulgaris* there are at least nine of these genes (*PvCRV1*, *PvFER1*, *PvFER2*, *PvHERK1A*, *PvHERK1C*, *PvMEDOS1A*, *PvMEDOS1C*, *PvCAD3*, and *PvTHE2*) (Figure 5A), nine genes in *L. japonicus* (*LjCRV1*, *LjCRV2*, *LjFER1*, *LjHERK1A*, *LjHERK2A*, *LjHERK2B*, *LjMEDOS4A*, *LjCAD1*, and *LjCAD3*) (Figure 5B), 10 genes in *G. max* (*GmFER1*, *GmFER2*, *GmHERK1C*, *GmHERK2*, *GmMEDOS3A*, *GmMEDOS4*, *GmCAD1*, *GmCAD3*, *GmTHE1*, and *GmTHE2*) and nine genes in *M. truncatula* (*MtFER1*, *MtFER2*, *MtHERK1A*, *MtHERK1C*, *MtMED1A*, *MtMED3C*, *MtMEDOS4B*, *MtCAD5*, and *MtTHE1*) (Figure 5C–D). We noticed that several of the 37 genes expressed in nodules are shared among the four legumes (Figure S5). *FER1* is expressed in nodules in all four legumes, whereas five genes were shared between three different legumes: *CAD3* is shared between *P. vulgaris*, *L. japonicus*, and *G. max*; *HERK1A* is shared between *P. vulgaris*, *L. japonicus*, and *M. truncatula*; while *L. japonicus*, *G. max*, and *M. truncatula* share *MEDOS4*; and *FER2* and *HERK1C* are shared between *P. vulgaris*, *G. max*, and *M. truncatula*. Moreover, seven genes were found in nodules in two legume pairs; *P. vulgaris* and *L. japonicus* nodules express *CRV1*, *THE2* is expressed in nodules of *P. vulgaris* and *G. max*, and *MEDOS1A* in *P. vulgaris* and *M. truncatula*. Furthermore, *HERK2A* and *CAD1* are expressed in *G. max* and *L. japonicus*, while *MEDOS3* and *THE1* are shared between *G. max* and *M. truncatula*. This comparative analysis also revealed four nodule-expressed genes that were exclusive to one legume (Figure S5).

These data together indicate that along with the highly conserved expression profiles of *CrRLK1L* genes in legumes, some of them are differentially expressed in nodules, suggesting a possible role of these genes in the nodulation process.

3.7. Expression of *CrRLK1L* Genes in *P. vulgaris* Nodules

To validate the expression profile of some *CrRLK1L* genes in nodules that we observed in the PvGEA data [60], as well as to describe their expression patterns during different stages of nodulation, we selected eight *P. vulgaris* *CrRLK1L* genes for further investigation: *PvFER1*, *PvFER2*, *PvHERK1A*, *PvHERK1C*, *PvMEDOS1A*, *PvMEDOS1C*, *PvTHE2*, and *PvCAD3*. Expression of these genes was measured at four stages of the *P. vulgaris*-*R. tropici* symbiosis: 5, 7, 14, and 21 days post-inoculation (dpi) of wild-type roots. The eight genes were differentially expressed at the different stages of nodulation, corroborating their presumed role during nodulation in common beans.

These eight genes displayed four different expression profiles. Three genes were suppressed in at least one of the nodulation steps analyzed (blue box, Figure 6). *PvFER1* and *PvCAD3* showed reduced transcript accumulation in inoculated roots at 7, 14, and 21 dpi compared to uninoculated roots, whereas no differences were observed at 5 dpi (Figure 6A,B). *PvMEDOS1A* was downregulated at 7 and 21 dpi in inoculated roots but was expressed at similar levels regardless of inoculation at 5 and 14 dpi (Figure 6C). Three genes were upregulated in the early stages (5 or 7 dpi) but then suppressed in the later stages (14 or 21 dpi) (purple box, Figure 6); *PvFER2* and *PvHERK1C* were upregulated in inoculated roots at 5 dpi, and *PvHERK1A* was upregulated at 5 and 7 dpi. At 21 dpi, however, *PvFER2*, *PvHERK1C*, and *PvHERK1A* were downregulated in inoculated roots compared to the controls, as was *PvHERK1C* at 14 dpi (Figure 6D–F). A third expression pattern was displayed by *PvTHE2*; transcripts of this gene showed increased accumulation in inoculated roots at 7 and 14 dpi relative to the controls but at 5 and 21 dpi, levels of transcript accumulation were similar to the controls (green box, Figure 6G). Finally, *PvMEDOS1C* showed fine-tuned changes in expression; relative to the controls, transcript accumulation for this gene was decreased at 5 and 21 dpi, increased at 7 dpi, and unchanged at 14 dpi (brown box, Figure 6H).

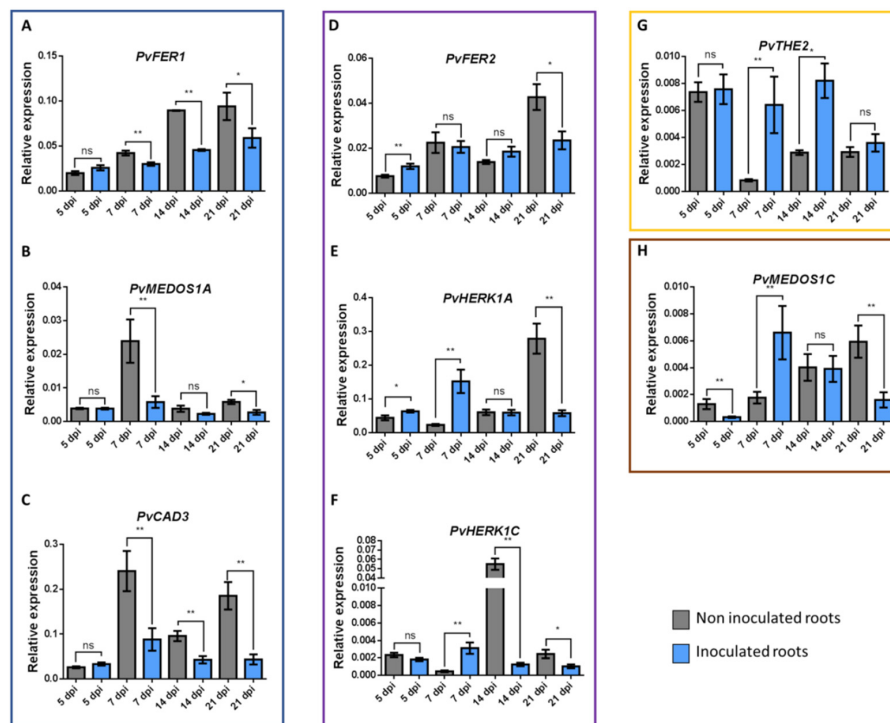


Figure 6. RT-qPCR expression analysis of eight *P. vulgaris* *CrRLK1L* genes. Relative expression profiles of eight *CrRLK1L* genes from *P. vulgaris* roots inoculated or not with *R. tropici*. Genes were classified into four groups according to their expression; the blue box indicates downregulated genes at the early and late time points: *PvFER1* (A), *PvMEDOS1A* (B), and *PvCAD3* (C); the purple box shows genes whose expression is upregulated at the early time points assessed, but downregulated later on: *PvFER2* (D), *PvHERK1A* (E), and *PvHERK1C* (F); the yellow box indicates the expression of *PvTHE2* (G), which was upregulated at the early and late time points; and the brown box displays *PvMEDOS1C* (H), showing variable expression at the different time points evaluated. The transcript accumulation of the selected genes was assessed by RT-qPCR and normalized according to *elongation factor 1α* (*ef1α*) gene expression. Blue bars represent inoculated roots, whereas gray bars indicate the expression levels in non-inoculated roots. The error bars represent standard deviation of the mean ($n = 6$). A Student's *t*-test was performed to evaluate significant differences, * represents $p \leq 0.05$, ** represent $p \leq 0.01$, ns represents non-significant difference.

These data indicate that the eight genes analyzed here are indeed differentially expressed in common bean roots at different stages of the nodulation process and probably perform different functions throughout the symbiotic process.

4. Discussion

4.1. Structural Features of CrRLK1L Genes

The RLK subfamily CrRLK1L has emerged as an important signaling component of numerous biological processes, including development, immune responses, and fertilization, among others. Previous studies have analyzed the phylogeny of *CrRLK1L* genes in *A. thaliana*, rice (*O. sativa*), cotton (*Gossypium hirsutum*), and pear (*Pyrus bretschneideri*), as well as their expression under different conditions [68–70]. However, it is important to extend these studies to other agro-ecologically important crops, such as legumes, which have the ability to fix nitrogen in association with the soil bacteria rhizobia. Furthermore, a comprehensive phylogenetic study of the CrRLK1L subfamily in a larger number of plant species will yield new information about the functions of these proteins and their evolutionary paths since their appearance in the plant kingdom. The usefulness of this bioinformatic approach is evident in the current study, in which we were able to analyze the CrRLK1L subfamily in model legumes and common bean, using different “*in silico*” approaches, and thereby elucidate its possible functions in the legume-rhizobia mutualistic interaction.

We identified 1050 CrRLK1L proteins from the 57 embryophytes included in this analysis, which fell into 11 phylogenetically distinct clades (Figure 1A, Figure S1, Table S2). Chlorophytes lack this plant-specific RLK subfamily, indicating that it arose during the transition from chlorophytes to embryophytes, which probably occurred about 500 million years ago (mya) [71]. A feature that differentiates embryophytes from other plants is their sexual reproduction [71,72] and, since some CrRLK1Ls are key regulators of fertilization [4,16,30,31,33,36], this may link the emergence of the CrRLK1L subfamily with the advent of embryophytes. We found that bryophyte CrRLK1Ls cluster together in a unique clade (TINIA), whereas the land plant proteins are distributed among the remaining ten clades (Figure 1A, Figure S1, Table S2). The number of CrRLK1Ls in these mosses varies from one to seven, revealing the first duplication events of the *CrRLK1L* lineage. Monocots and eudicots diverged around 150 mya, and have evolved along different evolutionary paths [73]. The subsequent eudicot radiation, dated around 100 mya, has been associated with polyploidization events [74]. Interestingly, there are more CrRLK1L proteins in eudicots than in monocots, demonstrating the different evolutionary fates of the genes of this subfamily in monocots and eudicots. Some eudicots have particularly large numbers of these proteins compared to other eudicots (Figure 1A–B, Table S1). This increase in the number of CrRLK1Ls could be associated with the appearance of the MEDOS clade (present in all eudicots but in only a few monocots) and with subsequent expansion of the MEDOS clade in eudicots.

Our phylogenetic analysis of CrRLK1Ls from four legumes, *A. thaliana*, and *P. patens*, revealed 155 genes distributed in 11 clades (Figure S2A). As expected, the MEDOS clade had the most members. We observed that the proteins in this clade are clustered on one chromosome in *A. thaliana* and *L. japonicus*, while in *P. vulgaris*, *G. max*, and *M. truncatula*, these genes form two clusters (Figure S3). Some reports indicate that in plants the expansion of gene subfamilies mainly occurred through dispersed, tandem, and whole-genome duplications [75–80]. In pear, most *CrRLK1L* genes arose by whole-genome duplication and some by dispersed gene duplications [70]. The tandem duplications we observed suggest that, in the analyzed legumes and in plants with a high number of *CrRLK1L* genes, the MEDOS genes arose from tandem duplications, whereas the other *CrRLK1L* genes probably arose from whole-genome or segmental duplications.

The exon-intron structure of genes has been associated with gene function, and it affects RNA splicing, RNA stability, and chromatin organization [81–83]. Exon-intron patterns have been used to reveal time evolution, constant variation, and their co-variations [84]. Our comparative analysis of the exon–intron distribution in *CrRLK1L* genes revealed that, compared to *A. thaliana*, legumes have more *CrRLK1L* genes with introns and more introns in each gene (Figure 4). Nevertheless, the expression patterns of the legume *CrRLK1L* genes were similar to those of the corresponding orthologs in *A. thaliana* (Figure 5). In pear, it has been proposed that the *CrRLK1L* genes have lost introns, but their expression patterns are similar to those of *A. thaliana* and rice genes [85]. Our observations suggest that the increase in the number of introns is associated with duplication and evolutionary events, but these have little or no effect on gene function and expression.

Analysis of the synteny of the *CrRLK1L* genes revealed homology between some gene pairs in the plants analyzed. In *P. vulgaris*, *L. japonicus*, and *A. thaliana*, four, one, and two syntenic gene pairs were identified, respectively; each gene observed was syntenic with only one additional gene (Figure S4). By contrast, 21 syntenic gene pairs were identified in *G. max*, and some genes have synteny with more than one other gene (Figure S4). The higher number of syntenic genes in *G. max* is probably because of the polyploidization event that occurred in this legume [86]. Compared to the number of syntenic *CrRLK1L* genes we observed between *P. vulgaris* and *G. max*, there were fewer between either of these species and *L. japonicus*, and even less between *P. vulgaris* and *A. thaliana* (Figure 3). There is a clear correlation between the degree of synteny and the time of divergence between species [74,76]. The degree of synteny also depends on the evolution of the genome; in angiosperms, whole-genome duplication and subsequent gene loss have driven plant evolution and have also reduced collinearity across species [77,87]. Our data are consistent with an early divergence between *P. vulgaris* and *G. max*, compared to *L. japonicus*, and an even longer divergence time between *P. vulgaris* and *A. thaliana*.

The characteristics of a protein are important for its activity and correspond with taxonomy, environmental adaptation, subcellular localization, and genome size [85,88]. From this perspective, the contrast between the characteristics of CrRLK1Ls from legumes versus *A. thaliana* denotes greater variability in the legume sequences and correlates with larger genomes (Table 2). A protein's iP reflects its amino acid composition and conformation and determines its activity [89]; the wider iP ranges and longer sequences of the legume CrRLK1Ls could reflect specialization of some of these proteins for different tissues or processes, possibly giving these plants better adaptability to environmental changes. In the five plant species studied here, we observed a high conservation of overrepresented motifs in all of the CrRLK1Ls (Figure 2). The conservation of these motifs, which are located in the malectin and kinase domains characteristic of this subfamily, indicates their importance for protein activity.

4.2. Differences and Similarities in the Expression of *CrRLK1L* Genes in Legumes and in *A. thaliana*

Previous studies in *A. thaliana* have reported that *CrRLK1L* genes participate in a variety of processes, such as development, cell communication, and plant-microbe interactions (Table 1), and that the functions of these genes correspond with their expression profiles (Figure 5E). A previous study comparing *CrRLK1L* gene expression in pear and *A. thaliana* reported that the expression profiles of some genes are conserved between these species; however, the expression of many other genes was lost or altered in pear compared to *A. thaliana* [70]. We performed a comparative in silico analysis of *CrRLK1L* gene expression profiles in four legumes and *A. thaliana* and observed that the expression patterns of most of the genes are conserved. *FER*, *HERK*, *THE*, *CRV*, and *MEDOS* showed similar expression profiles in the five species examined; these genes are expressed in almost all tissues. In *A. thaliana*, *ANX*, *BUPS*, and *CAP* are pollen-specific genes. These genes are not expressed at detectable levels in *P. vulgaris* or *L. japonicus*, at least in the tissues included in the databases (Figure 5). However, since there is no data available for expression of these

genes in pollen or pollen tubes, we propose that these genes could also be pollen-specific in these legumes, as they are in *A. thaliana*. The expression profiles of some of the legume *CAD* genes differed by two to ten-fold from those of the *AtCAD* genes in some tissues, suggesting additional functions for these genes in legumes.

Legumes are characterized by the ability to form nodules that house endosymbiotic rhizobia. This relationship generates a driving force between the two symbionts that leads them to co-evolve [90,91]. It has been reported that plant lipochitoooligosaccharide receptors acquired symbiotic functions before gene duplication [92]. In the four legumes analyzed here, some *CrRLK1L* genes showed transcript accumulation in nodules, suggesting that these genes have been recruited to the symbiotic process, in addition to any other roles they may have. We identified nine genes that were expressed in nodules in *P. vulgaris*, nine in *L. japonicus*, ten in *G. max*, and nine in *M. truncatula*. Among those genes, *FER1* is expressed in nodules of all four legumes, (Figure 5, Figure S5), five other nodule-expressed genes were shared between three of the four legumes, seven shared by different pairs of legumes, and four genes were expressed in nodules of only one of the legumes (Figure 5, Figure S5). These data may suggest that some *CrRLK1L* genes participate in the symbiotic process. Nonetheless, further functional analyses are needed to test this hypothesis.

4.3. Putative Roles of *CrRLK1L* Genes during Nodulation

We examined the expression profiles of eight *CrRLK1L* genes in *P. vulgaris* roots inoculated with *R. tropici* and found that these genes were differentially expressed at different stages of nodule development. Figure 7 provides a schematic summary of the nodulation process in *P. vulgaris* and the steps in which the eight genes presumably participate. At 5 dpi, the nodule primordia begin to emerge from the root epidermis, and the infection thread, filled with bacteria, penetrates the outer cortex of the root and branches [93–97]; at this point, *PvMEDOS1C* was downregulated in the inoculated roots relative to the control, whereas *PvFER2* and *PvHERK1A* were upregulated (Figure 7). By 7 dpi, many nodules have already emerged from the root epidermis, and some nodule primordia cells contain bacteria that have been released from the infection threads [95–99]; at this time, four genes (*PvHERK1A*, *PvHERK1C*, *PvTHE2*, and *PvMEDOS1C*) showed high expression, whereas three others (*PvFER1*, *PvMEDOS1A*, and *PvCAD3*) exhibited low expression in inoculated roots relative to the control (Figure 7). Some *CrRLK1Ls* have been reported to be important regulators of cell expansion, cell wall maintenance, and membrane integrity during cell growth [3,32,100–102]. The *P. vulgaris* *CrRLK1L* genes that are induced at 5 and 7 dpi could be supporting similar functions, since at these nodulation stages, there are high rates of cell division and expansion [97,103–107]. Likewise, internalization of the bacteria depends on growth and branching of the infection thread through the root cortex and subsequent release of the bacteria into the cells of the nodule primordia [97–99]. Some *CrRLK1L* genes have been reported to be regulators of immune responses [9,11,12,20,108], indicating that downregulation of some *CrRLK1L* genes at this stage of nodulation might inhibit pathogen responses during infection.

At 14 dpi, the bacteria within the infected cells differentiate into bacteroids and most of the nodules are matured, initiating nitrogen fixation [93,95,96,109–111]. At this stage, three genes, *PvFER1*, *PvCAD3*, and *PvHERK1C*, were downregulated and *PvTHE2* was upregulated in inoculated roots relative to non-inoculated ones (Figure 7). The downregulated genes could be associated with avoidance of immune responses, as earlier in the nodulation process. In addition, the downregulated genes could be involved in regulating nitrogen flow, considering that, in *A. thaliana*, *FER* has been reported to be a growth regulator that responds to the C/N ratio [14]. Downregulation of some *CrRLK1Ls* may be necessary to promote nitrogen fixation. The upregulation of *PvTHE2* suggests that this gene may be associated with other functions during this stage, such as nodule development. At 21 dpi, common bean nodules are fully developed and display high rates of nitrogen fixation [93,96,110,111]. *PvFER1*, *PvFER2*, *PvHERK1A*, *PvHERK1C*,

PvMEDOS1A, *PvMEDOS1C*, and *PvCAD3* were downregulated at 21 dpi (Figure 7). The downregulation of most of the *CrRLK1L* genes could be related to the end of nodule development and to deactivation of immune responses to maintain symbiosis and nitrogen fixation at the highest levels.

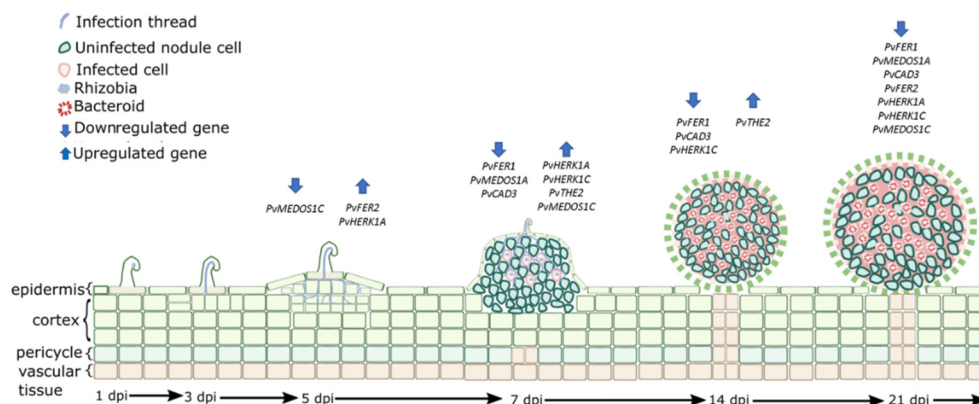


Figure 7. Putative roles of the eight *CrRLK1L* genes evaluated during nodule organogenesis in common bean. Graphic representation of the proposed roles of the eight *CrRLK1L* genes of common bean, based on the RT-qPCR data obtained in this study. Different stages corresponding to days post-inoculation (dpi) are observed. At the early stages, the bacteria and the root establish a molecular dialogue, which promotes curling of the root hair where the bacteria are enclosed in an infection chamber (not shown). One to three days before the bacteria become trapped, an infection thread (IT) is formed, which advances through the infected root hair cell, reaching the outer cortex of the root. Concurrently, cortex cells de-differentiate and divide. By 5 dpi, dividing cells in the outer cortex generate a nodule primordium, whereas the IT branches toward the primordium. By 1 nitrogen fixation rates.

ROS are important signaling molecules that participate in nodule organogenesis processes associated with *CrRLK1L*s. In plant cells, ROS are mainly produced through the activity of respiratory burst oxidase homologs (RBOHs in plants), which are called NADPH oxidases in mammals [112]. RBOH-dependent ROS production has been described as a conserved mechanism in *CrRLK1L* activity; FER, ANX1, and ANX2 promote phosphorylation, and thereby activation, of RBOH, inducing ROS-mediated polar growth in pollen tubes and root hairs in *A. thaliana* [10,32]. In *P. vulgaris* and *M. truncatula*, ROS signaling is essential for initiating root hair cell responses to the presence of rhizobia. Impairment of ROS production through downregulation of *Rbohs* in these species inhibits the progression of infection thread growth in *P. vulgaris* (*PvRbohA* and *PvRbohB*) and swelling of root hair tips in *M. truncatula* (*MtRbohB* and *MtRbohE*) [113–115]. In addition, previous studies have revealed *Rboh* promoter activity associated with cell division in the cortex and vascular bundles of nodules, suggesting a possible role of these oxidases in nodule development [114–117]. Similarly, *Rboh* genes are differentially expressed during nodulation in *P. vulgaris*, *L. japonicus*, and *M. truncatula* [114,115,117,118], as we observed for eight nodule-expressed *CrRLK1L* genes in common bean. Altogether, these data suggest that the *CrRLK1L* genes may be participating in the nodulation process through regulation of ROS signaling at specific stages of nodule organogenesis.

Phytohormones also appear to have roles in nodulation. For instance, studies in several legumes have reported that abscisic acid (ABA) is a negative regulator of nodulation [119,120] but that it also has some positive effects on the growth and functioning of nodules [121,122]. In pea (*Pisum sativum*) and soybean (*G. max*), brassinosteroid (BR) inhibits nodulation in some studies [123,124], whereas in peanut (*Arachis hypogaea*) and *P. vulgaris*, some studies show positive effects of BR on nodulation [125,126]. Jasmonic acid (JA) has both positive and negative effects on nodulation, depending on the legume species and the stage of nodule development at which it is applied [119,127–129]. Ethylene has mainly been

associated with negative regulation of nodulation [130,131]. RALF peptide hormones have been reported to be negative regulators of infection and nodule organogenesis in *M. truncatula* [45]. Some CrRLK1Ls are known to be RALF receptors [8,39,40,101]. In *A. thaliana*, the expression of *FER*, *THE*, and *HERK* is induced by BRs [4] and *FER* is a hormone response modulator, fine-tuning ethylene and BR signaling during hypocotyl growth [5] and suppressing ABA and JA signaling [6,9]. In the current study, we found that two *FER* genes, two *HER* genes, and one *THE* gene were differentially expressed at different stages of nodulation in common bean. These results, along with the previously described roles for these genes in regulating hormone signaling, allow us to speculate that these genes may participate in nodulation through the regulation of hormone signaling at several stages of nodule organogenesis. Nonetheless, experimental evidence is needed to test this hypothesis.

In this work, we examined eight differentially expressed *CrRLK1L* genes at different stages of nodulation in common bean. Based on our results, we postulate that these proteins are regulators in this process. Forthcoming reverse genetics experiments in common bean will expand our knowledge of the particular roles of these *CrRLK1L* genes in nodulation. Our analysis of previously published transcriptomic data [60,61,63,64] demonstrated that related *CrRLK1L* genes are expressed in nodules of other legumes. Based on the phylogenetic, syntenic, and expression profiling analyses reported here, we predict that *CrRLK1L* subfamily homologs in other legumes may have a conserved role in nodulation, as these genes presumably do in common bean.

5. Conclusions

In this study, we identified 1050 CrRLK1L proteins in 57 plant species, clustered into 11 clades, one of them specific to moss and clubmoss proteins. This receptor subfamily probably appeared with the emergence of land plants, since no homologous proteins were detected in chlorophytes. In silico analysis in legumes and *A. thaliana* revealed that these receptors have expanded mostly by whole-genome and isolated duplication, and in the case of the *MEDOS* clade, by tandem duplication. Moreover, this analysis revealed high conservation of gene and protein structure and high similarities in expression profiles, suggesting analogous functions. Remarkably, RT-qPCR quantification of transcript levels in *P. vulgaris* roots inoculated with *R. tropici* revealed that some *CrRLK1L* genes could have different roles at different stages of the nodulation process. Considering the genomic similarities observed, we speculate that these roles in nodule organogenesis could be conserved in other legumes.

Supplementary Materials: The following data are available online at <http://www.mdpi.com/2073-4425/11/7/793/s1>, Figure S1: Phylogenetic relationship among 1050 CrRLK1L proteins. Figure S2: Unrooted approximately maximum-likelihood phylogenetic tree of the CrRLK1L subfamily proteins in various species. Figure S3: Chromosomal location of *CrRLK1L* genes in various species. Figure S4: Synteny of all *CrRLK1L* genes of *P. vulgaris*, *L. japonicus*, *G. max*, *M. truncatula*, and *A. thaliana* and of the complete genome of each species. Figure S5: Venn diagram of *CrRLK1L*s genes expressed in nodules of *P. vulgaris*, *L. japonicus*, *G. max*, and *M. truncatula*. Table S1: Oligonucleotides designed for gene-specific detection by RT-qPCR. Table S2: Number of *CrRLK1L* subfamily genes in 62 different species. Table S3. ID list of 1050 *CrRLK1L* genes present in 57 species by clade. Table S4: List of syntenic gene pairs founded by species and between species.

Author Contributions: Conceptualized the study, C.Q. and J.S.-M.; analyzed the data, J.S.-M. and C.F.-G.; collected the vegetal tissues and performed the extraction and purification of the RNA and the cDNA synthesis, N.N., R.P. and C.F.-G.; performed the RT-qPCR experiments, J.S.-M.; drafted, edited, and revised the manuscript, C.Q., J.S.-M. and C.F.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Universidad Nacional Autónoma de México (UNAM), under a DGAPA grant (IN201118) to C.Q., and by Consejo Nacional de Ciencia y Tecnología in Mexico (CONACyT), under Fronteras de la Ciencia (FC) grant: FC-2016/1503 to C.Q., with the doctoral fellowships to J.S.-M. (463767) and to R.P. (823333) and with a postdoctoral fellowship to C.F.-G. (268158).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shiu, S.-H.; Bleecker, A.B. Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10763–10768. [[CrossRef](#)]
2. Schulze-Muth, P.; Irmeler, S.; Schröder, G.; Schröder, J. Novel type of receptor-like protein kinase from a higher plant (*Catharanthus roseus*): cDNA, gene, intramolecular autophosphorylation, and identification of a threonine important for auto- and substrate phosphorylation. *J. Biol. Chem.* **1996**, *271*, 26684–26689. [[CrossRef](#)]
3. Boisson-Dernier, A.; Kessler, S.A.; Grossniklaus, U. The walls have ears: The role of plant CrRLK1Ls in sensing and transducing extracellular signals. *J. Exp. Bot.* **2011**, *62*, 1581–1591. [[CrossRef](#)] [[PubMed](#)]
4. Guo, H.; Li, L.; Ye, H.; Yu, X.; Algreen, A.; Yin, Y. Three related receptor-like kinases are required for optimal cell elongation in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7648–7653. [[CrossRef](#)] [[PubMed](#)]
5. DesLauriers, S.D.; Larsen, P.B. Feronia is a Key Modulator of Brassinosteroid and Ethylene Responsiveness in Arabidopsis Hypocotyls. *Mol. Plant* **2010**, *3*, 626–640. [[CrossRef](#)]
6. Yu, F.; Qian, L.; Nibau, C.; Duan, Q.; Kita, D.; Levasseur, K.; Li, X.; Lu, C.; Li, H.; Hou, C.; et al. FERONIA receptor kinase pathway suppresses abscisic acid signaling in Arabidopsis by activating ABI2 phosphatase. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14693–14698. [[CrossRef](#)] [[PubMed](#)]
7. Mao, D.; Yu, F.; Li, J.; Van De Poel, B.; Tan, D.; Li, J.; Liu, Y.; Li, X.; Dong, M.; Chen, L.; et al. FERONIA receptor kinase interacts with S-adenosylmethionine synthetase and suppresses S-adenosylmethionine production and ethylene biosynthesis in Arabidopsis. *Plant Cell Environ.* **2015**, *38*, 2566–2574. [[CrossRef](#)]
8. Liao, H.; Tang, R.; Zhang, X.; Luan, S.; Yu, F. FERONIA Receptor Kinase at the Crossroads of Hormone Signaling and Stress Responses. *Plant Cell Physiol.* **2017**, *58*, 1143–1150. [[CrossRef](#)]
9. Guo, H.; Nolan, T.M.; Song, G.; Liu, S.; Xie, Z.; Chen, J.; Schnable, P.S.; Walley, J.W.; Yin, Y. FERONIA Receptor Kinase Contributes to Plant Immunity by Suppressing Jasmonic Acid Signaling in Arabidopsis thaliana. *Curr. Biol.* **2018**, *28*, 3316–3324. [[CrossRef](#)]
10. Duan, Q.; Kita, D.; Li, C.; Cheung, A.Y.; Wu, H.-M. FERONIA receptor-like kinase regulates RHO GTPase signaling of root hair development. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 17821–17826. [[CrossRef](#)]
11. Kessler, S.; Shimosato-Asano, H.; Keinath, N.F.; Wuest, S.E.; Ingram, G.C.; Panstruga, R.; Grossniklaus, U. Conserved Molecular Components for Pollen Tube Reception and Fungal Invasion. *Science* **2010**, *330*, 968–971. [[CrossRef](#)]
12. Shen, Q.; Bourdais, G.; Pan, H.; Robatzek, S.; Tang, D. Arabidopsis glycosylphosphatidylinositol-anchored protein LLG1 associates with and modulates FLS2 to regulate innate immunity. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 5749–5754. [[CrossRef](#)] [[PubMed](#)]
13. Zhu, S.; Estévez, J.M.; Liao, H.; Zhu, Y.; Yang, T.; Li, C.; Wang, Y.; Li, L.; Liu, X.; Pacheco, J.M.; et al. The RALF1–FERONIA Complex Phosphorylates eIF4E1 to Promote Protein Synthesis and Polar Root Hair Growth. *Mol. Plant* **2020**, *13*, 698–716. [[CrossRef](#)] [[PubMed](#)]
14. Xu, G.; Chen, W.; Song, L.; Chen, Q.; Zhang, H.; Liao, H.; Zhao, G.; Lin, F.; Zhou, H.; Yu, F. Feronia phosphorylates E3 ubiquitin ligase ATL6 to modulate the stability of 14-3-3 proteins in response to the carbon/nitrogen ratio. *J. Exp. Bot.* **2019**, *70*, 6375–6388. [[CrossRef](#)] [[PubMed](#)]
15. Huck, N.; Moore, J.M.; Federer, M.; Grossniklaus, U. The Arabidopsis mutant feronia disrupts the female gametophytic control of pollen tube reception. *Development* **2003**, *130*, 2149–2159. [[CrossRef](#)] [[PubMed](#)]
16. Escobar-Restrepo, J.-M.; Huck, N.; Kessler, S.A.; Gagliardini, V.; Gheyselinck, J.; Yang, W.-C.; Grossniklaus, U. The FERONIA Receptor-like Kinase Mediates Male-Female Interactions during Pollen Tube Reception. *Science* **2007**, *317*, 656–660. [[CrossRef](#)]
17. Rotman, N.; Gourgues, M.; Guitton, A.-E.; Faure, J.-E.; Berger, F. A Dialogue between the Sirène Pathway in Synergids and the Fertilization Independent Seed Pathway in the Central Cell Controls Male Gamete Release during Double Fertilization in Arabidopsis. *Mol. Plant* **2008**, *1*, 659–666. [[CrossRef](#)]
18. Ngo, Q.A.; Vogler, H.; Lituiev, D.S.; Nestorova, A.; Grossniklaus, U. A Calcium Dialog Mediated by the FERONIA Signal Transduction Pathway Controls Plant Sperm Delivery. *Dev. Cell* **2014**, *29*, 491–500. [[CrossRef](#)]

19. Kessler, S.A.; Lindner, H.; Jones, D.S.; Grossniklaus, U. Functional analysis of related Cr RLK 1L receptor-like kinases in pollen tube reception. *EMBO Rep.* **2014**, *16*, 107–115. [[CrossRef](#)]
20. Stegmann, M.; Monaghan, J.; Smakowska-Luzan, E.; Rovenich, H.; Lehner, A.; Holton, N.; Belkhadir, Y.; Zipfel, C. The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. *Science* **2017**, *355*, 287–289. [[CrossRef](#)]
21. Campos, W.F.; Dressano, K.; Ceciliato, P.H.O.; Guerrero-Abad, J.C.; Silva, A.L.; Fiori, C.S.; Canto, A.M.D.; Bergonci, T.; Claus, L.A.N.; Silva-Filho, M.C.; et al. Arabidopsis thaliana rapid alkalization factor 1-mediated root growth inhibition is dependent on calmodulin-like protein 38. *J. Biol. Chem.* **2017**, *293*, 2159–2171. [[CrossRef](#)] [[PubMed](#)]
22. Yu, F.; Li, J.; Huang, Y.; Liu, L.; Li, D.; Chen, L.; Luan, S. FERONIA Receptor Kinase Controls Seed Size in Arabidopsis thaliana. *Mol. Plant* **2014**, *7*, 920–922. [[CrossRef](#)] [[PubMed](#)]
23. Zhao, C.; Zayed, O.; Yu, Z.; Jiang, W.; Zhu, P.; Hsu, C.-C.; Zhang, L.; Tao, W.A.; Lozano-Durán, R.; Zhu, J. Leucine-rich repeat extensin proteins regulate plant salt tolerance in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 13123–13128. [[CrossRef](#)] [[PubMed](#)]
24. Yang, T.; Wang, L.; Li, C.; Liu, Y.; Zhu, S.; Qi, Y.; Liu, X.; Lin, Q.; Luan, S.; Yu, F. Receptor protein kinase FERONIA controls leaf starch accumulation by interacting with glyceraldehyde-3-phosphate dehydrogenase. *Biochem. Biophys. Res. Commun.* **2015**, *465*, 77–82. [[CrossRef](#)]
25. Li, C.; Liu, X.; Qiang, X.; Li, X.; Li, X.; Zhu, S.; Wang, L.; Wang, Y.; Liao, H.; Luan, S.; et al. EB1 nuclear accumulation negatively feeds back on FERONIA-mediated RALF1 signaling. *PLoS Biol.* **2018**, *16*, e2006340. [[CrossRef](#)]
26. Wang, L.; Yang, T.; Lin, Q.; Wang, B.; Li, X.; Luan, S.; Yu, F. Receptor kinase FERONIA regulates flowering time in Arabidopsis. *BMC Plant Biol.* **2020**, *20*, 26. [[CrossRef](#)]
27. Richter, J.; Ploderer, M.; Mongelard, G.; Gutierrez, L.; Hauser, M.-T. Role of CrRLK1L Cell Wall Sensors HERCULES1 and 2, THESEUS1, and FERONIA in Growth Adaptation Triggered by Heavy Metals and Trace Elements. *Front. Plant Sci.* **2017**, *8*, 1554. [[CrossRef](#)]
28. Yang, Z.; Xing, J.; Wang, L.; Liu, Y.; Qu, J.; Tan, Y.; Fu, X.; Lin, Q.; Deng, H.; Yu, F. Mutations of two FERONIA-like receptor genes enhance rice blast resistance without growth penalty. *J. Exp. Bot.* **2020**, *71*, 2112–2126. [[CrossRef](#)]
29. Jia, M.; Ding, N.; Zhang, Q.; Xing, S.; Wei, L.; Zhao, Y.; Du, P.; Mao, W.; Li, J.; Li, B.; et al. A FERONIA-Like Receptor Kinase Regulates Strawberry (*Fragaria × ananassa*) Fruit Ripening and Quality Formation. *Front. Plant Sci.* **2017**, *8*, 1099. [[CrossRef](#)]
30. Zhu, L.; Chu, L.-C.; Liang, Y.; Zhang, X.-Q.; Chen, L.-Q.; Ye, D. The Arabidopsis CrRLK1L protein kinases BUPS1 and BUPS2 are required for normal growth of pollen tubes in the pistil. *Plant J.* **2018**, *95*, 474–486. [[CrossRef](#)]
31. Boisson-Dernier, A.; Roy, S.; Kritsas, K.; Grobei, M.A.; Jaciubek, M.; Schroeder, J.I.; Grossniklaus, U. Disruption of the pollen-expressed FERONIA homologs ANXUR1 and ANXUR2 triggers pollen tube discharge. *Development* **2009**, *136*, 3279–3288. [[CrossRef](#)] [[PubMed](#)]
32. Boisson-Dernier, A.; Lituiev, D.S.; Nestorova, A.; Franck, C.M.; Thirugnanarajah, S.; Grossniklaus, U. ANXUR receptor-like kinases coordinate cell wall integrity with growth at the pollen tube tip via NADPH oxidases. *PLoS Biol.* **2013**, *11*, e1001719. [[CrossRef](#)] [[PubMed](#)]
33. Galindo-Trigo, S.; Blanco-Touriñán, N.; DeFalco, T.A.; Wells, E.S.; Gray, J.E.; Zipfel, C.; Smith, L.M. Cr RLK 1L receptor-like kinases HERK 1 and ANJEA are female determinants of pollen tube reception. *EMBO Rep.* **2019**, *21*, e48466. [[CrossRef](#)]
34. Hematy, K.; Höfte, H. Novel receptor kinases involved in growth regulation. *Curr. Opin. Plant Biol.* **2008**, *11*, 321–328. [[CrossRef](#)]
35. Schoenaers, S.; Balcerowicz, D.; Costa, A.; Vissenberg, K. The Kinase ERULUS Controls Pollen Tube Targeting and Growth in Arabidopsis thaliana. *Front. Plant Sci.* **2017**, *8*, 1942. [[CrossRef](#)] [[PubMed](#)]
36. Schoenaers, S.; Balcerowicz, D.; Breen, G.; Hill, K.; Zdanio, M.; Mouille, G.; Holman, T.J.; Oh, J.; Wilson, M.H.; Nikonorova, N.; et al. The Auxin-Regulated CrRLK1L Kinase ERULUS Controls Cell Wall Composition during Root Hair Tip Growth. *Curr. Biol.* **2018**, *28*, 722–732. [[CrossRef](#)]

37. Gachomo, E.W.; Jno Baptiste, L.; Kefela, T.; Saidel, W.M.; Kotchoni, S.O. The Arabidopsis CURVY1 (CVY1) gene encoding a novel receptor-like protein kinase regulates cell morphogenesis, flowering time and seed production. *BMC Plant Biol.* **2014**, *14*, 221. [[CrossRef](#)]
38. Richter, J.; Watson, J.M.; Stasnik, P.; Borowska, M.; Neuhold, J.; Berger, M.; Stolt-Bergner, P.; Schoft, V.; Hauser, M.-T. Multiplex mutagenesis of four clustered CrRLK1L with CRISPR/Cas9 exposes their growth regulatory roles in response to metal ions. *Sci. Rep.* **2018**, *8*, 12182. [[CrossRef](#)]
39. Haruta, M.; Sabat, G.; Stecker, K.; Minkoff, B.B.; Sussman, M.R. A Peptide Hormone and Its Receptor. *Science* **2014**, *343*, 408–411. [[CrossRef](#)]
40. Ge, Z.; Bergonci, T.; Zhao, Y.; Zou, Y.; Du, S.; Liu, M.-C.; Luo, X.; Ruan, H.; García-Valencia, L.E.; Zhong, S.; et al. Arabidopsis pollen tube integrity and sperm release are regulated by RALF-mediated signaling. *Science* **2017**, *358*, 1596–1600. [[CrossRef](#)]
41. Wu, J.; Kurten, E.L.; Monshausen, G.; Hummel, G.M.; Gilroy, S.; Baldwin, I.T. NaRALF, a peptide signal essential for the regulation of root hair tip apoplastic pH in *Nicotiana attenuata*, is required for root hair development and plant growth in native soils. *Plant J.* **2007**, *52*, 877–890. [[CrossRef](#)] [[PubMed](#)]
42. Covey, P.A.; Subbaiah, C.C.; Parsons, R.L.; Pearce, G.; Lay, F.T.; Anderson, M.A.; Ryan, C.A.; Bedinger, P.A. A pollen-specific RALF from tomato that regulates pollen tube elongation. *Plant Physiol.* **2010**, *153*, 703–715. [[CrossRef](#)] [[PubMed](#)]
43. Cao, J.; Shi, F. Evolution of the RALF Gene Family in Plants: Gene Duplication and Selection Patterns. *Evol. Bioinform.* **2012**, *8*, 271–292. [[CrossRef](#)]
44. Campbell, L.; Turner, S.R. A Comprehensive Analysis of RALF Proteins in Green Plants Suggests There Are Two Distinct Functional Groups. *Front. Plant Sci.* **2017**, *8*, 403. [[CrossRef](#)]
45. Combier, J.-P.; Küster, H.; Journet, E.-P.; Hohnjec, N.; Gamas, P.; Niebel, A. Evidence for the Involvement in Nodulation of the Two Small Putative Regulatory Peptide-Encoding Genes MtrALFL1 and MtrDVL1. *Mol. Plant Microbe Interact.* **2008**, *21*, 1118–1127. [[CrossRef](#)] [[PubMed](#)]
46. Doyle, J.J.; Luckow, M.A. The Rest of the Iceberg. Legume Diversity and Evolution in a Phylogenetic Context. *Plant Physiol.* **2003**, *131*, 900–910. [[CrossRef](#)] [[PubMed](#)]
47. Ferguson, B.; Mens, C.; Hastwell, A.; Zhang, M.; Su, H.; Jones, C.M.; Chu, X.; Gresshoff, P.M. Legume nodulation: The host controls the party. *Plant Cell Environ.* **2018**, *42*, 41–51. [[CrossRef](#)]
48. Goodstein, D.; Shu, S.; Howson, R.; Neupane, R.; Hayes, R.; Fazo, J.; Mitros, T.; Dirks, W.; Hellsten, U.; Putnam, N.H.; et al. Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res.* **2011**, *40*, D1178–D1186. [[CrossRef](#)]
49. Mun, T.; Bachmann, A.; Gupta, V.; Stougaard, J.; Andersen, S.U. Lotus Base: An integrated information portal for the model legume *Lotus japonicus*. *Sci. Rep.* **2016**, *6*, 39447. [[CrossRef](#)]
50. El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S.R.; Luciani, A.; Potter, S.C.; Qureshi, M.; Richardson, L.J.; Salazar, G.A.; Smart, A.; et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**, *47*, D427–D432. [[CrossRef](#)]
51. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
52. Larsson, A. AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **2014**, *30*, 3276–3278. [[CrossRef](#)] [[PubMed](#)]
53. Guindon, S.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O.; Dufayard, J.-F. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)] [[PubMed](#)]
54. Nguyen, L.-T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **2014**, *32*, 268–274. [[CrossRef](#)]
55. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
56. Bailey, T.L.; Bodén, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208. [[CrossRef](#)]

57. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. *The Proteomics Protocols Handbook*; Humana Press: Totowa, NJ, USA, 2005; pp. 571–608. [\[CrossRef\]](#)
58. Wolfe, D.; Dudek, S.; Ritchie, M.; Pendergrass, S.A. Visualizing genomic information across chromosomes with PhenoGram. *BioData Min.* **2013**, *6*, 18. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Wang, Y.; Tang, H.; DeBarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.-H.; Jin, H.; Marler, B.; Guo, H.; et al. MCSanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [\[CrossRef\]](#)
60. O'Rourke, J.A.; Iniguez, L.P.; Fu, F.; Bucciarelli, B.; Miller, S.S.; Jackson, S.A.; McClean, P.E.; Li, J.; Dai, X.; Zhao, P.X.; et al. An RNA-Seq based gene expression atlas of the common bean. *BMC Genom.* **2014**, *15*, 866. [\[CrossRef\]](#)
61. Benedito, V.A.; Torres-Jerez, I.; Murray, J.D.; Andriankaja, A.; Allen, S.; Kakar, K.; Wandrey, M.; Verdier, J.; Zuber, H.; Ott, T.; et al. A gene expression atlas of the model legume *Medicago truncatula*. *Plant J.* **2008**, *55*, 504–513. [\[CrossRef\]](#)
62. He, J.; Benedito, V.A.; Wang, M.; Murray, J.D.; Zhao, P.X.; Tang, Y.; Udvardi, M.K. The *Medicago truncatula* gene expression atlas web server. *BMC Bioinform.* **2009**, *10*, 441. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Schmid, M.; Davison, T.S.; Henz, S.R.; Pape, U.J.; Demar, M.; Vingron, M.; Schölkopf, B.; Weigel, D.; Lohmann, J.U. A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* **2005**, *37*, 501–506. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Libault, M.; Farmer, A.; Joshi, T.; Takahashi, K.; Langley, R.J.; Franklin, L.D.; He, J.; Xu, D.; May, G.; Stacey, G. An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. *Plant J.* **2010**, *63*, 86–99. [\[CrossRef\]](#)
65. Ortiz-Ramírez, C.; Hernandez-Coronado, M.; Thamm, A.; Catarino, B.; Wang, M.; Dolan, L.; Feijó, J.A.A.; Becker, J.D.D. A Transcriptome Atlas of *Physcomitrella patens* Provides Insights into the Evolution and Development of Land Plants. *Mol. Plant* **2016**, *9*, 205–220. [\[CrossRef\]](#)
66. Warnes, G.R.; Bolker, B.; Huber, W.; Lumley, T.; Maechler, M.; Magnusson, A.; Moeller, S. *Gplots: Various R Programming Tools for Plotting Data*. Available online: <https://CRAN.R-project.org/package=gplots> (accessed on 12 October 2019).
67. Islas, T.; Guillén, G.; Alvarado-Affantranger, X.; Lara-Flores, M.; Sánchez, F.; Villanueva, M.A. PvRACK1 Loss-of-Function Impairs Cell Expansion and Morphogenesis in *Phaseolus vulgaris* L. Root Nodules. *Mol. Plant Microbe Interact.* **2011**, *24*, 819–826. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Nguyen, Q.-N.; Lee, Y.-S.; Cho, L.-H.; Jeong, H.-J.; An, G.; Jung, K.-H. Genome-wide identification and analysis of *Catharanthus roseus* RLK1-like kinases in rice. *Planta* **2014**, *241*, 603–613. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Niu, E.; Cai, C.; Zheng, Y.; Shang, X.; Fang, L.; Guo, W. Genome-wide analysis of CrRLK1L gene family in *Gossypium* and identification of candidate CrRLK1L genes related to fiber development. *Mol. Genet. Genom.* **2016**, *291*, 1137–1154. [\[CrossRef\]](#)
70. Kou, X.; Qi, K.; Qiao, X.; Yin, H.; Liu, X.; Zhang, S.; Wu, J. Evolution, expression analysis, and functional verification of *Catharanthus roseus* RLK1-like kinase (CrRLK1L) family proteins in pear (*Pyrus bretschneideri*). *Genomic* **2017**, *109*, 290–301. [\[CrossRef\]](#)
71. Morris, J.L.; Puttick, M.N.; Clark, J.W.; Edwards, D.; Kenrick, P.; Pressel, S.; Wellman, C.H.; Yang, Z.; Schneider, H.; Donoghue, P.C.J. The timescale of early land plant evolution. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E2274–E2283. [\[CrossRef\]](#)
72. Niklas, K.J.; Kutschera, U. The evolution of the land plant life cycle. *New Phytol.* **2009**, *185*, 27–41. [\[CrossRef\]](#)
73. Chang, C.-C.; Chen, H.-L.; Li, W.-H.; Chaw, S.-M. Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *J. Mol. Evol.* **2004**, *58*, 424–441. [\[CrossRef\]](#)
74. Jiao, X.; Leebens-Mack, J.; Ayyampalayam, S.; Bowers, J.; McKain, M.R.; McNeal, J.R.; Rolf, M.; Ruzicka, D.R.; Wafula, E.K.; Wickett, N.J.; et al. A genome triplication associated with early diversification of the core eudicots. *Genome Biol.* **2012**, *13*, R3. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Lynch, M.; Conery, J.S. The Evolutionary Fate and Consequences of Duplicate Genes. *Science* **2000**, *290*, 1151–1155. [\[CrossRef\]](#)

76. Coghlan, A.; Eichler, E.E.; Oliver, S.G.; Paterson, A.H.; Stein, L. Chromosome evolution in eukaryotes: A multi-kingdom perspective. *Trends Genet.* **2005**, *21*, 673–682. [[CrossRef](#)] [[PubMed](#)]
77. Bowers, J.; Chapman, B.; Rong, J.; Paterson, A.H. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **2003**, *422*, 433–438. [[CrossRef](#)] [[PubMed](#)]
78. Freeling, M. Bias in Plant Gene Content Following Different Sorts of Duplication: Tandem, Whole-Genome, Segmental, or by Transposition. *Annu. Rev. Plant Biol.* **2009**, *60*, 433–453. [[CrossRef](#)] [[PubMed](#)]
79. Wang, Y.; Wang, X.; Tang, H.; Tan, X.; Ficklin, S.P.; Feltus, F.A.; Paterson, A.H. Modes of Gene Duplication Contribute Differently to Genetic Novelty and Redundancy, but Show Parallels across Divergent Angiosperms. *PLoS ONE* **2011**, *6*, e28150. [[CrossRef](#)]
80. Jami, S.K.; Clark, G.B.; Ayele, B.T.; Ashe, P.; Kirti, P.B. Genome-wide Comparative Analysis of Annexin Superfamily in Plants. *PLoS ONE* **2012**, *7*, e47801. [[CrossRef](#)]
81. Carle-Urioste, J.C.; Brendel, V.; Walbot, V. A combinatorial role for exon, intron and splice site sequences in splicing in maize. *Plant J.* **1997**, *11*, 1253–1263. [[CrossRef](#)]
82. Mattick, J.S.; Gagen, M.J. The evolution of controlled multitasked gene networks: The role of introns and other noncoding RNAs in the development of complex organisms. *Mol. Biol. Evol.* **2001**, *18*, 1611–1630. [[CrossRef](#)]
83. Schwartz, S.; Meshorer, E.; Ast, G. Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.* **2009**, *16*, 990–995. [[CrossRef](#)] [[PubMed](#)]
84. Zhu, L.; Zhang, Y.; Zhang, W.; Yang, S.; Chen, J.-Q.; Tian, D. Patterns of exon-intron architecture variation of genes in eukaryotic genomes. *BMC Genom.* **2009**, *10*, 47. [[CrossRef](#)] [[PubMed](#)]
85. Kiraga, J.; Mackiewicz, P.; Mackiewicz, D.; Kowalczyk, M.; Biecek, P.; Polak, N.; Smolarczyk, K.; Dudek, M.R.; Cebrat, S. The relationships between the isoelectric point and: Length of proteins, taxonomy and ecology of organisms. *BMC Genom.* **2007**, *8*, 163. [[CrossRef](#)] [[PubMed](#)]
86. Doyle, J.J.; Egan, A.N. Dating the origins of polyploidy events. *New Phytol.* **2009**, *186*, 73–85. [[CrossRef](#)] [[PubMed](#)]
87. Zhao, T.; Schranz, M.E. Network approaches for plant phylogenomic synteny analysis. *Curr. Opin. Plant Biol.* **2017**, *36*, 129–134. [[CrossRef](#)]
88. Schwartz, R.; Ting, C.S.; King, J. Whole Proteome pI Values Correlate with Subcellular Localizations of Proteins for Organisms within the Three Domains of Life. *Genome Res.* **2001**, *11*, 703–709. [[CrossRef](#)]
89. Weiller, G.F.; Caraux, G.; Silvester, N. The modal distribution of protein isoelectric points reflects amino acid properties rather than sequence evolution. *Proteomics* **2004**, *4*, 943–949. [[CrossRef](#)]
90. Yoder, J.B. Understanding the coevolutionary dynamics of mutualism with population genomics. *Am. J. Bot.* **2016**, *103*, 1742–1752. [[CrossRef](#)] [[PubMed](#)]
91. De Moura, G.G.D.; Remigi, P.; Masson-Boivin, C.; Capela, D. Experimental Evolution of Legume Symbionts: What Have We Learnt? *Genes* **2020**, *11*, 339. [[CrossRef](#)]
92. De Mita, S.; Streng, A.; Bisseling, T.; Geurts, R. Evolution of a symbiotic receptor through gene duplications in the legume-rhizobium mutualism. *New Phytol.* **2013**, *201*, 961–972. [[CrossRef](#)]
93. Patriarca, E.J.; Tatè, R.; Ferraioli, S.; Iaccarino, M. Organogenesis of Legume Root Nodules. *Int. Rev. Cytol.* **2004**, *234*, 201–262. [[CrossRef](#)] [[PubMed](#)]
94. Monahan-Giovanelli, H.; Pinedo, C.A.; Gage, D.J. Architecture of Infection Thread Networks in Developing Root Nodules Induced by the Symbiotic Bacterium *Sinorhizobium meliloti* on *Medicago truncatula*. *Plant Physiol.* **2006**, *140*, 661–670. [[CrossRef](#)]
95. Ferguson, B.; Indrasumunar, A.; Hayashi, S.; Lin, M.-H.; Lin, Y.-H.; Reid, D.E.; Gresshoff, P.M. Molecular Analysis of Legume Nodule Development and Autoregulation. *J. Integr. Plant Biol.* **2010**, *52*, 61–76. [[CrossRef](#)] [[PubMed](#)]
96. Popp, C.; Ott, T. Regulation of signal transduction and bacterial infection during root nodule symbiosis. *Curr. Opin. Plant Biol.* **2011**, *14*, 458–467. [[CrossRef](#)] [[PubMed](#)]
97. Xiao, T.T.; Schilderink, S.; Moling, S.; Deinum, E.E.; Kondorosi, E.; Franssen, H.; Kulikova, O.; Niebel, A.; Bisseling, T. Fate map of *Medicago truncatula* root nodules. *Development* **2014**, *141*, 3517–3528. [[CrossRef](#)] [[PubMed](#)]

98. Limpens, E.; Ivanov, S.; Van Esse, G.W.; Voets, G.; Fedorova, E.; Bisseling, T. Medicago N₂-Fixing Symbiosomes Acquire the Endocytic Identity Marker Rab7 but Delay the Acquisition of Vacuolar Identity. *Plant Cell* **2009**, *21*, 2811–2828. [[CrossRef](#)]
99. Gavrin, A.; Chiasson, D.; Ovchinnikova, E.; Kaiser, B.N.; Bisseling, T.; Fedorova, E.E. VAMP721a and VAMP721d are important for pectin dynamics and release of bacteria in soybean nodules. *New Phytol.* **2016**, *210*, 1011–1021. [[CrossRef](#)]
100. Cheung, A.Y.; Wu, H.-M. THESEUS 1, FERONIA and relatives: A family of cell wall-sensing receptor kinases? *Curr. Opin. Plant Biol.* **2011**, *14*, 632–641. [[CrossRef](#)]
101. Mecchia, M.A.; Santos-Fernandez, G.; Duss, N.N.; Somoza, S.C.; Boisson-Dernier, A.; Gagliardini, V.; Martinez-Bernardini, A.; Fabrice, T.; Ringli, C.; Muschietti, J.; et al. RALF4/19 peptides interact with LRX proteins to control pollen tube growth in Arabidopsis. *Science* **2017**, *358*, 1600–1603. [[CrossRef](#)]
102. Vogler, H.; Santos-Fernandez, G.; Mecchia, M.A.; Grossniklaus, U. To preserve or to destroy, that is the question: The role of the cell wall integrity pathway in pollen tube growth. *Curr. Opin. Plant Biol.* **2019**, *52*, 131–139. [[CrossRef](#)]
103. Kuppasamy, K.T.; Ivashuta, S.; Bucciarelli, B.; Vance, C.P.; Gantt, J.S.; VandenBosch, K.A. Knockdown of CELL DIVISION CYCLE16 reveals an inverse relationship between lateral root and nodule numbers and a link to auxin in Medicago truncatula. *Plant Physiol.* **2009**, *151*, 1155–1166. [[CrossRef](#)] [[PubMed](#)]
104. Suzuki, T.; Yano, K.; Ito, M.; Umehara, Y.; Sukanuma, N.; Kawaguchi, M. Positive and negative regulation of cortical cell division during root nodule development in Lotus japonicus is accompanied by auxin response. *Development* **2012**, *139*, 3997–4006. [[CrossRef](#)] [[PubMed](#)]
105. Guan, D.; Stacey, N.; Liu, C.-W.; Wen, J.; Mysore, K.S.; Torres-Jerez, I.; Vernié, T.; Tadege, M.; Zhou, C.; Wang, Z.-Y.; et al. Rhizobial infection is associated with the development of peripheral vasculature in nodules of Medicago truncatula. *Plant Physiol.* **2013**, *162*, 107–115. [[CrossRef](#)] [[PubMed](#)]
106. Sogawa, A.; Yamazaki, A.; Yamasaki, H.; Komi, M.; Manabe, T.; Tajima, S.; Hayashi, M.; Nomura, M. SNARE Proteins LjVAMP72a and LjVAMP72b Are Required for Root Symbiosis and Root Hair Formation in Lotus japonicus. *Front. Plant Sci.* **2019**, *9*, 1992. [[CrossRef](#)] [[PubMed](#)]
107. Foucher, F.; Kondorosi, E. Cell cycle regulation in the course of nodule organogenesis in Medicago. *Plant Mol. Biol.* **2000**, *43*, 773–786. [[CrossRef](#)] [[PubMed](#)]
108. Zipfel, C.; Oldroyd, G.E.D. Plant signalling in symbiosis and immunity. *Nature* **2017**, *543*, 328–336. [[CrossRef](#)]
109. Mergaert, P.; Uchiumi, T.; Alunni, B.; Evanno, G.; Cheron, A.; Catrice, O.; Mausset, A.-E.; Barloy-Hubler, F.; Galibert, F.; Kondorosi, A.; et al. Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5230–5235. [[CrossRef](#)]
110. Ott, T.; Sullivan, J.; James, E.K.; Flegmetakis, E.; Günther, C.S.; Gibon, Y.; Ronson, C.; Udvardi, M.K. Absence of Symbiotic Leghemoglobins Alters Bacteroid and Plant Cell Differentiation During Development of Lotus japonicus Root Nodules. *Mol. Plant-Microbe Interact.* **2009**, *22*, 800–808. [[CrossRef](#)]
111. Wang, L.; Rubio, M.C.; Xin, X.; Zhang, B.; Fan, Q.; Wang, Q.; Ning, G.; Becana, M.; Duanmu, D. CRISPR/Cas9 knockout of leghemoglobin genes in Lotus japonicus uncovers their synergistic roles in symbiotic nitrogen fixation. *New Phytol.* **2019**, *224*, 818–832. [[CrossRef](#)]
112. Torres, M.A.; Jones, J.D.; Dangl, J.L. Pathogen-induced, NADPH oxidase-derived reactive oxygen intermediates suppress spread of cell death in Arabidopsis thaliana. *Nat. Genet.* **2005**, *37*, 1130–1134. [[CrossRef](#)]
113. Lohar, D.P.; Haridas, S.; VandenBosch, K.A.; Gantt, J.S. A transient decrease in reactive oxygen species in roots leads to root hair deformation in the legume-rhizobia symbiosis. *New Phytol.* **2006**, *173*, 39–49. [[CrossRef](#)] [[PubMed](#)]
114. Marino, D.; Andrio, E.; Danchin, E.G.J.; Oger, E.; Gucciardo, S.; Lambert, A.; Puppo, A.; Pauly, N. A Medicago truncatula NADPH oxidase is involved in symbiotic nodule functioning. *New Phytol.* **2010**, *189*, 580–592. [[CrossRef](#)] [[PubMed](#)]
115. Montiel, J.; Nava, N.; Cárdenas, L.; Sanchez-Lopez, R.; Arthikala, M.-K.; Santana, O.; Sánchez, F.; Quinto, C. A Phaseolus vulgaris NADPH Oxidase Gene is Required for Root Infection by Rhizobia. *Plant Cell Physiol.* **2012**, *53*, 1751–1767. [[CrossRef](#)] [[PubMed](#)]

116. Arthikala, M.-K.; Montiel, J.; Sanchez-Lopez, R.; Nava, N.; Cárdenas, L.; Quinto, C. Respiratory Burst Oxidase Homolog Gene A Is Crucial for Rhizobium Infection and Nodule Maturation and Function in Common Bean. *Front. Plant Sci.* **2017**, *8*, 2003. [[CrossRef](#)]
117. Montiel, J.; Fonseca-García, C.; Quinto, C. Phylogeny and Expression of NADPH Oxidases during Symbiotic Nodule Formation. *Agriculture* **2018**, *8*, 179. [[CrossRef](#)]
118. Montiel, J.; Arthikala, M.-K.; Cardenas, L.; Quinto, C. Legume NADPH Oxidases Have Crucial Roles at Different Stages of Nodulation. *Int. J. Mol. Sci.* **2016**, *17*, 680. [[CrossRef](#)]
119. Khadri, M.; Tejera, N.A.; Plá, C.L. Alleviation of Salt Stress in Common Bean (*Phaseolus vulgaris*) by Exogenous Abscisic Acid Supply. *J. Plant Growth Regul.* **2006**, *25*, 110–119. [[CrossRef](#)]
120. Ding, Y.; Kaló, P.; Yendrek, C.; Sun, J.; Liang, Y.; Marsh, J.F.; Harris, J.M.; Oldroyd, G.E.D. Abscisic Acid Coordinates Nod Factor and Cytokinin Signaling during the Regulation of Nodulation in *Medicago truncatula*. *Plant Cell* **2008**, *20*, 2681–2695. [[CrossRef](#)]
121. Liang, Y.; Mitchell, D.M.; Harris, J.M. Abscisic acid rescues the root meristem defects of the *Medicago truncatula* latd mutant. *Dev. Biol.* **2007**, *304*, 297–307. [[CrossRef](#)]
122. Biswas, B.; Chan, P.K.; Gresshoff, P.M. A Novel ABA Insensitive Mutant of *Lotus japonicus* with a Wilty Phenotype Displays Unaltered Nodulation Regulation. *Mol. Plant* **2009**, *2*, 487–499. [[CrossRef](#)]
123. Hunter, W.J. Influence of Root-Applied Epibrassinolide and Carbenoxolone on the Nodulation and Growth of Soybean (*Glycine max* L.) Seedlings. *J. Agron. Crop Sci.* **2001**, *186*, 217–221. [[CrossRef](#)]
124. Ferguson, B.; Ross, J.J.; Reid, J.B. Nodulation Phenotypes of Gibberellin and Brassinosteroid Mutants of Pea1. *Plant Physiol.* **2005**, *138*, 2396–2405. [[CrossRef](#)] [[PubMed](#)]
125. Vardhini, B.V.; Rao, S.S.R. Effect of brassinosteroids on nodulation and nitrogenase activity in groundnut (*Arachis hypogaea* L.). *Plant Growth Regul.* **1999**, *28*, 165–167. [[CrossRef](#)]
126. Upreti, K.; Murti, G. Effects of Brassinosteroids on Growth, Nodulation, Phytohormone Content and Nitrogenase Activity in French Bean Under Water Stress. *Biol. Plant.* **2004**, *48*, 407–411. [[CrossRef](#)]
127. Sun, J.; Cardoza, V.; Mitchell, D.M.; Bright, L.; Oldroyd, G.; Harris, J.M. Crosstalk between jasmonic acid, ethylene and Nod factor signaling allows integration of diverse inputs for regulation of nodulation. *Plant J.* **2006**, *46*, 961–970. [[CrossRef](#)]
128. Nakagawa, T.; Kawaguchi, M. Shoot-applied MeJA Suppresses Root Nodulation in *Lotus japonicus*. *Plant Cell Physiol.* **2006**, *47*, 176–180. [[CrossRef](#)]
129. Suzuki, A.; Suriyagoda, L.; Shigeyama, T.; Tominaga, A.; Sasaki, M.; Hiratsuka, Y.; Yoshinaga, A.; Arima, S.; Agarie, S.; Sakai, T.; et al. *Lotus japonicus* nodulation is photomorphogenetically controlled by sensing the red/far red (R/FR) ratio through jasmonic acid (JA) signaling. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 16837–16842. [[CrossRef](#)]
130. Guinel, F.C.; Geil, R.D. A model for the development of the rhizobial and arbuscular mycorrhizal symbioses in legumes and its use to understand the roles of ethylene in the establishment of these two symbioses. *Can. J. Bot.* **2002**, *80*, 695–720. [[CrossRef](#)]
131. Gresshoff, P.M.; Lohar, D.; Chan, P.-K.; Biswas, B.; Jiang, Q.; Reid, D.E.; Ferguson, B.; Stacey, G. Genetic analysis of ethylene regulation of legume nodulation. *Plant Signal. Behav.* **2009**, *4*, 818–823. [[CrossRef](#)]

