



**UNIVERSIDAD NACIONAL DE LA PLATA**

**FACULTAD DE CIENCIAS EXACTAS**

**DEPARTAMENTO DE CIENCIAS BIOLÓGICAS**

---

**Trabajo de Tesis Doctoral:**

***Bases moleculares de la interacción simbiótica eficiente  
entre *Phaseolus vulgaris* y *Rhizobium etli****

---

**Tesista: Joaquín Clúa**

**Director/a: María Eugenia Zanetti**

**Codirector/a: Flavio Antonio Blanco**

**Año: 2018**



Universidad Nacional de La Plata  
Facultad de Ciencias Exactas  
Departamento de Ciencias Biológicas

*Tesis para optar al título de Doctor de la Facultad de Ciencias  
Exactas*

# **“Bases moleculares de la interacción simbiótica eficiente entre *Phaseolus vulgaris* y *Rhizobium etli*”**

Lic. Joaquín Clúa  
Director: Dr. María Eugenia Zanetti  
Co director: Dr. Flavio Antonio Blanco



Instituto de Biotecnología y Biología Molecular  
UNLP - CONICET  
Año 2018

El presente trabajo de Tesis para optar al grado de Doctor de la Facultad de Ciencias Exactas de la Universidad Nacional de La Plata ha sido realizado en el Instituto de Biotecnología y Biología Molecular, bajo la dirección de la Dr. María Eugenia Zanetti y la co-dirección del Dr. Flavio Antonio Blanco.

# *Agradecimientos*

A mis padres Marita y Javier por su apoyo incondicional y el estímulo que siempre me dieron para estudiar y mejorar.

A la Dra. María Eugenia Zanetti y al Dr. Flavio Blanco por la dirección del presente trabajo. No solamente me formaron como profesional sino que también potenciaron mi gusto por la ciencia y la investigación.

A mis compañeros y amigos del laboratorio por todos los buenos momentos compartidos: Melisse, Marcos, Karen, Priscila, Virginia, Carolina R., Mauricio, Efen, Soledad, Claudio, Ana, Carolina G., Micaela y Carla.

Al Dr. Federico Ariel (Laboratorio de Epigenética y ARNs no codificantes del IAL) por la gran ayuda con los experimentos de CHIP y haberme aceptado en su laboratorio para realizar algunos de estos experimentos.

A los CPAs del Instituto, especialmente a Claudio, Paula, Silvana, Luciana y Ulises por sus aportes al trabajo de investigación y en el trabajo diario dentro del laboratorio.

A todos los investigadores y becarios del IBBM por la generosidad y buena predisposición para colaborar.

A mis hermanos Magdalena e Ignacio por su apoyo, compañía y amistad.

A mis amigos de toda la vida, que son muchos para nombrar, pero su amistad y apoyo fueron fundamentales durante estos años de doctorado.

Al CONICET por la beca doctoral que posibilitó la realización de la presente tesis.

# Índice

<b>RESUMEN</b> .....	<b>1</b>
<b>ABREVIATURAS</b> .....	<b>4</b>
<b>INTRODUCCIÓN</b> .....	<b>6</b>
1.El nitrógeno en la agricultura .....	7
2.Fijación biológica de nitrógeno .....	10
3.Importancia de las leguminosas .....	11
4. <i>Phaseolus vulgaris</i> .....	13
5.Interacción simbiótica entre rizobios y leguminosas .....	14
6.Vía de señalización de la nodulación .....	18
7.Funciones de los factores de transcripción NF-Y y SIN1 en la nodulación .....	22
8.La asociación simbiótica preferencial establecida entre variedades de <i>Phaseolus vulgaris</i> y cepas de <i>Rhizobium etli</i> .....	26
<b>HIPÓTESIS Y OBJETIVOS</b> .....	<b>30</b>
1.Hipótesis de trabajo y objetivo general .....	31
2.Objetivos específicos .....	32
<b>CAPÍTULO I: Análisis transcriptómico de <i>P. vulgaris</i> en respuesta a cepas de <i>R. etli</i></b> .....	<b>33</b>
<b>RESULTADOS</b> .....	<b>34</b>
1.Diseño experimental y generación del material biológico .....	34
2.Construcción de las bibliotecas para RNA-seq, secuenciación y controles de calidad .....	36
3.Relación entre los perfiles transcriptómicos .....	44
4.Análisis general de genes con expresión diferencial .....	46
5.Genes diferenciales en respuesta a la inoculación con <i>R. etli</i> .....	49
6.Identificación de las respuestas moleculares comunes y específicas frente a las cepas de <i>R. etli</i> de tipo <i>nodC-<math>\alpha</math></i> o $-\delta$ .....	55
7.Análisis de los genes con expresión diferencial específicos de cepas tipo <i>nodC-<math>\alpha</math></i> y $-\delta$ .....	61
8.Cambios transcripcionales de los componentes principales de la vía de transducción de señales de la nodulación en respuesta a las distintas cepas de <i>R. etli</i> .....	68
<b>DISCUSIÓN</b> .....	<b>72</b>
<b>CAPÍTULO II: Identificación de sitios de unión de los factores de transcripción NF-YA1, C1 y SIN1</b> .....	<b>81</b>
<b>RESULTADOS</b> .....	<b>82</b>
1.Diseño experimental .....	82
2.Identificación de targets putativos de PvNF-YC1 .....	83
3.PvNF-YC1 se encuentra asociado a elementos de los promotores de <i>PvERN1</i> y <i>PvCYCP</i> que contienen secuencias <i>CCAAT</i> .....	87

4.Construcción de bibliotecas para ChIP-seq .....	92
<b>DISCUSIÓN .....</b>	<b>96</b>

**CAPITULO III: Análisis funcional de NIPK (NF-YC1 *Interacting Protein Kinase*).....100**

<b>RESULTADOS .....</b>	<b>101</b>
1.Identificación de una proteína quinasa que interacciona con PvNF-YC1.....	101
2.NIPK interacciona con NF-YC1 en levadura e <i>in planta</i> .....	104
3.Localización subcelular de NIPK .....	108
4.Patrón de expresión de <i>NIPK</i> .....	112
5.Silenciamiento postranscripcional de NIPK.....	114
6.NIPK es requerido para la organogénesis del nódulo.....	116
7.NIPK es requerido para la formación de ITs .....	119
8.La especificidad de cepa no se ve afectada en raíces con niveles disminuidos de NIPK.....	120
<b>DISCUSIÓN .....</b>	<b>122</b>

**CONCLUSIONES Y PERSPECTIVAS.....127**

**MATERIALES Y MÉTODOS.....130**

<b>1. MATERIAL BIOLÓGICO .....</b>	<b>131</b>
1.1. Semillas .....	131
1.2. Cepas de <i>E. coli</i> .....	131
1.3. Cepas de <i>Agrobacterium</i> .....	131
1.4. Cepas de <i>R. etli</i> .....	132
1.5. Medio de cultivo para <i>E. coli</i> , <i>A. rhizogenes</i> y <i>A. tumefaciens</i> .....	132
1.6. Medios de cultivo para rizobios .....	133
1.7. Medios de cultivo para levaduras .....	133
1.8. Antibióticos .....	134
1.9. Medio Fahraeus .....	135
1.10. Medio de preservación de microorganismos .....	135
1.11. Esterilización y germinación de las semillas .....	136
1.12. Crecimiento de plántulas .....	136
1.13. Crecimiento de <i>A. rhizogenes</i> K599 .....	136
1.14. Obtención de plantas compuestas .....	137
1.15. Inoculación de plantas de poroto .....	138
1.16. Generación de células de <i>E. coli</i> DH5 $\alpha$ competentes químicas .....	139
1.17. Transformación de células de <i>E. coli</i> competentes químicas .....	139
1.18. Generación de células de <i>Agrobacterium rhizogenes</i> K599 y <i>Agrobacterium tumefaciens</i> GV3101 electrocompetentes .....	140
1.19. Transformación de células de <i>Agrobacterium</i> electrocompetentes .....	140
<b>2. MANIPULACIÓN DE ÁCIDOS NUCLEICOS .....</b>	<b>141</b>
2.1. Minipreparación de DNA plasmídico para secuenciación.....	141
2.2. Minipreparación de DNA plasmídico de rutina .....	141
2.3. Reacción en Cadena de la Polimerasa (PCR) .....	142
2.4. Electroforesis de DNA y RNA en geles de agarosa .....	142

2.5. Extracción de RNA .....	143
2.6. Síntesis de cDNA .....	144
2.7. PCR cuantitativa en tiempo real (qRT-PCR) .....	145
2.8. Construcción de vectores mediante sistema GATEWAY .....	145
<b>3. CARACTERIZACIÓN FENOTÍPICA DE PLANTAS COMPUESTAS .....</b>	<b>148</b>
3.1. Cinéticas de nodulación .....	148
3.2. Determinación del tamaño de los nódulos .....	149
3.3. Ensayos de coinoculación y análisis de polimorfismo del gen <i>nodC</i> .....	149
3.4. Análisis de hilos de infección (ITs) .....	150
<b>4. ANÁLISIS DE LA EXPRESIÓN DE PROTEÍNAS MEDIANTE WESTERN BLOT.....</b>	<b>151</b>
4.1. Extracción de proteínas totales .....	151
4.2. Separación de proteínas por SDS-PAGE .....	151
4.3. Transferencia de las proteínas a membrana de nitrocelulosa .....	151
4.4. Inmunodetección y revelado por quimioluminiscencia .....	152
<b>5. LOCALIZACIÓN SUBCELULAR E INTERACCIÓN PROTEÍNA-PROTEÍNA .....</b>	<b>153</b>
5.1. Ensayos de localización subcelular y complementación bimolecular por fluorescencia (BiFC) en <i>N. benthamiana</i> .....	153
5.2. Ensayos de co-inmunoprecipitación .....	153
5.3. Observaciones macroscópicas y microscópicas generales .....	154
5.4. Ensayo de doble híbrido de levadura .....	155
5.5. Actividad $\beta$ -galactosidasa .....	156
<b>6. ANÁLISIS BIOINFORMÁTICOS GENERALES .....</b>	<b>157</b>
6.1. Análisis general de secuencias y diseño de primers .....	157
6.2. Bases de datos de secuencias genómicas .....	157
6.3. BLAST .....	157
6.4. Alineamientos múltiples de secuencias proteicas .....	158
6.5. Análisis filogenético .....	158
6.6. Identificación de dominios funcionales en NIPK .....	158
<b>7. RNA-SEQ .....</b>	<b>159</b>
7.1. Material biológico y construcción de las bibliotecas .....	159
7.2. Procesamiento y análisis de la información de secuenciación .....	160
7.3. Análisis de <i>clustering</i> jerárquico y PCA .....	163
7.4. Análisis de enriquecimiento de GO ( <i>gene ontology</i> ) .....	164
7.5. Diagramas de Venn .....	164
<b>8. CHIP-PCR Y CHIP-SEQ .....</b>	<b>165</b>
8.1. Fijación y colección de tejido .....	165
8.2. Aislamiento de núcleos y sonicación .....	165
8.3. Preparación de <i>Dynabeads</i> conjugadas a <i>anti</i> -FLAG o IgG .....	166
8.4. Inmunoprecipitación .....	166
8.5. <i>Reverse crosslink</i> .....	167
8.6. Precipitación del DNA .....	167
8.7. Análisis de datos .....	168

<b>9. PRIMERS UTILIZADOS.....</b>	<b>168</b>
9.1. Diseño .....	168
9.2. Secuencias .....	168

<b>MATERIAL SUPLEMENTARIO .....</b>	<b>171</b>
-------------------------------------	------------

<b>BIBLIOGRAFÍA .....</b>	<b>175</b>
---------------------------	------------

#### **DVD ADJUNTO**

1. Tabla S1. Valores de expresión de los genes de NAG12
2. Tabla S2. Valores de expresión de los genes de Alubia
3. Tabla S3. Valores de expresión de los genes de NAG12 por réplica
4. Tabla S4. Valores de expresión de los genes de Alubia por réplica
5. Tabla S5. Análisis de diferencialidad para los genes de NAG12
6. Tabla S6. Análisis de diferencialidad para los genes Alubia
7. Tabla S7. Genes diferenciales respecto a YEM por grupo



# Resumen

Las leguminosas establecen una asociación simbiótica fijadora de nitrógeno con bacterias del suelo conocidas como rizobios. Cuando la planta crece en condiciones limitantes de nitrógeno comienza un intercambio de señales químicas con especies de rizobios compatibles presentes en la rizósfera, que permite un reconocimiento mutuo muy específico. La bacteria se adhiere a la punta de los pelos radicales, y la planta genera una estructura tubular llamada hilo de infección que permite el avance de la bacteria hacia las células del córtex. Allí se activan las divisiones celulares formando un nuevo órgano llamado nódulo, donde los rizobios son internalizados al citoplasma de las células vegetales y se lleva a cabo la fijación de nitrógeno atmosférico a formas fácilmente asimilables por la planta.

*Phaseolus vulgaris* (poroto) es una leguminosa de origen americano que ha diversificado genéticamente en dos centros geográficos independientes, el Andino y el Mesoamericano. El estudio molecular de la diversidad de rizobios presentes en nódulos ha permitido determinar que *Rhizobium etli* es la especie que ocupa predominantemente los nódulos de *P. vulgaris*. A su vez, los cultivares de origen Mesoamericano nodulan preferentemente y de manera más eficiente con cepas que poseen el alelo tipo  $\alpha$  en el gen *nodC*, mientras que las plantas de origen Andino nodulan más eficientemente con cepas con un alelo tipo  $\delta$ . Este sistema representa un excelente modelo experimental para el estudio de los determinantes moleculares que participan en el reconocimiento específico entre ambos simbioses y lleva al establecimiento de una interacción preferencial y más eficiente.

Estudios previos de nuestro laboratorio permitieron identificar y caracterizar a nivel funcional distintos genes implicados en esta respuesta. Uno de ellos codifica para la

## *Resumen*

subunidad C1 del factor de transcripción NF-Y, el cual es un regulador transcripcional requerido tanto para la infección bacteriana y la organogénesis del nódulo, como para la nodulación preferencial por cepas *nodC-α*. La subunidad NF-YC1 interacciona con las subunidades NF-YA1 y NF-YB7 formando un trímero funcional en la simbiosis. A su vez, NF-YC1 interacciona físicamente con un factor de transcripción de la familia GRAS denominado SIN1, el cual es requerido para la infección y organogénesis del nódulo, así como también para el desarrollo de raíces laterales, siendo un excelente ejemplo de la interconexión de programas transcripcionales en el desarrollo de dos órganos diferentes.

En la presente Tesis doctoral se abordó la caracterización de las bases moleculares de una interacción simbiótica eficiente mediante el análisis transcriptómico por secuenciación masiva de ARN de raíces de plantas de poroto de origen Mesoamericano y Andino inoculadas con cepas tipo *nodC-α*, *nodC-δ* o con medio de cultivo a modo de control. Este análisis permitió identificar genes que se expresen de manera diferencial frente a las cepas más eficientes, y además, distinguir las respuestas moleculares específicas de cada cultivar frente a cada cepa. Estos genes codifican proteínas involucradas en diferentes procesos de percepción y señalización, regulación de la transcripción, procesos redox y en modificaciones de la pared celular y la membrana plasmática, los cuales son requeridos para la iniciación y progresión de los hilos de infección.

En una segunda etapa de este trabajo de Tesis se identificaron genes que son regulados de manera directa por los factores de transcripción NF-YC1, NF-YA1 y SIN en etapas tempranas de la simbiosis entre *P. vulgaris* y *R. etli*. Para ello se utilizaron plantas que expresan de manera ectópica y constitutiva estos factores de transcripción como fusiones traduccionales al epítipo FLAG, las cuales se utilizaron en experimentos de

## *Resumen*

inmunoprecipitación de cromatina seguida de PCR (ChIP-PCR) para identificar genes regulados directamente por dichos factores transcripcionales. Esta aproximación experimental permitió identificar genes que participan en la activación de las divisiones mitóticas y de la infección rizobiana.

Por último, para profundizar en el conocimiento sobre la vía de transducción de señales que involucra a NF-YC1, en nuestro laboratorio se realizó un *screening* de una biblioteca de doble híbrido de levadura para identificar proteínas que interaccionen físicamente con NF-YC1. Uno de los clones identificados codifica una proteína quinasa, a la cual denominamos NF-YC1 *Interacting Protein Kinase* (NIPK). La interacción física entre NF-YC1 y NIPK fue confirmada mediante ensayos de doble híbrido, complementación bimolecular de fluorescencia y ensayos de co-inmunoprecipitación. Posteriormente se caracterizó la localización subcelular de NIPK detectándose tanto en el citoplasma como en la membrana plasmática. La caracterización fenotípica de plantas de *P. vulgaris* silenciadas en NIPK utilizando dos RNA de interferencia (RNAi) dirigidos a diferentes regiones del mRNA, revelaron que NIPK sería requerido para la organogénesis y desarrollo de los nódulos, así como también la frecuencia de la infección rizobiana.

El conjunto de resultados obtenidos en este trabajo de Tesis no sólo contribuyen a comprender en mayor detalle los mecanismos moleculares que gobiernan el establecimiento de una simbiosis fijadora de nitrógeno eficiente entre *P. vulgaris* y *R. etli*, sino que también sienta las bases para futuros proyectos de investigación y desarrollo que permitan optimizar la fijación de nitrógeno en una planta de interés agronómico y alto valor social para los pueblos del Noroeste Argentino.

# Abreviaturas

- aa** (aminoácido)
- AC** (*activation domain*)
- ARFLP** (*amplified restriction fragment length polymorphism*)
- BD** (*binding domain*)
- BiFC** (*bimolecular fluorescence complementation*)
- CCaMK** (*calcium and calmodulin-binding kinase*)
- cDNA** (DNA complementario)
- ChIP** (*chromatin immunoprecipitation*)
- CNGC15** (*cyclic nucleotide - gated channel 15*)
- CoIP** (co-inmunoprecipitación)
- CSP** (*common symbiosis pathway*)
- CYC** (*cyclin*)
- DCL2** (*Dicer-like 2*)
- DMI1** (*doesn't make infections 1*)
- DMI2** (*doesn't make infections 2*)
- dpi** (días post inoculación)
- EIN3** (*Ethylene insensitive 3 family protein*)
- ERF** (*Ethylene responsive factor*)
- ERN1** (*ERF Required for Nodulation 1*)
- ERN2** (*ERF Required for Nodulation 2*)
- FBN** (Fijación Biológica de Nitrogeno)
- FDR** (*false discovery rate*)
- FPKM** (*Fragments Per Kilobase of transcript per Million mapped reads*)
- GA** (*gibberellic acid*; ácido giberélico)
- GED** (genes con expresión diferencial)
- GFP** (*green fluorescent protein*)
- GO** (*Gene Ontology*)
- GRAS** (*Gibberellic Acid-Insensitive (GAI), Repressor of GAI, and Scarecrow*)
- HF** (epitope His-FLAG)
- HFM** (*histone fold motive*)
- hpi** (horas post inoculación)
- IT** (*infection thread*)
- LHK1** (*L. japonicus Histidine Kinase 1*)
- LOB** (*Lateral organ boundaries*)
- LRR** (*Leucine rich repeat*)
- LYK3** (*LysM domain receptor-like kinase 3*)
- LysM** (*lysine motif*)
- Mt** (Millones de toneladas)

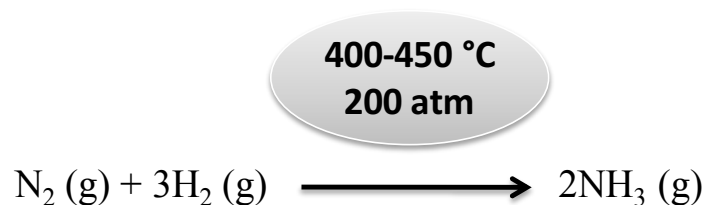
## *Abreviaturas*

**NFP** (*Nod Factor Perception*)  
**NFR1** (*Nod factor receptor 1*)  
**NFR5** (*Nod factor receptor 5*)  
**NFs** (*Nod Factors*)  
**NF-Y** (*Nuclear factor Y*)  
**NIN** (*Nodule inception*)  
**NIPK** (*NF-YC1 Interacting Protein Kinase*)  
**NSP1** (*Nodulation signaling pathway 1*)  
**NSP2** (*Nodulation signaling pathway 2*)  
**ORF** (*open reading frame*)  
**p35S** (*Cauliflower mosaic virus 35S promoter*)  
**pb** (pares de bases)  
**PCA** (análisis de componentes principales)  
**PCDB** (*plasmodesmata callose-binding protein 1*)  
**QDO** (quadruple drop-out)  
**RIN** (*RNA integrity number*)  
**RNAi** (RNA de interferencia)  
**RNA-seq** (secuenciación masiva de RNA)  
**SD** (*standard deviation*)  
**SEA** (*Singular Enrichment Analysis*)  
**SEM** (*standard error of the mean*)  
**SIN1** (*Scarecrow-like13 Involved in Nodulation*)  
**siRNA** (pequeños RNA de interferencia)  
**SYMRK** (*Symbiosis Receptor Kinase*)  
**TDO** (*triple drop-out*)  
**TEM** (*transmission electron microscopy*)  
**UTR** (*untranslated region*)  
**YEM** (*Yeast Extract Mannitol*)  
**YFP** (*yellow fluorescent protein*)

# INTRODUCCIÓN

## **1. El nitrógeno en la agricultura**

El nitrógeno es un elemento esencial para todos los seres vivos ya que es requerido para la síntesis de aminoácidos, nucleótidos y vitaminas. A pesar de que el 78% de la atmósfera está compuesta por nitrógeno (N<sub>2</sub>), sólo un limitado número de microorganismos es capaz de asimilarlo de manera directa (Vitousek et al., 2002). El resto de los microorganismos, plantas y animales dependen de fuentes de nitrógeno combinado para incorporarlo al metabolismo celular. En sistemas agrícolas, la disponibilidad de fuentes asimilables de nitrógeno en los suelos suele ser un factor limitante del crecimiento vegetal y en consecuencia de su productividad (Smil, 1999). Para suplir esta deficiencia, durante siglos se recurrió a la incorporación de fertilizantes de origen natural como el guano y el salitre provenientes principalmente de yacimientos ubicados en Bolivia, Perú y Chile (Galloway et al., 2013). Sin embargo, el comienzo del siglo XX marcó un punto de inflexión en el mundo de la agricultura cuando el químico alemán Fritz Haber desarrolló la síntesis de amonio a partir de hidrógeno y nitrógeno, lo que le valió en 1918 el premio Nobel de Química por “su contribución a mejorar los estándares de la agricultura y el bienestar de la humanidad” (Figura 1). Posteriormente, Carl Bosch logró llevar el desarrollo de Haber a escala industrial mediante la química de alta presión, por lo que también obtuvo el premio Nobel de Química en 1931.



**Figura 1. Proceso Haber para la obtención de amoníaco.**

El proceso de Haber permite la síntesis de amoníaco (NH<sub>3</sub>) a partir de nitrógeno (N<sub>2</sub>) e hidrógeno (H<sub>2</sub>) molecular. La reacción de síntesis requiere un gran aporte de energía; 400-450 °C de temperatura y una presión de 200 atm.

## Introducción

A partir de ese momento el hombre fue capaz de producir fertilizantes nitrogenados en una escala sin precedentes. A nivel global, la fijación de nitrógeno por fuentes antropogénicas ha igualado a la natural. En sistemas agrícolas los datos son aun más impactantes. Se estima que anualmente se fijan 120 millones de toneladas (Mt) de nitrógeno mediante el proceso de Haber y Bosch, mientras que la fijación de nitrógeno por fuentes naturales es de sólo 63 Mt anuales (Tabla 1) (Fowler et al., 2013).

**Tabla 1.** Fijación global de nitrógeno en millones de toneladas (Mt) de nitrógeno

<b>Natural</b>	
FBN* Terrestre	58
Luz	5
FBN Oceánica	140
<b>Total</b>	<b>203</b>
<b>Antropogénica</b>	
Combustión	30
Haber-Bosch	120
FBN Agricultura	60
<b>Total</b>	<b>210</b>

\* Fijación Biológica de Nitrógeno

La disponibilidad de fertilizantes químicos permitió suplir la deficiencia de nitrógeno de los suelos y reponerlo luego de cada cosecha, aumentando fuertemente el rendimiento de los cultivos. La cantidad de seres humanos alimentados por hectárea de suelo pasó de 1,9 en el año 1908 a 4,3 en el 2008, en gran medida gracias a la utilización de fertilizantes. Su importancia en la agricultura es tal que se estima que el proceso de Haber-Bosch es responsable de la producción de alimentos que consume la mitad de la humanidad (Erisman et al., 2008). Sin embargo, la utilización indiscriminada de fertilizantes trae aparejados graves problemas ambientales y socio-económicos. Sólo una fracción del amonio ( $\text{NH}_4^+$ ) de los fertilizantes es incorporado por las plantas, mientras que el resto es convertido en distintos compuestos nitrogenados (Fowler et al.,



## *Introducción*

2013). Una porción del nitrógeno retorna a la atmósfera en forma de  $N_2O$ , un gas de efecto invernadero 298 veces más potente que el  $CO_2$  (Reay et al., 2012). Además, parte del nitrógeno es convertido en distintos óxidos de nitrógeno que generan lluvia ácida, afectando la productividad de los suelos (Galloway et al., 2008). Los microorganismos nitrificantes convierten el amonio en nitratos ( $NO_3^-$ ), los cuales lixivian hacia las napas y cursos de agua provocando su eutrofización (Gruber and Galloway, 2008). En este proceso se produce una gran proliferación de algas y plantas en la superficie del agua que disminuye la llegada de la luz solar al lecho, impidiendo que microorganismos y plantas lleven a cabo el proceso de fotosíntesis, y en consecuencia, oxigenen el agua. Por este motivo se produce un gran daño ecológico debido a la pérdida de biodiversidad (Erisman et al., 2013). Por otra parte, la ingesta de agua con altos contenidos de nitratos está asociada con diversas patologías, como un aumento en el riesgo de contraer distintos tipos de cáncer y el síndrome del bebé azul, una patología que afecta la hemoglobina y en consecuencia el transporte de oxígeno en la sangre. Es por este motivo que los niños adquieren un color azulado, y de no ser tratados pueden sufrir fallas respiratorias y digestivas. Por todo esto, la Organización Mundial de la Salud sugiere el límite de 10 mg de nitratos por litro de agua potable ([http://www.who.int/water\\_sanitation\\_health/dwq/gdwq3\\_es\\_full\\_lowres.pdf](http://www.who.int/water_sanitation_health/dwq/gdwq3_es_full_lowres.pdf)).

Además de los daños medioambientales mencionados, el uso indiscriminado de fertilizantes trae aparejado problemas sociales y económicos. La producción industrial de fertilizantes es un proceso no sustentable que requiere la utilización de grandes cantidades de energía para alcanzar las altas presiones y temperaturas requeridas para acelerar la reacción de síntesis. Aproximadamente el 1 % de la energía mundial es destinada a este proceso (Erisman et al., 2008). Consecuentemente, la utilización de fertilizantes en cultivos extensivos tiene un alto costo que el pequeño y mediano

## Introducción

productor no pueden afrontar, aumentando las ventajas competitivas de los grandes *pooles* de siembra.

La población mundial es de 7.400 millones de personas y las proyecciones indican que continuará creciendo de manera exponencial en los próximos años (Gerland et al., 2014). Unos de los grandes retos que debe abordar la ciencia, incluyendo la biotecnología moderna, es cómo generar la cantidad de alimento necesario para alimentar a la humanidad de una manera sustentable (Tilman, 2002). En este contexto, la fijación biológica de nitrógeno surge como una alternativa promisoriosa, debido a su bajo costo de implementación y la sustentabilidad del proceso (Duhamel and Vandenkoornhuyse, 2013).

## 2. Fijación biológica de nitrógeno

La fijación biológica de nitrógeno (FBN) sólo es llevada a cabo por un grupo especializado de organismos procariotas capaces de convertir el nitrógeno gaseoso ( $N_2$ ) en amoníaco ( $NH_3$ ). Colectivamente, este grupo de microorganismos son denominados diazótrofos (*azoe*: nitrógeno; *trofos*: alimentación), y comprende diversos grupos filogenéticos como las cianobacterias, actinomicetes, proteobacterias y algas verde-azuladas (cianofíceas) (Vitousek et al., 2002). Todos los diazótrofos descritos hasta el momento catalizan la reacción de fijación de nitrógeno por medio del complejo enzimático denominado nitrogenasa:



La estequiometría de la reacción muestra un alto consumo de energía en la forma de ATP y equivalentes de reducción. Además, el complejo nitrogenasa es muy sensible al

## *Introducción*

oxígeno, por lo que la reducción de N<sub>2</sub> mediante la FBN debe realizarse en un ambiente anaerobio o microaerobio (Zumft and Mortenson, 1975).

Los diazótrofos son capaces de reducir el nitrógeno atmosférico en diversas condiciones ecológicas: como microorganismos de vida libre, de manera asociativa con otros organismos y de manera simbiótica principalmente con plantas leguminosas. En sistemas terrestres, la contribución a la fijación biológica de nitrógeno global por parte de microorganismos de vida libre es relativamente baja, dado que la disponibilidad de fuentes de carbono y energía para realizar la costosa reducción de N<sub>2</sub> es escasa (Smil, 1999). Aquellos diazótrofos que fijan nitrógeno de manera asociativa con plantas contribuyen de manera más significativa. Especies del género *Azospirillum* forman una íntima asociación con cereales de gran interés agronómico como el maíz, el arroz y el trigo. Sin embargo, la principal fuente de formas asimilables de nitrógeno en los suelos es dada por la asociación simbiótica fijadora de nitrógeno entre plantas leguminosas y un grupo de diazótrofos llamados colectivamente rizobios (Smill et al., 1999). La presencia de señales químicas provenientes de los rizobios induce la formación de un nuevo órgano en la raíz de la planta llamado nódulo, donde la bacteria es alojada. El nódulo le provee al rizobio un ambiente con baja presión de oxígeno y el suministro de una fuente de carbono que le permite a la bacteria realizar la conversión del N<sub>2</sub> atmosférico a formas asimilables para la planta (Oldroyd, 2013).

### **3. Importancia de las leguminosas**

Las leguminosas (*Fabaceae*) constituyen la tercera gran familia de plantas superiores comprendiendo 670-750 géneros y 18.000-19.000 especies, de las cuales el 88% es capaz de formar nódulos fijadores de nitrógeno por medio de la asociación simbiótica con rizobios, lo que les otorga excelentes propiedades agronómicas y nutricionales. Por

## Introducción

este motivo ocupan el segundo lugar en cuanto a su importancia agronómica, después de los cereales (*Gramineae*) (Graham and Vance, 2003). Según datos de la FAO, la producción de leguminosas en 2013 fue de 399 Mt (Foyer et al., 2016). Entre las especies más importantes desde el punto de vista agronómico podemos mencionar soja (*Glycine max*), poroto (*Phaseolus vulgaris*), maní (*Arachis hypogaeae*), arveja (*Pisum sativum*), alfalfa (*Medicago sativa*) y garbanzo (*Cicer arietinum*) (Figura 2).



**Figura 2. Principales leguminosas de interés agronómico.**

Soja (*Glycine max*) (A), poroto (*Phaseolus vulgaris*) (B), maní (*Arachis hypogaeae*) (C), arveja (*Pisum sativum*) (D), alfalfa (*Medicago sativa*) (E) y garbanzo (*Cicer arietinum*) (F).

El cultivo de leguminosas aporta al suelo 5-7 Mt de nitrógeno asimilable, lo que representa un ahorro a nivel global de 8-12 miles de millones de dólares por la sustitución de fertilizantes (Foyer et al., 2016). Sumado a esto, el uso de leguminosas en la rotación con otros cultivos mejora la productividad, reduce la incidencia de patógenos e incrementa la asimilación de otros nutrientes esenciales como el fósforo. Desde el punto de vista nutricional, las leguminosas se destacan fundamentalmente por el aporte de proteínas, con valores superiores al 25 % del peso total, mientras que alimentos como arroz, trigo y maíz (cereales) poseen un 12 %, la carne un 20 %, el pescado un 18 % y la leche un 3 %. Estos parámetros dan cuenta de la enorme importancia social del cultivo de leguminosas, principalmente en países subdesarrollados, donde constituyen

## *Introducción*

un sustituto de bajo costo de la carne y la principal fuente de proteínas de la dieta. Por estos motivos, las Naciones Unidas proclamó el 2016 como el año internacional de las legumbres (<http://www.fao.org/pulses-2016/es/>), bajo el eslogan “semillas nutritivas para un futuro sostenible”. Esta iniciativa tuvo como objetivo difundir y sensibilizar a la opinión pública sobre las ventajas de las leguminosas en la alimentación y la agricultura.

### **4. *Phaseolus vulgaris***

*Phaseolus* es un género de la familia *Fabaceae* donde se agrupan unas 90 especies que representan algunos de los cultivos más ancestrales de América. Junto con el maíz y la mandioca fueron los cultivos dominantes en las regiones de altitud media-baja de América por milenios. Los cultivos del género *Phaseolus*, comúnmente denominados porotos o frijoles, son las leguminosas de grano más importante para el consumo humano con una producción anual que excede las 23 millones de toneladas métricas, de las cuales 22,57 corresponden a la especie *Phaseolus vulgaris* (Broughton et al., 2003). Sumado a su gran importancia económica y al aporte de nitrógeno a los suelos a través de la fijación simbiótica de nitrógeno, *P. vulgaris* es un cultivo con una notable importancia social, aportando el 15 % de las calorías diarias y alcanzando hasta un 36 % de las proteínas de la dieta en regiones de América y África. Además, sus granos aportan una gran cantidad de hierro, fósforo, magnesio, manganeso, y en menor medida, zinc, cobre y calcio (Schmutz et al., 2014).

La producción de poroto en Argentina oscila entre 280.00 y 300.000 toneladas anuales (Cámara de Legumbres de la República Argentina, CLERA, [www.clera.com.ar](http://www.clera.com.ar)), provenientes casi exclusivamente del Noroeste del país. En el ámbito regional (Noroeste Argentino), la producción de porotos genera la mayor fuente de divisas agroindustriales

## *Introducción*

de exportación en esta región, aun por encima de la soja y de los cítricos. Prácticamente la totalidad de la producción argentina de poroto se destina a la exportación. Existen dos mercados y producciones bien distintas. Por un lado el poroto negro, dirigido principalmente a países latinoamericanos y africanos y por otra parte el mercado de poroto alubia, de alta calidad y valor, destinado a países europeos del Mediterráneo. Argentina es formadora de precios de poroto alubia a escala mundial y de poroto negro en el contexto del Mercosur. En esta variedad, Argentina depende fuertemente del mercado brasileño, que absorbe más de la mitad de los volúmenes producidos (Sistema Nacional Argentino de Vigilancia y Monitoreo de Plagas).

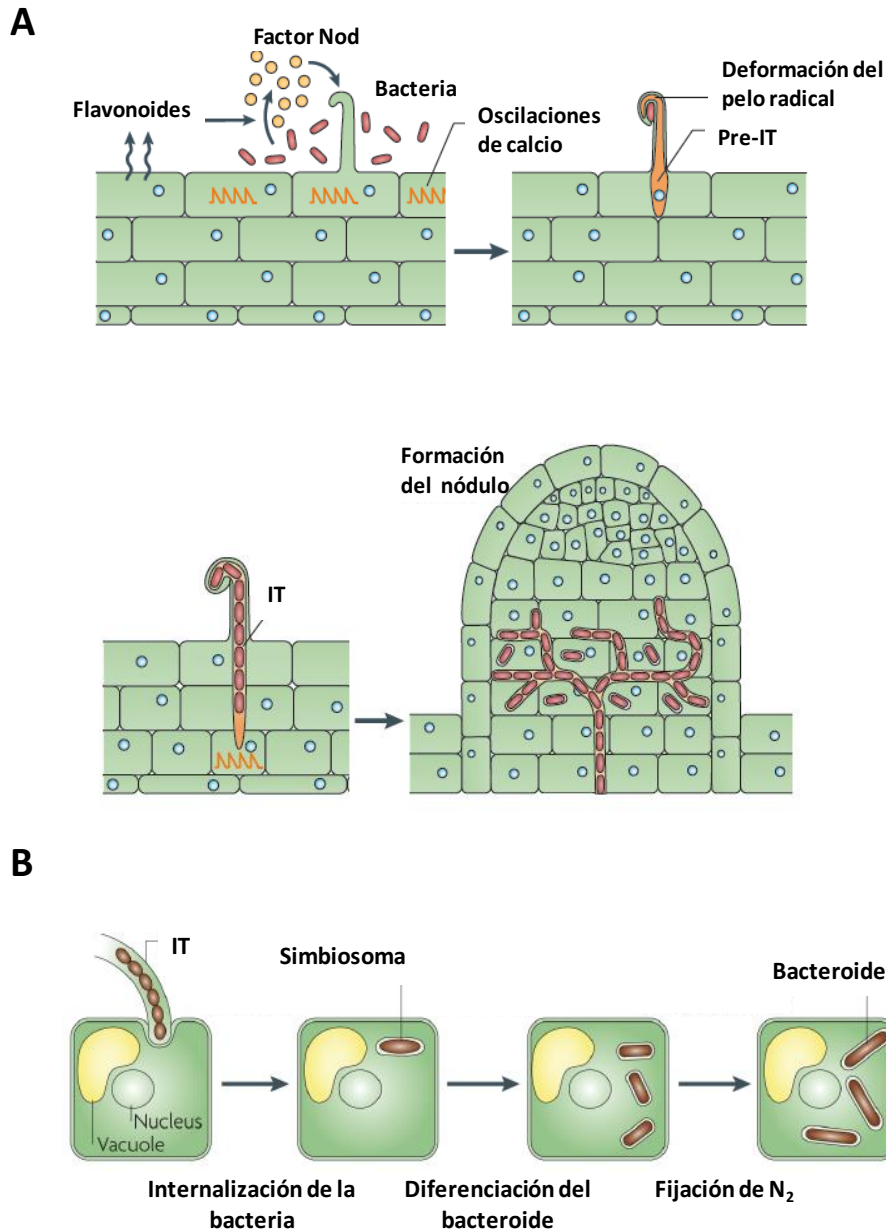
### **5. Interacción simbiótica entre rizobios y leguminosas**

La simbiosis mutualista fijadora de nitrógeno entre plantas leguminosas y rizobios surgió hace aproximadamente 100 millones de años en un grupo de angiospermas pertenecientes al clado *Fabidae*, otorgándoles a estas plantas la ventaja adaptativa de obtener fuentes de nitrógeno asimilables en suelos pobres en estos nutrientes (Martin et al., 2017; Werner et al., 2014). Cuando la planta sensa la limitación de nitrógeno en el suelo exuda flavonoides e isoflavonoides a través de sus raíces, los cuales son percibidos por el rizobio activando el factor de transcripción NodD (Perret et al., 2000). Este regulador transcripcional promueve la expresión de los genes de nodulación, que participan en la síntesis y secreción de los factores de nodulación (NFs; *Nod Factors*), oligómeros de N acetyl glucosamina N-acetilados que a pesar de ser sintetizados por la bacteria actúan de manera similar a las hormonas vegetales; son suficientes para iniciar la organogénesis del nódulo y los cambios morfológicos y fisiológicos que tienen lugar durante la infección rizobiana (Geurts et al., 2012).

## *Introducción*

El primer paso directo en la asociación simbiótica es la adhesión de la bacteria a la epidermis de la raíz, predominantemente en la punta de los pelos radicales. Lectinas de la planta promueven este proceso mediante la unión a polisacáridos específicos localizados en la superficie de la bacteria. La producción de NFs por parte de las bacterias adheridas a la punta del pelo radical interrumpe su crecimiento polar simétrico, cambiando su eje de crecimiento y formando un bucle que atrapa a las bacterias entre las paredes celulares del pelo radical curvado denominado bolsillo de infección (*infection pocket*). Los rizobios allí atrapados comienzan a dividirse generando un foco de infección, que eventualmente dará lugar a la formación del hilo de infección (IT; *infection thread*). El IT es una estructura tubular hueca que se genera por la degradación de la pared celular vegetal en el foco de infección, la cual se invagina junto con la membrana plasmática y orienta su crecimiento hacia el córtex de la raíz. El lumen del IT puede considerarse aún el exterior de la célula, ya que en condiciones normales durante la elongación del IT las bacterias no entran en contacto con el citoplasma de la célula vegetal. Los rizobios se multiplican en la matriz glicoproteica del IT, acompañando su crecimiento hacia las capas más internas de tejido de la raíz donde se está formando el nódulo (Oldroyd and Downie, 2008; Oldroyd et al., 2011) (Figura 3A).

Cuando los NFs son sensados en las células epidérmicas, se activan las divisiones mitóticas de las células del córtex radical de manera coordinada con el proceso de infección, iniciando la organogénesis del nódulo (Oldroyd et al., 2011). Tanto el tipo celular específico que da comienzo a las primeras divisiones celulares, como la morfología y ontogenia del nódulo maduro dependen del tipo de leguminosa. Las leguminosas de climas templados como *M. truncatula* y *P. sativum* desarrollan nódulos de tipo indeterminado, mientras que las leguminosas de climas tropicales y



**Figura 3. Infección y formación del nódulo.**

(A) Los procesos de infección y organogénesis del nódulo ocurren coordinadamente y se encuentran altamente regulados. La interacción comienza cuando los flavonoides son exudados por la raíz en condiciones de deficiencia de nitrógeno. Estos son reconocidos por los rizobios que en respuesta sintetizan y secretan los Factores Nod (NFs). Los NFs activan la vía de transducción de señales simbiótica, que incluye la generación de oscilaciones de calcio (curvas naranja) inicialmente en las células epidérmicas, generando la deformación del pelo y la formación del hilo de infección (IT) y, posteriormente en las células corticales, promoviendo la mitosis y formación del primordio nodular (Adaptado de Oldroyd et al., 2013). (B) El IT avanza hasta el primordio de nódulo donde los rizobios son internalizados y rodeados por una membrana de origen vegetal formando estructuras de tipo organelas denominadas simbiosomas. La bacteria en este entorno se diferencia a bacteroide y comienza a fijar N<sub>2</sub> (Adaptado de (Jones et al., 2007)).



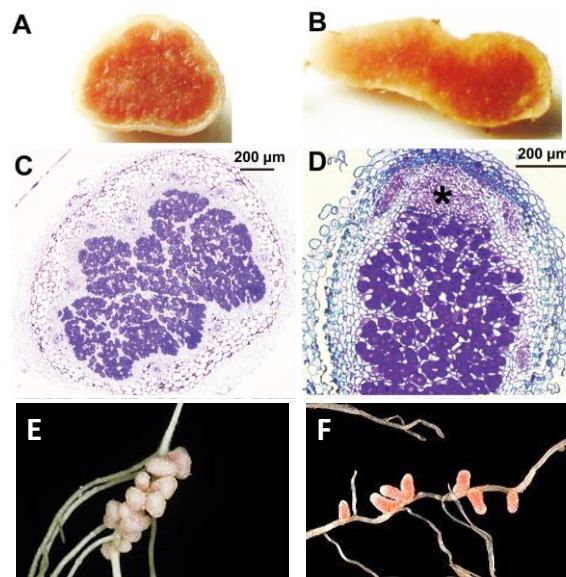
## Introducción

subtropicales como *P. vulgaris*, *G. max* y *Lotus japonicus* forman nódulos de tipo determinado. Los primeros poseen un meristema persistente en la zona apical que provee de nuevas células a los tejidos del nódulo a lo largo de su etapa de crecimiento y se caracterizan por su forma cilíndrica (Figura 4). El meristema es originado a partir de la tercera capa celular del córtex, mientras que algunas capas de la base del nódulo se originan de la división de células del córtex interno, periciclo y endodermis (Xiao et al., 2014). Por el contrario, los nódulos de tipo determinado se originan a partir de unas pocas divisiones celulares del córtex externo, pierden su meristema en estadios tempranos del desarrollo y se caracterizan por su forma esférica (Ferguson et al., 2010; Xiao et al., 2014) (Figura 4).

El inicio de la infección bacteriana y la organogénesis del nódulo ocurren de manera concomitante, pero distanciados por varias capas celulares. Se ha demostrado que estos eventos pueden ser separados genéticamente; es decir que puede llevarse a cabo la organogénesis del nódulo en ausencia de rizobios, y puede producirse la infección sin que haya formación de nódulos. Sin embargo, en presencia del rizobio estos dos procesos están finamente regulados y coordinados. El hilo de infección avanza desde la epidermis hacia el nódulo en formación o primordio del nódulo, donde las bacterias son internalizadas al citoplasma de la célula vegetal y rodeadas por una membrana derivada de la membrana citoplasmática vegetal denominada membrana del simbiosoma. Dentro de estas estructuras subcelulares tipo organelas, las bacterias se diferencian a bacteroides (Figura 3B). Los bacteroides formados en nódulos indeterminados y determinados presentan características distintivas. En los primeros, los simbiosomas contienen sólo un bacteroide que lleva a cabo la endoreduplicación de su genoma, generando alrededor de 24 cromosomas, en contraste con uno o dos de la bacteria de vida libre. La diferenciación del bacteroide es irreversible, por lo que se cree que son

## Introducción

incapaces de retornar a la vida libre. Por el contrario, en nódulos determinados, los simbiosomas contienen dos o más bacteroides que mantienen el número de cromosomas en uno o dos, de manera similar a la bacteria en estado saprofítico. En este caso, la diferenciación de los bacteroides es reversible, por lo que pueden volver a la vida libre. En última instancia, los bacteroides son capaces de transformar el  $N_2$  atmosférico en amonio gracias a la fuente de carbono y energía en forma de dicarboxilatos que le suministra la planta, tanto en nódulos determinados como indeterminados (Oldroyd et al., 2011).



**Figura 4. Nódulos determinados e indeterminados.**

**A-B.** Corte de nódulos fijadores de nitrógeno de *Vigna unguiculata* de tipo determinado (**A**) y *Leucaena leucocephala* de tipo indeterminado (**B**). **C-D.** Secciones ultrafinas de nódulos de *V. unguiculata* (**C**) y *L. leucocephala* (**D**). El asterisco indica la zona meristemática de los nódulos indeterminados de *L. leucocephala* (**D**). **E-F.** Nódulos fijadores de nitrógeno de *P. vulgaris* (de tipo determinado (**E**)) y *M. truncatula* de tipo indeterminado (**F**). Adaptado de (Li et al., 2013).

## 6. Vía de señalización de la nodulación

En los últimos veinte años, la comunidad científica abocada al estudio de la simbiosis fijadora de nitrógeno entre rizobios y leguminosas ha logrado enormes avances en la comprensión de los mecanismos de percepción por parte de la planta de los NFs

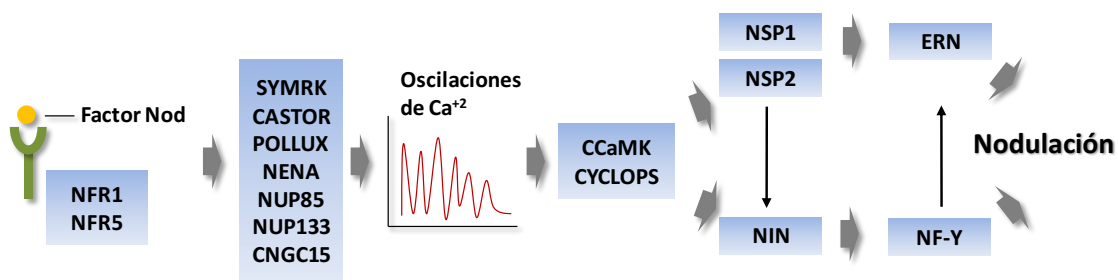
## Introducción

secretados por los rizobios y de las vías de señalización que tienen lugar en las células de la raíz luego de dicho reconocimiento. Estos avances fueron posibles principalmente debido a la identificación y caracterización genética de mutantes con pérdida o ganancia de función en las leguminosas modelo *L. japonicus* y *M. truncatula*.

Los NFs secretados por los rizobios son reconocidos por receptores de membrana de tipo Serina/Treonina quinasa con dominios extracelulares LysM (*lysin motif*), los cuales fueron denominados NFR1 (*Nod factor receptor 1*) y NFR5 (*Nod factor receptor 5*) en *L. japonicus* (Radutoiu et al., 2003; Madsen et al., 2003) (Figura 5). Mutantes de pérdida de función en estos genes tienen fenotipos no nodulantes y carecen de las respuestas tempranas típicas de la simbiosis nodular como la deformación de los pelos radicales luego de la inoculación con rizobios. Lo mismo ocurre en *M. truncatula* en mutantes de los genes ortólogos de NFR1 y NFR5 denominados LYK3 (*LysM domain receptor-like kinase 3*) y NFP (*Nod Factor Perception*), respectivamente (Limpens et al., 2003; Arrighi et al., 2006). La activación de estos receptores por los NFs desencadena oscilaciones de calcio en el núcleo y su periferia denominados *calcium spiking*, los cuales promueven los cambios transcripcionales requeridos para la infección y organogénesis del nódulo (Figura 5). Aún se desconocen los detalles del mecanismo por el cual la señal generada en la membrana plasmática como consecuencia de la percepción de los NFs es transmitida hacia el núcleo generando el *calcium spiking*. Sin embargo, se han identificado varios genes que son requeridos para que se produzcan las oscilaciones de los niveles de calcio. Un receptor de tipo quinasa con dominios LRR (*Leucine rich repeat*) denominado SYMRK (*Symbiosis Receptor Kinase*) en *L. japonicus* (Stracke et al., 2002) o DMI2 (*doesn't make infections*) en *M. truncatula* (Endre et al., 2002), el cual actúa corriente abajo de los receptores de los NFs, es esencial para la activación de dichas oscilaciones. Mutantes en SYMRK son incapaces

## Introducción

de nodular o formar hilos de infección, aunque retienen algunas respuestas frente a la inoculación con rizobios, como la deformación de pelos radicales. El ligando del dominio extracelular LRR de SYMRK aun no ha sido identificado, aunque se ha reportado que el dominio quinasa intracelular interacciona físicamente con la proteína 3-hidroxi-metilglutaril-CoA reductasa (Kevei et al., 2007). Dado que esta enzima interviene en la síntesis de mevalonato, se especula que este metabolito podría actuar como un segundo mensajero que transmita la percepción de los NFs hacia el núcleo (Figura 5).



**Figura 5. Vía de transducción de señales activada por el factor Nod.**

Los receptores de tipo LysM-RLK (*LysM domain receptor-like kinase*) NFR1 y NFR5 perciben los factores Nod por medio de su dominio extracelular LysM y transmiten la señal al citoplasma por medio de los dominios citoplasmáticos tipo quinasa, desencadenando oscilaciones de calcio en la región perinuclear. Para que éstas tengan lugar son requeridos SYMRK (una LRR-RLK), CASTOR y POLLUX (dos canales de potasio), NENA, NUP85 y NUP133 (componentes del poro nuclear), y CNGC15 (*cyclic nucleotide-gated channel*). Las oscilaciones de calcio serían decodificadas por la proteína quinasa dependiente de calcio y calmodulina (CCaMK), que fosforila al factor de transcripción CYCLOPS, iniciando una activación jerárquica de factores de transcripción que involucra a NSP1, NSP2, NIN, ERN y NF-Y, y finaliza con la reprogramación celular requerida para la nodulación.

Las oscilaciones de calcio producidas en el núcleo y su periferia se producen a partir del reservorio de calcio ubicado entre la membrana nuclear interna y externa, que forma un continuo con el retículo endoplasmático. La proteína calcio ATPasa MCA8 ubicada en la membrana nuclear es responsable de bombear el calcio nuevamente hacia el reservorio (Capoen et al., 2011). Además son requeridos los canales de potasio

## Introducción

denominados CASTOR y POLLUX en *L. japonicus* (Charpentier et al., 2008) y DMI1 (*doesn't make infections*) en *M. truncatula* (Ané et al., 2004), así como también tres componentes de las nucleoporinas específicos de la nodulación llamados NUP85, NUP133 y NENA (Saito et al., 2007; Kanamori et al., 2006; Groth et al., 2010), cuya función posiblemente esté relacionada con la localización de componentes de la vía de transducción de señales en la membrana nuclear. Recientemente se han identificado tres canales activados por nucleótidos cíclicos denominados CNGC15 *a*, *b* y *c* (*cyclic nucleotide-gated channel*) como los responsables de la liberación de calcio hacia el nucleoplasma, los cuales se localizan en la envoltura nuclear e interaccionan con el canal de potasio DMI1 (Charpentier et al., 2016) (Figura 5).

Una vez generadas las oscilaciones de calcio, estas son decodificadas por una proteína quinasa dependiente de calcio y calmodulina específica de plantas, denominada CCaMK en *L. japonicus* (*calcium and calmodulin-binding kinase*) (Levy et al., 2004) y DMI3 en *M. truncatula* (Mitra et al., 2004). Cuando es activada, CCaMK fosforila el factor de transcripción CYCLOPS (IPD3 en *M. truncatula*), iniciando una activación jerárquica de factores de transcripción que desencadenan los cambios transcripcionales requeridos para la simbiosis (Yano et al., 2008; Horváth et al., 2011; Messinese et al., 2007; Ovchinnikova et al., 2011). Mutaciones de ganancia de función de CCaMK o CYCLOPS resultan en la generación de nódulos espontáneos (es decir, en ausencia de rizobios) (Gleason et al., 2006a; Singh et al., 2014). Cuando CYCLOPS es fosforilado, se une al promotor del factor de transcripción NIN (*Nodule inception*), induciendo su expresión (Singh et al., 2014). NIN es requerido tanto para la formación de hilos de infección como para la organogénesis del nódulo (Schauser et al., 1999; Marsh et al., 2007). La regulación del promotor de NIN es compleja, ya que también requiere la presencia de dos factores de transcripción de la familia GRAS, designados NSP1

## Introducción

(*Nodulation signaling pathway 1*) y NSP2 (*Nodulation signaling pathway 2*). A su vez, NSP1 activa la transcripción del factor de transcripción ERN1 (*ERF Required for Nodulation 1*) (Middleton et al., 2007; Cerri et al., 2012), el cual es requerido para la inducción de nodulinas tempranas (genes que se inducen de manera temprana al producirse la interacción simbiótica). De manera secuencial, NIN induce la expresión de los factores de transcripción de tipo NF-Y (*Nuclear factor Y*), NF-YA1 y NF-YB1 (Soyano et al., 2013), los cuales promueven la reactivación del ciclo celular, reiniciando las divisiones mitóticas de las células del córtex requeridas para la organogénesis del nódulo (Figura 5).

Una gran incógnita que permanece hoy en día es cómo es transmitida la información generada en las células de la epidermis por la presencia de los NFs (que no difunden hacia el interior del tejido) hacia las células del córtex. Se ha propuesto que las citoquininas desempeñarían una función fundamental en este mecanismo, ya que mutaciones de ganancia de función en el receptor de citoquininas de *L. japonicus* LHK1 (*L. japonicus Histidine Kinase 1*) dan lugar a la formación de nódulos espontáneos. Por el contrario, mutantes de pérdida de función de LHK1 son incapaces de formar nódulos, aunque exhiben hilos de infección, algunos de los cuales progresan hacia el córtex, mientras otros se detienen o se desvían en las células epidérmicas (Gonzalez-Rizzo et al., 2006; Murray et al., 2007; Tirichine et al., 2007).

## 7. Funciones de los factores de transcripción NF-Y y SIN1 en la nodulación

Entre los factores de transcripción involucrados en la vía de transducción de señales activada por rizobios, los miembros de la familia NF-Y ocupan un lugar central y son requeridos tanto para la infección como para la organogénesis del nódulo (Zanetti et al., 2010; Combier et al., 2006; Laporte et al., 2013; Laloum et al., 2014; Baudin et al.,

## *Introducción*

2015). Los factores NF-Ys están compuestos por tres subunidades denominadas NF-YA, NF-YB y NF-YC y regulan la transcripción mediante su unión a secuencias *CCAAT* (*CCAAT-box*) ubicadas en las regiones promotoras de los genes sobre los que actúan. Mientras que en animales y levaduras cada subunidad del heterotrímero está codificada por un único gen, en plantas ha ocurrido una gran expansión y diversificación de estos genes, dando lugar a familias génicas que comprenden entre 8 y 39 miembros para cada subunidad según la especie (Laloum et al., 2013). Aunque las tres subunidades del heterotrímero son requeridas para la unión al *CCAAT-box*, la subunidad NF-YA es la que determina la especificidad de la unión y establece contacto directo con el DNA (Nardini et al., 2013; Romier et al., 2003). Las proteínas NF-YA se caracterizan por poseer dos dominios altamente conservados. El dominio A1 forma una alfa hélice involucrada en la interacción con las subunidades NF-YB y NF-YC, y el dominio A2 es requerido para la unión al *CCAAT-box* (Mantovani, 1999). Por otra parte, las subunidades B y C presentan dominios tipo histonas que permiten su dimerización y unión a NF-YA, y de manera análoga al dímero de histonas H2A y H2B, la unión inespecífica al ADN. Mientras que la familia NF-YB se caracteriza por la presencia de un dominio HFM (*histone fold motive*) semejante al que se encuentra en la histona H2B, la familia NF-YC presenta un dominio HFM relacionado con la histona H2A (Zanetti et al., 2017). Estudios de localización subcelular realizados en modelos animales y vegetales han determinado que mientras NF-YA se localiza en el núcleo, las subunidades NF-YB y NF-YC se localizan tanto en núcleo como citoplasma, siendo requerida su dimerización para la importación al núcleo (Laloum et al., 2014; Kahle et al., 2005). Una vez dentro del núcleo, el dímero puede unirse al *CCAAT-box* por medio de la interacción con NF-YA, o bien puede conformar trímeros funcionales con otros factores de transcripción y reconocer elementos reguladores diferentes. Recientemente,

## Introducción

Gnesutta y colaboradores han reportado que en *Arabidopsis thaliana* el factor de transcripción CONSTANS es capaz de formar un trímero funcional con las subunidades AtNF-YB2 y AtNF-YC3, reconociendo elementos *CCACA* (Gnesutta et al., 2017).

En el contexto de la nodulación, estudios genéticos y moleculares de distintas subunidades del complejo NF-Y lo han posicionado en un lugar central dentro de las redes transcripcionales activadas por NFs. En *L. japonicus* se ha reportado que el factor de transcripción NIN, un regulador positivo de la infección y organogénesis del nódulo, regula la expresión de LjNF-YA1 y LjNF-YB1 (Soyano et al., 2013). De manera consistente, en *M. truncatula* la inducción del gen *MtNF-YA1* en respuesta a los NFs es dependiente de NIN (Laloum et al., 2014). A su vez, las subunidades MtNF-YA1, MtNF-YC2 y MtNF-YB16 reconocen elementos *CCAAT* en el promotor del factor de transcripción ERN1 activando su expresión (Laloum et al., 2014; Baudin et al., 2015).

Más allá de su posicionamiento genético en la vía de transducción de señales, diversos estudios funcionales han demostrado el rol de los factores de transcripción NF-Y en la nodulación. El primer reporte surgió del estudio de la subunidad NF-YA1 de *M. truncatula* (MtNF-YA1) (Combiér et al., 2006). El silenciamiento postranscripcional de *MtNF-YA1* mediado por RNAi mostró que es requerido para la organogénesis del nódulo a través del control de la persistencia del meristema. La expresión del gen MtNF-YA1 se encuentra limitada a la zona meristemática mediante su regulación postranscripcional por el microRNA169a (miR169a). Consistentemente, el silenciamiento de NF-YA1 de *L. japonicus*, LjNF-YA1, (el mejor homólogo de MtNF-YA1), también muestra defectos en la organogénesis del nódulo a pesar de no poseer un meristema persistente, ya que regula las divisiones celulares y la expresión de la ciclina B1 (CYCB1; *cyclin* B1) (Soyano et al., 2013). Posteriormente, Laporte y colaboradores demostraron mediante el estudio de una mutante nula que MtNF-YA1 interviene en las



## Introducción

etapas tempranas de la interacción simbiótica, siendo requerido para la progresión de los hilos de infección (Laporte et al., 2014). Este fenotipo es más pronunciado cuando se silencian de manera simultánea MtNF-YA1 y su parálogo más cercano MtNF-YA2, mostrando que ambas subunidades presentan redundancia génica parcial (Laloum et al., 2014).

En *P. vulgaris*, la subunidad PvNF-YC1 controla no sólo la infección y organogénesis del nódulo, sino también la respuesta relativa de la planta frente a la inoculación con diferentes cepas de rizobios. Plantas silenciadas en PvNF-YC1 muestran una menor frecuencia de hilos de infección que plantas control, y aquellos que se forman se encuentran muy comprometidos en su progresión, abortando antes de alcanzar el córtex radical. Sumado a esto, el silenciamiento de PvNF-YC1 afecta drásticamente tanto el número de nódulos formados como su desarrollo, posiblemente mediante la regulación de genes del ciclo celular de transición G2/M. Más allá de estar implicado en la fisiología básica de la nodulación, los niveles de expresión de NF-YC1 determinan el éxito competitivo y la eficiencia de nodulación de distintas cepas de rizobio (Zanetti et al., 2010) (fenómeno que se explica en detalle en la siguiente sección).

Hasta el momento no hay reportes donde se le asigne a NF-YB una función directa en la nodulación, aunque hay evidencias que sugieren que está implicado en este proceso. Soyano y colaboradores determinaron que el factor de transcripción NIN regula la expresión de LjNF-YA1 y LjNF-YB1, y que ambas subunidades interaccionan *in planta* (Soyano et al., 2013). Consistentemente, el ortólogo de LjNF-YB1 en *M. truncatula* (MtNF-YB16) forma un trímero con las subunidades MtNF-YA1/2 y MtNF-YC1/2, y a su vez, en *P. vulgaris*, se forma el trímero NF-Y entre las subunidades PvNF-YB7, PvNF-YC1 y PvNF-YA1 (Baudin et al., 2015). Estos resultados sugieren que

## *Introducción*

subunidades particulares de cada familia conforman un trímero simbiótico evolutivamente conservado en leguminosas.

En *P. vulgaris*, PvNF-YC1 es capaz de interactuar con un factor de transcripción de la familia GRAS (*Gibberellic Acid-Insensitive (GAI), Repressor of GAI, and Scarecrow*) denominado SIN1 (*Scarecrow-like13 Involved in Nodulation*) (Battaglia et al., 2014). Estudios funcionales realizados mediante genética reversa han determinado que SIN1 participa en la infección rizobiana y la organogénesis del nódulo, así como también en el desarrollo de las raíces laterales. Estos resultados sugieren una interesante vinculación genética entre estos dos programas de desarrollo. Cabe destacar que, al igual que en el caso de *PvNF-YC1*, el silenciamiento de *SIN1* afecta la activación transcripcional en respuesta al rizobio de genes del ciclo celular de la transición G2/M (Battaglia et al., 2014).

## **8. La asociación simbiótica preferencial establecida entre variedades de *Phaseolus vulgaris* y cepas de *Rhizobium etli***

Los análisis filogenéticos han sugerido que *P. vulgaris* surgió en América hace aproximadamente 4 millones de años (Delgado-Salinas et al., 2006); sin embargo intensos debates han tenido lugar acerca de la región geográfica exacta donde esto ha ocurrido. Estudios más recientes basados en análisis genómicos proponen la región Mesoamericana como el centro de origen de *P. vulgaris*, a partir del cual ocurrió una radiación de la población hacia la región Andina hace aproximadamente 165.000 años, generando dos *pooles* génicos silvestres aislados geográfica y genéticamente (Bitocchi et al., 2012, 2017; Schmutz et al., 2014; Rendón-Anaya et al., 2017). A partir de estos *pooles* genéticos silvestres, hace aproximadamente 8.000 años, *P. vulgaris* fue domesticado de manera independiente en lo que hoy es México y la región Andina de

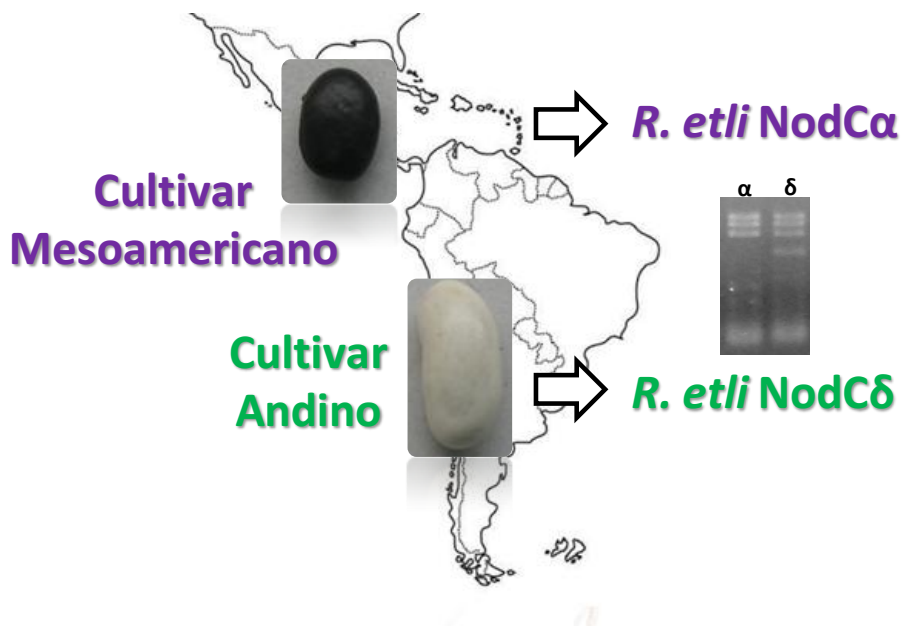
## Introducción

Sudamérica (Bitocchi et al., 2013; Gepts et al., 1986; Mamidi et al., 2013). La existencia de variedades de *P. vulgaris* silvestres y domesticadas provenientes de dos centros geográficos independientes hacen de esta especie un modelo ideal para el estudio de aspectos evolutivos de su interacción simbiótica con bacterias fijadora de nitrógeno, más aun teniendo en cuenta que en los últimos años se ha generado una gran cantidad de información genómica y transcriptómica de variedades correspondientes a cada uno de los *pooles* genéticos. Actualmente se encuentra disponible la secuencia genómica de cultivares Andinos (Schmutz et al., 2014) y Mesoamericanos (Vlasova et al., 2016), así como también un atlas de la expresión génica (O'Rourke et al., 2014).

*Rhizobium etli* es la especie de rizobio predominantemente asociada a las variedades de *P. vulgaris* tanto silvestres como domesticadas (Nour et al., 2001; Aguilar et al., 2001, 1998). El estudio de las poblaciones de *R. etli* basado en el polimorfismo del gen *nodC*, realizado mediante la técnica de ARFLP (*amplified restriction fragment lenght polimorfism*), permitió determinar que las poblaciones de *R. etli* presentes en suelos del centro de diversificación Mesoamericano y Andino son diferentes. En suelos Mesoamericanos predominan las cepas con el polimorfismo del gen *nodC* tipo  $\alpha$  (cepas tipo  $-\alpha$  de acá en adelante), mientras que en suelos Andinos predominan cepas que llevan el alelo *nodC* tipo  $\delta$  (cepas tipo  $-\delta$ ), lo que se correlaciona con el tipo de alelo presente en los microsimbiontes aislados de nódulos de cultivares de cada centro de diversificación (Aguilar et al., 2004) (Figura 6). Sumado a esto, las variedades Mesoamericanas desarrollan una mayor cantidad de nódulos cuando son inoculadas con cepas tipo *nodC*- $\alpha$  que con cepas tipo *nodC*- $\delta$ , mientras que en las variedades Andinas ocurre el fenómeno inverso (Meschini et al., 2008). Estos datos sugieren fuertemente que planta y bacteria han co-evolucionado de manera independiente en cada centro de diversificación desarrollando mecanismos de reconocimiento mutuo. Más aun, en

## Introducción

experimentos de co-inoculación con cepas tipo  $-\alpha$  y  $-\delta$ , las variedades Mesoamericanas nodulan casi exclusivamente con cepas tipo  $-\alpha$  (Aguilar et al., 2004). El estudio de las bases moleculares de este mecanismo de alta especificidad permitió determinar que la expresión del factor de transcripción PvNF-YC1 es determinante para la selección de cepa. Las cepas tipo  $-\alpha$  son capaces de inducir fuertemente la expresión de PvNF-YC1 en cultivares Mesoamericanos determinando su éxito competitivo frente a cepas tipo  $-\delta$ . De manera consistente, la expresión ectópica de PvNF-YC1 en raíces transgénicas de *P. vulgaris* iguala la capacidad competitiva de ambas cepas en experimentos de co-inoculación, sugiriendo que PvNF-YC1 sería parte de los mecanismos moleculares que determinan la preferencia de cepa en la formación de nódulos (Zanetti et al., 2010).



**Figura 6. Interacción simbiótica entre *P. vulgaris* y *R. etli*.**

*P. vulgaris* ha diversificado en dos centros genéticos aislados, la región Mesoamericana y la Andina. Los cultivares de origen Mesoamericano y Andino establecen una asociación simbiótica más eficiente con cepas de *R. etli* portadoras del polimorfismo en el gen *nodC* de tipo  $-\alpha$  y  $-\delta$ , respectivamente. A su vez, las variedades Mesoamericanas son noduladas preferencialmente por cepas que portan el alelo *nodC*-  $\alpha$ .

## *Introducción*

Por lo expuesto anteriormente, la interacción simbiótica fijadora de nitrógeno entre *P. vulgaris* y *R. etli* surge como un excelente modelo para el estudio de las bases moleculares de esta sofisticada interacción. No solamente se trata de un cultivo de interés agronómico y regional, sino que desde el punto de vista básico *P. vulgaris* cuenta con una gran biodiversidad de variedades silvestres y domesticadas que han co-evolucionado con cepas de *R. etli* de manera independiente en los distintos centros de diversificación genética. Esta característica pone a este sistema biológico en un lugar único para comprender los mecanismos por los cuales la planta reconoce y distingue diferentes cepas de rizobios que muestran una diferente eficiencia en la formación de nódulos.

# **HIPÓTESIS Y OBJETIVOS**

## **1. Hipótesis de trabajo y objetivo general**

A escala global la agricultura es una práctica no sustentable que depende fuertemente de la incorporación de fertilizantes sintéticos para lograr los rendimientos que la población mundial demanda. En este contexto, la fijación biológica de nitrógeno surge como una alternativa sustentable y se posiciona como un tema de estudio científico sumamente relevante.

Si bien en los últimos años el estudio de las leguminosas modelo *M. truncatula* y *L. japonicus* ha permitido dilucidar los principales componentes de la vía de transducción de señales requeridos para la nodulación, estamos lejos de conocer la totalidad de sus componentes y menos aun cuáles son los mecanismos regulatorios subyacentes. La mayor parte de los trabajos publicados en el área han aplicado modelos experimentales donde se analiza la respuesta de una planta modelo frente a una cepa de rizobio de referencia. Por este motivo, el entendimiento de los mecanismos de reconocimiento y selectividad que presenta la planta frente a distintos microorganismos en un ambiente tan diverso y complejo como la rizósfera ha quedado relegado. Como se ha mencionado en la introducción, la interacción *P. vulgaris* – *R. etli* es un excelente modelo para este tipo de estudios debido a la co-evolución que ha ocurrido entre ambos simbioses en dos centros de diversificación geográficamente aislados. Sumado a esto, en nuestro laboratorio se ha identificado y estudiado funcionalmente el factor de transcripción NF-Y, el cual es determinante para la especificidad de cepa observada en cultivares de poroto Mesoamericanos, siendo éste un excelente punto de partida para estudios moleculares.

La hipótesis general de este trabajo postula que las variedades Mesoamericanas y Andinas de *P. vulgaris* son capaces de reconocer y discriminar entre cepas de rizobio de distinto linaje evolutivo desencadenando cambios transcripcionales diferenciales que

## *Hipótesis y objetivos*

determinan el grado de eficiencia de la interacción simbiótica, siendo el complejo NF-Y y otras proteínas asociadas reguladores importantes de dicha reprogramación de la expresión génica.

El objetivo del presente trabajo es caracterizar los mecanismos moleculares y vías de señalización que controlan la reprogramación transcripcional que determinan una asociación simbiótica más eficiente entre variedades de *P. vulgaris* y *R. etli*.

### **2. Objetivos específicos**

Para contrastar la hipótesis planteada se propusieron los siguientes objetivos específicos:

1. Analizar el transcriptoma de variedades Andinas y Mesoamericanas en respuesta a cepas de *R. etli* representativas de ambos centros de diversificación.
2. Identificar genes regulados a nivel transcripcional por el complejo NF-Y y el factor de transcripción SIN1 en etapas tempranas de la asociación simbiótica entre *P. vulgaris* y *R. etli*.
3. Caracterizar funcionalmente nuevos componentes moleculares que participan de la vía de transducción de señales vinculada al factor de transcripción NF-YC1.



# CAPÍTULO I

## Resultados y discusión

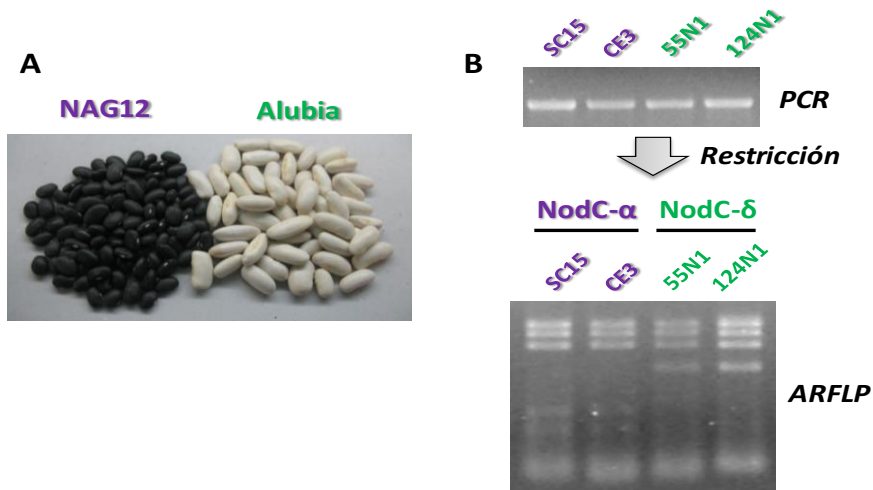
**Análisis transcriptómico de *P. vulgaris* en respuesta a cepas de *R. etli***

## RESULTADOS

---

### 1. Diseño experimental y generación del material biológico

Con el objetivo de caracterizar el transcriptoma asociado a una interacción simbiótica eficiente, se utilizaron las variedades de *P. vulgaris* NAG12 y Alubia como representantes de los *pooles* genéticos Mesoamericanos y Andinos, respectivamente. La variedad NAG12 se caracteriza por el color negro de su semilla, mientras que la variedad Alubia presenta un grano de color blanco y un tamaño relativo mayor (Figura 7A). El comportamiento de ambos cultivares durante la interacción simbiótica con diferentes cepas de *R. etli* fue caracterizado previamente en nuestro laboratorio, revelando que NAG12 forma un mayor número de nódulos con cepas que portan el alelo *nodC- $\alpha$*  que con aquellas que llevan el alelo *nodC- $\delta$* , mientras que Alubia nodula más eficientemente con la cepa *nodC- $\delta$*  (Meschini et al., 2008; Mazziotta et al., 2013). Respecto al microsimbionte, se seleccionaron las cepas de *R. etli* SC15 y CE3 como representativas de cepas portadoras del alelo *nodC- $\alpha$* , mientras que las cepas 55N1 y 124N1 se utilizaron como representativas de cepas con el alelo *nodC- $\delta$* . Previo a la realización de los ensayos se verificó la presencia de los alelos esperados en el gen *nodC* para cada una de las cepas utilizadas mediante un análisis de polimorfismo de fragmentos de restricción de productos amplificados (ARFLP; Figura 7B).

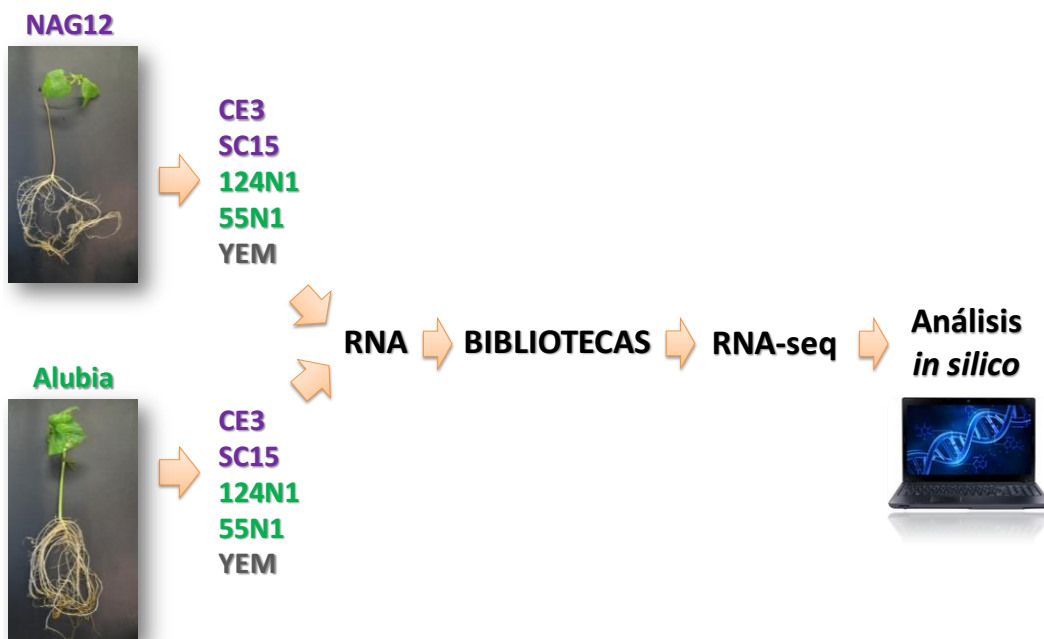


**Figura 7. Sistema biológico utilizado.**

**A.** Se utilizaron las variedades NAG12 y Alubia como representantes de cultivares Mesoamericanos y Andinos de *P. vulgaris*, respectivamente. **B.** Polimorfismo del gen *nodC* para las cepas tipo  $\alpha$  (SC15 y CE3) y tipo  $\delta$  (55N1 y 124N1). Un fragmento del gen *nodC* fue amplificado por PCR (panel superior) y luego digerido con la enzima de restricción *Hinf*I. El polimorfismo de los fragmentos de restricción de productos amplificados (ARFLP) fue analizado mediante electroforesis en gel de agarosa (panel inferior).

Nuestra hipótesis de trabajo propone que plantas y bacterias provenientes del mismo centro de diversificación genética han desarrollado mecanismos de reconocimiento mutuo en las primeras etapas de la interacción simbiótica que desencadenan cambios transcripcionales específicos en cada combinación cultivar x cepa, los cuales determinan la eficiencia de la asociación simbiótica. Por este motivo, nuestro diseño experimental consistió en inocular los cultivares NAG12 y Alubia con cada una de las cepas seleccionadas (SC15, CE3, 55N1, 124N1) o con el medio de crecimiento bacteriano (YEM) como control de inoculación (Figura 8). La recolección del tejido se llevó a cabo a las 24 horas post inoculación (hpi), momento en el cual aún no son visibles los hilos de infección, pero sí se producen respuestas morfológicas como la deformación de los pelos radicales y se inducen las respuestas moleculares asociadas a la infección y organogénesis del nódulo (Meschini et al., 2008; Blanco et al., 2009; Zanetti et al., 2010). Este momento que precede la infección bacteriana constituye una

etapa crítica en la interacción simbiótica ya que determina la compatibilidad entre los simbiontes y condiciona la competitividad entre las cepas (Meschini et al., 2008; Gage and Margolin, 2000). Con el fin de llevar adelante un análisis estadístico robusto y estimar la reproducibilidad de los datos obtenidos de la secuenciación, se generaron 3 réplicas biológicas por condición en cada una de las cuales se combinó tejido de raíz proveniente de 5 plantas para homogeneizar la variación biológica de diferentes individuos.



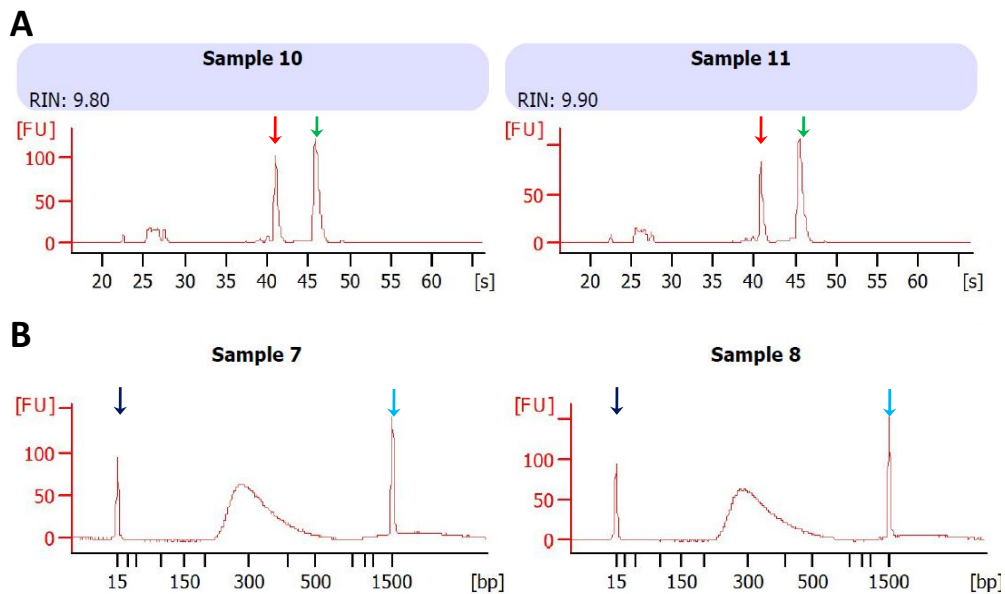
**Figura 8. Esquema del diseño experimental utilizado.**

Plantas de *P. vulgaris* NAG12 y Alubia fueron inoculadas con el medio de crecimiento bacteriano (YEM) como control o con las cepas de *R. etli* SC15, CE3, 55N1 y 124N1. Al cabo de 24 hs se colectó el tejido, se extrajo el RNA total de tres réplicas biológicas por condición, se purificó el RNA poliadenilado y se construyeron las bibliotecas para secuenciación masiva de RNA (RNA-seq). Los datos generados fueron analizados *in silico* para identificar genes con expresión diferencial.

## 2. Construcción de las bibliotecas para RNA-seq, secuenciación y controles de calidad

Para llevar a cabo la construcción de las bibliotecas para la secuenciación masiva se realizaron extracciones de RNA total a partir de las muestras descritas en la sección

anterior. La integridad, calidad y concentración del RNA obtenido fue analizado mediante electroforesis capilar en un *Agilent Technologies 2100 Bioanalyzer*. En todos los casos se obtuvieron muestras de RNA con un número de integridad de RNA (RIN: *RNA integrity number*) superior o igual a 8, que es el valor mínimo recomendado por la compañía proveedora del kit para la construcción de bibliotecas (Illumina). Ejemplos de los electroferogramas obtenidos para las muestras de RNA se ilustran en la Figura 9A.

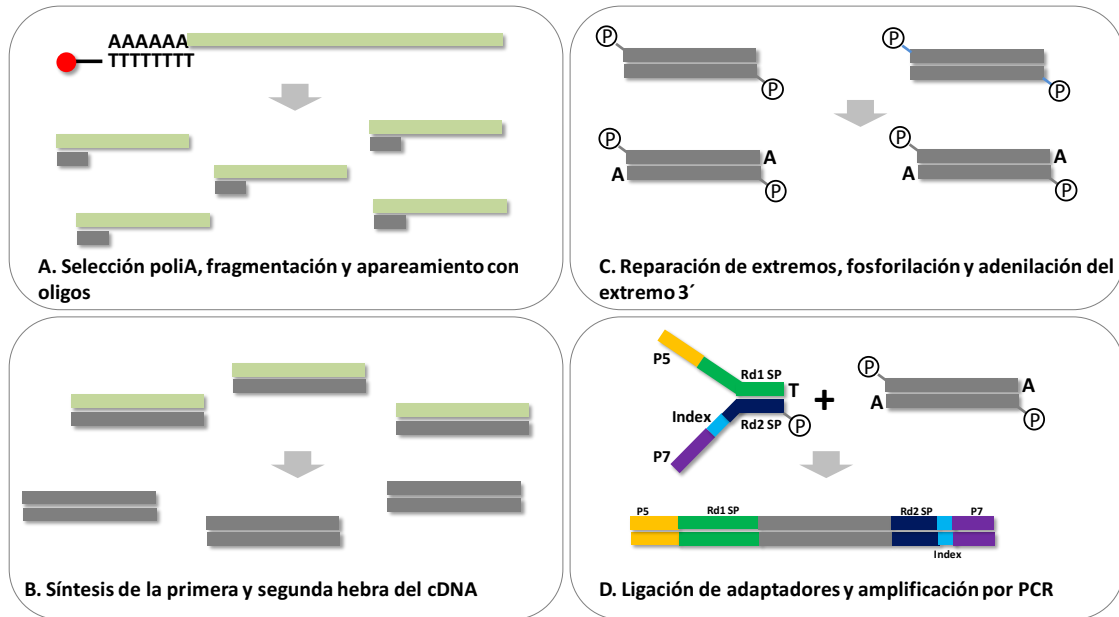


**Figura 9. Control de calidad del RNA y las bibliotecas de cDNA generadas.**

**A.** Electroferogramas del RNA total extraído. El valor de RIN (*RNA integrity number*) indica la calidad e integridad del RNA. Las flechas rojas y verdes indican la posición de los RNA ribosomales 18S y 28S, respectivamente. **B.** Electroferograma mostrando la concentración y distribución de tamaños de las bibliotecas de cDNA, que como era esperado, se encuentran centrados alrededor de los 300 pb. Los picos indicados con flechas azules y celestes corresponden al marcador inferior y superior, respectivamente. Tanto el RNA como las bibliotecas de cDNA fueron analizadas mediante electroforesis capilar en el equipo *Agilent Technologies 2100 Bioanalyzer*. Se muestran ejemplos de dos muestras de RNA y dos bibliotecas de cDNA.

A continuación se procedió a la síntesis de las bibliotecas de DNA complementario (cDNA) requeridas para RNA-seq utilizando la tecnología de Illumina. La construcción de las mismas involucra la purificación de los RNA poliadenilados, su fraccionamiento y síntesis de cDNA. Posteriormente, se realiza la reparación de los extremos del cDNA,

la adenilación de los extremos 3', se ligan adaptadores específicos y por último, la biblioteca es amplificada por PCR (Figura 10).



**Figura 10. Esquema de la síntesis de cDNA para RNAseq.**

RNAs poliadenilados fueron purificados con bolillas magnéticas conjugadas a oligo dT. El RNA purificado fue tratado con cationes divalentes y altas temperaturas para generar fragmentos de aproximadamente 200 pb, que luego fueron incubados con oligos aleatorios (A). Síntesis de la primera y segunda hebra de cDNA (B). Los fragmentos de cDNA fueron procesados por reparación de extremos, fosforilación en el extremo 5' y adenilación de los extremos 3' (C). Por último, se ligaron adaptadores específicos y se amplificaron las bibliotecas por PCR (D). Las regiones P5 y P7 son necesarias para la adhesión de los fragmentos a las celdas durante la secuenciación. Rd1SP y Rd3SP son regiones complementarias a los *primers* de secuenciación. La secuencia del *index* permite identificar de qué muestra proviene el fragmento secuenciado.

El análisis cuali y cuantitativo de las 30 bibliotecas generadas fue realizado mediante electroforesis capilar en el equipo *Agilent Technologies 2100 Bioanalyzer*. Los electroferogramas obtenidos mostraron que todas las bibliotecas presentaron una distribución de tamaños con un pico máximo de entre 200 y 300 nucleótidos, tal como se ilustra en los dos ejemplos presentados en la Figura 9B. Esta distribución de tamaño es coincidente con lo recomendado por los fabricantes del kit (Illumina). Por último, las muestras fueron enviadas al servicio de secuenciación de la Universidad de California

## Capítulo I

Davis, donde fueron *pooleadas* de a 12 y posteriormente secuenciadas en la plataforma HiSeq 4000 con lecturas simples de 90 pb. Cada biblioteca generó más de 20 millones de lecturas, con más del 92% de las bases presentando un valor de Q (*Phred quality score*) mayor o igual a 30, resultando en un valor de Q promedio superior a 38 (Tabla 2). El valor Q es el método más utilizado para indicar la calidad y precisión de la secuenciación. Dado que el valor de Q30 indica un *score* de calidad equivalente a una precisión del 99,9%, los resultados muestran que los datos de secuenciación fueron confiables. Igualmente se decidió realizar un análisis de calidad más exhaustivo mediante la herramienta FastQC disponible en la plataforma Galaxy, la cual calcula los siguientes parámetros: valor medio de Q para cada base, índice de calidad por secuencia, el contenido de GC por secuencia, el porcentaje de cada nucleótido para cada posición en la secuencia, el contenido de bases no identificadas (N) por secuencia, la distribución de tamaños de las secuencias, cantidad de secuencias duplicadas, y la existencia de secuencias y Kmers sobrerrepresentadas (Andrews 2010). Para cada una de las muestras los parámetros calculados fueron los esperados para el tipo de secuenciación realizado.

Como puede apreciarse en la Tabla 2, de las 30 bibliotecas generadas sólo fueron secuenciadas 24. La justificación de este número se encuentra en el limitado presupuesto, el cual sólo alcanzaba para solventar los costos de dos líneas del secuenciador, y a su vez en que las bibliotecas fueron construidas con 12 *indexes* distintos, limitando a 12 la cantidad máxima de muestras por línea. Como consecuencia se realizó la secuenciación masiva de todas las muestras del cultivar NAG12, pero del cultivar Alubia sólo se secuenciaron dos réplicas de las muestras inoculadas con YEM, SC15, CE3, 124N1, y sólo una réplica con 55N1. Está previsto realizar la secuenciación

## Capítulo I

de las muestras restantes en el futuro cercano para mejorar la robustez estadística del análisis en el cultivar Andino.

**Tabla 2. Resultados de la secuenciación masiva de RNA para cada muestra.**

Muestra	# de lecturas	Lecturas alineadas	% de bases $\geq Q30$	Promedio de Score de calidad
ALUBIA 124N1 #2	28,208,329	88.50%	93.26	38.59
ALUBIA124N1 #3	37,432,552	88.50%	92.71	38.43
ALUBIA 55N1 #3	32,325,155	88.00%	93.02	38.52
ALUBIA CE3 #1	29,760,286	93.40%	95.91	39.28
ALUBIA CE3 #3	22,334,239	88.80%	93.03	38.52
ALUBIA SC15 #1	39,750,228	88.30%	92.93	38.49
ALUBIA SC15 #2	34,875,352	93.40%	95.82	39.26
ALUBIA YEM #1	39,565,180	88.30%	92.98	38.51
ALUBIA YEM #2	36,105,074	88.30%	93.32	38.61
NAG12 124N1 #1	37,002,663	89.20%	95.85	39.26
NAG12 124N1 #2	25,133,179	85.30%	93.32	38.6
NAG12 124N1 #3	37,398,117	88.00%	95.93	39.29
NAG12 55N1 #1	29,632,617	84.80%	93.1	38.54
NAG12 55N1 #2	39,102,099	87.70%	96	39.31
NAG12 55N1 #3	26,047,265	90.30%	95.9	39.27
NAG12 CE3 #1	41,795,672	89.60%	95.97	39.29
NAG12 CE3 #2	33,533,384	88.10%	95.89	39.27
NAG12 CE3 #3	39,333,995	82.80%	93.35	38.61
NAG12 SC15 #1	23,555,214	84.70%	93.15	38.56
NAG12 SC15 #2	24,565,550	88.30%	95.89	39.28
NAG12 SC15 #3	22,814,435	89.30%	95.68	39.21
NAG12 YEM #1	26,507,459	89.30%	95.87	39.27
NAG12 YEM #2	42,519,480	93.30%	96.02	39.32
NAG12 YEM #3	28,448,831	85.40%	93.26	38.59

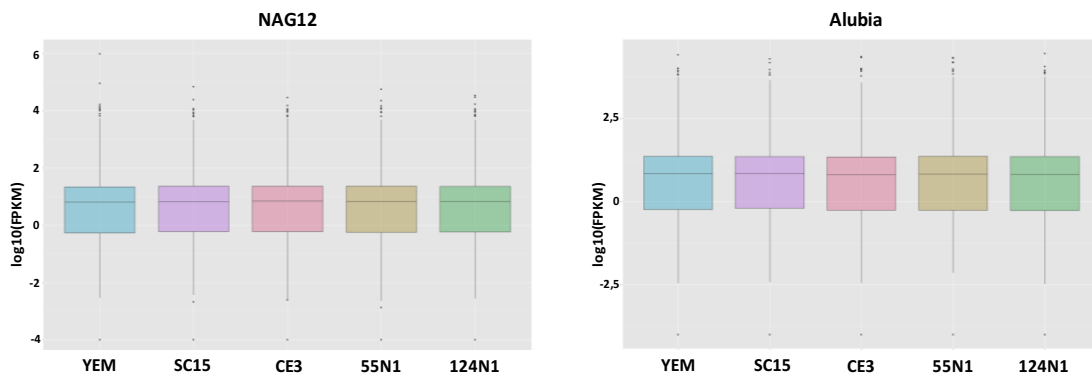
El análisis bioinformático de los datos de RNA-seq obtenidos fue realizado a través de la plataforma Galaxy aplicando el protocolo Tuxedo (Trapnell et al., 2012). El primer paso consistió en el alineamiento de las lecturas de RNA-seq al genoma de referencia de *P. vulgaris* (*Phaseolus vulgaris* v2.1, DOE-JGI y USDA-NIFA, <http://phytozome.jgi.doe.gov/>) utilizando el programa *TopHat2* (Kim et al., 2013). Este programa mapea las lecturas al genoma de referencia por medio del algoritmo *Bowtie* e identifica sitios de *splicing* utilizando para ello lecturas que mapean en dos exones



contiguos (*exon junctions*). Mediante esta herramienta el 88,4 % de las lecturas en promedio fue alineada al genoma de referencia de *P. vulgaris* (Tabla 2).

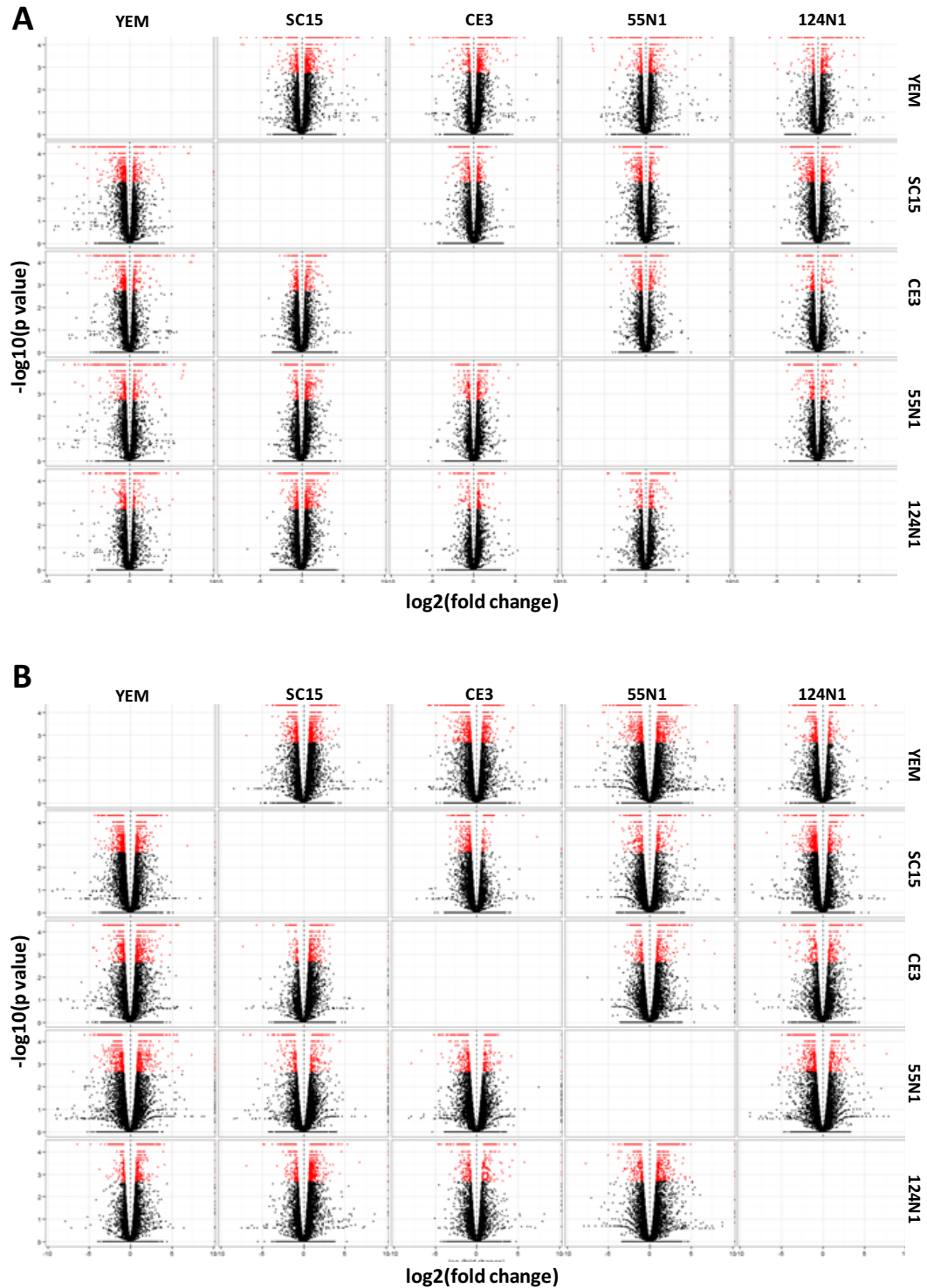
A partir de los archivos generados por *TopHat2* conteniendo las lecturas alineadas al genoma de *P. vulgaris* se realizó el ensamblado de los transcritos y se estimó su abundancia en FPKM (*Fragments Per Kilobase of transcript per Million mapped reads*) utilizando el programa *Cufflinks*. Los archivos de *Cufflinks* de cada muestra fueron fusionados mediante la herramienta *Cuffmerge*, la cual proporciona una base uniforme para calcular la expresión de genes y transcritos en cada condición. Por último, se realizó un análisis con *Cuffdiff* para identificar genes con expresión diferencial (Trapnell et al., 2011).

A partir de los resultados del análisis por *Cuffdiff* se utilizó la herramienta *Cummerbund* para determinar la mediana y la dispersión de los valores de FPKM obtenidos para cada muestra mediante gráficas de *boxplot* (Figura 11) (Trapnell et al., 2013). Los resultados muestran que la distribución de los valores de FPKM fue similar entre las distintas muestras. Por otro lado, se realizaron diagramas *volcano* para visualizar la presencia de genes con expresión diferencial y significancia estadística ( $p < 0,05$ ) en las comparaciones múltiples a una escala global (Figura 12).



**Figura 11. Distribución de los valores de FPKM obtenidos a partir de los datos de RNA-seq.**

*Box plots* de las distribuciones de los valores de FPKM en cada muestra. Los análisis fueron realizados con *CummeRbund* en lenguaje R.

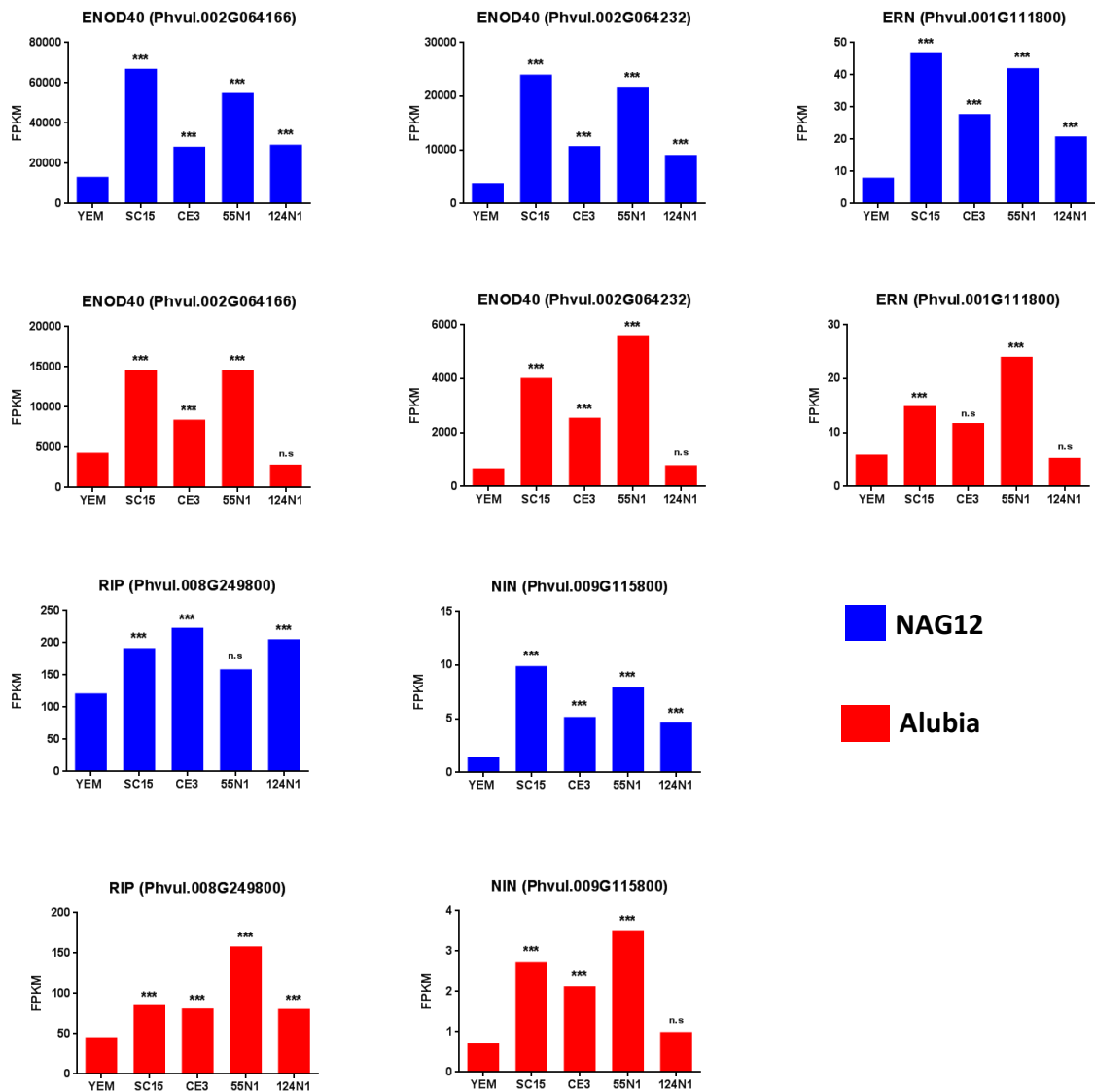


**Figura 12.** Cambios en el transcriptoma de muestras de raíz inoculadas con YEM y las cepas de *R. etli* SC15, CE3, 55N1 y 124N1.

Se muestran los cambios transcripcionales en el cultivar NAG12 (A) y Alubia (B). Cada punto representa un gen. El eje x corresponde al logaritmo en base 2 ( $\log_2$ ) del cambio del valor absoluto entre las muestras comparadas, mientras que en el eje y se representa el  $-\log_{10}$  del valor de probabilidad ( $p$ ). Los puntos con color rojo refieren a genes con diferencias de expresión estadísticamente significativas ( $p < 0,05$ ). Los puntos ubicados en la derecha de cero en el eje x son genes regulados positivamente mientras que los genes a la izquierda de cero corresponden a genes regulados negativamente.

## *Capítulo I*

En conclusión, esta serie de análisis demuestran que los datos de secuenciación generados superan los requerimientos de calidad mínimos y permiten la identificación de genes con expresión diferencial entre las distintas condiciones. Para corroborar que efectivamente haya ocurrido la reprogramación transcripcional desencadenada por la interacción simbiótica con rizobios en las muestras sometidas a RNA-seq, se analizaron los valores de FPKM de los genes marcadores de la nodulación a etapas tempranas *ENOD40*, *ERN*, *RIP* y *NIN* (Figura 13). Los resultados muestran que las cepas de *R. etli* SC15, CE3, 55N1 y 124N1 fueron capaces de incrementar los niveles de transcritos de los genes marcadores de las etapas tempranas de la nodulación tanto en el cultivar NAG12 como en Alubia (Figura 13).



**Figura 13. Niveles de expresión de genes marcadores de etapas tempranas de la nodulación.**

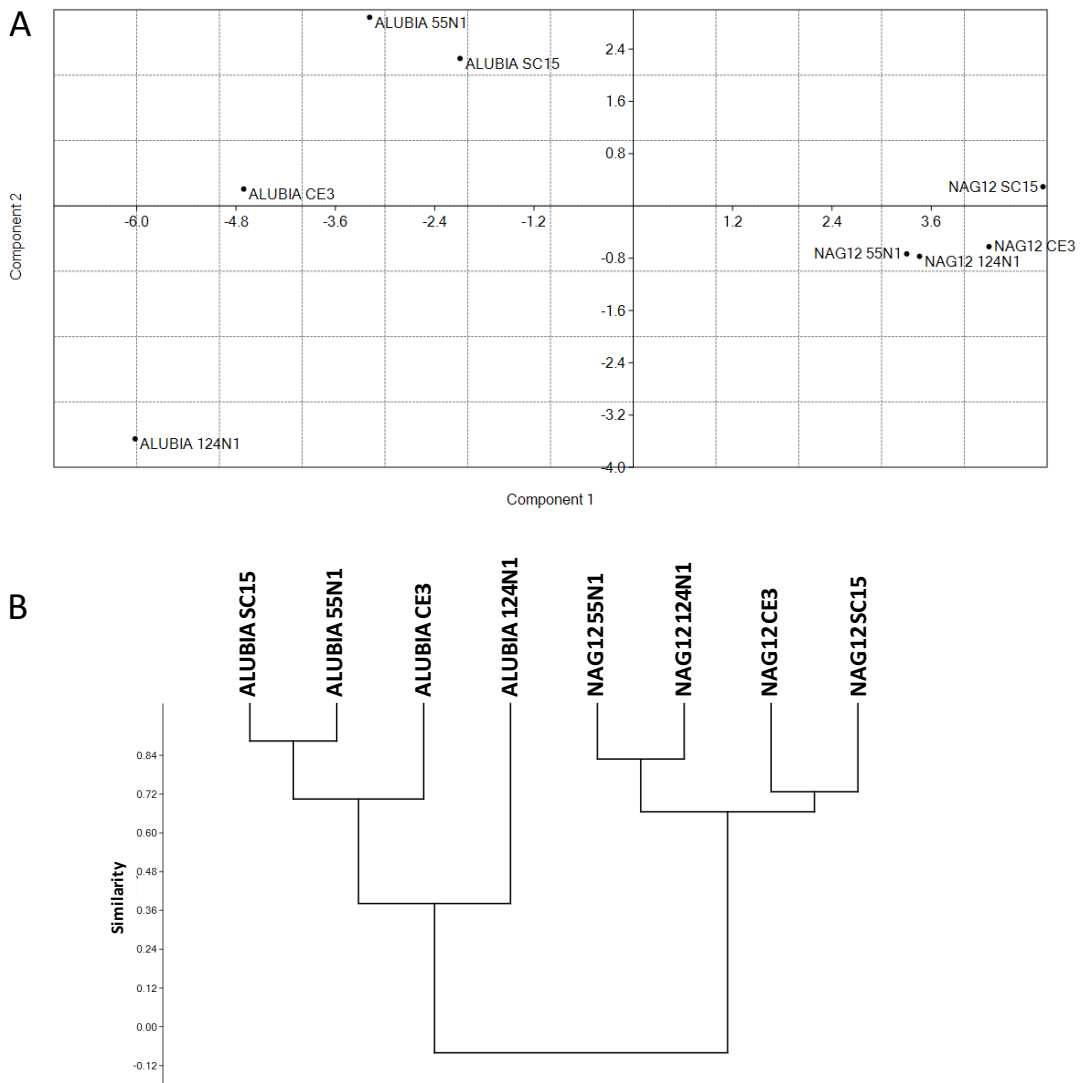
Se muestran los valores promedio de las réplicas biológicas en FPKM (*Fragments Per Kilobase Million*) para los genes *ENOD40*, *ERN*, *RIP* y *NIN* en el cultivar NAG12 (barras azules) y Alubia (barras rojas). Los asteriscos indican diferencias significativas respecto a YEM (\*,  $q < 0,05$ ; \*\*,  $q < 0,01$ ; \*\*\*,  $q < 0,001$ ).

### 3. Relación entre los perfiles transcriptómicos

Uno de los aspectos evolutivos interesantes de este trabajo, apunta a caracterizar cómo han divergido los cambios en los perfiles transcriptómicos inducidos en respuesta al rizobio en dos cultivares que han sufrido 165.000 años de aislamiento geográfico y genético. Para alcanzar este objetivo, se realizó un análisis de componentes principales

## Capítulo I

(PCA) de los datos de expresión obtenidos para reducir la dimensionalidad de los datos (Figura 14A). El componente 1 y 2 explican el 76,23 % y el 15,28 % de la varianza, respectivamente. Por lo tanto, mediante la reducción de las variables a estos dos componentes se puede explicar el 91,51 % de la varianza de los datos. El PCA muestra que los perfiles transcriptómicos de las plantas inoculadas tienden a agruparse por el cultivar y no por el tipo de cepa utilizada en la inoculación (Figura 14A).



**Figura 14. Los perfiles transcriptómicos de plantas inoculadas se agrupan principalmente según el cultivar.**

**A.** Análisis de componentes principales (PCA) utilizando el método de correlación. **B.** Análisis de *clustering* jerárquico por el método de correlación. Para realizar ambos análisis se utilizaron los valores del log<sub>2</sub> de los FPKM de cada gen normalizados por el valor en YEM.

## Capítulo I

De manera consistente, el análisis de *clustering* jerárquico también muestra que los transcriptomas de plantas inoculadas se agrupan por cultivar (Figura 14B). En las muestras NAG12 puede observarse que los perfiles transcriptómicos de plantas inoculadas con cepas de *R. etli nodC- $\alpha$*  (SC15 y CE3) y *nodC- $\delta$*  (55N1 y 124N1) se agrupan en clados separados. Sin embargo, en el cultivar Alubia no sólo no se observa esta correlación, sino que además los perfiles con mayor porcentaje de similitud están dados por la inoculación con SC15 (*nodC- $\alpha$* ) y 55N1 (*nodC- $\delta$* ). Esto podría deberse a que el cultivar NAG12 forma parte del *pool* genético ancestral originado en Mesoamérica, por lo que estas plantas junto con los microsimbiontes presentes en suelos Mesoamericanos han co-evolucionado por alrededor de 4 millones de años, mientras que el *pool* genético Andino surgió hace tan sólo 165.000 años. El mayor tiempo de co-existencia entre los cultivares Mesoamericanos y la microbiota propia de estos suelos podría explicar la capacidad fisiológica de estas plantas de discriminar entre cepas tipo *nodC- $\alpha$*  y *- $\delta$*  al momento de nodular, lo cual se correlaciona con una respuesta transcripcional distintiva.

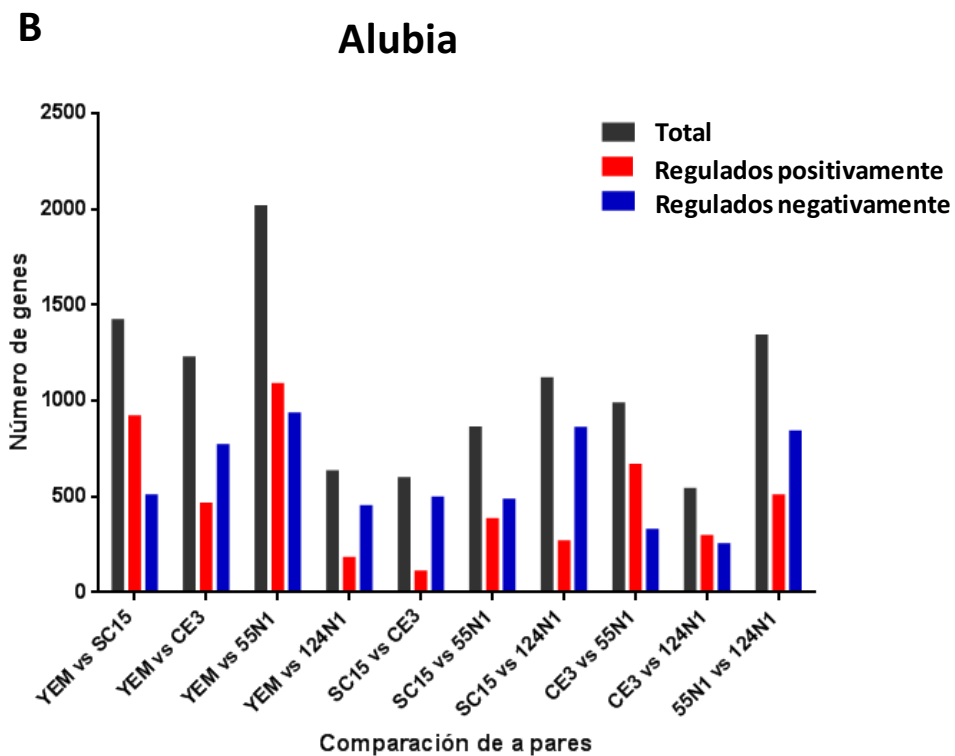
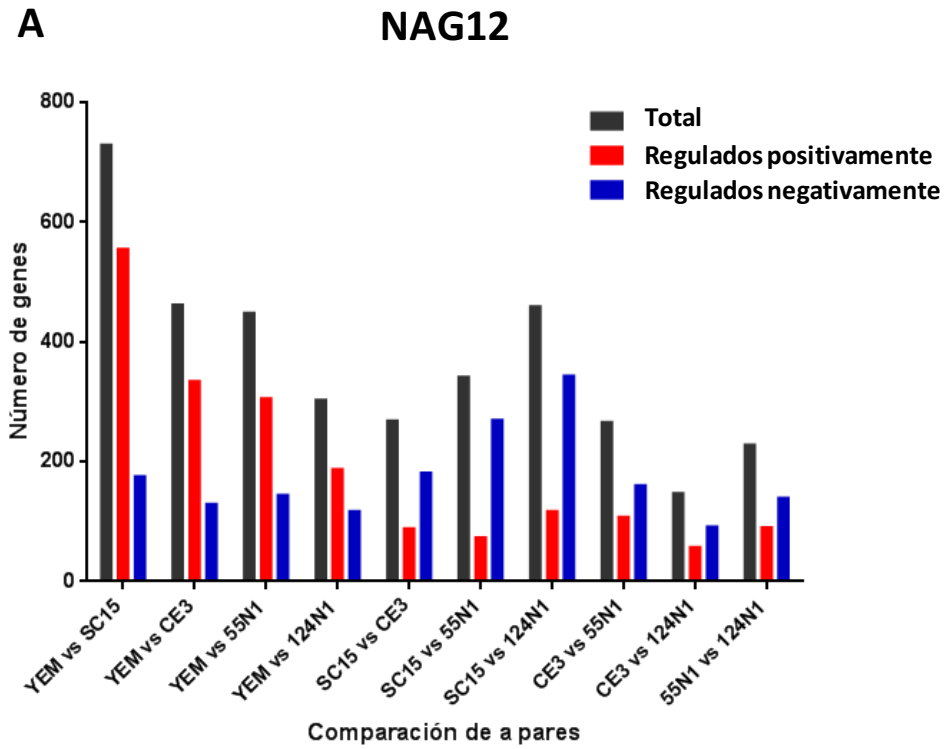
### 4. Análisis general de genes con expresión diferencial

Con el objetivo de identificar genes con expresión diferencial (GEDs) en cada cultivar de *P. vulgaris*, se realizaron comparaciones sistemáticas de a pares entre todas las condiciones utilizando la herramienta *Cufdiff*. Entre los *outputs* que genera *Cufdiff*, se encuentran tablas con los valores de FPKM para cada gen en cada condición (Tablas Suplementarias S1 y S2) y en cada réplica biológica (Tablas Suplementarias S3 y S4), así como también, las comparaciones sistemáticas de a pares para todos los genes y los valores de p y q (valor de p ajustado por el *false discovery rate* (FDR)) que determinan si las diferencias son estadísticamente significativas (Tablas Suplementarias S5 y S6).

## Capítulo I

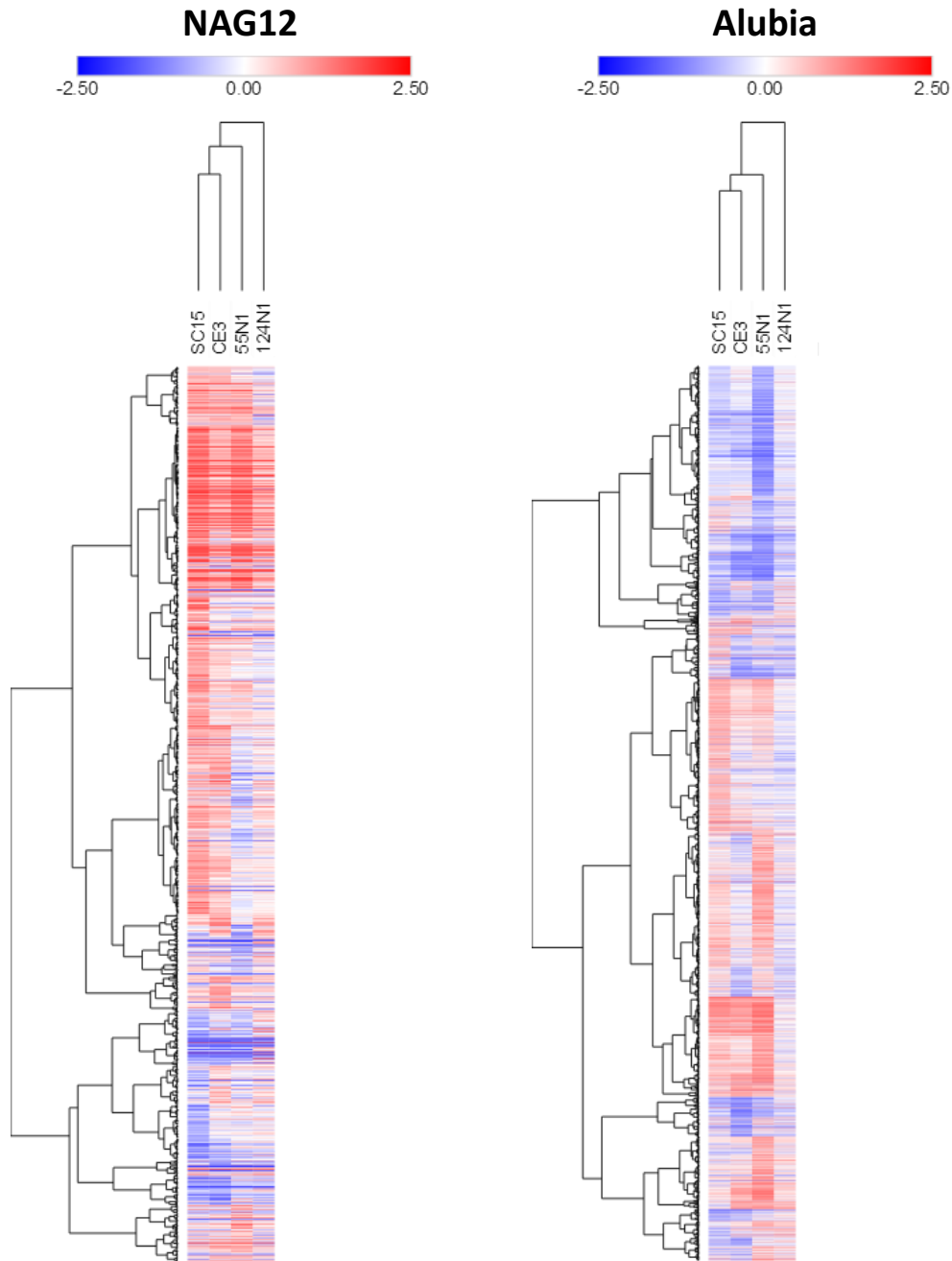
Con el fin de definir un criterio de diferencialidad robusto sólo fueron considerados GEDs aquellos genes con un  $\log_2$  del *fold change* mayor a 1 o menor a -1, con un valor de  $p < 0,05$  y que tuviera un valor mayor o igual a 1 FPKM en al menos una de las muestras. Para el caso de NAG12 se encontraron 1230 genes con expresión diferencial en al menos una de las comparaciones de a pares, observándose la mayor cantidad de GEDs (728) al comparar el transcriptoma de raíces inoculadas con SC15 respecto a YEM. De los 728 GEDs, 554 se encuentran regulados positivamente y 174 negativamente (Figura 15A). Para el caso de Alubia la cantidad total de GEDs fue de 3754, tres veces superior a la de NAG12. La mayor cantidad de GEDs en Alubia fue observada al comparar la inoculación con la cepa 55N1 respecto a YEM, donde se encontraron 1082 DEGs regulados positivamente y 929 negativamente (Figura 15B). En ambos cultivares, la menor cantidad de GEDs se observa en la comparación entre CE3 y 124N1, con sólo 146 y 534 GEDs en NAG12 y Alubia, respectivamente (Figuras 15A y 15B).

El agrupamiento de los GEDs de cada cultivar en base a su patrón de expresión se muestra a través de un *heat map* (Figura 16). El análisis por *clustering* jerárquico en cada uno de los cultivares muestra que los perfiles transcripcionales de los GEDs de las muestras inoculadas con SC15 y CE3 (ambas *nodC- $\alpha$* ) se encuentran agrupadas en el mismo clado, mientras que los perfiles de las muestras inoculadas con 55N1 y 124N1 son más divergentes (Figura 16).



**Figura 15. Genes con expresión diferencial (GEDs) en comparaciones de a pares.** Se muestra el número de genes con expresión diferencial para cada una de las comparaciones indicadas en el eje X para NAG12 (A) y Alubia (B). Las barras grises, rojas y azules indican GEDs totales, regulados positivamente y negativamente, respectivamente.





**Figura 16. Análisis de *clustering* jerárquico.**

Se muestra un mapa de calor (*heat map*) para los GEDs. El análisis de *clustering* jerárquico en base a los valores del logaritmo en base 2 de los FPKM de los genes diferenciales normalizados por los valores de YEM muestra el agrupamiento según las cepas utilizadas en los cultivares NAG12 y Alubia. El análisis fue realizado con el software Morpheus utilizando la correlación de Pearson y el *Average linkage method*.

## 5. Genes diferenciales en respuesta a la inoculación con *R. etli*

Luego de haber analizado los GEDs a nivel general, se decidió realizar una clasificación funcional de aquellos genes que cambian sus niveles estacionarios de mRNA en

## Capítulo I

respuesta a la inoculación con al menos una de las cepas de *R. etli* respecto a YEM, para comprender cuáles son los procesos celulares que sufren una reprogramación durante las etapas tempranas del establecimiento de la simbiosis (Tabla Suplementaria S7). Para cada cultivar se realizó un análisis de enriquecimiento de GO (*Gene Ontology*) a partir de los GEDs regulados positiva y negativamente utilizando la herramienta *Singular Enrichment Analysis* (SEA) disponible en la plataforma AgriGO (Tian et al., 2017) (Figuras 17, 18 y 19). Esta herramienta permite identificar categorías de GO estadísticamente sobrerrepresentadas respecto al genoma de *P. vulgaris*. En NAG12 puede verse que dentro de los GEDs regulados positivamente por el rizobio, las categorías sobrerrepresentadas son respuesta a estímulos (*response to stimulus*), respuesta al estrés (*response to stress*), respuesta a estímulos químicos (*response to chemical stimulus*), respuesta al estrés oxidativo (*response to oxidative stress*) y óxido reducción (*oxidation reduction*) (Figura 17 y 18). Esta última categoría también se encuentra sobrerrepresentada dentro de los GEDs regulados negativamente, sugiriendo que luego de la inoculación se produce una reprogramación de los procesos redox de la célula. En Alubia, estas 5 categorías muestran un patrón similar, aunque en este cultivar puede observarse una mayor variedad de las categorías de GO sobrerrepresentadas, tanto entre los genes regulados positiva como negativamente en respuesta a la inoculación (Figura 17 y 19). En los primeros se encuentran significativamente enriquecidas las categorías óxido reducción, respuesta a estímulos bióticos (*response to biotic stimulus*), respuesta al estrés oxidativo (*response to oxidative stress*), respuesta de defensa (*defense response*), y aquellas relacionadas con la modificación de la pared celular, el metabolismo de glucanos y el movimiento basado en microtúbulos.

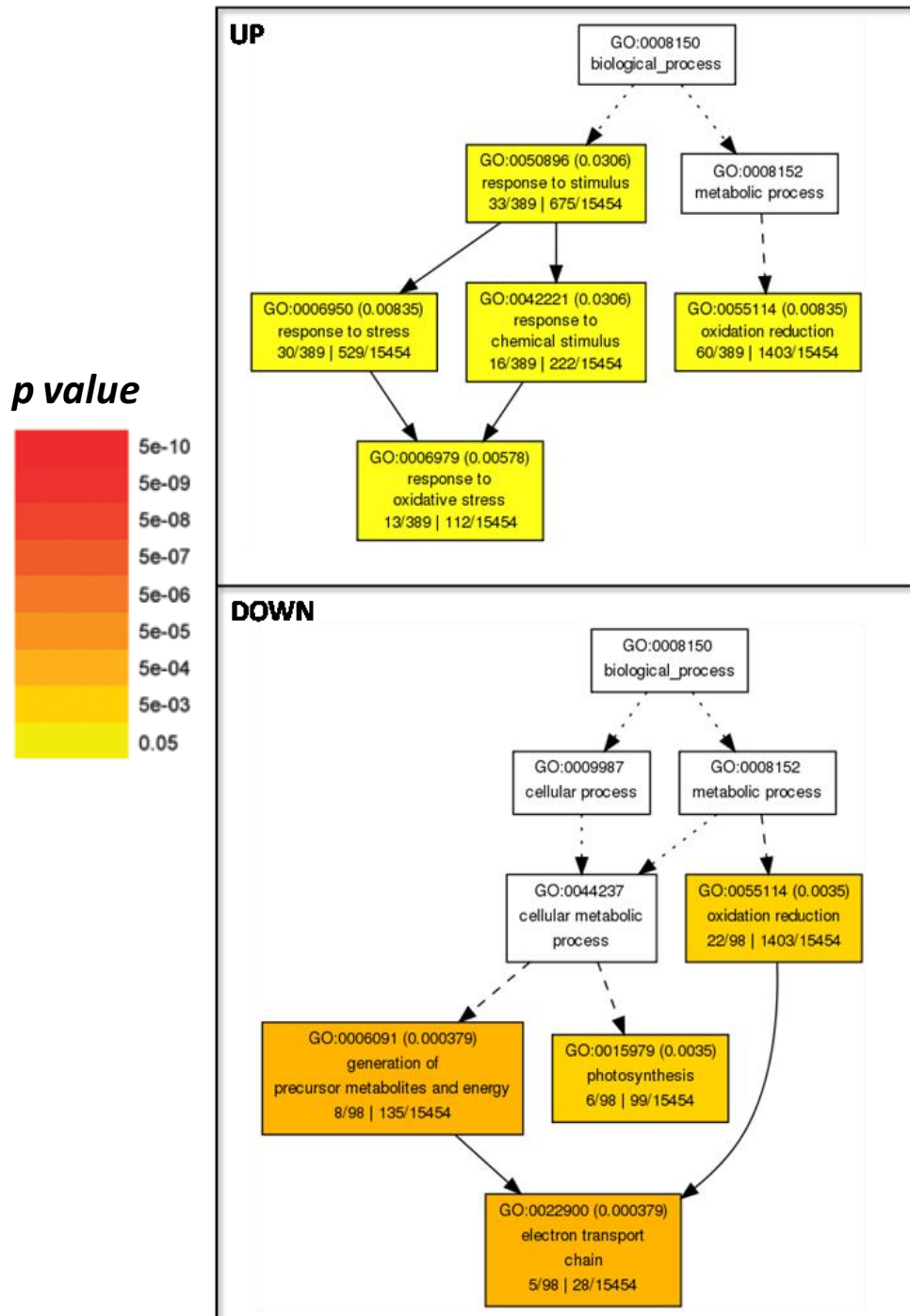
## Capítulo I

Información GO		UP				DOWN		UP		DOWN	
GO Term	Description	NAG12	NAG12	ALUBIA	ALUBIA	FDR	Num	FDR	Num	FDR	Num
GO:0006979	response to oxidative stress	Yellow	Grey	Red	Red	0.0058	13	---	---	2.80E-10	34
GO:0055114	oxidation reduction	Yellow	Grey	Red	Red	0.0084	60	0.0035	22	5.30E-05	128
GO:0006950	response to stress	Yellow	Grey	Red	Red	0.0084	30	---	---	8.80E-10	77
GO:0050896	response to stimulus	Yellow	Grey	Red	Red	0.031	33	---	---	1.10E-09	89
GO:0042221	response to chemical stimulus	Yellow	Grey	Red	Red	0.031	16	---	---	8.80E-10	46
GO:006091	generation of precursor metabolites and energy	Grey	Orange	Orange	Orange	---	---	0.00038	8	---	---
GO:0022900	electron transport chain	Grey	Orange	Orange	Orange	---	---	0.00038	5	---	---
GO:0015979	photosynthesis	Grey	Orange	Orange	Orange	---	---	0.0035	6	---	---
GO:0006952	defense response	Grey	Orange	Orange	Orange	---	---	---	---	4.00E-05	31
GO:0006073	cellular glucan metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	0.00036	20
GO:0044042	glucan metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	0.00036	20
GO:0009607	response to biotic stimulus	Grey	Orange	Orange	Orange	---	---	---	---	0.00071	14
GO:0034622	cellular macromolecular complex assembly	Grey	Orange	Orange	Orange	---	---	---	---	0.0013	18
GO:0071555	cell wall organization	Grey	Orange	Orange	Orange	---	---	---	---	0.0013	17
GO:0007017	microtubule-based process	Grey	Orange	Orange	Orange	---	---	---	---	0.0013	19
GO:0071554	cell wall organization or biogenesis	Grey	Orange	Orange	Orange	---	---	---	---	0.0013	22
GO:0034621	cellular macromolecular complex subunit organization	Grey	Orange	Orange	Orange	---	---	---	---	0.0063	18
GO:0042545	cell wall modification	Grey	Orange	Orange	Orange	---	---	---	---	0.012	14
GO:0044264	cellular polysaccharide metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	0.02	22
GO:0005976	polysaccharide metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	0.047	23
GO:0005975	carbohydrate metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	0.049	68
GO:0007018	microtubule-based movement	Grey	Orange	Orange	Orange	---	---	---	---	0.05	12
GO:0080090	regulation of primary metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0009889	regulation of biosynthetic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0006355	regulation of transcription, DNA-dependent	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0010556	regulation of macromolecule biosynthetic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0051252	regulation of RNA metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0031326	regulation of cellular biosynthetic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0045449	regulation of transcription	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0010468	regulation of gene expression	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0019219	regulation of nucleobase, nucleoside, nucleotide and nu	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0051171	regulation of nitrogen compound metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0045
GO:0060255	regulation of macromolecule metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0053
GO:0031323	regulation of cellular metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.0072
GO:0019222	regulation of metabolic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.009
GO:0050794	regulation of cellular process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.021
GO:0006350	transcription	Grey	Orange	Orange	Orange	---	---	---	---	---	0.021
GO:0032774	RNA biosynthetic process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.021
GO:0006351	transcription, DNA-dependent	Grey	Orange	Orange	Orange	---	---	---	---	---	0.021
GO:0050789	regulation of biological process	Grey	Orange	Orange	Orange	---	---	---	---	---	0.029

**Figura 17. Comparación de los términos de GO diferenciales.**

En la tabla se muestran los términos de GO con su descripción y un *heatmap* que representa los valores de FDR. Para cada término de GO se indican la cantidad de genes diferenciales que pertenecen a esta categoría.

La presencia de genes regulados positivamente pertenecientes a las categorías relacionadas a la modificación de la pared celular y el movimiento a través de microtúbulos sugieren una participación en la activación de los cambios fisiológicos y morfológicos requeridos para la infección bacteriana, tales como el redireccionamiento del crecimiento de los pelos radicales y la formación del hilo de infección. Por otro lado, entre los genes regulados negativamente se encuentran enriquecidas las categorías óxido-reducción, así como también diversas categorías relacionadas a la regulación de la transcripción (Figura 17 y 19). Esto sugiere que en la célula se estaría generando una reprogramación transcripcional y cambios en los procesos redox involucrados en las etapas tempranas de la interacción simbiótica.



**Figura 18. Categorías de GO que cambian significativamente en respuesta al rizobio en NAG12.**

Se muestran las categorías de GO de los genes regulados positiva (UP) o negativamente (DOWN) en respuesta al rizobio, que cambian significativamente. El color de los cuadros indica el grado de significancia estadística en base al valor de p (cuanto más cálido el color menor es el valor de p). En cada caja se indica el identificador de GO y el valor de p en la primera línea. En la segunda se describe el término de GO, y en la tercera línea se muestra la cantidad de genes que tienen ese término de GO respecto a todos los ingresados, y luego de la barra vertical, como es esa relación en el genoma. Las flechas continuas, a trazos y de puntos significan que existen, uno o ningún término enriquecido en los extremos que conectan.

UP

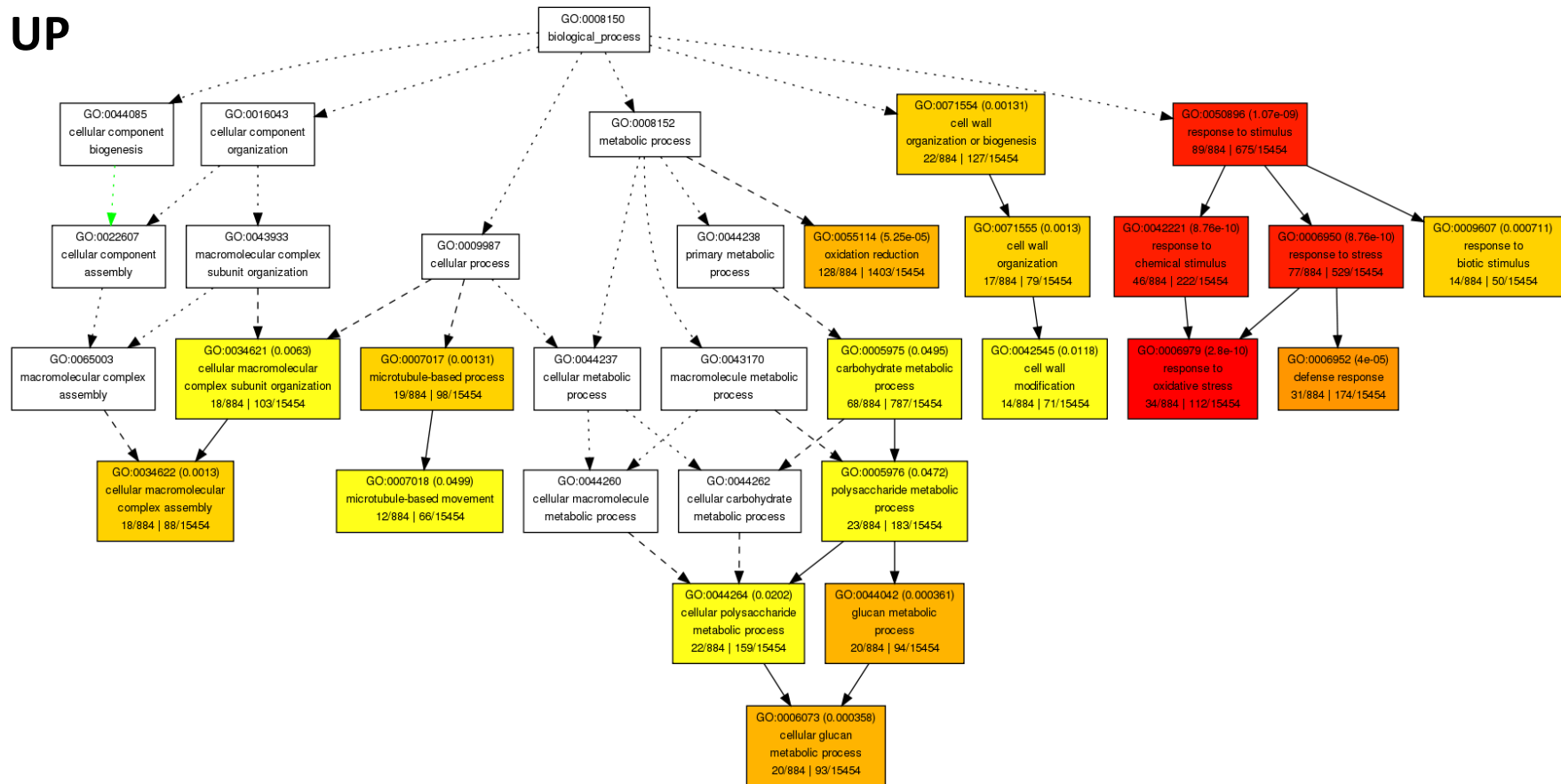
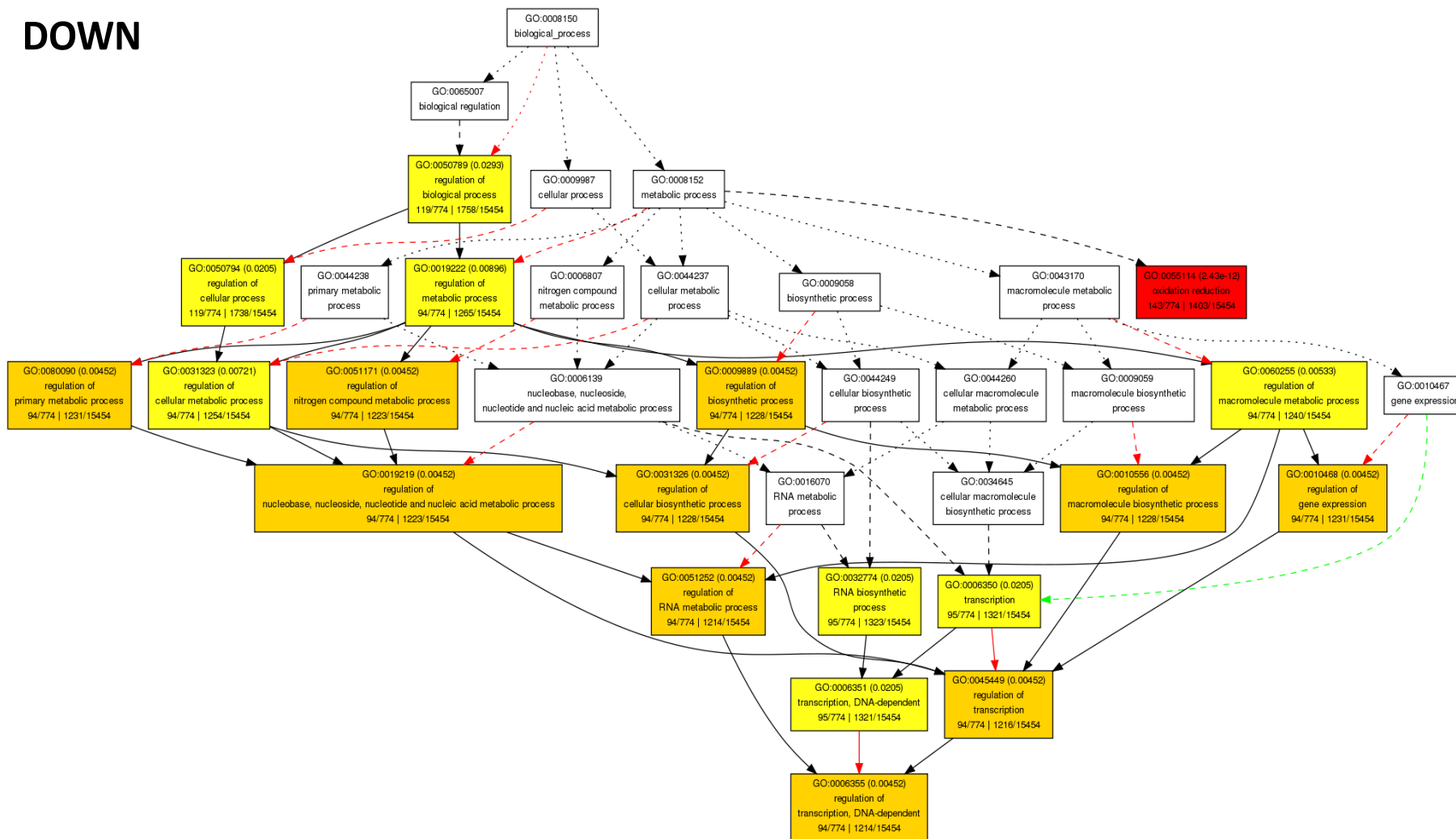


Figura 19. Categorías de GO que cambian significativamente en respuesta al rizobio en *Alubia* (CONTINÚA EN LA PÁGINA SIGUIENTE)

# DOWN



**Figura 19. Categorías de GO que cambian significativamente en respuesta al rizobio en Alubia.**

Se muestran las categorías de GO de los genes regulados positiva (UP) o negativamente (DOWN) en respuesta al rizobio, que cambian significativamente. El color de los cuadros indica el grado de significancia estadística en base al p (cuanto más cálido el color menor es el valor de p). En cada caja se indica el identificador de GO y el valor de p en la primera línea. En la segunda se describe el término de GO, y en la tercera línea se muestra la cantidad de genes que tienen ese término de GO respecto a todos los ingresados, y luego de la barra vertical, como es esa relación en el genoma. Las flechas continuas, a trazos y de puntos significan que existen dos, uno o ningún término enriquecido en los extremos que conectan. El color rojo y verde de las líneas indica que es un regulador positivo y negativo, respectivamente.

En conclusión, estos resultados sugieren que a tiempos tempranos del establecimiento de la interacción simbiótica se producen cambios en la expresión de genes relacionados al estrés y a procesos de óxido reducción, los cuales se conservan en ambos cultivares. Sin embargo, los GEDs en Alubia mostraron un enriquecimiento significativo en categorías de GO muy diversas, sugiriendo que a 24 hpi las células de la raíz ya están sometidas a una fuerte reprogramación de la transcripción, el metabolismo, las respuestas de defensa y los cambios fisiológicos asociados a la membrana y pared celular que son requeridos para la infección y la organogénesis del nódulo. Es posible que este tipo de respuestas se encuentren desfasadas temporalmente en NAG12 y se activen a tiempos más tardíos o más tempranos de la interacción.

## **6. Identificación de las respuestas moleculares comunes y específicas frente a las cepas de *R. etli* de tipo *nodC- $\alpha$* o *- $\delta$***

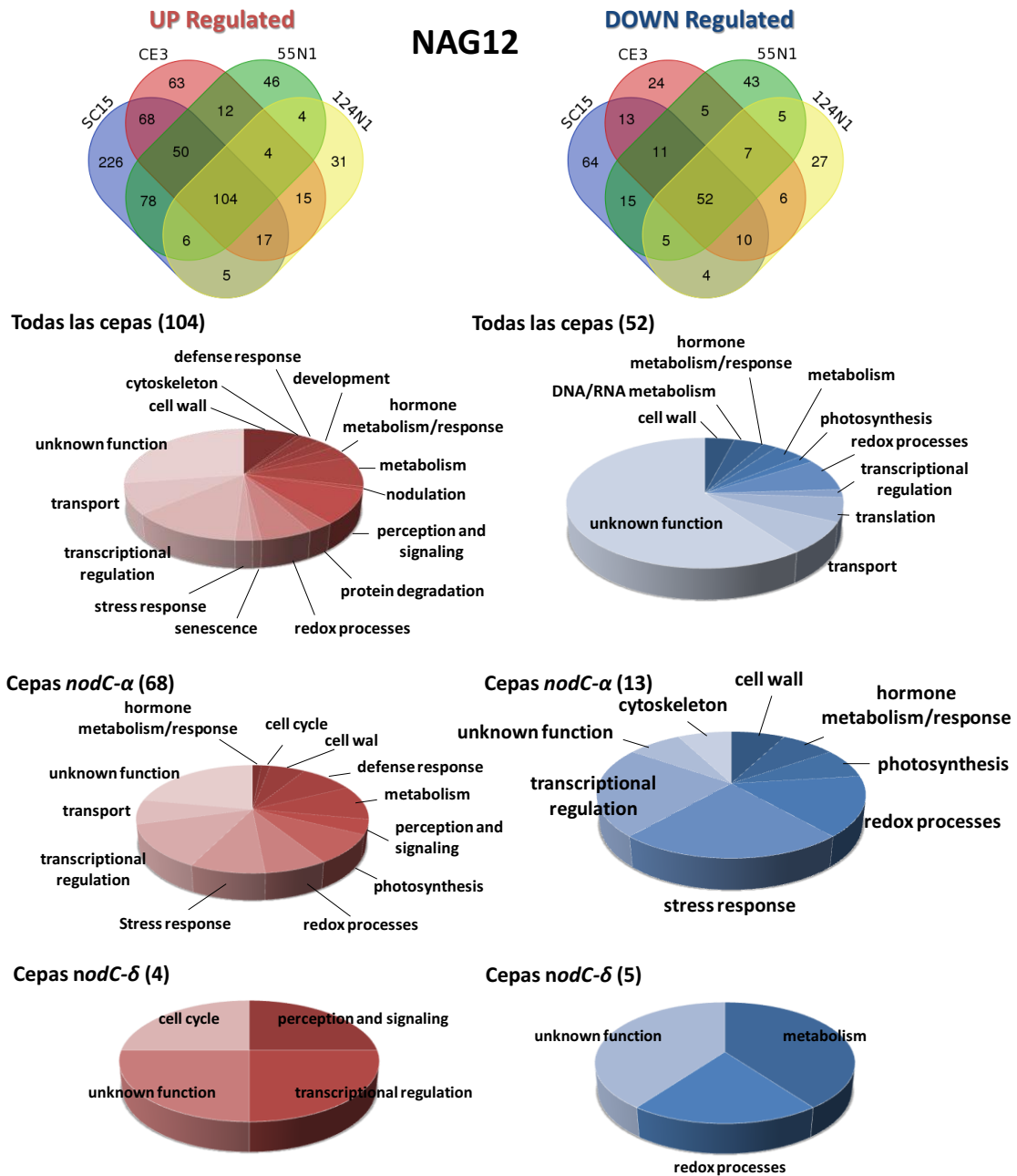
A continuación se decidió estudiar el grado de superposición de las respuestas moleculares que disparan cada una de las cepas en cada uno de los cultivares analizados en este trabajo. Para esto se utilizaron las listas de genes regulados positiva o negativamente en respuesta a cada cepa en cada cultivar y se calcularon las intersecciones entre las listas de elementos (Tabla Suplementaria S7), así como también, los diagramas de Venn asociados, mediante la herramienta web *Calculate and draw custom Venn diagrams* (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) (Figuras 20 y 21). Para el caso del cultivar NAG12, los resultados muestran un grupo de 156 genes que cambian sus niveles estacionarios en respuesta a las cuatro cepas de rizobios, de los cuales 104 son regulados positivamente y 52 negativamente (Figura 20). Este núcleo de genes formaría parte de una respuesta molecular común a tiempos tempranos de la inoculación frente a distintas cepas de rizobios. Con el fin de comprender los procesos

## Capítulo I

biológicos en los cuales participarían estos genes, se realizó una clasificación funcional manual en base a su anotación. Las categorías dominantes dentro de los genes regulados positivamente fueron función desconocida (*unknown function*), regulación transcripcional (*transcriptional regulation*) y percepción y señalización (*perception and signaling*), lo cual es consistente con la activación de las vías de transducción de señales desencadenadas por el rizobio y la reprogramación transcripcional de las células vegetales requerida para la simbiosis. De manera consistente con la bibliografía, dentro de aquellos genes regulados positivamente por las cuatro cepas se encuentran genes marcadores de la nodulación como *PvENOD40*, *PvERN1*, *PvNIN* y *PvNF-YA9*. Las categorías más numerosas entre los genes regulados negativamente por las cuatro cepas fueron función desconocida, procesos rédox (*redox processes*) y transporte (*transport*) (Figura 20).

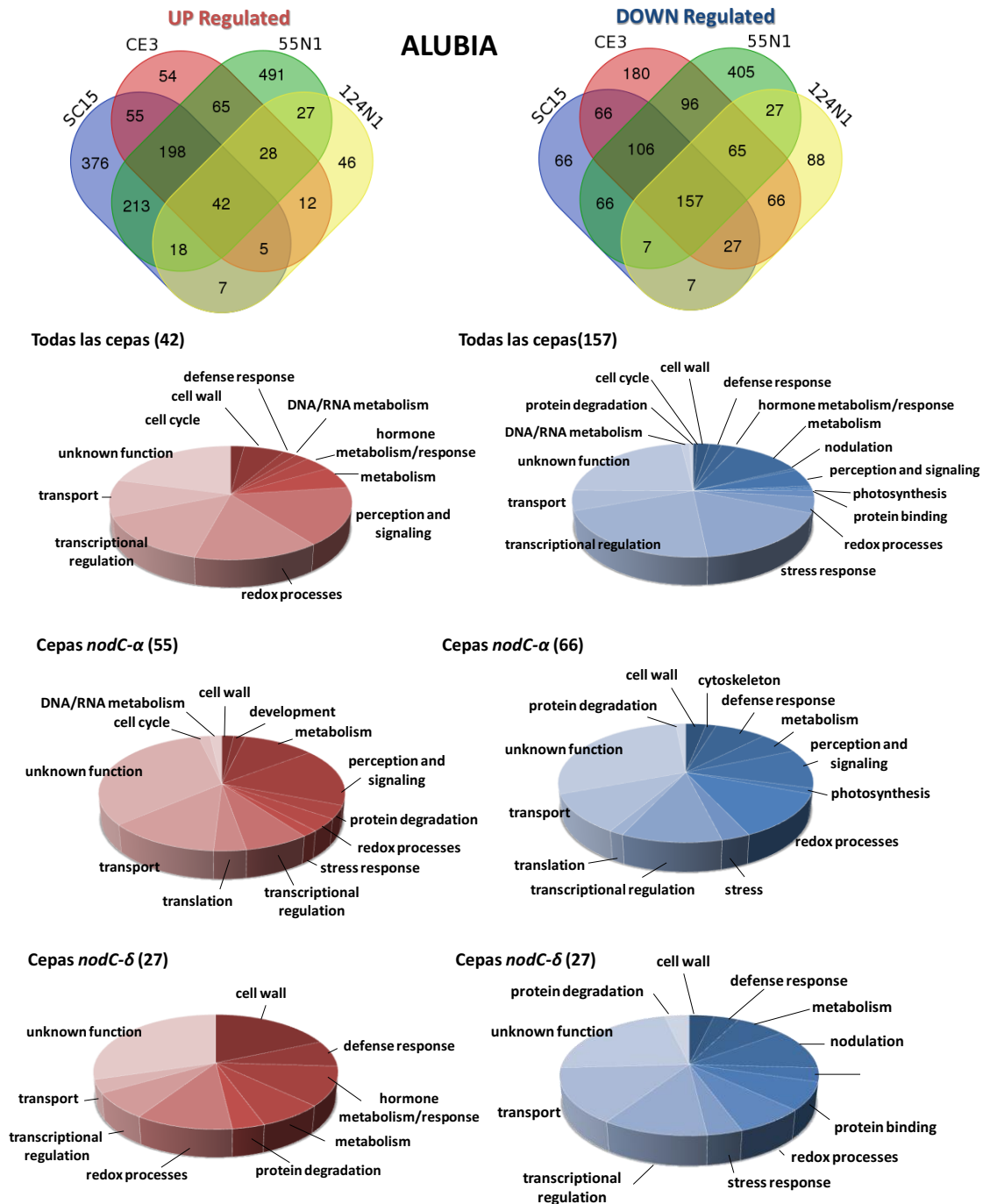
Para el caso de Alubia, el número de GEDs comunes a las cuatro cepas fue 199, de los cuales 42 se encontraron regulados positivamente y 157 negativamente. Similarmente a lo observado en NAG12, las categorías dominantes en los genes regulados positivamente en el cultivar Alubia fueron función desconocida (*unknown function*), regulación transcripcional (*transcriptional regulation*), percepción y señalización (*perception and signaling*) y procesos rédox (*redox processes*) (Figura 21). Sin embargo los genes marcadores de las etapas tempranas de la nodulación *PvNIN* y *PvENOD40* se encontraron inducidos en respuesta a SC15, CE3 y 55N1, pero no en respuesta a la cepa 124N1, sugiriendo que esta última podría generar una respuesta más lenta respecto de las otras cepas, al menos en el cultivar Alubia. Las categorías dominantes en los genes regulados negativamente fueron función desconocida (*unknown function*), regulación transcripcional (*transcriptional regulation*), respuesta a estrés (*stress response*) y metabolismo (*metabolism*) (Figura 21).





**Figura 20. Grado de superposición y clasificación de los genes con expresión diferencial en NAG12.**

Los genes con expresión diferencial (GEDs) en plantas inoculadas con las diferentes cepas de *R. etli* respecto a YEM fueron utilizados para generar diagramas de Venn mostrando la superposición de las respuestas moleculares. A partir de los diagramas de Venn se identificaron y clasificaron funcionalmente los GEDs comunes a las respuestas frente a las cuatro cepas, comunes a las cepas tipo *nodC-α* y aquellos comunes a las cepas tipo *nodC-δ*. Para el análisis se estudiaron los genes regulados positiva (rojo) y negativamente (azul) por separado.



**Figura 21. Grado de superposición y clasificación de los genes con expresión diferencial en Alubia.**

Los genes con expresión diferencial (GED) en plantas inoculadas con las diferentes cepas de *R. etli* respecto a YEM fueron utilizados para generar diagramas de Venn mostrando la superposición de las respuestas moleculares. A partir de los diagramas de Venn se identificaron y clasificaron funcionalmente los GED comunes a las respuestas frente a las cuatro cepas, comunes a las cepas tipo  $\alpha$  y aquellos comunes a las cepas tipo  $\delta$ . Para el análisis se estudiaron los genes regulados positiva (rojo) y negativamente (azul) por separado.

## Capítulo I

Con el fin de explorar los mecanismos moleculares asociados a una nodulación eficiente se analizaron aquellos genes que están regulados de manera diferencial en respuesta a las dos cepas tipo *nodC-α* (SC15 y CE3) o las dos cepas tipo *nodC-δ* (55N1 y 124N1), las cuales nodulan de manera más eficiente en el cultivar NAG12 o Alubia, respectivamente (Meschini et al., 2008; Zanetti et al., 2010). En el cultivar NAG12 la respuesta molecular común a las cepas tipo *nodC-α* involucra una mayor cantidad de genes que la respuesta de las cepas tipo *nodC-δ*. Las primeras modulan los niveles de transcritos de 81 genes (68 positivamente y 13 negativamente), mientras que las cepas *nodC-δ* modulan sólo los niveles de transcritos de 9 genes comunes (4 positivamente y 5 negativamente), sugiriendo que aquellas cepas que nodulan más eficientemente en NAG12 generarían una alteración común del transcriptoma más pronunciada, al menos a las 24 hpi (Figura 20). Tanto entre los genes modulados positiva como negativamente por cepas tipo *nodC-α*, se encuentran genes pertenecientes a las categorías pared celular (*cell wall*), rédox (*redox*), respuesta a estrés (*stress response*), regulación transcripcional (*transcriptional regulation*) y respuesta y metabolismo de hormonas (*hormone metabolism/response*). Sólo los genes regulados positivamente muestran categorías relacionadas a la respuesta de defensa (*defense response*), el metabolismo (*metabolism*), transporte (*transport*), el ciclo celular (*cell cycle*) y percepción y señalización (*perception and signaling*). La única categoría particular de los genes regulados negativamente es citoesqueleto (*cytoskeleton*) (Figura 20).

Al igual que en NAG12, en Alubia puede observarse que existe una mayor superposición de la respuesta molecular entre las dos cepas tipo *nodC-α* (121 genes; 55 regulados positivamente y 66 negativamente) respecto de las dos cepas tipo *nodC-δ* (54 genes; 27 regulados positivamente y 27 regulados negativamente), aunque la diferencia es sustancialmente menor que la observada en NAG12 (Figura 21). Entre los genes

## Capítulo I

regulados positivamente por las cepas tipo *nodC-α* dominan las categorías función desconocida (*unknown function*), percepción y señalización (*perception and signaling*), metabolismo (*metabolism*) y transporte (*transport*), mientras que entre los genes regulados negativamente predominan las categorías función desconocida (*unknown function*), procesos rédox (*redox processes*), percepción y señalización (*perception and signaling*), regulación transcripcional (*transcriptional regulation*) y respuesta de defensa (*defense response*). El hecho de que la categoría respuesta de defensa se encuentre regulada específicamente entre los genes reprimidos por las cepas *nodC* tipo- $\alpha$ , sugiere que las respuestas de defensa de la planta no jugarían un rol determinante en la menor eficiencia de estas cepas respecto a las cepas tipo *nodC-δ*. Por su parte, las cepas 55N1 y 124N1 inducen la expresión de genes pertenecientes principalmente a las categorías función desconocida (*unknown function*), pared celular (*cell wall*), rédox (*redox*) y respuesta y metabolismo de hormonas (*hormone metabolism/response*), mientras que reprimen la expresión de genes vinculados al transporte (*transport*), nodulación (*nodulation*) y de función desconocida (*unknown function*). Interesantemente, la categoría respuesta y metabolismo de hormonas (*hormone metabolism/response*) sólo está presente entre los genes activados por el rizobio (Figura 21).

En conclusión, el análisis transcriptómico realizado ha permitido identificar genes que son modulados específicamente por cepas tipo *nodC-α* y  $-\delta$  en NAG12 y Alubia, los cuales podrían formar parte de los mecanismos moleculares asociados a una nodulación eficiente. La clasificación funcional de los genes identificados muestra que estos mecanismos podrían ser muy diversos e involucrar cambios en el desarrollo, la pared celular y la modulación y/o activación de vías de transducción de señales.

## 7. Análisis de los genes con expresión diferencial específicos de cepas tipo *nodC-α* y *-δ*

Con el objetivo de comprender cómo los cambios en los niveles estacionarios de genes individuales podrían estar influenciando la eficiencia de la interacción simbiótica frente a los dos tipos de cepas, se realizó un análisis más detallado de aquellos genes incluidos en categorías funcionales asociados a la regulación de la expresión génica, ya que los productos de estos genes podrían estar activando o reprimiendo programas genéticos específicos de respuestas a cepas tipo *nodC-α* o *-δ*. Para esto, a partir de los identificadores de cada uno de los genes pertenecientes a la categoría de regulación transcripcional, se realizó una búsqueda en la base de datos de factores de transcripción de *P. vulgaris* (<http://www.multiomics.in/PvTFDB/>) (Bhawna et al., 2016). Llamativamente, sólo las cepas tipo *nodC-α* modulan la expresión específica de factores de transcripción en NAG12, mientras que no se encontraron GEDs pertenecientes a esta categoría que respondan frente a ambas cepas de tipo *nodC-δ*. Las cepas tipo *nodC-α* inducen la expresión de miembros de las familias ERF (Phvul.001G160100), DBB (Phvul.011G053700), MYB (Phvul.005G087400), CO-like (Phvul.011G106300; Phvul.005G153500) y C2H2 (Phvul.003G147100), y reprimen la expresión de miembros de las familias MYB-related (Phvul.011G062100; Phvul.008G240700) y MICK (Phvul.003G182700) (Figura 22). En el cultivar Alubia, sí se observó una modulación de la expresión de factores de transcripción específicos en respuesta a las cepas tipo *nodC-δ*. Estas cepas indujeron la expresión de genes de la familia WRKY (Phvul.004G105800) y bZIP (Phvul.004G045800), y reprimieron la expresión de un miembro de la familia MYB-related (Phvul.007G074500) y uno de la familia ARR-B (Phvul.010G119700). De manera similar a NAG12, las cepas tipo *nodC-α* modificaron los niveles d mRNAs de diversos factores de transcripción; 2 miembros de la familia

## Capítulo I

NAC (Phvul.011G003900; Phvul.003G229600), 3 miembros de la familia ERF (Phvul.007G086600; Phvul.008G019600; Phvul.001G084000) y 1 miembro de las familias DBB (Phvul.011G040800), ZF-HD (Phvul.007G114300) y MYB-related (Phvul.006G025400) (Figura 22). En conclusión, tanto en NAG12 como en Alubia, las cepas que nodulan de manera más eficiente cada cultivar serían capaces de desencadenar cambios en los niveles estacionarios de mRNA de factores de transcripción pertenecientes a diversas familias. Estos factores de transcripción podrían estar orquestando la activación o represión de programas genéticos de la planta que determinen la nodulación más eficiente de estas cepas.

		NAC	ERF	DBB	ZF-HD	MYB-related	MYB	WRKY	bZIP	ARR-B	CO-like	C2H2	MICK
NAG12	<i>nodC-α</i>	x	●	●	x	●●	●	x	x	x	●●	●	●
	<i>nodC-δ</i>	x	x	x	x	x	x	x	x	x	x	x	x
ALUBIA	<i>nodC-α</i>	●●	●●●	●	●	●	x	x	x	x	x	x	x
	<i>nodC-δ</i>	x	x	x	x	●	x	●	●	●	x	x	x

**Figura 22. Clasificación de los factores de transcripción con expresión diferencial en cepas *nodC-α* y *-δ*.**

Para cada cultivar se muestran los factores de transcripción que se inducen (círculos rojos) o reprimen (círculos azules) en raíces inoculadas con cepas *nodC-α* o *-δ*. En la parte superior de la figura se indican las familias de factores de transcripción. La cantidad de círculos indica la cantidad de miembros de esa familia con expresión diferencial.

Para caracterizar con mayor profundidad la función del producto de los genes diferenciales correspondientes a otras categorías, se realizó una búsqueda de dominios y motivos proteicos mediante una búsqueda por homología de secuencia en Pfam (<http://pfam.xfam.org/>). Para el caso de genes que codifican enzimas, su actividad enzimática y el proceso metabólico en el cual estarían implicadas se analizó utilizando la plataforma *Plant Metabolic Network* (<https://www.plantcyc.org/>).

Los genes pertenecientes a la categoría de percepción y señalización podrían participar del reconocimiento de moléculas conservadas o específicas de las cepas tipo *nodC-α* o *-δ*.

## Capítulo I

$\delta$ , y/o transmitir una señal que active vías de transducción de señales específicas. En NAG12, las cepas tipo *nodC- $\alpha$*  modularon positivamente la expresión de una proteína con dominio *dirigent protein*, una proteína con dominio quinasa y una proteína fosfatasa 2C (Figura 23). El único gen perteneciente a la categoría percepción y señalización que es modulado por cepas tipo *nodC- $\delta$*  (Phvul.002G323404) fue anotado como una TIR-NBS-LRR, pero el análisis manual realizado en Pfam no mostró la presencia de ningún dominio conocido, por lo cual el mismo no fue incluido en la Figura 23. El caso de Alubia resultó mucho más complejo ya que las cepas *nodC- $\alpha$*  modularon la expresión de 14 genes pertenecientes a la categoría percepción y señalización, presentando un diverso repertorio de dominios de reconocimiento de moléculas de señalización, tales como dominios de unión a galacturonanos (*Wall-associated receptor kinase galacturonan-binding*), dominios LRR (*Leucine rich repeat*), dominios tipo malectina (*Malectin like*) y dominios tipo PAN (*PAN-like domain*). Muchos de estos dominios se encuentran en proteínas que también contienen dominios tipo proteína quinasa (*Protein kinase*). También se encuentran proteínas con dominios *NB-ARC domain* y *TIR domain*, los cuales están presentes en los genes de resistencia (R) que reconocen efectores bacterianos iniciando la inmunidad inducida por efector (McDowell and Woffenden, 2003; Cui et al., 2015). Llamativamente, uno de estos genes se encuentra regulado positivamente en respuesta a las cepas SC15 y CE3, sugiriendo que algún efector común a estas cepas *nodC- $\alpha$*  podría ser reconocido por el sistema inmune de la planta. Por su parte, las cepas tipo *nodC- $\delta$*  55N1 y 124N1 sólo modulan un gen en común perteneciente a esta categoría, el cual codifica una proteína con dominios quinasa, LRR y un dominio lectina de leguminosas (*legume lectin domain*) (Figura 23). Este dominio lectina podría participar del reconocimiento específico de carbohidratos presentes específicamente en estas cepas (Díaz et al., 1989; Brewin and Kardailsky, 1997).

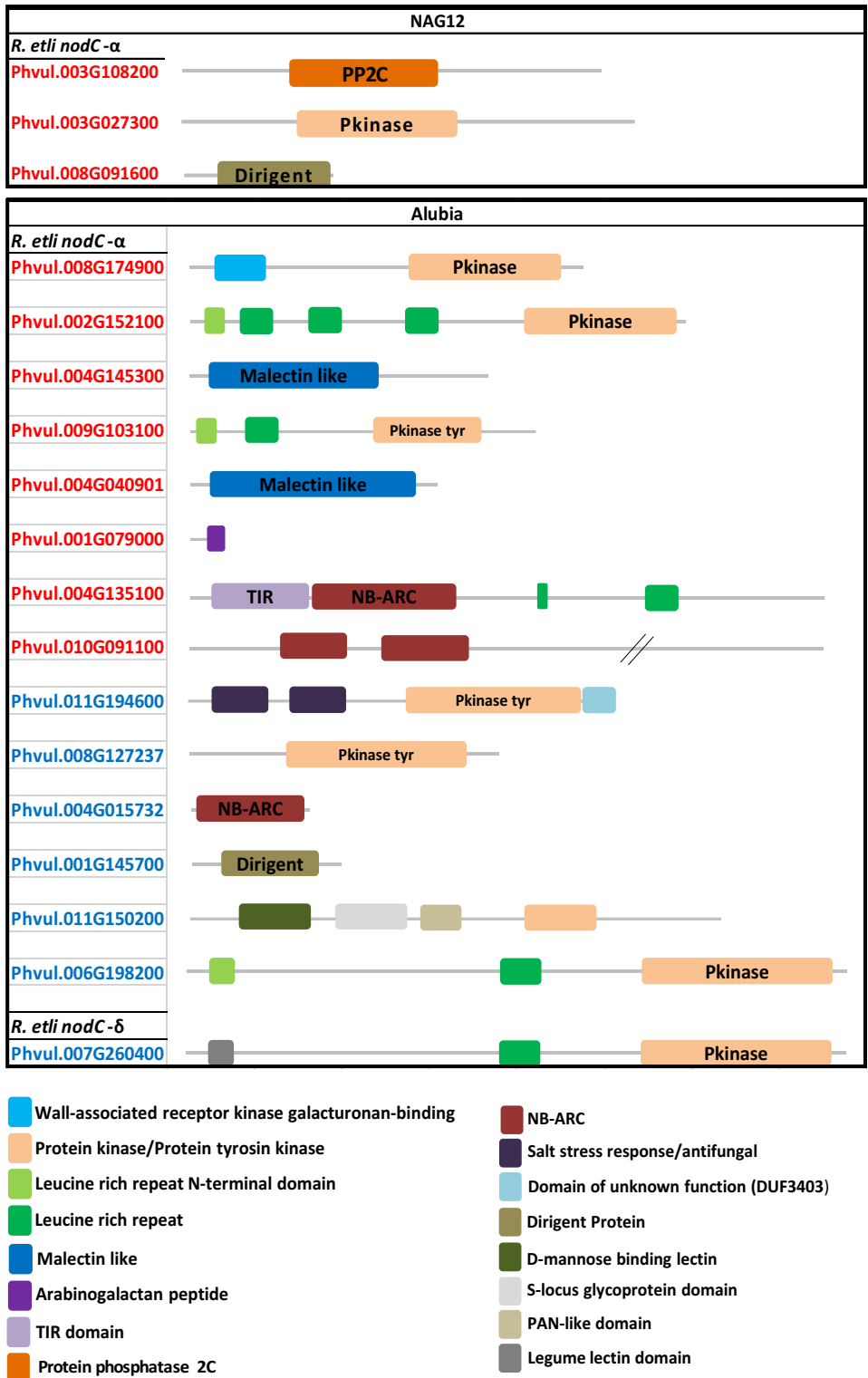


Figura 23. Genes con expresión diferencial en cepas *nodC-α* y *-δ* pertenecientes a la categoría percepción y señalización.

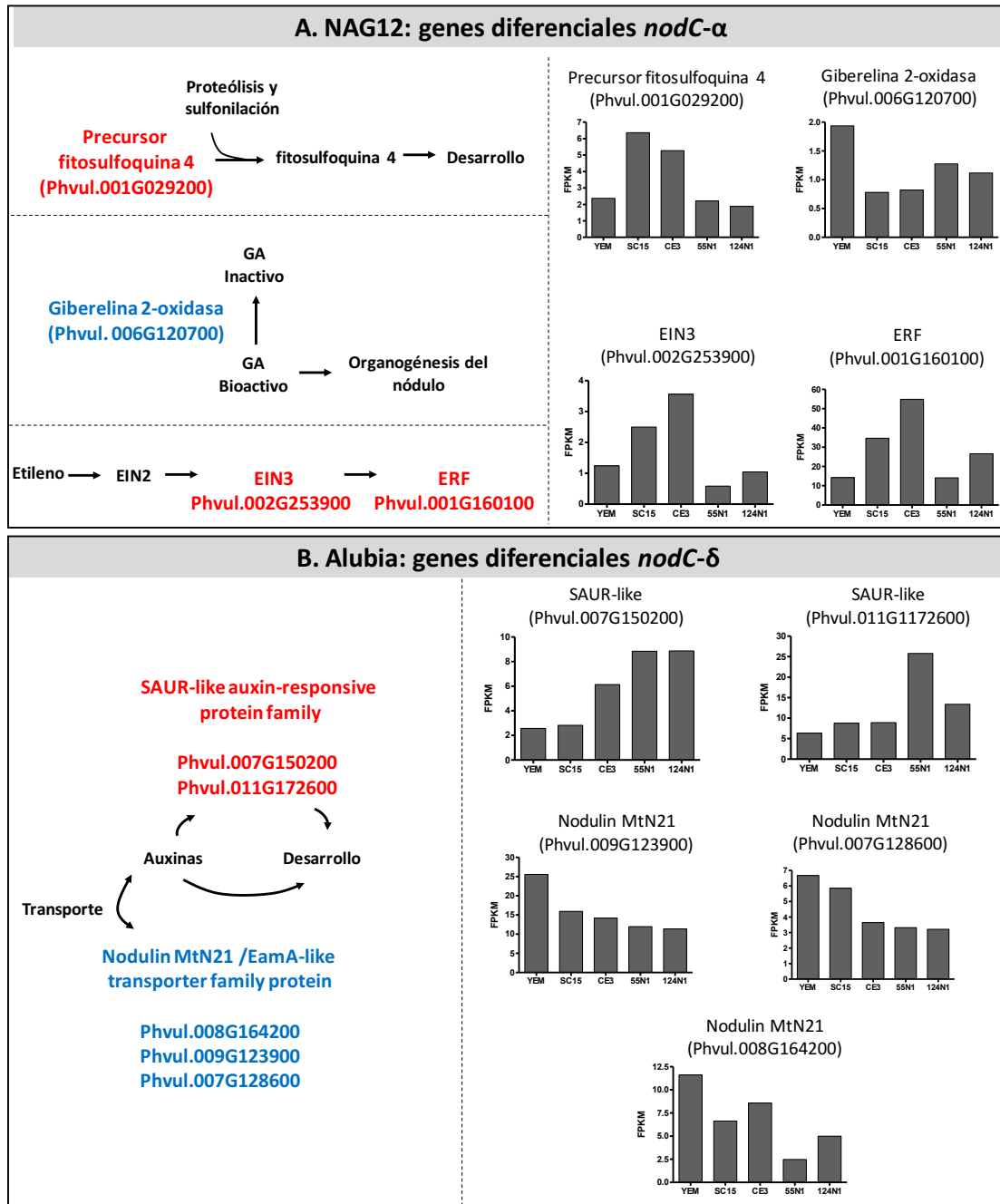
Para cada cultivar se muestran los productos de los genes pertenecientes a la categoría percepción y señalización que se inducen (identificador de gen en rojo) o reprimen (identificador de gen en azul) en raíces inoculadas con cepas *nodC-α* o *-δ*. Las cajas de colores indican los dominios conservados que fueron identificados en cada proteína.



## Capítulo I

Por último, se realizó un análisis detallado de los genes implicados en el desarrollo, el ciclo celular y aquellos involucrados en la respuesta y/o metabolismo de hormonas, ya que estos procesos podrían vincularse de manera directa con la organogénesis del nódulo. En NAG12, las cepas tipo *nodC-α* modulan la expresión común de genes relacionados con el ácido giberélico y el etileno, así como también la inducción del precursor del factor de crecimiento vegetal fitosulfoquina (PSK) (Phvul.001G029200). PSK es un pentapéptido disulfatado que se genera por el clivaje proteolítico del precursor en el apoplasto, que luego es percibido por receptores tipo quinasa en la membrana plasmática, activando vías de transducción de señales vinculadas al crecimiento celular, el movimiento polar del tubo polínico y la modulación de la respuesta inmune en función del estado metabólico (Sauter, 2015) (Figura 24A). En base a esta información, el gen del precursor de la fitosulfoquina 4 es un gran candidato para caracterizar a nivel funcional, ya que podría estar involucrado en procesos como la formación y progresión de los hilos de infección, así como también la organogénesis del nódulo y la modulación de las respuestas de defensa requeridas para la infección.

El gen Phvul.006G120700 codifica una giberelina 2-oxidasa y es reprimido ante la inoculación tanto con SC15 como por CE3 (Figura 24A). Estas enzimas catabolizan formas bioactivas del ácido giberélico (GA) (GA1 y GA4) y de sus precursores inmediatos (GA9, GA12, GA20 y GA53) a formas inactivas mediante la hidroxilación  $2\beta$ . Dado que el flujo de formas bioactivas de GA está dada por el balance entre su síntesis y degradación, y que el GA es generalmente considerado un modulador positivo de la formación y desarrollo de nódulos (Hayashi et al., 2014), la disminución de los niveles de Phvul.006G120700 podría estar aumentando los niveles de formas bioactivas de GA y promoviendo la infección y/o desarrollo del nódulo de una manera más eficiente (Figura 24A).



**Figura 24. Genes con expresión diferencial en los cultivares NAG12 y Alubia frente a las cepas *nodC-α* y *-δ*, respectivamente, vinculados al desarrollo y el metabolismo y respuesta a hormonas.**

Para cada gen diferencial en NAG12 en respuesta a cepas tipo *nodC-α* (A) o diferencial en Alubia en respuesta a cepas tipo *-δ* (B), se esquematiza el proceso biológico en el que estaría involucrado y se muestran los niveles de FPKM para cada gen en cada una de las muestras. El color rojo y azul indica si el gen se induce o reprime diferencialmente respecto a YEM, respectivamente, en respuesta al tipo de cepa indicada.

Recientemente se ha reportado que se produce un pico máximo en la producción de etileno en *L. japonicus* a las 24 hpi con *M. loti* (Reid et al., 2018). Entre los genes

## Capítulo I

regulados positivamente por las cepas tipo  $-\alpha$  se encuentra el gen Phvul.002G253900, que codifica el factor de transcripción EIN3 (*Ethylene insensitive 3 family protein*), el cual es un componente clave de la respuesta a etileno que promueve la expresión de factores de transcripción de la familia ERF (Guinel, 2015). Interesantemente, entre los factores de transcripción inducidos por estas cepas también se encuentra un miembro de la familia ERF (*Ethylene responsive factor*) (Figura 24A).

En Alubia, las cepas tipo  $-\alpha$  activaron la expresión de genes con funciones potencialmente relacionadas al proceso de simbiosis. El gen Phvul.006G073100 codifica una proteína de la familia LOB (*Lateral organ boundaries*), los cuales son factores de transcripción específicos de plantas implicados en el desarrollo de órganos vegetales (Majer and Hochholdinger, 2011). El gen Phvul.008G129300 codifica la proteína DCL2 (*Dicer-like 2*), la cual está involucrada en la producción de pequeños RNA de interferencia (siRNA) (Mallory and Vaucheret, 2006). Por último, el gen Phvul.004G101800 codifica una proteína PSF2 que es requerida para la iniciación de la replicación del DNA (Riera et al., 2017). Las cepas *nodC*- $\delta$  indujeron la expresión de dos genes que codifican miembros de la familia *SAUR-like auxin-responsive protein family* (Phvul.007G150200; Phvul.011G172600), la cual es la mayor familia génica implicada en la respuesta temprana de auxinas (Ren and Gray, 2015) (Figura 24B). Esta hormona es clave en la organogénesis del nódulo, ya que su acumulación local es un prerrequisito para el desarrollo del nódulo. De hecho, la inhibición de transportadores de auxinas genera pseudonódulos en ausencia de rizobios (Mathesius et al., 1998; Hirsch et al., 1989). Llamativamente, las cepas tipo *nodC*- $\delta$  reprimen la expresión de 3 genes que codifican miembros de la familia de nodulinas *MtN21/EamA-like transporter family protein* (los tres miembros de la categoría nodulación) (Figura 24B). Específicamente, los miembros de esta familia codifican transportadores bidireccionales de aminoácidos y

de auxinas. Plantas de *A. thaliana* con pérdida de función en el gen AtWAT1, uno de los miembros de la familia *MtN21/EamA-like transporter family protein*, mostraron deficiencias en la producción y transporte de auxinas (Denancé et al., 2014). Estos resultados sugieren que en el cultivar Alubia las cepas tipo *nodC-δ* podrían estar modulando los niveles de auxinas de manera local aumentando la expresión de miembros de la familia SAUR-like (Figura 24B).

### **8. Cambios transcripcionales de los componentes principales de la vía de transducción de señales de la nodulación en respuesta a las distintas cepas de *R. etli***

El estudio de la interacción simbiótica fijadora de nitrógeno entre rizobios y leguminosas ha permitido identificar los componentes claves de la percepción del factor Nod, que como se mencionó en la introducción, involucran receptores de membrana, proteínas citoplasmáticas y proteínas nucleares (Oldroyd et al., 2011). Con el fin de evaluar la respuesta transcripcional de estos componentes de la vía en respuesta a las distintas cepas, se analizó la expresión de cada uno de los genes listados en la Tabla 3 y se determinó si los mismos presentaban una expresión diferencial en al menos alguna condición. Los genes resaltados en verde son aquellos que se expresan de manera diferencial en alguna de las condiciones analizadas y fueron incluidos en la Figura 25. Como era de esperarse, no todos los genes de la vía muestran una expresión diferencial a 24 hpi, que fue el tiempo analizado en este trabajo. Esto puede deberse simplemente a una cuestión temporal, o a que los mismos no sean regulados a nivel transcripcional o postranscripcional. De hecho, Reynoso y colaboradores han demostrado previamente en *M. truncatula* que componentes de la vía de la nodulación como CCaMK y CYCLOPS no muestran cambios en los niveles estacionarios de RNA o en su asociación a

polisomas en respuesta al rizobio (Reynoso et al., 2012, 2013; Levy et al., 2004; Messinese et al., 2007).

**Tabla 3. Genes canónicos de la vía de transducción de señales de la nodulación**

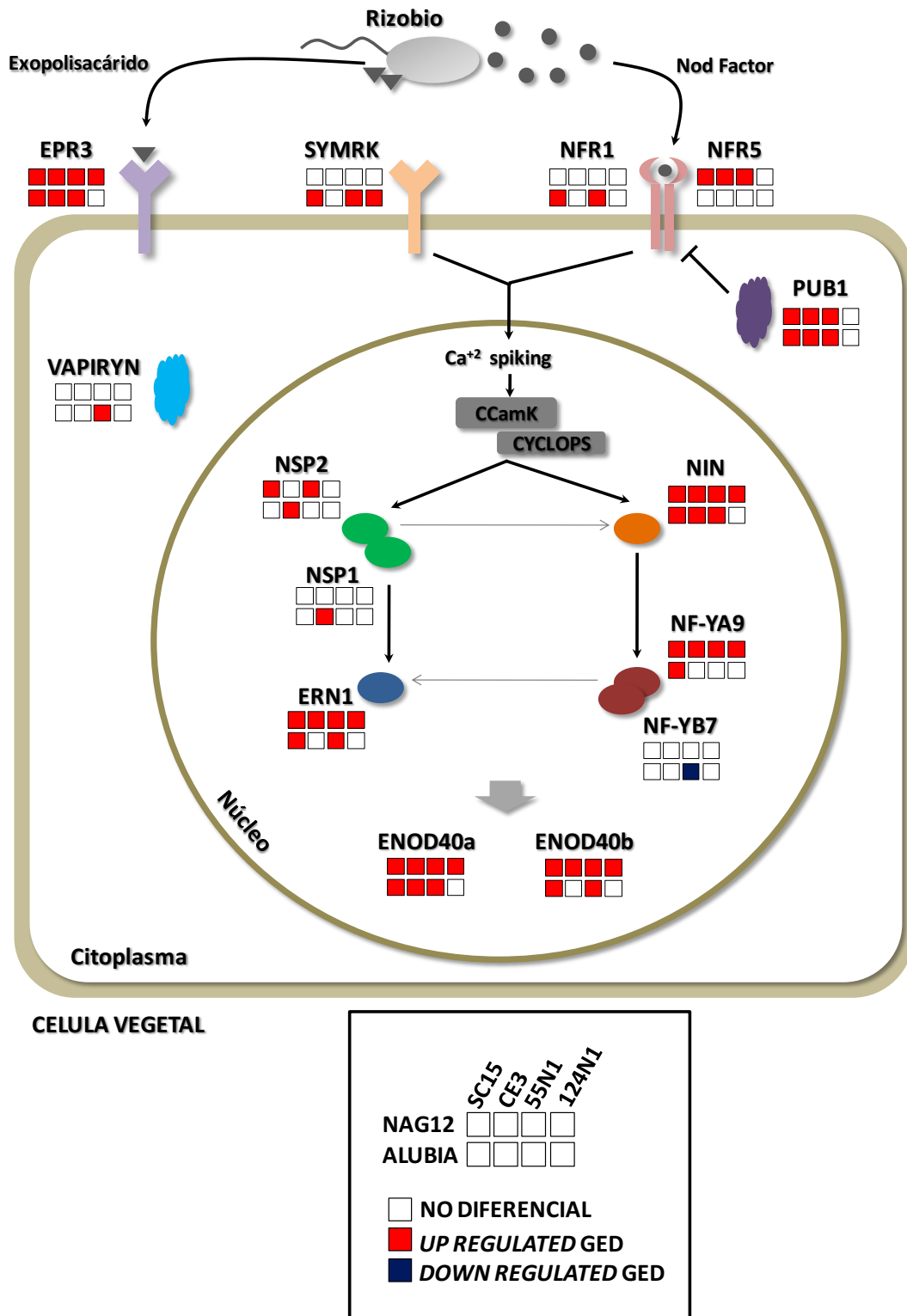
GEN	ID	GEN	ID
ERN1	Phvul.001G111800	VAPIRIN	Phvul.001G113900
NIN	Phvul.009G115800	NFR1	Phvul.008G211200
NF-YC1	Phvul.006G093200	NFR5	Phvul.002G025500
NF-YA1	Phvul.001G196800	HMGR	Phvul.002G118900
NF-YA9	Phvul.007G267100	FLOT2/4	Phvul.009G090700
EPR3	Phvul.002G059500	NUP85	Phvul.001G074000
ENOD40	Phvul.002G064166	NUP133	Phvul.008G284900
ENOD40	Phvul.002G064232	NENA	Phvul.002G163300
RIP	Phvul.008G249800	CASTOR	Phvul.009G263100
SYMRK	Phvul.011G148700	POLLUX	Phvul.006G218700
SYMREM1	Phvul.002G287608	NSP1	Phvul.010G129400
PUB1	Phvul.008G222100	NSP2	Phvul.009G122700
MCA8	Phvul.001G172300	LHK1	Phvul.002G324900
DELLA	Phvul.001G230500	NF-YB7	Phvul.003G086800
CCaMK	Phvul.011G186900	CNGC15A	Phvul.005G133900
CYCLOPS	Phvul.002G128600		

De los componentes canónicos de la vía, los genes *EPR3*, *NIN*, *ENOD40a/b*, *ERN1* y *PUB1*, son los que muestran una expresión diferencial prácticamente con todas las cepas en ambos cultivares. El receptor de exopolisacáridos *EPR3*, el factor de transcripción *NIN* y el RNA no codificante *ENOD40a* muestran el mismo patrón; se encuentran diferencialmente inducidos por las cuatro cepas en ambos cultivares, con la excepción de *R. etli* 124N1 en el cultivar Alubia. *ERN1* y *ENOD40b* no muestran expresión diferencial en Alubia inoculado con CE3 y 124N1, y *PUB1* no se induce únicamente por 124N1 (Figura 25). La inducción de estos genes a 24 hpi, sugiere que a tiempos tempranos de la interacción simbiótica las células de la raíz de los cultivares NAG12 y Alubia son capaces de percibir las moléculas de señalización extracelular del rizobio mediante receptores de membrana, iniciando la activación jerárquica de factores de transcripción requerida para la expresión de nodulinas. Llamativamente, la cepa

## Capítulo I

124N1, a pesar de ser una cepa eficiente en Alubia, no dispara la inducción de algunos estos componentes a las 24 hpi. De hecho, en Alubia sólo induce la expresión diferencial del receptor de membrana *SYMRK* (Figura 25).

A diferencia de los genes mencionados anteriormente, algunos componentes de la vía sólo aparecen como diferenciales específicamente en alguno de los cultivares. El receptor del Factor Nod *NFR5*, se induce en NAG12 sólo por SC15, CE3 y 55N1. Por el contrario, *VAPIRYN*, *SYMRK*, *NFRI*, *NSP1* y *NF-YB7* muestran una expresión diferencial exclusivamente en el cultivar Alubia. *VAPIRYN* y *NF-YB7* se encuentran regulados positiva y negativamente, respectivamente, en el cultivar Alubia en respuesta a la cepa 55N1. Llamativamente, estudios funcionales de *NF-YB7* realizados en NAG12, mostraron que el silenciamiento de este gen mejora la competitividad de 55N1 en este cultivar, y disminuye la formación de nódulos por SC15 (los resultados mencionados fueron realizados en el transcurso de esta tesis aunque no fueron incluidos). El hecho de que la cepa 55N1 genere la represión de *NF-YB7* en Alubia podría contribuir a su mayor eficiencia simbiótica respecto a cepas *nodC- $\alpha$* . Por su parte, *SYMRK* y *NFRI* muestran una inducción en respuesta tanto a cepas tipo *nodC- $\alpha$*  como *- $\delta$* , sugiriendo que su activación no estaría involucrada en los mecanismos moleculares responsables de la mayor eficiencia simbiótica de las cepas tipo *nodC- $\delta$* . Es más, *NSP1* sólo se encuentra activado por el rizobio en Alubia en respuesta a CE3, una cepa *nodC- $\alpha$*  (Figura 25).



**Figura 25. Activación de la vía de transducción de señales de la nodulación en las distintas interacciones planta-microorganismo.**

Para cada uno de los genes de la vía de transducción de señales de la nodulación se indica si se encuentra entre los genes diferenciales *up regulated* (cuadrado rojo), *down regulated* (cuadrado azul) o no es diferencial (cuadrado blanco), en el cultivar NAG12 (cuadrados de la fila superior) o Alubia (cuadrados de la fila inferior), para cada una de las cepas de *R. etli* utilizadas (de izquierda a derecha; SC15, CE3, 55N1 O 124N1).

## Capítulo I

Por último, los genes *NSP2* y *NF-YA9* muestran un patrón más variable. *NSP2* se encuentra inducido en NAG12 en respuesta a SC15 y 55N1, y en Alubia sólo en respuesta a CE3. En NAG12, *NF-YA9* se induce diferencialmente en respuesta a las cuatro cepas, mientras que en Alubia sólo lo hace en respuesta a SC15 (Figura 25).

En resumen, los resultados muestran que en ambos cultivares se generan cambios en los niveles estacionarios de RNA de los componentes canónicos de la vía de transducción de señales de la nodulación en respuesta a las cepas estudiadas. Estos cambios fueron evidentes tanto en el caso de los receptores de membrana como en los componentes posicionados río abajo en la vía, con la excepción del cultivar Alubia en respuesta a *R. etli* 124N1. Por otro lado, no se observa una correlación entre la eficiencia simbiótica de las cepas y su capacidad de activar los diversos componentes de la vía de transducción de señales, sugiriendo que los mecanismos moleculares de compatibilidad cultivar/cepa se posicionan de manera paralela a esta vía, y posiblemente involucren los genes activados o reprimidos específicamente descritos en la sección anterior.

## DISCUSIÓN

---

La producción agrícola depende fuertemente de la utilización de grandes cantidades de fertilizantes sintéticos. Estos productos químicos aportan al suelo la cantidad de nitrógeno necesaria para que los cultivos alcancen los niveles de producción requeridos para alimentar a la población mundial. Sin embargo, la producción y aplicación masiva de fertilizantes genera serios problemas ambientales como la eutrofización y contaminación de cursos de agua, y un aumento en la producción de gases de efecto invernadero (Erisman et al., 2008). Por este motivo, conocer los mecanismos biológicos



## Capítulo I

por los cuales las leguminosas son capaces establecer una asociación simbiótica fijadora de nitrógeno eficiente es de suma importancia para lograr una agricultura sustentable (Galloway et al., 2013). Sin embargo, el enorme progreso alcanzado en la comprensión de estos mecanismos no se ha trasladado a la agricultura, donde la fijación de nitrógeno se considera subóptima en términos de la capacidad teórica del proceso. Esto se debe principalmente a una pobre comprensión de los aspectos ecológicos de la interacción entre las leguminosas y las complejas comunidades microbianas que habitan los suelos; típicamente un gramo de suelo contiene entre  $10^6$ - $10^9$  células bacterianas que comprenden más de 10.000 especies diferentes (Gans et al., 2005). En este contexto, la eficiencia de fijación de nitrógeno y la capacidad competitiva de diferentes especies y cepas de rizobios, así como los mecanismos de selección de la planta, determinan la eficiencia global de la simbiosis, y en consecuencia, el rendimiento de los cultivos. A pesar que los rizobios utilizados como inoculantes en la agricultura fijan nitrógeno muy eficientemente, suelen tener un bajo impacto en el rendimiento ya que no son capaces de competir con las especies de rizobio autóctonas de los suelos. Los estudios transcriptómicos pueden contribuir a identificar genes claves para la eficiencia o especificidad de cepa, como lo ha demostrado previamente el uso de la técnica de hibridación sustractiva, la cual que permitió identificar NF-YC1 como un gen determinante de la competitividad observada en NAG12 (Meschini et al., 2008; Zanetti et al., 2010).

El estudio transcriptómico realizado en este trabajo de tesis ha permitido caracterizar los cambios moleculares que ocurren en *P. vulgaris* en respuesta a cepas de *R. etli* que exhiben distinta eficiencia en la formación de nódulos, y además, difieren en la capacidad competitiva de ocupar nódulos en el cultivar Mesoamericano NAG12. Si bien con el surgimiento de las técnicas de secuenciación masiva de RNA un gran número de

## *Capítulo I*

trabajos científicos han abordado el estudio de los cambios que ocurren a nivel transcriptómico cuando las raíces de las leguminosas son inoculadas con rizobios, la mayoría de estos trabajos han sido realizados utilizando una cepa de rizobio de referencia y una variedad particular de la especie vegetal compatible (Libault et al., 2010b, 2010a; Lohar et al., 2006; Kouchi et al., 2004; O'Rourke et al., 2014). Esto ha llevado a que se preste poca atención a los cambios transcriptómicos que ocurren en respuesta a la inoculación con cepas de rizobio con distinta eficiencia simbiótica, o las diferencias que ocurren entre distintas variedades de una misma especie vegetal.

El análisis general de los genes con expresión diferencial en NAG12 y Alubia mostró que los genes pertenecientes a las categorías de estrés oxidativo y óxido-reducción se encuentran sobrerrepresentados. Estos resultados son consistentes con diversos estudios transcriptómicos y funcionales donde muestran que a tiempos tempranos de la interacción simbiótica se genera una inducción en la producción de especies reactivas de oxígeno (ROS) que posteriormente es reprimida (Damiani et al., 2016; Cárdenas and Quinto, 2008; Rubio et al., 2004; Santos et al., 2001; Ramu et al., 2002). Se especula que el pico en la producción de ROS podría estar regulando la expresión génica y/o el estado redox de proteínas involucrados en la deformación de los pelos radicales, la progresión de los hilos de infección y el desarrollo del nódulo (Lohar et al., 2007; Montiel et al., 2016; Puppo et al., 2013). El análisis de enriquecimiento de GO no sólo reveló esta similitud entre las respuestas de ambos cultivares, sino que puso en evidencia que en el caso de Alubia también se encuentran enriquecidas otras categorías que incluyen la regulación de la expresión génica, la respuestas de defensa y la remodelación de la pared celular, entre otras. Esta diversidad también se evidencia en el número de genes con expresión diferencial. En el cultivar Alubia se identificaron 3754 genes con expresión diferencial en al menos una condición, mientras que en NAG12

## Capítulo I

sólo se identificaron 1230. Estos resultados podrían sugerir que los cambios transcripcionales de ambos cultivares en respuesta al rizobio han divergido notablemente. Sin embargo, hay que tener en cuenta que el número diferente de muestras (triplicados en las muestras de NAG12 y duplicados en el caso de Alubia) puede incidir sobre la significancia estadística que define el número de DEGs. De acuerdo con esta hipótesis, el análisis de componentes principales mostró que los perfiles transcriptómicos de plantas inoculadas muestran una mayor similitud en función del genotipo del cultivar, y no de la cepa utilizada en la inoculación. Estos resultados son coincidentes con los trabajos de Heat et al., 2012 y Burghardt et al., 2017, donde se examinó cómo varía la expresión génica de distintas variedades de *M. truncatula* en función de distintos microsimbiontes (Heath et al., 2012; Burghardt et al., 2017). Ambos trabajos, al igual que los encontrados en esta Tesis, muestran que la variación genética de las plantas tuvo la mayor influencia en los perfiles transcriptómicos.

Por su parte, el análisis de los genes diferenciales en respuesta a las cepas SC15, CE3, 55N1 y 124N1 respecto al control, permitió la identificación de grupos de genes que son modulados por las cuatro cepas, que son modulados sólo por cepas tipo *nodC- $\alpha$*  (SC15 y CE3) o *- $\delta$*  (55N1 y 124N1), o bien que son específicos de alguna cepa en particular. Es interesante notar que en ambos cultivares el número de GEDs que es común a las cuatro cepas es mayor que el número de GEDs específicos del tipo de cepa. Esto cobra sentido teniendo en cuenta que a pesar de diferencias cuantitativas en la eficiencia simbiótica, las cuatro cepas son capaces de inducir la formación de nódulos fijadores de nitrógeno, tanto en NAG12 como en Alubia. Esto también fue evidenciado en los patrones de expresión de los genes de la vía de transducción de señales activada por factores Nod. Las cuatro cepas regulan positivamente componentes de esta vía, aunque la intensidad

## Capítulo I

de la respuesta no se correlaciona con la eficiencia simbiótica, siendo el caso de la interacción entre 124N1 y Alubia el ejemplo más claro. Esta cepa, a pesar de establecer una simbiosis eficiente con Alubia, sólo regula positivamente la expresión de un componente de la vía (SYMRK) al tiempo estudiado en este trabajo. Es más, en NAG12, a pesar de establecer una interacción poco eficiente, 124N1 regula positivamente los genes *ERP3*, *NIN*, *ERN1*, *NIN*, *NF-YA9* y *ENOD40a/b*, incluso a pesar que la cantidad de GEDs respecto a YEM en este cultivar es aproximadamente la mitad que en Alubia. Un análisis similar de los genes de la vía simbiótica fue realizado con tres especies vegetales diferentes del género *Lupinus* (*L. albus*, *L. luteus* y *L. mariae-josephae*) inoculadas con cepas de *Bradyrhizobium japonicum* y *B. valentinum* (Keller et al., 2018). *L. albus* desarrolla nódulos en respuesta a ambas cepas aunque los nódulos ocupados por *B. valentinum* no se desarrollan completamente y fijan nitrógeno muy ineficientemente. Por su parte, *L. luteus* y *L. mariae-josephae* solo son capaces de desarrollar nódulos con las cepas compatibles *B. japonicum* y *B. valentinum*, respectivamente. Tanto en *L. luteus* como en *L. mariae-josephae*, se encontraron diferencias significativas en la expresión de varios componentes de la vía tales como SYMRK, PUB1 y NIN al comparar los transcriptomas de nódulos de plantas inoculadas con cepas compatibles e incompatibles. Llamativamente, el mismo análisis en *L. luteus* mostró que no existen diferencias significativas en la expresión de ninguno de los componentes de la vía a pesar que los nódulos formados por *B. valentinum* son muy ineficientes. Estos resultados contrastan con los de nuestro trabajo, donde se encontraron diferencias en la expresión de estos componentes, incluso aunque las cuatro cepas utilizadas inducen la formación de nódulos completamente desarrollados y fijadores de nitrógeno. Evidentemente, los mecanismos de compatibilidad y

## Capítulo I

reconocimiento entre rizobios y leguminosas dependen de la especie vegetal y de su historia evolutiva con los rizobios del suelo.

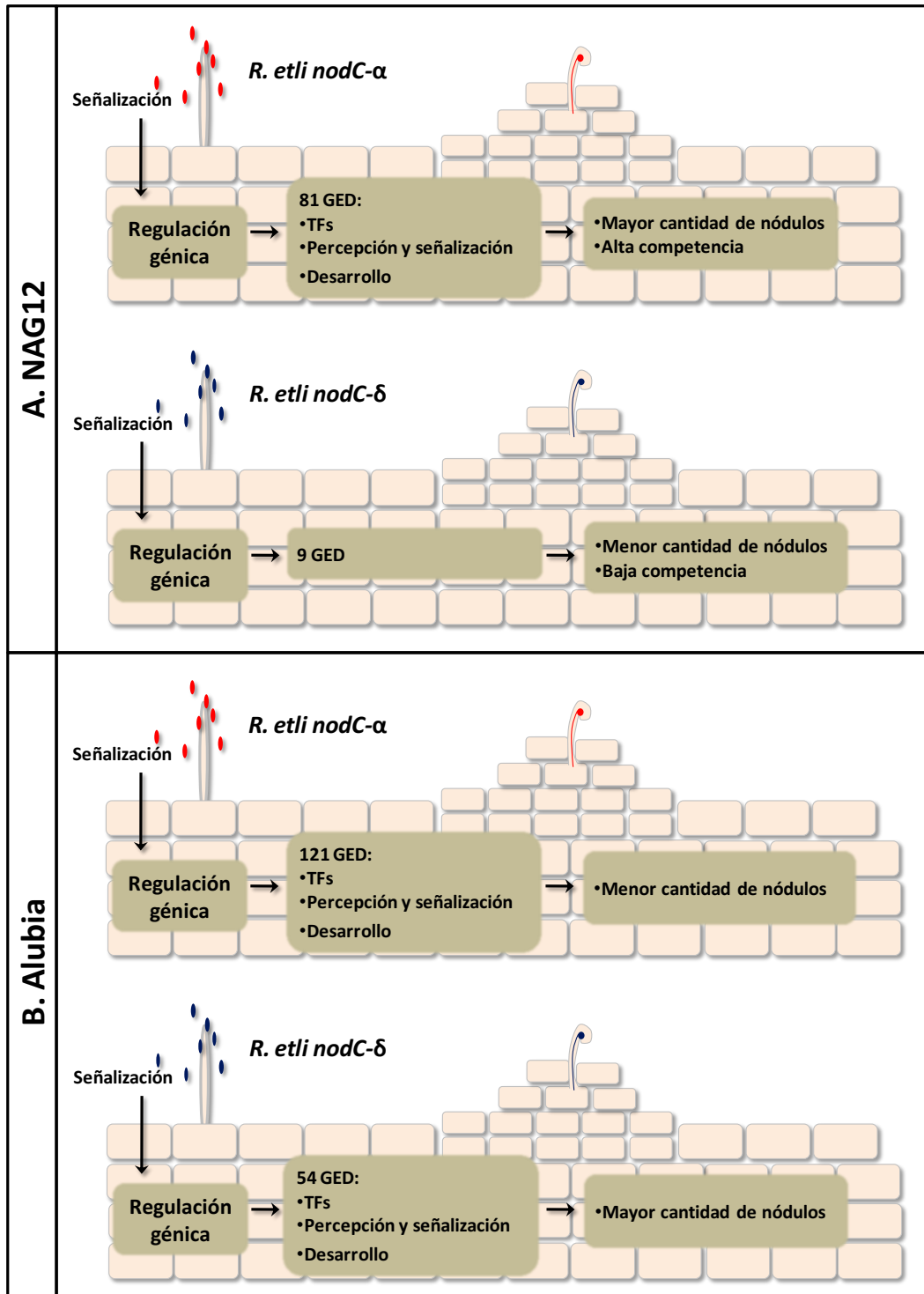
El conjunto de genes que son modulados específicamente por cepas tipo *nodC- $\alpha$*  o *- $\delta$*  son los principales candidatos a formar parte de los mecanismos moleculares de la planta responsables del establecimiento de una asociación simbiótica eficiente entre *P. vulgaris* y *R. etli*. Para el caso del cultivar NAG12, hemos identificado 81 GEDs en respuesta a las cepas eficientes *nodC- $\alpha$* , mientras que en respuesta a las cepas menos eficientes *nodC- $\delta$*  sólo se expresan diferencialmente 9 (Figura 26A). Esta fuerte diferencia en el número de GEDs podría explicarse por la existencia de vías de señalización que determinan la mayor eficiencia simbiótica de las cepas tipo *- $\alpha$*  en NAG12, y la capacidad de este cultivar de discriminar y nodular preferencialmente con estas cepas cuando las proporciones de cepas tipo *- $\alpha$*  y *- $\delta$*  en el medio son semejantes. En concordancia con esta hipótesis, la inoculación con cepas tipo *nodC- $\alpha$*  desencadenó en NAG12 cambios en los niveles de transcritos de factores de transcripción y genes que codifican proteínas vinculadas a la percepción de señales y los posteriores mecanismos de señalización. Sumado a esto, las cepas tipo *nodC- $\alpha$*  modularon la expresión diferencial de genes relacionados con programas de desarrollo que podrían promover la infección y/o la organogénesis del nódulo de manera más eficiente.

En Alubia, la respuesta molecular común a las cepas tipo *nodC- $\delta$*  comprende la regulación diferencial de 54 genes, 6 veces más de lo que ocurre en NAG12 (Figura 26B). Entre los genes diferenciales se encuentran 4 factores de transcripción y una proteína quinasa con dominios LRR y *Legume lectin*, que podrían estar activando y/o regulando vías de transducción de señales asociadas a una nodulación más eficiente. Sumado a esto, 5 de los GEDs comunes están relacionados con el transporte y la respuesta a auxinas. Teniendo en cuenta que el aumento local de auxinas es requerido

## Capítulo I

para la organogénesis del nódulo (Suzaki et al., 2013; Bensmihen, 2015), una regulación diferencial en las concentraciones locales de auxinas por las cepas tipo *nodC*- $\delta$  podría traducirse en una mayor eficiencia en la formación del nódulo. La ausencia de componentes moleculares comunes en las interacciones más eficientes (NAG12/*R. etli nodC*- $\alpha$  y Alubia/*R. etli nodC*- $\delta$ ) sugiere que los mecanismos moleculares asociados a la eficiencia serían propios de cada cultivar y habrían evolucionado independientemente. Esto es coherente con la fuerte importancia del cultivar por sobre otros factores que se desprende del análisis PCA.

Estas diferencias entre cultivares podrían estar relacionadas con la historia evolutiva de ambos cultivares. NAG12 es un cultivar Mesoamericano que ha co-evolucionado con la microbiota de los suelos de esta región desde el surgimiento de la especie *P. vulgaris* hace aproximadamente 4 millones de años (Delgado-Salinas et al., 2006). Esta larga historia co-evolutiva ha permitido el surgimiento de mecanismos moleculares de compatibilidad entre cultivares Mesoamericanos y cepas de *R. etli nodC*- $\alpha$  (Aguilar et al., 2004). No sólo los cultivares Mesoamericanos desarrollan una mayor cantidad de nódulos con estas cepas, sino que también nodulan de manera preferencial incluso en presencia de cepas - $\delta$  (Meschini et al., 2008; Zanetti et al., 2010) (Figura 26A). Por el contrario, el cultivar Alubia ha divergido evolutivamente de su antecesor Mesoamericano tan sólo hace 165.000 años en la región Andina (Bitocchi et al., 2012; Bitocchi et al., 2017; Schmutz et al., 2014; Rendón-Anaya et al., 2017). Este período de tiempo fue suficiente para que los cultivares Andinos desarrollen nódulos de manera más eficientes con las cepas tipo tipo *nodC*- $\delta$  predominantes en estos suelos, aunque no se observa actualmente un mecanismo de selección de tipo de cepa (Meschini et al., 2008; Mazziotta et al., 2013) (Figura 26B).



**Figura 26. Respuesta molecular y fisiológica de *P. vulgaris* en respuesta a cepas de *R. etli nodC-α* y *-δ*.**

En NAG12 se identificaron 81 y 9 genes con expresión diferencial (GED) en respuesta a cepas de *R. etli nodC-α* y *-δ*, respectivamente. Estas diferencias podrían determinar la mayor formación de nódulos por cepas *-α*, así como también la alta competitividad de estas cepas en experimentos de co-inoculación (A). En Alubia se identificaron 121 y 54 genes con expresión diferencial en respuesta a cepas de *R. etli nodC-α* y *-δ*, respectivamente. El cultivar Alubia, si bien forma más nódulos con cepas *nodC-δ* que con *-α*, no es capaz de nodular de manera preferencial con estas cepas en experimentos de co-inoculación (B).

## Capítulo I

Por último, el análisis transcriptómico realizado contribuye a complementar el conocimiento generado sobre los mecanismos de reconocimiento que determinan la compatibilidad entre ambos simbioses, es decir, los mecanismos que le permiten distinguir entre cepas que son o no capaces de nodular una determinada especie vegetal. Estos mecanismos históricamente fueron concebidos como mecanismos de “llave y cerradura” secuenciales. Una leguminosa secreta al medio flavonoides e isoflavonoides que sólo pueden activar el factor de transcripción nodD de los rizobios compatibles. En respuesta, estos sintetizan y liberan Factores Nod que son reconocidos por receptores compatibles ubicados en la membrana plasmática de las células vegetales, activando los programas genéticos requeridos para la formación del hilo de infección y organogénesis del nódulo. Una vez que comienza la infección bacteriana a través del hilo de infección, moléculas de señalización del rizobio como lipo y exopolisacáridos, y proteínas secretadas por los sistemas de secreción tipo III y IV, pueden ser reconocidas por la planta favoreciendo o abortando la infección (Perret et al., 2000; Via et al., 2016). El sistema biológico estudiado en esta tesis permitió caracterizar los mecanismos moleculares que determinan diferencias más sutiles en la interacción simbiótica entre cepas de rizobios y cultivares de *P. vulgaris* compatibles. En última instancia, estas diferencias tendrían un impacto en la eficiencia de la fijación de nitrógeno y por lo tanto en el rendimiento del cultivo de *P. vulgaris* y/o en el contenido de proteína de sus semillas.



# **CAPÍTULO II**

## Resultados y discusión

**Identificación de sitios de unión de los factores de transcripción NF-YA1, C1 y SIN1**

## RESULTADOS

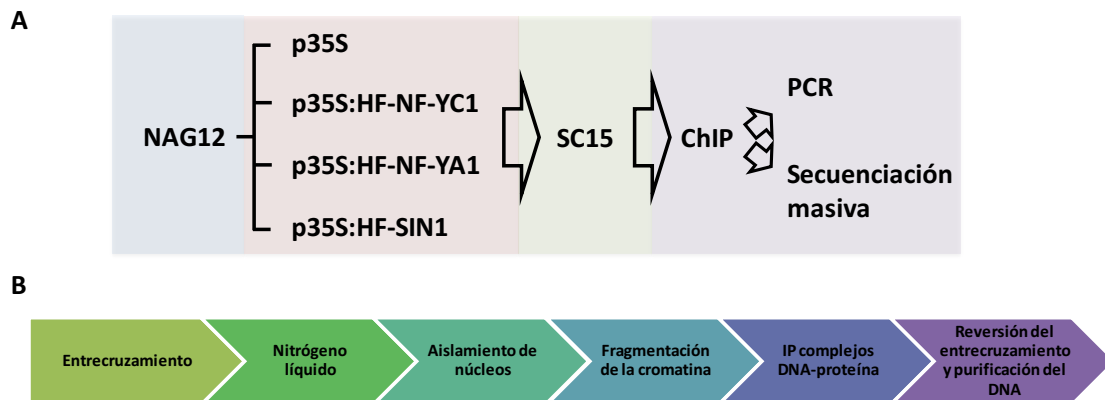
---

### 1. Diseño experimental

En este capítulo de la Tesis nos planteamos como objetivo identificar y caracterizar las regiones promotoras y los elementos regulatorios del DNA de aquellos genes que serían regulados de manera directa por los factores de transcripción PvNF-YC1, PvNF-YA1 y SIN1. Para ello se realizaron fusiones traduccionales de cada uno de los factores de transcripción al epítopo His-FLAG (HF) para luego utilizar la técnica de inmunopurificación de cromatina seguida de PCR (ChIP-PCR) o bien de secuenciación masiva de fragmentos de DNA (ChIP-seq). La técnica de ChIP se basa en co-inmunoprecipitar regiones de la cromatina que se encuentran asociadas a factores de transcripción (u otras proteínas) mediante el uso de anticuerpos, en este caso, anticuerpos anti-FLAG conjugados a bolillas magnéticas.

Se generaron plantas compuestas de *P. vulgaris* variedad NAG12 que sobreexpresan cada una de las proteínas de fusión bajo el control del promotor del virus del mosaico del coliflor (CaMV 35S promoter; p35S), el cual es un promotor fuerte y cuasi-constitutivo. Como control se transformaron plantas con el vector vacío. Las plantas compuestas se generan por transformación mediada por *Agrobacterium rhizogenes* y consisten en una parte aérea salvaje, no transformada, y un sistema de raíces transgénicas, en el cual cada raíz representa un evento de transformación independiente. El vector utilizado (pK7WG2D,1) para la transformación posee un cassette para la expresión del gen reportero GFP bajo el promotor *rolD*, permitiendo la visualización y selección de las raíces transgénicas bajo luz UV. Las plantas generadas fueron inoculadas con la cepa de *R. etli* SC15 y al cabo de 24 hpi se inició el protocolo de ChIP (Figura 27A). Este comienza con la fijación de las interacciones proteína-DNA por

entrecruzamiento mediante la utilización de formaldehído, seguido de la maceración del tejido en nitrógeno líquido. A partir de este tejido se realiza el aislamiento de los núcleos y la fragmentación de la cromatina mediante sonicación. Posteriormente, se lleva a cabo la inmunopurificación de los factores de transcripción HF-NF-YC1, HF-NF-YA1 y HF-SIN1 utilizando bolillas magnéticas conjugadas con el anticuerpo anti-FLAG o anti-IgG como control. Finalmente, se purifican los fragmentos de DNA (Figura 27B), los cuales pueden ser utilizado para reacciones de PCR (ChIP-PCR) o bien para la construcción de bibliotecas para secuenciación masiva de DNA (ChIP-seq).



**Figura 27. Diseño experimental y técnica de ChIP.**

**A.** Plantas de poroto variedad NAG12 fueron transformadas con el vector vacío como control, o con las construcciones que portan los factores de transcripción NF-YC1, NF-YA1 o SIN1 fusionados al epítipo His-FLAG (HF) bajo el control del promotor p35S. Las plantas compuestas fueron inoculadas con *R. etli* SC15. Al cabo de 24 hpi el tejido radical fue utilizado para realizar la técnica de inmunopurificación de cromatina (ChIP) seguida de PCR (ChIP-PCR) o para la construcción de las bibliotecas de secuenciación masiva (ChIP-seq). **B.** La técnica de ChIP requiere de la fijación de las interacciones proteína-DNA por medio de formaldehído, la pulverización del tejido en nitrógeno líquido, el aislamiento de los núcleos celulares, la lisis de los núcleos y fragmentación de la cromatina por sonicación y la inmunopurificación de los complejos proteína-DNA. Por último, se revierte el entrecruzamiento y se purifican los fragmentos de DNA.

## 2. Identificación de *targets* putativos de PvNF-YC1

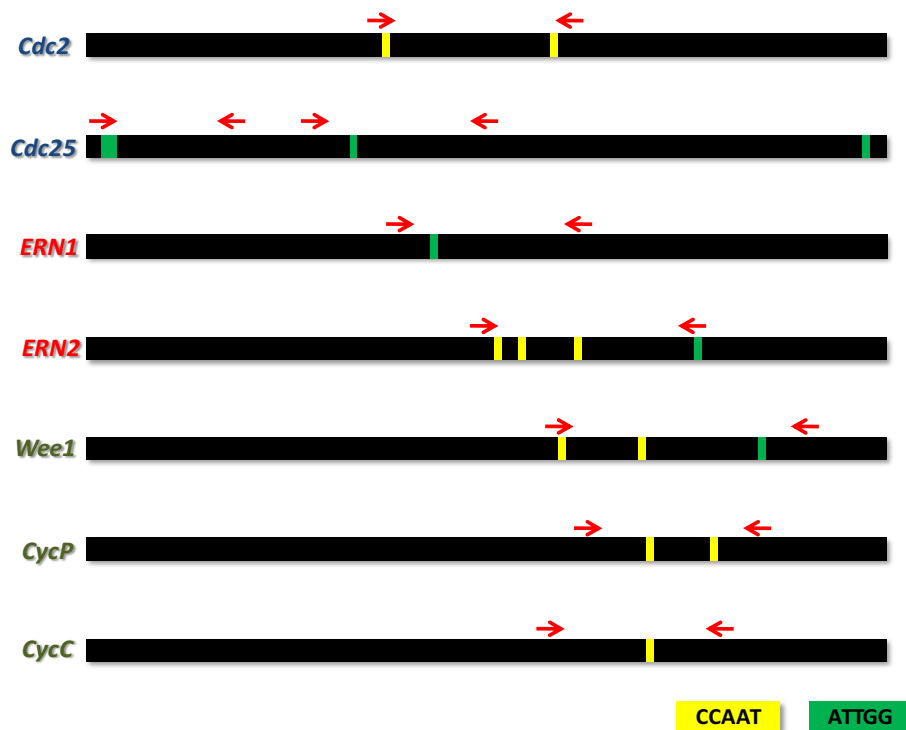
Previamente a la caracterización de los sitios de unión de los factores de transcripción PvNF-YC1, PvNF-YA1 y SIN1 a nivel genómico mediante ChIP-seq, se decidió abordar experimentos de ChIP seguido de PCR (ChIP-PCR) para identificar regiones

## Capítulo II

regulatorias que se encuentren asociadas a PvNF-YC1. Esta estrategia encuentra su fundamento en que hasta el momento se desconocen los genes blancos (*targets*) directos de estos factores de transcripción en *P. vulgaris* que permitan validar de manera robusta el éxito de las reacciones de inmunoprecipitación de cromatina. Los resultados publicados por Zanetti et al., 2010 sugieren que al igual que en sistemas animales, PvNF-YC1 podría regular genes del ciclo celular de la transición G2/M. Sumado a esto, estudios realizados en *M. truncatula* han demostrado que la subunidad MtNF-YA1, la MtNF-YC2 y la MtNF-YB16 se unen a elementos *CCAAT* presentes en la región promotora del factor de transcripción *MtERN1* (Laloum et al., 2014; Baudin et al., 2015). En base a esta información, decidimos evaluar si PvNF-YC1 se encuentra asociado a elementos *CCAAT* presentes en los promotores de *PvERN1* y de varios genes del ciclo celular de transición G2/M, tales como *PvCYCB*, *PvCDC2* y *PvCDC25*, que fueron reportados como *targets* putativos de PvNF-YC1 (Zanetti et al., 2010), así como también, otros genes del ciclo celular que presenten elementos *CCAAT* en sus promotores.

El ortólogo putativo del gen *MtERN1* en *P. vulgaris* fue identificado mediante un TBLASTN contra el genoma de referencia de *P. vulgaris* depositado en Phytozome, el cual es de origen Andino. Dado que los resultados del BLAST mostraron dos genes con un alto grado de identidad decidimos continuar el análisis con ambos genes, los cuales fueron denominados *PvERN1* (*Phvul.001G111800*) y *PvERN2* (*Phvul.004G081200*). Para identificar los genes del ciclo celular de transición G2/M se realizaron búsquedas por BLAST y palabra clave en el genoma de *P. vulgaris*. Para todos estos genes y *PvERN1*, *PvERN2*, *PvCYCB*, *PvCDC2* y *PvCDC25*, se analizaron las regiones promotoras hasta 1 Kb *upstream* del sitio de inicio de la transcripción (tanto en el genoma Andino y Mesoamericano) en busca de elementos *CCAAT* (Tabla 4). Se decidió

utilizar este criterio ya que en *M. truncatula*, MtNF-YA1 y MtNF-YC2 reconocen sitios *CCAAT* ubicados a aproximadamente 500 pb del sitio de inicio de la transcripción. Además, se tuvo en cuenta de manera cualitativa el perfil de expresión de cada gen en tejido de nódulos y raíz a partir de datos transcriptómicos depositados en Phytozome. A partir del análisis presentado en la Tabla 4 se seleccionaron aquellos genes que presentan regiones *CCAAT* tanto en el genoma Andino y Mesoamericano y se expresan en nódulos y/o raíz. En base a estos criterios se seleccionaron las regiones promotoras de *PvERN1*, *PvERN2*, *PvCDC25*, *PvCDC2*, *PvWEE* (*Phvul.001G204900*), *PvCYCP* (*Phvul.008G049600*), *PvCYCC* (*Phvul.007G143100*) para estudiar su asociación a PvNF-YC1. A partir de las secuencias promotoras se diseñaron *primers* flanqueando los elementos *CCAAT* para realizar ensayos de ChIP-PCR (Figura 28).



**Figura 28.** Elementos *CCAAT* y *primers* diseñados para identificar los posibles *targets* de PvNF-YC1.

Se esquematiza 1 Kb de la región promotora de los posibles *targets* de PvNF-YC1 (barra negra), así como los elementos *CCAAT* en orientación sentido (barras amarillas) o antisentido (*ATTGG*, barras verdes). Las flechas rojas indican la posición de los *primers* diseñados para determinar el enriquecimiento de estas regiones en las muestras de ChIP mediante PCR cuantitativa (qPCR).

## Capítulo II

**Tabla 4. Identificación y análisis de genes del ciclo celular y *PvERN*.**

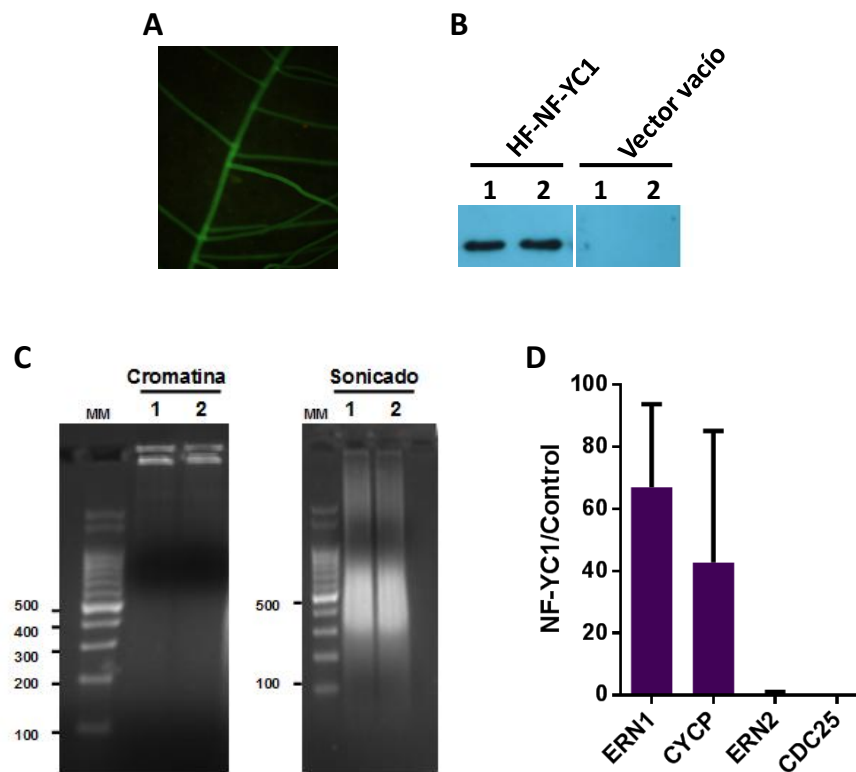
En la tabla se muestra el identificador de cada gen analizado junto con su nombre y la cantidad de elementos *CCAAT* o *ATTGG* (secuencia invertida) presentes en su promotor, a partir de los datos de los genomas Andino y Mesoamericano. También se muestran los niveles de expresión de cada gen en nódulos y raíz de plantas de origen Andino obtenidos de Phytozome.

Gen	Nombre	ANDINO						MESOAMERICANO			
		CCAAT	Posición CCAAT	ATTGG	Posición ATTGG	Expresión NÓDULOS	Expresión RAÍZ	CCAAT	Posición CCAAT	ATTGG	Posición ATTGG
Phvul.006G122800	cdc2	2	153/471	1	958	alta	alta	2	418/629	0	
Phvul.001G180100	cdc25	0		3	473/774/786	alta	baja	0		4	33/670/971/983
Phvul.001G119700	cycB	3	801/819/866	0		baja	alta	0		0	
Phvul.001G111800	ERN1	0		1	572	alta	alta	0		1	567
Phvul.004G081200	ERN2	3	121/184/222	0		media	muy alta	3	392/455/493	1	237
Phvul.001G204900	wee1	2	308/409	1	164	alta	muy alta	2	413/313	1	164
Phvul.008G049600	cycP	2	182/266	0		media	media	2	215/299	0	
Phvul.007G143100	cycC	1	311	1	707	muy alto	baja	1	308	0	
Phvul.002G015500	cycA	2	249/286	0		baja	baja	2	353/390	0	
Phvul.002G096900	cycA/ cycB	3	69/454/748	0		baja	baja	3	905/620/236	0	
Phvul.002G032500	cycA/ cycB	1	553	0		alta	alta	1	559	0	
Phvul.002G047100	cyc	1	612	1	8	baja	baja	1	423	1	179
Phvul.002G274800	cycB	2	47/558	0		baja	media	4	894/383/142/12	1	134
Phvul.004G152000	cdkA / cdkB	1	114	1	656	alta	alta	1	127	1	645
Phvul.007G054800	cdkA / cdkB	2	31/329	1	139	media	media	1	833	2	160/934
Phvul.007G179300	cdkA / cdkB	2	665/885	0		muy alta	alta	2	902/683	0	
Phvul.007G257900	cdkA / cdkB	2	574/918	2	614/836	muy alta	alta	1	54	0	
Phvul.008G144200	cdkA / cdkB	1	211	1	500	baja	alta	1	453	1	740
Phvul.011G054400	wee1	1	33	0		alto	muy alto	1	226	0	
Phvul.001G232800	cdc20	1	394	0		medio	medio	1	620	0	
Phvul.002G037700	cdc20	2	350/701	0		medio	alto	2	392/743	1	781
Phvul.001G096600	cdc25	1	283	2	466/935	media	alta	1	834	1	224
Phvul.002G112500	cdc25	1	573	0		media	muy alta	2	347/824	0	
Phvul.002G154800	cdc25	2	302/537	1	14	media	muy alta	3	72/633/863	3	101/201/345
Phvul.002G301900	cdc25	1	121	0		muy alto	muy alto	1	248	2	224/301
Phvul.003G100900	cdc25	5	75/326/411/655/944	0		media	media	1	858	1	165
Phvul.003G114400	cdc25	2	209/432	1	105	media	media	2	657/880	2	553/31
Phvul.003G148100	cdc25	1	484	2	497/807	media	media	3	326/488/510	1	79
Phvul.003G174700	cdc25	1	490	1	784	media	media	2	143/750	0	
Phvul.003G211900	cdc25	3	85/872/983	0		medio	alta	1	308	0	
Phvul.003G271000	ERN	1	661	0		alta	muy alta	3	12/65/951	3	
Phvul.002G117600	ERN	1	37	0		alto	muy alto	1	188	0	
Phvul.004G065600	ERN	2	605/673	0		media	media	1	495	3	223/553/792
Phvul.001G091000	cycA/ cycB	1	783	0		baja	baja	0		2	231/747
Phvul.001G149200	cycA/ cycB	1	555	3	176/429/927	alta	muy alta	0		1	287
Phvul.002G229000	cdc25	2	384/642	0		media	muy alta	0		3	264/356/662
Phvul.001G024900	cycA	3	288/344/683	0		baja	baja	2	796/740	1	714

### 3. PvNF-YC1 se encuentra asociado a elementos de los promotores de *PvERN1* y *PvCYCP* que contienen secuencias *CCAAT*

Raíces de plantas compuestas de *P. vulgaris* NAG12 que sobreexpresan HF-PvNF-YC1, o plantas control transformadas con el vector vacío, fueron utilizadas como material de partida para realizar ensayos de ChIP-PCR (Figura 29A). Previo a estos ensayos, se analizó la expresión de la proteína de fusión HF-NF-YC1 mediante *Western blot* utilizando anticuerpos anti-FLAG conjugados a la peroxidasa HRP (*horseradish peroxidase*). El revelado mediante quimioluminiscencia mostró una banda de aproximadamente 33 kD, consistente con la masa molecular de HF-NF-YC1 (Figura 29B). A continuación se procedió al asilamiento de los núcleos y la posterior fragmentación de la cromatina. La sonicación se ajustó para obtener una población de fragmentos de DNA con un tamaño promedio de 500 pb (Figura 29C). Por último, se realizó la inmunoprecipitación de HF-NF-YC1 y se purificó el DNA. El DNA obtenido fue utilizado como molde para reacciones de qPCR utilizando los *primers* diseñados para cada gen con el objetivo de determinar la asociación de PvNF-YC1 a las regiones promotoras de los genes seleccionados. Adicionalmente se utilizó un juego de *primers* diseñado a partir de una región genómica que no presenta secuencias *CCAAT* para normalizar los datos (*primers* qNIPK; Tabla 5) y corregir las variaciones técnicas producto del ChIP, tales como la diferencia de concentración de DNA entre muestras, la eficiencia de inmunoprecipitación y la variación en la recuperación del DNA. Los resultados del ChIP-PCR de muestras HF-NF-YC1 mostraron que las regiones con elementos *CCAAT* de los genes *PvERN1* y *PvCYCP* presentan un enriquecimiento relativo al control (plantas transformadas con el vector vacío) de 66,7 y 42,7 veces, respectivamente. Sin embargo, las regiones promotoras de los genes *PvERN2* y *PvCDC25* no mostraron un enriquecimiento significativo (Figura 29D). No se pudo

evaluar la asociación de PvNF-YC1 a los promotores del resto de los genes propuestos debido a la escasa cantidad de DNA obtenida para cada una de las muestras luego del protocolo de ChIP.



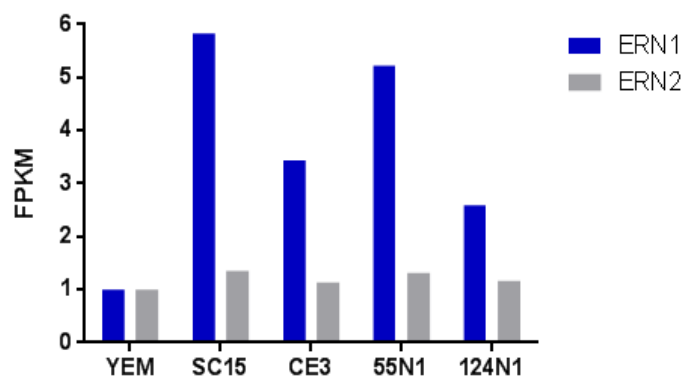
**Figura 29. Experimento de ChIP-PCR.**

Las raíces transgénicas fueron identificadas por la expresión de GFP en la lupa de fluorescencia (A) y la expresión de HF-NF-YC1 fue corroborada mediante un análisis por western blot en dos muestras de tejido de raíz colectado en experimentos independientes (1 y 2) (B). Luego del aislamiento de los núcleos se determinó el estado de fragmentación de la cromatina antes y después de la sonicación mediante una electroforesis en gel de agarosa (C). Grado de enriquecimiento de las regiones *CCAAT* de *PvERN1*, *PvCYCP*, *PvERN2* y *PvCDC25* en muestras de ChIP obtenidas a partir de raíces HF-NF-YC1 respecto a raíces transformadas con el vector vacío (Control) (D). Las regiones de interés fueron cuantificadas por qPCR y relativizadas a un fragmento de la cromatina que no posee regiones *CCAAT* utilizado como control. Las barras representan la media y SEM de dos réplicas técnicas. Se muestra los valores de un experimento representativo de dos experimentos independientes.

El fuerte enriquecimiento de la región *CCAAT* del promotor de *PvERN1* sugiere que este gen podría ser un target directo de PvNF-YC1 a tiempos tempranos de la interacción simbiótica, lo cual es consistente con lo reportado previamente en *M. truncatula* (Baudino et al, 2015). Estudios previos han mostrado que *ERN1* es inducido



fuertemente en etapas tempranas de la interacción entre *P. vulgaris* y la cepa de *R. etli* SC15, así como también en la establecida entre *M. truncatula* y *S. melilota* (Blanco et al., 2009; Zanetti et al., 2010; Cerri et al., 2012). Si bien Cerri y colaboradores han demostrado que *ERN2* también se induce tempranamente luego de la inoculación con rizobio en *M. truncatula*, nuestros resultados de ChIP-PCR sugieren que *PvERN2* no sería un *target* directo de PvNF-YC1 en poroto, por lo que es posible especular que *PvERN2* no muestre una inducción en los niveles de mRNA 24 hs luego de la inoculación. Para contrastar esta hipótesis, se analizaron los niveles de expresión de *PvERN1* y *PvERN2* en nuestros datos de secuenciación masiva reportados en los resultados del Capítulo 1 (Figura 30). Los resultados obtenidos son compatibles con la hipótesis propuesta; mientras que los niveles de mRNA de *PvERN1* mostraron una fuerte acumulación respecto a YEM luego de la inoculación con las cuatro cepas de rizobio, los niveles de *PvERN2* se mantuvieron relativamente constantes.



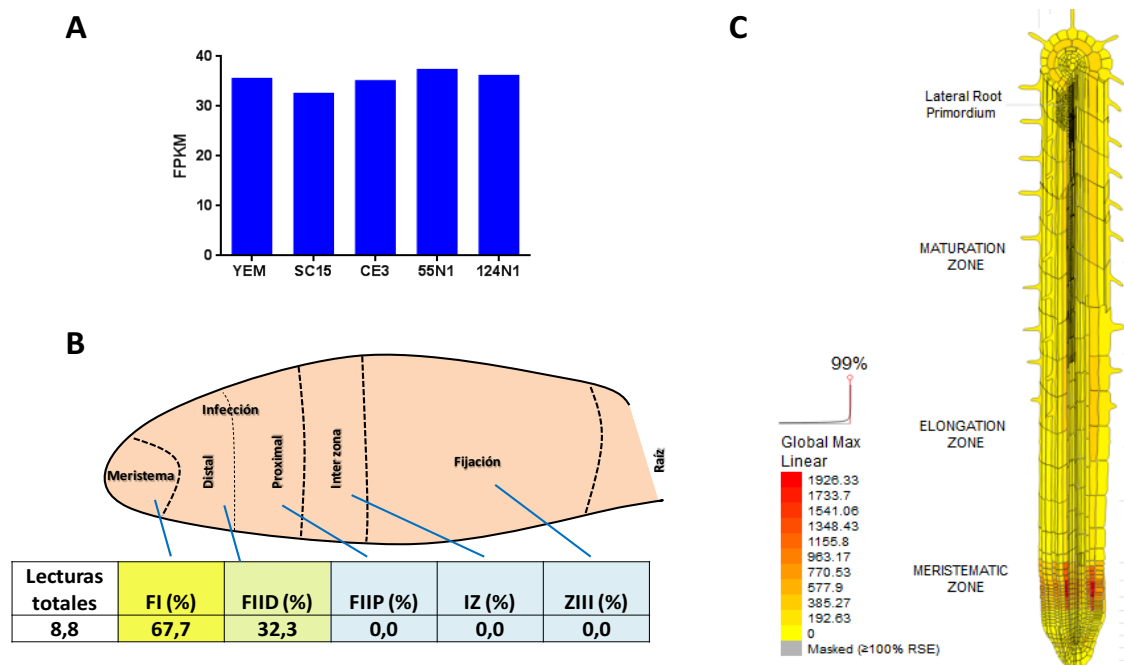
**Figura 30. Niveles de expresión de *PvERN1* y *PvERN2*.**

Niveles de expresión de *PvERN1* (barras azules) y *PvERN2* (barras grises) en el cultivar NAG12 frente a la inoculación con YEM (control) o con las cepas de *R. etli* SC15, CE3, 55N1 y 124N1. Los niveles de expresión están expresados en FPKM y los datos fueron normalizados por los valores de expresión de YEM.

Al igual que en *PvERN1*, las regiones *CCAAT* presentes en el promotor de *PvCYCP* se encuentran enriquecidas en el inmunoprecipitado de HF-NF-YC1 respecto al control,

## Capítulo II

sugiriendo que PvNF-YC1 podría estar regulando la expresión de *PvCYCP*. Sin embargo, los datos de secuenciación masiva muestran que no hay un aumento en los niveles estacionarios del mRNA *PvCYCP* en respuesta al rizobio (Figura 31A). Esto podría deberse a un efecto de dilución, ya que los datos de expresión utilizados corresponden a raíz completa y posiblemente *PvCYCP* presente un patrón de expresión específico de tipo celular. Teniendo en cuenta estos resultados, se decidió investigar su patrón de expresión en *M. truncatula* (Roux et al., 2014) y *A. thaliana* (<http://bar.utoronto.ca/eplant/>) haciendo uso de datos de secuenciación masiva de tipos celulares específicos disponibles. Para esto, los ortólogos putativos de *PvCYCP* en estas especies fueron identificados mediante un TBLASTN contra los respectivos genomas.



**Figura 31. Niveles de expresión de *CYCP* en *P. vulgaris*, *M. truncatula* y *A. thaliana*.**  
**A.** Niveles de expresión de *PvCYCP* en raíces de *P. vulgaris* inoculadas con YEM (control), y las cepas de *R. etli* SC15, CE3, 55N1 y 124N1. Los valores están expresados en FPKM. **B.** Niveles de expresión del ortólogo putativo de *PvCYCP* de *M. truncatula* (Medtr7g081130) en las distintas zonas del nódulo; meristema (FI), zona de infección distal (FIID) y proximal (FIIP), interzona (IZ) y zona de fijación (ZIII). **C.** Expresión del ortólogo putativo de *PvCYCP* (AT2G44740) en raíces de *A. thaliana*.

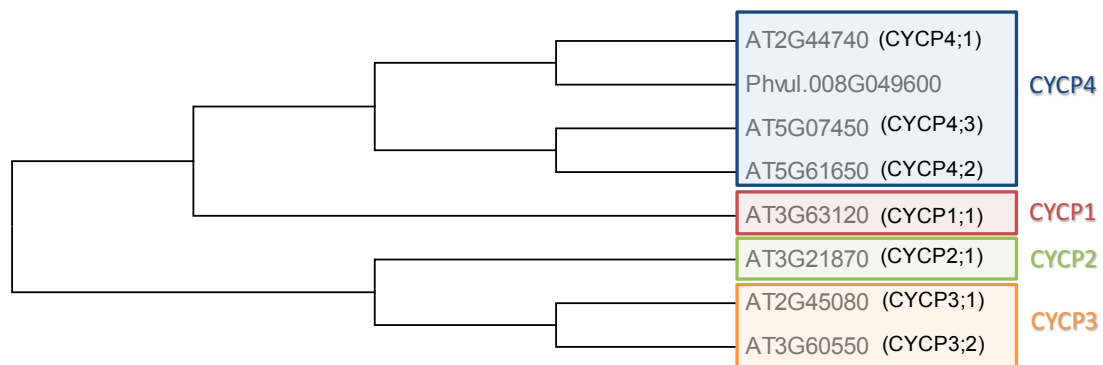
## Capítulo II

El patrón de expresión del ortólogo putativo de *PvCYCP* en nódulos de *M. truncatula* muestra claramente que esta ciclina se expresa principalmente en la zona meristemática (67,7%) y en menor medida en la zona de infección distal adyacente (32,3%), lo cual es consistente con el rol de las ciclinas en la regulación del ciclo celular (Figura 31B). De manera análoga, el ortólogo de *PvCYCP* en *A. thaliana* (*CYCP4;1*) se expresa fuertemente en la zona meristemática de la raíz (Figura 31C).

Las ciclinas de tipo P fueron identificadas por primera vez en *A. thaliana* mediante un *screening* de una biblioteca de doble híbrido de levadura en busca de proteínas que interaccionen físicamente con la quinasa dependiente de ciclina CDKA;1 (Torres Acosta et al., 2004). Fueron designadas con la letra P ya que presentan homología con la ciclina PHO80 de *Saccharomyces cerevisiae*, la cual regula diversos procesos celulares en función del estado nutricional de la célula (Kaffman et al., 1994; Smets et al., 2010). Sin embargo, en la bibliografía también se las puede encontrar con la designación de ciclinas tipo U debido a inconsistencias en la nomenclatura (Wang et al., 2004). En *A. thaliana* se identificaron 7 miembros pertenecientes a la familia CYCP, los cuales fueron agrupados en 4 clases dependiendo de sus relaciones evolutivas (Torres Acosta et al., 2004). Para determinar la clase a la cual pertenece *PvCYCP*, se realizó un alineamiento múltiple a partir de la secuencia proteica de *PvCYCP* y las 7 ciclinas de tipo P descritas en *A. thaliana*, el cual fue utilizado para construir un árbol filogenético (Figura 32). Los resultados muestran que *PvCYCP* pertenece a la clase 4 y que el mejor homólogo es *AtCYCP4;1*, por lo que decidimos denominarla *PvCYCP4;1*.

En resumen, mediante la técnica de CHIP-PCR se ha logrado determinar que HF-NF-YC1 se encuentra asociado de manera específica a elementos *CCAAT* presentes en los promotores de *PvERN1* y *PvCYCP4;1*, lo cual no solamente permite contar con un control positivo para las reacciones de CHIP-seq, sino que también vincula la subunidad

C1 del factor de transcripción NF-Y con la regulación de *PvERN1*, estableciendo un paralelismo con los mecanismos regulatorios que ocurren en la leguminosa modelo *M. truncatula* y que son esenciales para la infección rizobiana. Sumado a esto, la identificación de *PvCYCP4;1* como un posible blanco de PvNF-YC1 abre las puertas a futuros estudios genéticos y moleculares que exploren la función de esta ciclina en la nodulación.



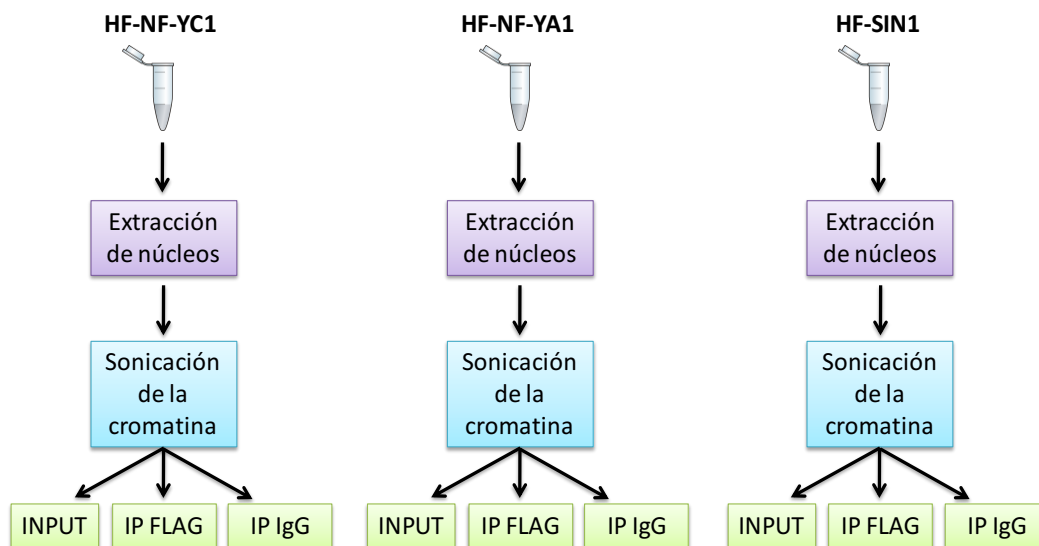
**Figura 32. Análisis filogenético de CYCP.**

Se muestra un árbol filogenético generado por el método de *neighbor-joining* utilizando un *bootstrap* de 1000, a partir de un alineamiento múltiple de las secuencias proteicas de las siete ciclinas tipo P de *A. thaliana* y la ciclina P de *P. vulgaris* identificada en este trabajo (Phvul.008G049600). En colores se indican las 4 clases de ciclinas tipo P.

#### 4. Construcción de bibliotecas para ChIP-seq

Habiendo identificado blancos directos de PvNF-YC1 que permitan evaluar el éxito de las reacciones de ChIP, se decidió iniciar los experimentos de ChIP-seq. Como se mencionó en la sección de “diseño experimental” de este capítulo, se utilizó como material de partida tejido de raíces transgénicas expresando las fusiones traduccionales HF-PvNF-YC1, HF-PvNF-YA1 o HF-PvSIN1 24 hpi con SC15. A partir de este tejido se realizó la extracción de núcleos y se llevó a cabo la sonicación de la cromatina de todas las muestras en simultáneo en condiciones controladas que permiten obtener fragmentos con un tamaño promedio de 200 pb. El material sonicado de cada muestra fue dividido en tres alícuotas. Una de ellas se conservó como *INPUT*, y las otras dos

fueron utilizadas en reacciones de inmunoprecipitación con bolillas magnéticas conjugadas al anticuerpo anti-FLAG (IP FLAG) o IgG (IP IgG) como control (Figura 33). Por último se realizó la elución de los factores de transcripción, se revirtió el *cross-linking* y se purificó el DNA. A diferencia del protocolo realizado para los experimentos de ChIP-PCR, no fue utilizado el tejido de plantas transformadas con el vector vacío, sino que se decidió utilizar como control negativo la inmunoprecipitación con IgG. Esta decisión se basa en que a la hora de realizar la secuenciación masiva de DNA es recomendable secuenciar un control negativo que provenga del mismo *INPUT* que la muestra de interés.

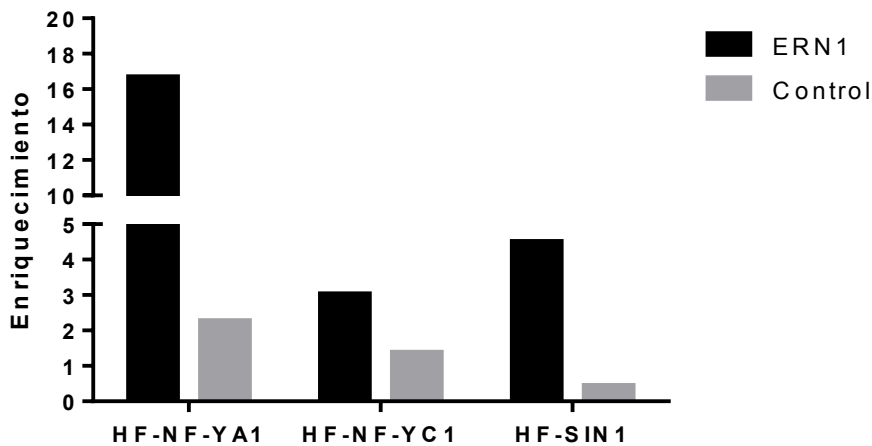


**Figura 33. Diseño experimental del ChIP-seq.**

Se muestra el esquema del diseño experimental seguido para el experimento de ChIP-seq. A partir de tejido de raíces que expresan las proteínas de fusión HF-NF-YC1, HF-NF-YA1 y HF-SIN1 se realizó la extracción de núcleos seguido de la sonicación de la cromatina, la cual fue dividida en tres alícuotas; una sin tratar (INPUT), y las restantes que fueron incubadas con bolillas magnéticas conjugadas con anticuerpo anti-FLAG (IP FLAG) o con IgG (IP IgG).

A partir de las muestras de DNA obtenidas del *INPUT*, IP FLAG e IP IgG de cada factor de transcripción, se realizaron reacciones de PCR en tiempo real para cuantificar el grado de enriquecimiento de la región promotora de *PvERN1* y de esta manera

evaluar el rendimiento de las reacciones de ChIP como se describió en la sección anterior. Si bien hasta el momento sólo demostramos que PvNF-YC1 se encuentra asociado a dicho *loci*, se evaluó también el grado de enriquecimiento para las muestras de HF-NF-YA1 y HF-SIN1, debido a que ambas proteínas interactúan físicamente con PvNF-YC1 (Battaglia et al, 2014, Baudin et al, 2015) y podrían formar complejos transcripcionales que actúen sobre los mismos genes *targets*. Como muestra la Figura 34, el IP FLAG de los tres factores de transcripción mostró un enriquecimiento de la región promotora de *PvERN1* de entre 3 y 16 veces respecto al IP IgG (barras negras), que en todos los casos fue superior al control negativo utilizando una región genómica no relacionada (Control, barras grises). Estos resultados indican que las reacciones de inmunoprecipitación fueron realizadas con éxito (Figura 34).



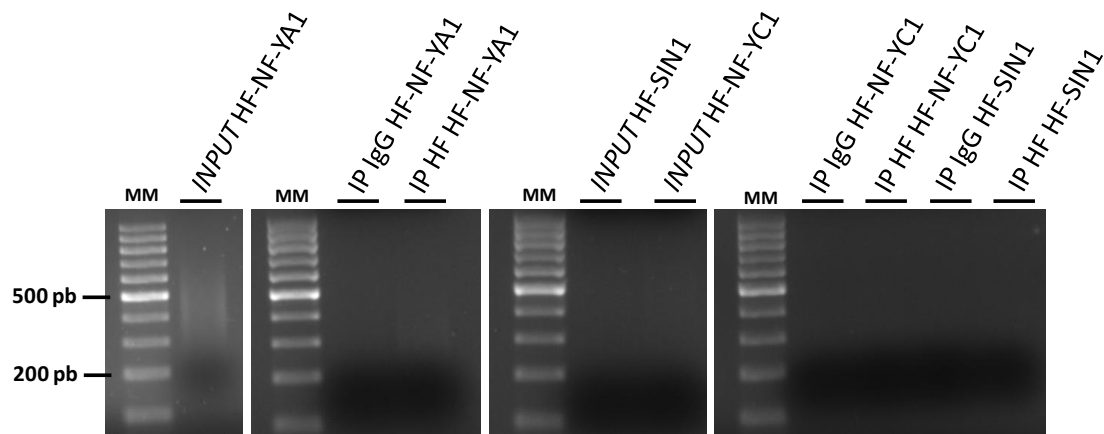
**Figura 34. Enriquecimiento de la región promotora de ERN1 en muestras de ChIP.**

Se muestra el grado de enriquecimiento del promotor de ERN1 (barras negras) y el control (barras grises), en muestras de HF-NF-YA1, HF-NFYC1 y HF-SIN1 inmunoprecipitadas con anti-FLAG (IP FLAG) respecto a muestras inmunoprecipitadas con IgG (IP IgG).

En base a los resultados anteriores, se procedió a llevar adelante la construcción de las bibliotecas de DNA para secuenciación masiva. Para esto se utilizó el mismo kit descrito en el Capítulo 1 (*TruSeq RNA Library Prep*). Teniendo en cuenta que en este caso el material de partida consistió en muestras de DNA, se procedió directamente con

## Capítulo II

el paso de reparación de extremos y se continuó el protocolo como se describió en el Capítulo 1. Para estimar la concentración y distribución de tamaño de las bibliotecas de DNA sintetizadas, se realizó una electroforesis en gel de agarosa a partir de una alícuota de las mismas (Figura 35). Lamentablemente, la baja intensidad de las bandas de DNA reveladas con bromuro de etidio sugiere una baja concentración de las bibliotecas. Los mejores resultados se obtuvieron con las muestras correspondientes al IP FLAG, IP IgG e *INPUT* de HF-NF-YA1, siendo el *INPUT* la muestra más concentrada. Como era de esperarse, en los tres casos puede observarse que la distribución de tamaños del DNA generado se encuentra entre 200 y 500 pb. Sin embargo, para determinar de manera cuantitativa si la concentración de las bibliotecas supera el umbral recomendado para llevar a cabo la secuenciación masiva de DNA, será necesario analizar las muestras mediante electroforesis capilar en el equipo *Agilent Technologies 2100 Bioanalyzer*. Desafortunadamente, estas mediciones exceden los plazos del presente trabajo de tesis aunque serán realizadas en un futuro cercano por otros miembros del laboratorio.



**Figura 35. Evaluación de la concentración y distribución de tamaños de las bibliotecas de DNA sintetizadas.**

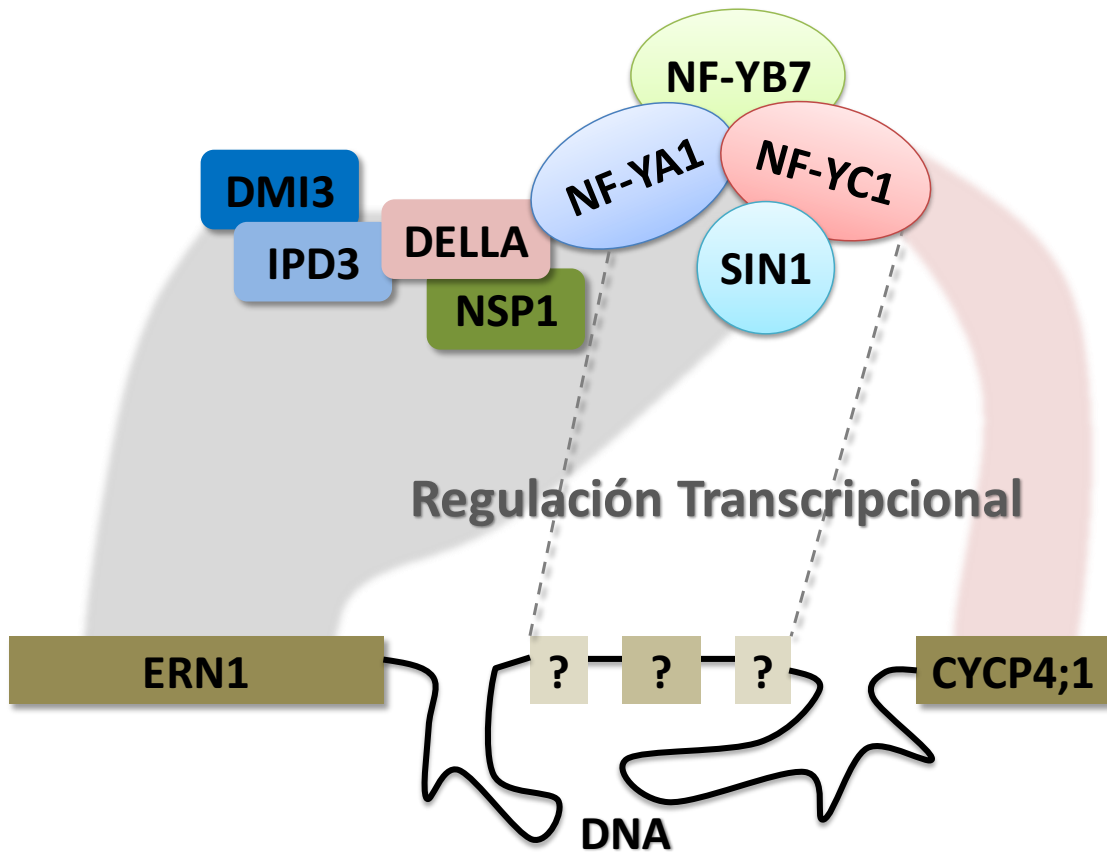
Electroforesis de DNA en gel de agarosa 1,2 % (p/v) de las bibliotecas sintetizadas reveladas con bromuro de etidio. Se indican las bandas de 500 y 200 pb del marcador de peso molecular (MM).

## DISCUSIÓN

---

Los factores de transcripción de tipo NF-Y y GRAS comprenden familias génicas numerosas en plantas que poseen 30 (Petroni et al., 2012) y 33 (Tian et al., 2004) miembros en *A. thaliana*, respectivamente. Estos reguladores transcripcionales están implicados en múltiples procesos biológicos como el desarrollo, arquitectura de la raíz, respuesta a distintos tipos de estrés e interacciones bióticas, entre otras (Zanetti et al., 2016; Sun et al., 2012). En los últimos años, diversos miembros de estas familias pasaron a ocupar un lugar central en la vía de transducción de señales requerida para el establecimiento de la interacción simbiótica fijadora de nitrógeno de plantas leguminosas con rizobios. Estudios realizados en *M. truncatula* han evidenciado que MtNF-YA1 es requerido para la infección y organogénesis del nódulo, y que es un regulador positivo de ERN1 (Combier et al., 2006; Laporte et al., 2013; Laloum 2014). Sumado a esto, se ha identificado el trímero simbiótico compuesto por las subunidades MtNF-YA1/B16/C2, los cuales se asocian al promotor de *ERN1* (Baudin et al., 2015). De manera consistente, el estudio de la subunidad NF-YA1 en las leguminosas de nódulos determinados *P. vulgaris* (Tesis Doctoral de Carolina Rípodas) y soja (Soyano et al., 2013) han revelado que esta subunidad es requerida para la infección y organogénesis del nódulo. En *P. vulgaris*, el rol de la subunidad NF-YC1 trasciende la regulación del programa de desarrollo del nódulo, ya que también es parte del mecanismo molecular por el que cultivares Mesoamericanos de *P. vulgaris* reconocen cepas de *R. etli* más eficientes (Zanetti et al., 2010). Al igual que en *M. truncatula*, se ha identificado la formación del trímero simbiótico compuesto por las subunidades PvNF-YA1/B7/C1 (Baudin et al., 2015) (Figura 36).





**Figura 36. Esquema de las interacciones entre factores de transcripción relacionados con NF-Y.**

Se esquematiza las distintas interacciones proteína-proteína que han sido reportadas entre factores de transcripción vinculados a NF-Y. Se ha reportado mediante estudios en *L. japonicus* y *M. truncatula* que DELLA, NSP1, NF-YA y NF-YC regulan la expresión del gen *ERN1*. Consistentemente, en este trabajo también se ha vinculado a los factores de transcripción de *P. vulgaris* NF-YA1, NF-YC1 y SIN1 con este locus. Sumado a esto, NF-YC1 mostró estar asociado al promotor de la ciclina *CYCP4;1*.

En cuanto a los factores de transcripción de tipo GRAS, NSP1 y NSP2 cumplen un rol clave en la simbiosis ya que son fundamentales para la inducción de los genes tempranos de la nodulación, incluyendo *ERN1* (Smit et al., 2005; Kaló et al., 2005; Heckmann et al., 2006; Baudin et al., 2015). Para que esto ocurra, es necesaria la formación del heterodímero NSP1/2 (Hirsch and Oldroyd, 2009; Hirsch et al., 2009). Llamativamente, en 2016 dos trabajos han reportado que otro miembro de la familia GRAS denominado MtDELLA, es capaz de interactuar físicamente con los factores de transcripción IPD3 (CYCLOPS), NSP2 y NF-YA1 promoviendo la formación de un

## Capítulo II

complejo transcripcional que regula positivamente la expresión de ERN1 y ENOD11 (Fonouni-Farde et al., 2016a, 2016b; Jin et al., 2016). Sumado a esto, en nuestro laboratorio se ha caracterizado funcionalmente un factor de transcripción de tipo GRAS denominado PvSIN1, que interacciona físicamente con PvNF-YC1 y que no sólo es requerido para la organogénesis del nódulo y progresión de los hilos de infección, sino que también está involucrado en el desarrollo de las raíces laterales (Figura 36). Estos tres factores de transcripción, PvNF-YC1, PvNF-YA1 y SIN1, son requeridos también para la activación de genes de la transición de fase G2/M del ciclo celular en etapas tempranas de la interacción simbiótica (Zanetti et al, 2010, Battaglia et al, 2014), los cuales participan presumiblemente de la activación de las divisiones celulares que dan lugar a la formación del primordio del nódulo.

Como puede apreciarse, el gran número de factores de transcripción involucrados en la vía de transducción de señales de la nodulación, así como sus múltiples interacciones, ponen de manifiesto la necesidad de dilucidar las redes transcripcionales que están regulando y con qué grado de superposición lo hacen. Llamativamente, hasta el momento no se han publicado trabajos que aborden la caracterización de estas redes a escala genómica, por lo que creemos que las bibliotecas de DNA que hemos generado a partir de la técnica de CHIP serán de gran valor para la comunidad científica abocada al estudio de la nodulación. Promisoriamente, hemos determinado por CHIP-PCR que PvNF-YC1 se encuentra asociado al promotor de la ciclina PvCYCP4;1, confirmando que este factor de transcripción podría estar regulando de manera directa genes del ciclo celular que promueven la organogénesis del nódulo. El estudio del patrón de expresión de los ortólogos de PvCYCP4;1 en *M. truncatula* y *A. thaliana* sugieren que esta ciclina podría funcionar en la regulación del ciclo celular a nivel de tipo celular, como el meristema de nódulos y/o raíz. En línea con esta hipótesis, en arroz se determinó que

## Capítulo II

CYCP4;1 es capaz de activar las divisiones celulares de un pequeño grupo de células de la cara abaxial del esclerénquima promoviendo la erección de las hojas (Sun et al., 2015). En base a esta información, será interesante realizar estudios funcionales mediante genética reversa para determinar si CYCP4;1 cumple algún rol en la organogénesis del nódulo (Figura 36).

También hemos mostrado que, de manera análoga a lo que ocurre en *M. truncatula* (Laloum 2014, Baudin et al., 2015), PvNF-YC1 y PvNF-YA1 se encuentran asociados al promotor de *PvERN1*, y que en *P. vulgaris* SIN1 también sería reclutado a este *locus*. Será interesante corroborar estos resultados a partir de los datos de CHIP-seq, y por sobre todo, esbozar las redes transcripcionales reguladas por estos factores de transcripción (Figura 36).

# CAPÍTULO III

## Resultados y discusión

Análisis funcional de NIPK (NF-YC1

*Interacting Protein Kinase)*

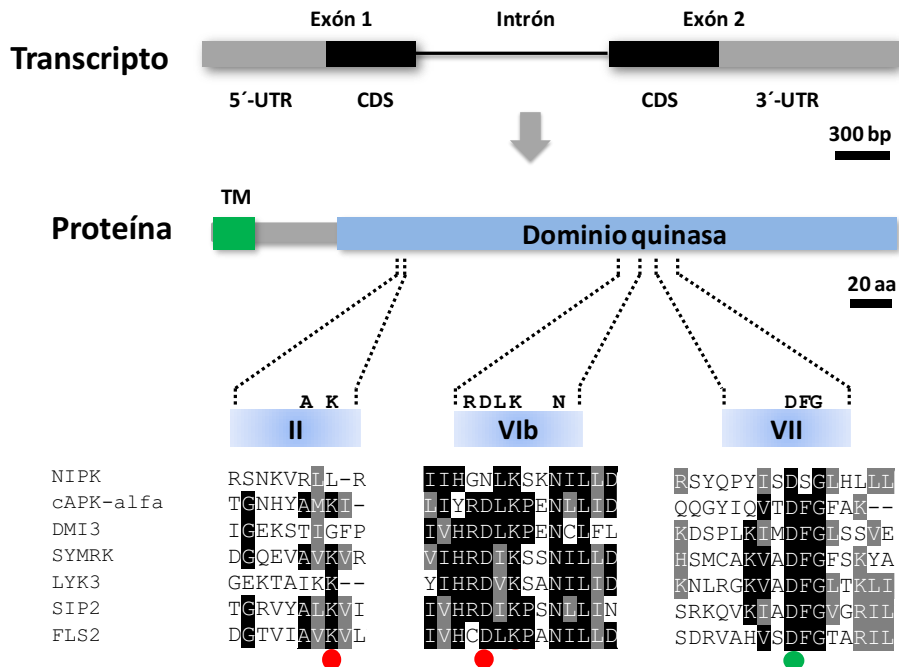
## RESULTADOS

---

### 1. Identificación de una proteína quinasa que interacciona con PvNF-YC1

Mediante el *screening* de una biblioteca de doble híbrido de levadura realizada en nuestro laboratorio (Battaglia et al., 2014) se identificaron una serie de clones que interaccionan físicamente con PvNF-YC1. Entre ellos, se identificó el factor de transcripción de tipo GRAS PvSIN1 mencionado en el capítulo anterior. En el presente trabajo se decidió abordar el estudio de otro de los clones identificados. Mediante un BLASTN contra la base de datos genómica de *Phaseolus vulgaris* v1.0 (<https://phytozome.jgi.doe.gov/>) se determinó que la secuencia identificada corresponde al gen *Phvul011G181900* ubicado en el cromosoma 11. El análisis del modelo génico reveló que el transcrito primario posee una longitud de 3.757 pares de bases (pb), conteniendo un intrón de 1035 pb y extremos 5' y 3' no traducidos (UTR; *untranslated region*) de 677 y 980 pb, respectivamente. El transcrito presenta un marco abierto de lectura de 1065 pares de bases (pb) (Figura 37). Por medio del análisis bioinformático de la secuencia proteica predicha de 353 aminoácidos (aa) se identificó un dominio transmembrana en el extremo N-terminal y un dominio quinasa de 285 aa en el extremo C-terminal, de acuerdo a los algoritmos usados por TMHMM y UNIPROT, respectivamente (Figura 37). En base a estos resultados nombramos al gen *Phvul011G181900* NIPK (*NF-YC1 Interacting Protein Kinase*). Con el objetivo de realizar un análisis detallado del dominio quinasa de NIPK, se realizó un alineamiento múltiple incluyendo la secuencia de aminoácidos del dominio quinasa de NIPK, así como también dominios quinasas de proteínas vegetales con actividad fosfotransferasa reportada, y de la proteína cPKA-alfa como secuencia de referencia para identificar los motivos característicos de esta familia de proteínas (Hanks et al., 1988). El análisis reveló que NIPK posee los once subdominios característicos de las proteínas quinasa, lo

que es consistente con la clasificación funcional establecida por UNIPROT (Figura Suplementaria S1). Sin embargo, dos de los tres aminoácidos reportados como indispensables para la actividad catalítica, no se encuentran conservados en NIPK (Figura 37 y Suplementaria S1) (Hanks, 2003).

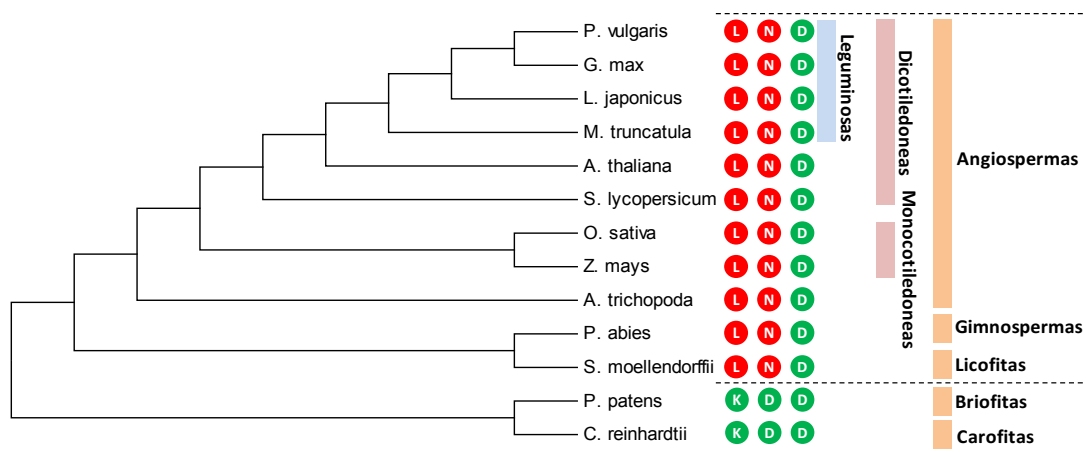


**Figura 37. NIPK codifica una pseudoquinasa.**

En la parte superior de la figura se muestra la representación esquemática del transcripto primario del gen *NIPK*. Las barras grises representan los extremos 5' y 3' no traducidos y las barras negras la secuencia codificante, la cual se encuentra interrumpida por un intrón (línea negra). El transcripto maduro codifica una proteína con un dominio transmembrana (TM) en el extremo N-terminal (barra verde) y un dominio quinasa en el extremo C-terminal (barra celeste). En la parte inferior de la figura se muestra el alineamiento múltiple de la secuencia proteica de NIPK, cAPK-alfa, DMI3 (Gleason et al., 2006b), SYMRK (Yoshida and Parniske, 2005), LYK3 (Jayaraman et al., 2017), SIP2 (Chen et al., 2012) y FLS2 (Lu et al., 2010; Sun et al., 2013), correspondiente a los motivos II, VIb y VII. Los círculos indican la posición de los tres aminoácidos requeridos para la actividad catalítica de las proteínas quinasa. El color verde y rojo de los círculos indica si el aminoácido se encuentra conservado o no en NIPK, respectivamente, en base a la secuencia considerada canónica.

La lisina del subdominio II requerida para anclar y orientar el ATP se encuentra sustituida por una leucina en NIPK, y el ácido aspártico del subdominio VIb que funciona como la base catalítica en la reacción fosfotransferasa es reemplazada por una asparagina. El único aminoácido de estos 3 conservado en NIPK es el ácido aspártico

del subdominio VII que es requerido para anclar y orientar el ATP junto con la lisina del subdominio II. Estos resultados sugieren que NIPK pertenece al subgrupo de proteínas quinasas que no poseen actividad catalítica debido a las sustituciones no conservativas en los motivos canónicos del dominio quinasa, conocidas como pseudoquinasas (Boudeau et al., 2006). Interesantemente, este patrón en los tres aminoácidos requeridos para la actividad catalítica se encuentra conservado en los ortólogos putativos de NIPK de especies evolutivamente tan distantes como *Selaginella moellendorffii* (una licófito), *Picea abies* (una gimnosperma), o *Amborella trichopoda*, considerada la angiosperma más elemental (Figura 38 y Suplementaria S2). Por el contrario, los ortólogos putativos de NIPK de la carófito *Chlamydomonas reinhardtii* y de la briófito *Physcomitrella patens* presentan el patrón canónico en estos tres aminoácidos, sugiriendo que la posible pérdida de actividad catalítica de NIPK surgió evolutivamente en las plantas vasculares.



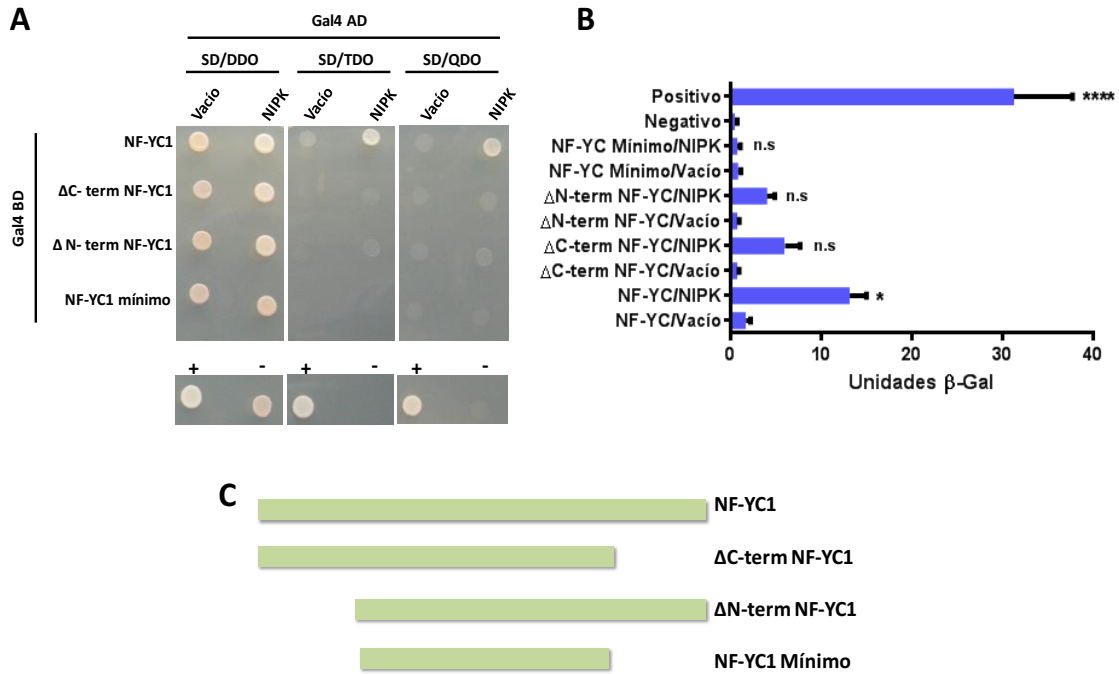
**Figura 38. Análisis evolutivo del dominio quinasa de NIPK.**

Se muestra un árbol filogenético de PvNIPK y los ortólogos putativos en las especies vegetales *Medicago truncatula*, *Arabidopsis thaliana*, *Lotus japonicus*, *Glycine max*, *Amborella trichopoda*, *Solanum lycopersicum*, *Zea mays*, *Oryza sativa*, *Selaginella moellendorffii*, *Picea abies*, *Physcomitrella patens* and *Chlamydomonas reinhardtii*. El árbol filogenético fue generado con el programa MEGA6 utilizando el método neighbor-joining. Como *input* fue utilizado un alineamiento múltiple de las secuencias proteicas realizado con Clustal Omega. Para cada especie vegetal, se indica el aminoácido correspondiente a los tres residuos requeridos para la actividad catalítica. El color verde o rojo de los círculos indica si el aminoácido presente se corresponde o no con el canónico, respectivamente.

## 2. NIPK interacciona con NF-YC1 en levaduras e *in planta*

Con el objetivo de verificar la interacción entre NF-YC1 y NIPK detectada mediante el *screening* de la biblioteca de doble híbrido de levadura, se realizó un experimento de doble híbrido utilizando los ORFs (*open reading frame*) completos de ambas proteínas. *NF-YC1* fue fusionada al dominio de unión al DNA de GAL4 (BD; *binding domain*), mientras que *NIPK* fue fusionada al dominio de activación (AD; *activation domain*). En este sistema, al producirse la interacción física entre las proteínas de interés se restituye la actividad del activador transcripcional GAL4, permitiendo la expresión de los genes marcadores *ADE3*, *HIS* y *lacZ*. Cepas diploides que coexpresan las fusiones traduccionales BD-NF-YC1 y AC-NIPK fueron capaces de crecer en medios selectivos que carecen de leucina, triptofano y adenina (TDO; *triple drop-out*), o leucina, triptófano, adenina e histidina (QDO; *quadruple drop-out*), mientras que la cepa control que expresa BD-NF-YC1 y porta el vector vacío pGADT7 no mostró crecimiento en estos medios (Figura 39A). Los medios sin triptofano y leucina seleccionan la presencia de los plásmidos que expresan los BD y AD de GAL4 o sus fusiones, mientras que los medio carentes de adenina e histina permiten seleccionar aquellas levaduras donde se produjo una interacción entre PvNF-YC1 y NIPK. La interacción positiva fue confirmada y cuantificada mediante el ensayo enzimático de actividad  $\beta$ -galactosidasa (Figura 39B). La cepa diploide NF-YC1/NIPK presentó una actividad  $\beta$ -galactosidasa 8 veces mayor que el control NF-YC1/vector vacío. La delección de los extremos N- y C-terminal de NF-YC1 que flanquean el dominio de tipo histona comprometen parcialmente la interacción con NIPK, sugiriendo que ambas regiones son requeridas para la formación del dímero (Figura 39 A, B y C).





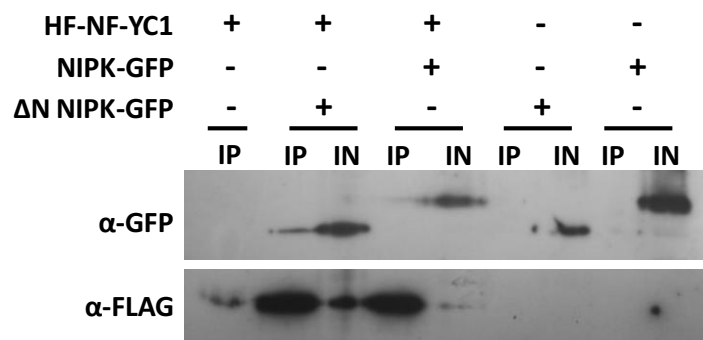
**Figura 39. NIPK interacciona con PvNF-YC1 en levadura.**

**A.** Interacción en levadura por el método de doble híbrido. Se muestra el crecimiento de levaduras transformadas con los vectores pGBKT7 y pGADT7 en medio sintético definido (SD) suplementado con aminoácidos, pero que carece de leucina y triptofano (SD/DDO), leucina, triptofano y adenina (SD/TDO) o leucina, triptofano, adenina e histidina (SD/QDO). El plásmido pGBKT7, que contiene el dominio de unión al DNA de GAL4 (BD), fue fusionado al ORF completo de NF-YC1 o versiones truncadas que carecen del extremo N-terminal (ΔN), C-terminal (ΔC) o ambos (NF-YC1 mínimo). Éste plásmido porta el gen *TRP1* que permite la selección por complementación de la auxotrofia por triptofano. El plásmido pGADT7, que contiene el dominio de activación de GAL4, fue fusionado a NIPK y seleccionado por la complementación de la auxotrofia por leucina. Como control positivo (+) y negativo (-) se utilizaron cepas diploides que coexpresan las proteínas p53 y AgT, o p53 y Lam, respectivamente. **B.** Cuantificación de la actividad β-Galactosidasa. La magnitud de la interacción entre las proteínas fue cuantificada mediante la actividad β-Galactosidasa utilizando ONPG como sustrato. Una unidad β-Galactosidasa fue definida como la cantidad de enzima que hidroliza 1 μmol of ONPG a o-nitrofenol y D-galactosa por minuto por célula (Unidad Miller). Las barras representan la media y SEM de tres experimentos independientes. Los asteriscos indican diferencias significativas respecto al control en base a un análisis de la varianza (ANOVA) seguido de un test de Tukey para comparaciones múltiples (\*, p<0,05; \*\*, p<0,01; \*\*\*, p<0,001; \*\*\*\*, p<0,0001). **C.** Esquema de las deleciones de NF-YC1 utilizadas en los experimentos de doble híbrido.

Con el objetivo de determinar si la interacción entre NIPK y NF-YC1 se produce *in planta*, se llevó a cabo un ensayo de co-inmunoprecipitación (CoIP) en hojas de *Nicotiana benthamiana*. Para esto se generaron vectores binarios que permiten la

expresión de las fusiones traduccionales His-FLAG-NF-YC1 (HF-NF-YC1) y NIPK-GFP o una versión truncada que contiene sólo el dominio quinasa ( $\Delta$ N NIPK-GFP) bajo el promotor p35S. Cepas de *Agrobacterium tumefaciens* portando los vectores correspondientes se utilizaron para expresar las proteínas de manera transiente en hojas de *N. benthamiana* mediante co-transformación por agroinfiltración. Tres días luego de la infiltración se colectó tejido y llevó a cabo la extracción de proteínas totales (*input*, IN). A partir de este extracto se realizó la inmunoprecipitación con anticuerpos anti-FLAG y la presencia de NIPK-GFP o su versión truncada en la fracción eluida (IP) fue evaluada con anticuerpos anti-GFP.

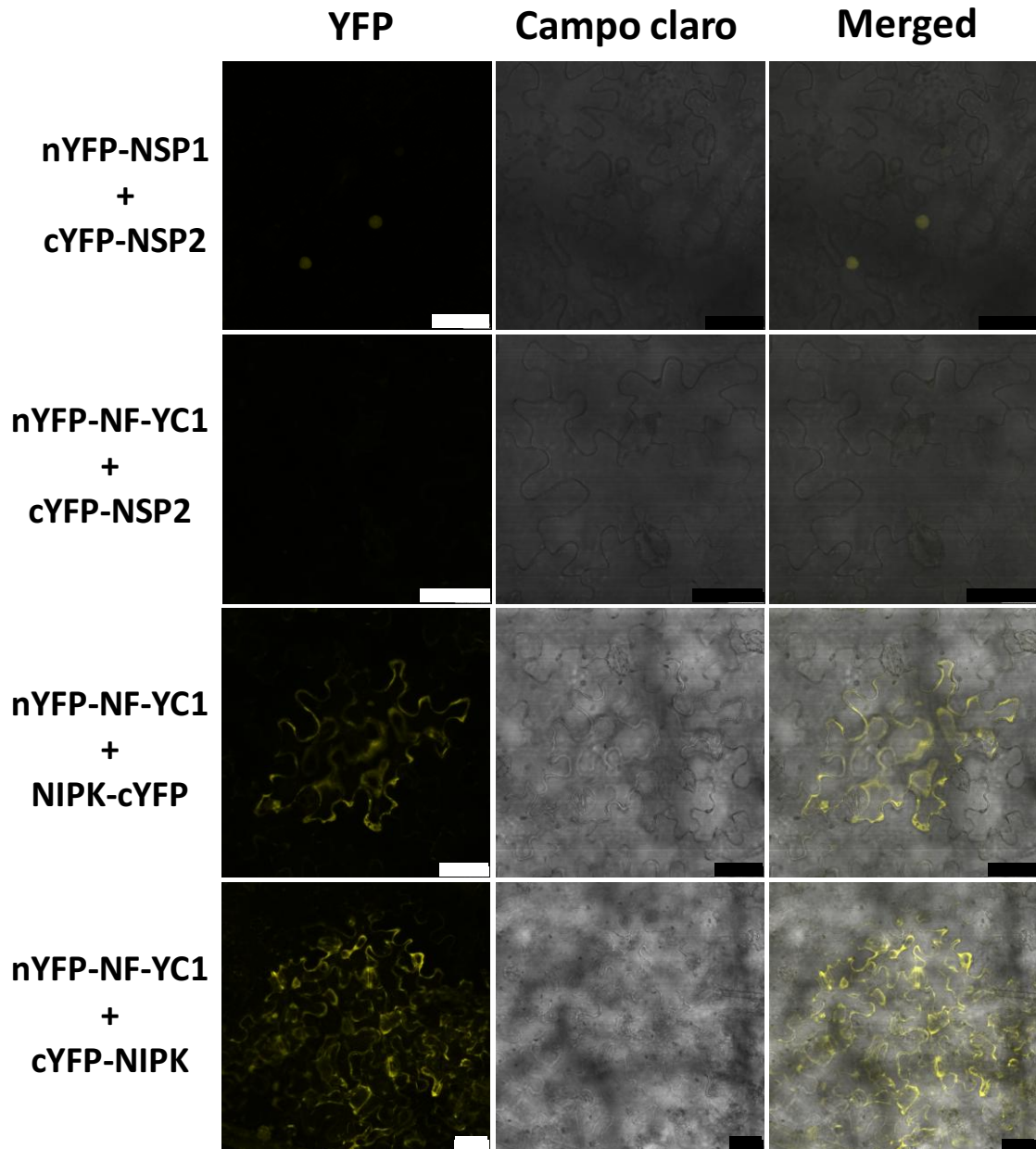
En las fracciones IP donde se coexpresó HF-NF-YC1 y NIPK-GFP o HF-NF-YC1 y  $\Delta$ N NIPK-GFP, se detectaron bandas de proteínas de 67 kD y 59 kD, correspondientes al peso molecular teórico estimado para NIPK-GFP y  $\Delta$ N NIPK-GFP, respectivamente, las cuales no fueron detectadas en las IP de los controles que no expresan HF-NF-YC1 (Figura 40). Estos resultados sugieren que NF-YC1 y NIPK interactúan *in planta*, y que el dominio quinasa de NIPK es suficiente para la formación del heterodímero.



**Figura 40. Interacción entre NIPK y NF-YC1 en un ensayo de co-inmunoprecipitación.** Hojas de *N. benthamiana* fueron agroinfiltradas con cepas de *A. tumefaciens* GV3101 portando las construcciones que se indican en la parte superior de la figura; His-FLAG-NF-YC1 (HF-NF-YC1), NIPK-GFP y una versión truncada de NIPK que carece del extremo amino terminal ( $\Delta$ N NIPK-GFP). Fracciones del extracto crudo o *input* (IN) y el eluido de la inmunoprecipitación con  $\alpha$ -FLAG (IP) fueron corridas en una electroforesis en gel de poliacrilamida desnaturalizante (SDS-PAGE 12 % (p/v)), y las proteínas de fusión fueron detectadas mediante un Western blot utilizando anticuerpos anti GFP ( $\alpha$ -GFP) o anti FLAG ( $\alpha$ -FLAG) conjugados a la enzima peroxidasa.

### Capítulo III

Con el objetivo de determinar la localización subcelular donde se produce la interacción entre NIPK y NF-YC1 se realizó un experimento de complementación bimolecular de fluorescencia (BiFC; *bimolecular fluorescence complementation*) en células epidérmicas de hojas de *N. benthamiana*. El principio de esta técnica se basa en fusionar las proteínas de interés a las regiones C- y N-terminal de la proteína fluorescente amarilla (YFP; *yellow fluorescent protein*). De producirse la interacción física de las dos proteínas, se restituye la fluorescencia de la YFP, la cual puede visualizarse mediante microscopía confocal de fluorescencia. Para realizar el ensayo de BiFC se coexpresó NF-YC1 fusionado a la mitad N-terminal de YFP (nYFP-NF-YC1) junto con NIPK fusionada a la mitad C-terminal de la YFP en posición tanto C- como N-terminal (NIPK-cYFP y cYFP-NIPK, respectivamente). Dos días luego de la agroinfiltración, pudo verse la señal correspondiente a la longitud de onda de emisión de la YFP en el citoplasma de las células epidérmicas (Figura 41). Dicha señal no fue detectada en el control donde se coexpresaron nYFP-NF-YC1 y cYFP-NSP2. En el caso del control positivo, donde se expresaron los factores de transcripción NSP1 y NSP2, los cuales interaccionan fuertemente en hojas de *N. benthamiana* (Hirsch et al, 2009), pudo visualizarse una intensa señal en el núcleo (Figura 41). Para determinar la especificidad de la interacción entre NF-YC1 y NIPK, se evaluó la interacción física entre NIPK y el ortólogo de NF-YC1 en *M. truncatula* (MtNF-YC2), así como también una versión mutada (MtNF-YC2\*\*) que no es capaz de interaccionar con las subunidades A y B por BiFC (Baudin et al., 2015). Los resultados mostraron que NIPK no interacciona con MtNF-YC2 o MtNF-YC2\*\*, sugiriendo que la interacción NIPK/NF-YC1 es altamente específica (Figura Suplementaria S3). En resumen, los resultados sugieren que la interacción entre NF-YC1 y NIPK es específica y se produce en el citoplasma, lo cual es compatible con la localización citoplasmática y nuclear reportada para NF-YC1.



**Figura 41. Ensayo de complementación bimolecular de fluorescencia.**

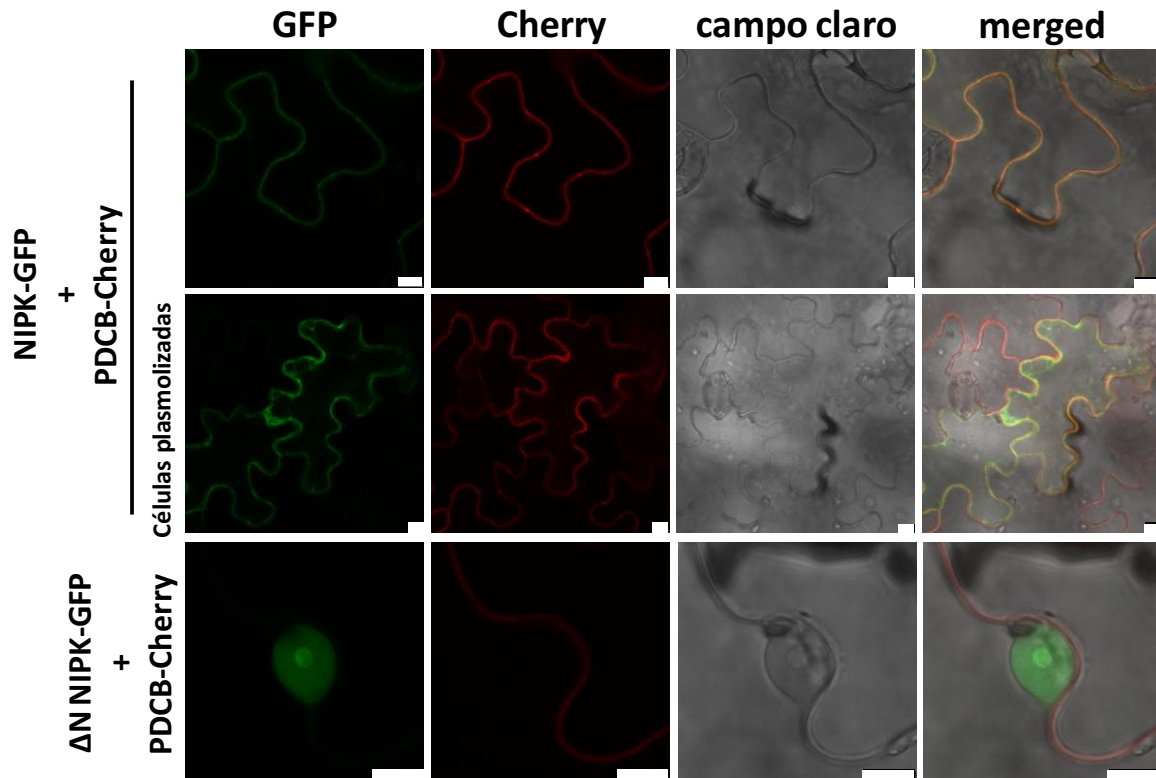
Hojas de *N. benthamiana* fueron agroinfiltradas con nYFP-NF-YC1 junto con cYFP-NIPK, NIPK-cYFP o cYFP-NSP2 (control negativo). Como control positivo fueron coexpresados nYFP-NSP1 y cYFP-NSP2. En la figura se muestran imágenes representativas de microscopía confocal láser correspondientes a la señal de la YFP en los paneles izquierdos, al campo claro en los paneles del medio y la superposición de ambas imágenes en los paneles derechos (merged). Barras de la escala: 50  $\mu$ m.

### 3. Localización subcelular de NIPK

Teniendo en cuenta que NIPK interacciona con NF-YC1 en el citoplasma, nos propusimos determinar la localización subcelular de NIPK con el fin de establecer si es

### Capítulo III

compatible con los resultados de la interacción. Para esto, las construcciones NIPK-GFP o  $\Delta$ N NIPK-GFP fueron introducidas en células epidérmicas de *N. benthamiana* mediante agroinfiltración, junto con la proteína PCDB (*plasmodesmata callose-binding protein 1*) fusionada a la proteína fluorescente roja *Cherry* como marcador de membrana plasmática (Simpson et al., 2009). Las imágenes de microscopía confocal mostraron que NIPK-GFP colocaliza con PDCB-Cherry (Figura 42, paneles superiores), sugiriendo que NIPK se localiza en la membrana plasmática. Sin embargo, cuando se expresaron ambas proteínas y las células fueron plasmolizadas, pudo observarse que NIPK-GFP no se localiza solamente en la membrana plasmática, sino también en el citoplasma, fracción donde no se observa la señal correspondiente a PDCB-Cherry (Figura 42, paneles del medio). La plasmólisis, al reducir el volumen de la vacuola, aumenta el espacio citoplasmático libre, permitiendo distinguir entre la fracción de NIPK-GFP que se encuentra en la membrana y en el citoplasma. Para dilucidar si el dominio transmembrana de NIPK predicho bioinformáticamente es el elemento funcional que determina su localización subcelular, expresamos la construcción  $\Delta$ N NIPK-GFP, la cual carece de dicho dominio. Las imágenes muestran que  $\Delta$ N NIPK-GFP se localiza en núcleo y nucléolo, sugiriendo que el dominio transmembrana es funcional (Figura 42, paneles inferiores).

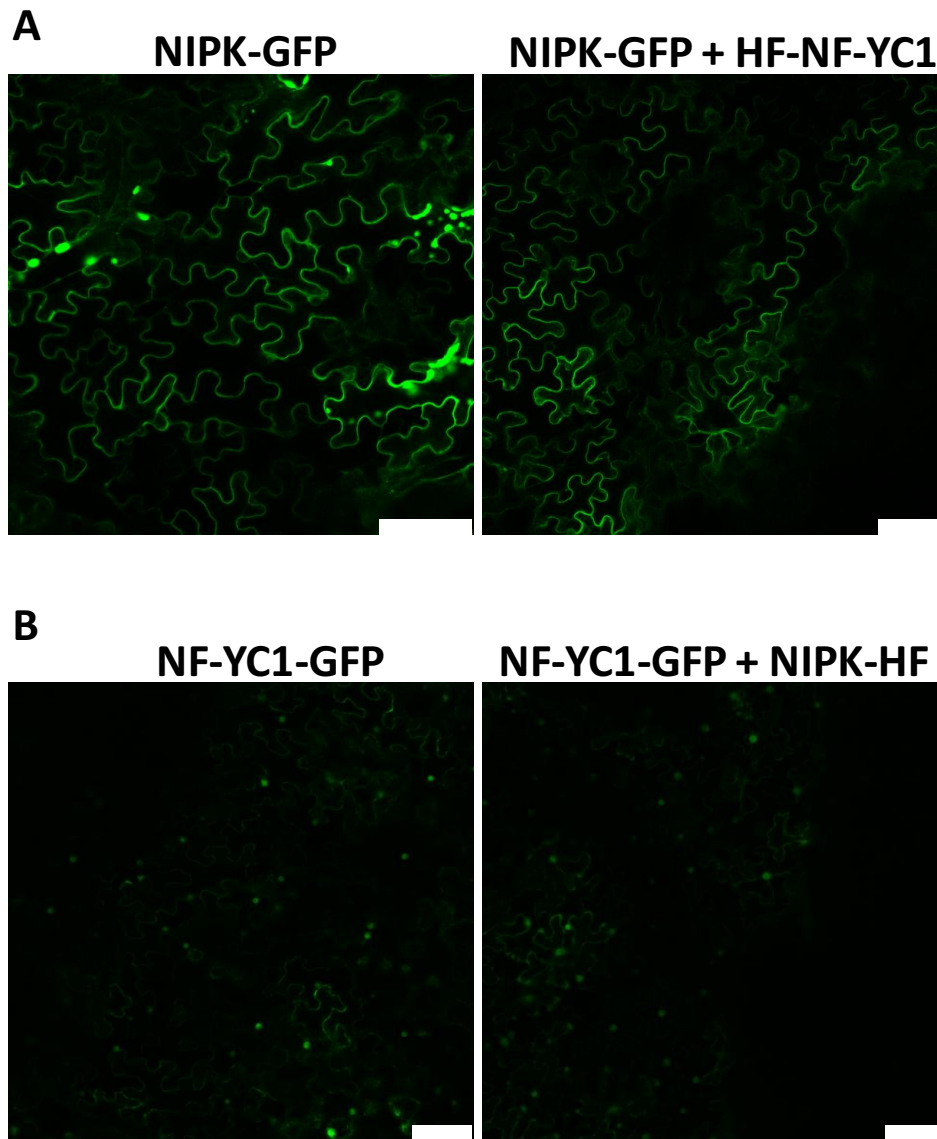


**Figura 42. Localización subcelular de NIPK.**

Se muestran imágenes de microscopía confocal de células epidérmicas de hojas de *N. benthamiana*, correspondientes a la longitud de onda de emisión de GFP y Cherry, el campo claro y la superposición de los tres canales (merged). Se realizó la co-agroinfiltración con las construcciones PDCB-Cherry como marcador de membrana plasmática y NIPK-GFP o  $\Delta N$  NIPK-GFP. Los paneles de la segunda fila muestran células epidérmicas plasmolizadas. Barras de escala: 10  $\mu\text{m}$ .

En base a los resultados de localización subcelular obtenidos y teniendo en cuenta que la interacción NIPK/NF-YC1 tiene lugar en el citoplasma nos planteamos dos preguntas. Primero si NIPK podría estar regulando el tránsito de NF-YC1 hacia el núcleo reteniéndolo en el citoplasma. Segundo, si como resultado de la interacción, NIPK es trasladado al núcleo y luego dissociado de NF-YC1. Para intentar responder estas preguntas se expresaron en *N. benthamiana* las proteínas de fusión NF-YC1-GFP, NIPK-FH, NIPK-GFP y HF-NF-YC1 en distintas combinaciones. La coexpresión de NIKP-GFP y HF-NF-YC1 muestra que la localización subcelular de NIPK no se ve afectada por la sobreexpresión de NF-YC1 (Figura 43A). De manera consistente con lo

reportado por Zanetti et al., 2010, la fusión traduccional GFP-NF-YC1 muestra una localización nuclear y citoplasmática que no es afectada por NIPK-HF (Figura 43B). Estos resultados sugieren que la interacción entre NIPK y NF-YC1 es un proceso dinámico donde la disociación del complejo ocurriría en el citoplasma previo a la importación nuclear de NF-YC1.



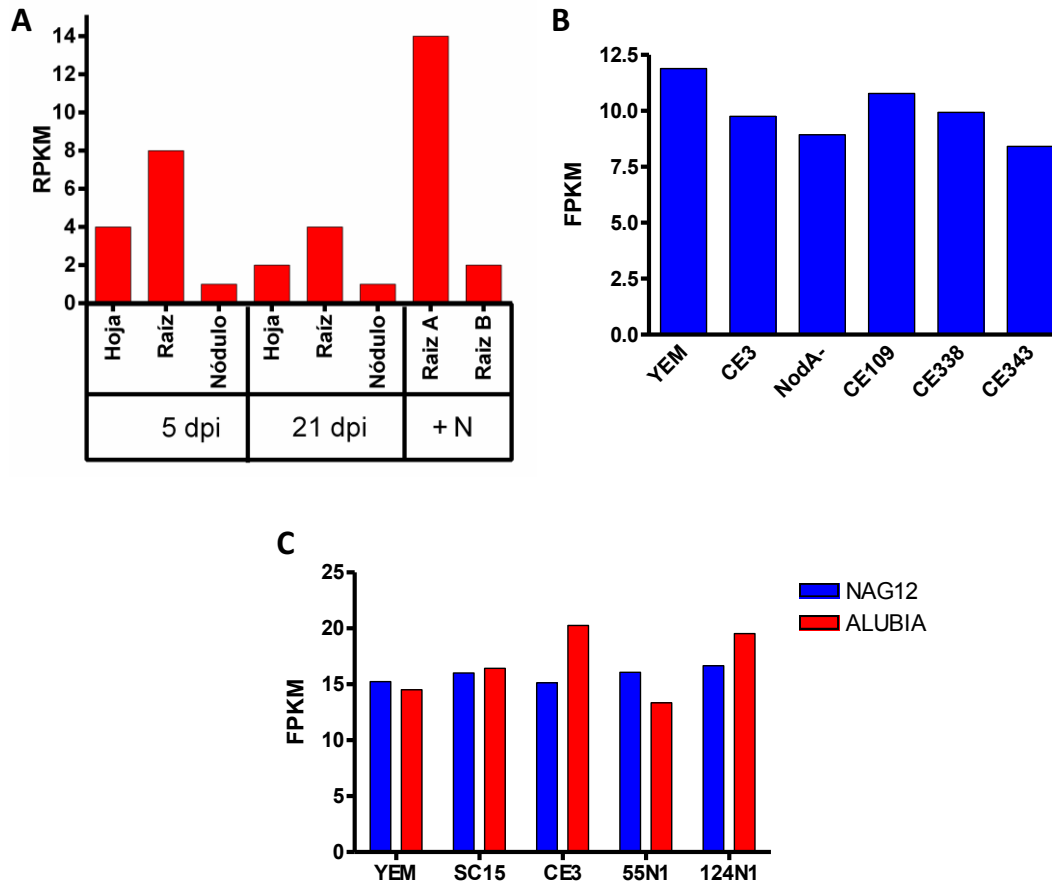
**Figura 43. Localización subcelular de NIPK y NF-YC1 en el contexto de la interacción.** Las proteínas de fusión NIPK-GFP, NIPK-HF, NF-YC1-GFP y HF-NF-YC1 fueron expresadas en *N. benthamiana* en las combinaciones indicadas. Se muestran imágenes de microscopía confocal correspondientes a la longitud de onda de emisión de GFP. Barras de escala: 100  $\mu$ m.

#### 4. Patrón de expresión de *NIPK*

Teniendo en cuenta que nuestra hipótesis plantea la posibilidad de que *NIPK* cumpla algún rol en el establecimiento de la asociación simbiótica con rizobios, decidimos analizar su patrón de expresión en distintos órganos de la planta tanto en condiciones de inoculación con rizobios como condiciones control (Figura 44). En primer lugar se observó el patrón de expresión de *NIPK* en distintos órganos de *P. vulgaris* haciendo uso de los datos de expresión depositados en el atlas de expresión génica de poroto (<http://plantgrn.noble.org/PvGEA/>) (O'Rourke et al., 2014). Los resultados muestran que *NIPK* se expresa en hojas, raíces (a las cuales se le quitaron los nódulos) y nódulos de plantas de 5 y 21 días *post* inoculación (dpi), aunque los valores de expresión máximos se producen en la raíz. En raíces fertilizadas con nitrógeno (+N), la expresión de *NIPK* es mayor en el estadio del desarrollo equivalente al de plantas de 21 dpi (Raíz A), que en plantas en el estadio en que se despliega la segunda hoja trifoliada (Raíz B) (Figura 44A). Para determinar si la expresión de *NIPK* es dependiente de alguna de las moléculas de señalización del rizobio (o del rizobio) a tiempos tempranos de la interacción simbiótica (24 hpi), se hizo uso de los datos generados en el laboratorio (Dalla Via et al., 2015). En dicho trabajo se caracterizó el transcriptoma de raíces de *P. vulgaris* NAG12 24 hpi con *R. etli* CE3 *wild type*, o cepas mutantes deficientes en la síntesis de Factor Nod (NodA-), de exopolisacáridos (CE338), de lipopolisacáridos (CE109), o de lipo y exopolisacáridos (CE343). Las cepas NodA-, CE109 y CE343 son incapaces de generar nódulos. Los datos de expresión muestran que *NIPK* no cambia sus niveles estacionarios de mRNA en respuesta a *R. etli* CE3 respecto a YEM, o a alguna de las cepas mutantes (Figura 44B). Teniendo en cuenta que el patrón de expresión de *NF-YC1* es dependiente de la eficiencia simbiótica de la cepa de *R. etli* utilizada, se determinó el patrón de expresión de *NIPK* haciendo uso de los datos



transcriptómicos generados en esta tesis (Capítulo 1). Los resultados muestran que tanto en el cultivar NAG12 como en Alubia, *NIPK* no muestra cambios significativos en sus niveles estacionarios de mRNA en raíces inoculadas con SC15, CE3, 55N1 o 124N1, respecto a plantas control inoculadas con YEM (Figura 44C).



**Figura 44. Patrón de expresión de *NIPK* en *P. vulgaris*.**

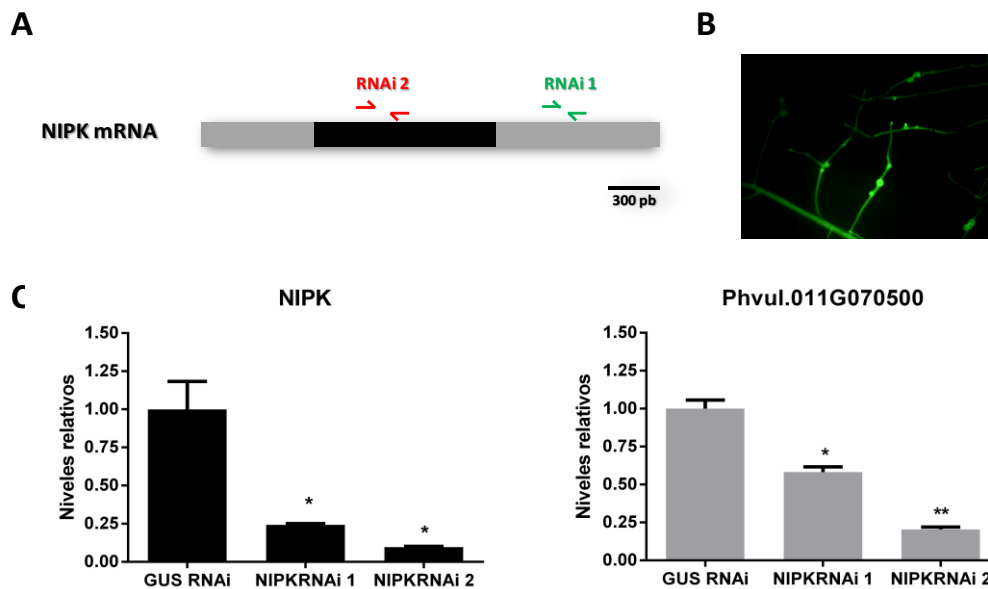
**A.** Valores de expresión de *NIPK* (expresados en RPKM) en hoja, raíz (a la cual se le quitaron los nódulos) y nódulos de plantas de *P. vulgaris* de origen Andino de 5 y 21 días post inoculación (dpi) con rizobio. También se muestran los valores de expresión en raíces de plantas fertilizadas con nitrógeno (+N) en un estadio del desarrollo equivalente al de plantas de 21 dpi (Raíz A) y plantas en que se desplegó la segunda hoja trifoliada (Raíz B). Los datos fueron tomados del atlas de expresión de *P. vulgaris* (O'Rourke et al., 2014). **B.** Valores de expresión de *NIPK* (en FPKM; *Fragments Per Kilobase of transcript per Million mapped reads*) en raíces de *P. vulgaris* NAG12 24 horas *post* inoculación (hpi) con YEM (control), *R. etli* CE3 *wild type* y cepas mutantes deficientes en la síntesis de Factor Nod (NodA-), de exopolisacáridos (CE338), de lipopolisacáridos (CE109), o de lipo y exopolisacáridos (CE343). Datos extraídos de Dalla Via et al., 2015. **C.** Valores de expresión de *NIPK* (en FPKM) en raíces de *P. vulgaris* NAG12 (barras azules) y Alubia (barras rojas) 24 hpi con YEM (control) y *R. etli* SC15, CE3, 55N1 o 124N1. Los datos fueron generados como parte de esta tesis (Capítulo 1).

En resumen, el análisis realizado sugiere que *NIPK* se expresa en raíces y nódulos de *P. vulgaris*, aunque sus niveles de expresión no muestran cambios significativos en los niveles estacionarios de mena en ninguna de las condiciones simbióticas analizadas.

### 5. Silenciamiento postranscripcional de *NIPK*

Con el objetivo de caracterizar funcionalmente a *NIPK* en el contexto de la nodulación, se decidió llevar a cabo el silenciamiento postranscripcional mediado por RNA de interferencia (RNAi) de *NIPK*, utilizando el sistema de generación de raíces transgénicas de poroto mediado por *Agrobacterium rhizogenes* (Estrada-navarrete et al., 2006). Por este sistema se obtienen plantas con la parte aérea salvaje y las raíces transgénicas. Se optó por este método de transformación dado que no existe un sistema de transformación estable eficiente para *P. vulgaris*, como tampoco colecciones de mutantes disponibles. Para realizar el silenciamiento postranscripcional se diseñaron dos RNAi a partir de distintas regiones del transcripto de *NIPK*, uno de 100 pb en base a la secuencia 3'-UTR (*NIPK* RNAi 1), y el otro de 311 pb en base a la secuencia del dominio quinasa (*NIPK* RNAi 2) (Figura 45A). La secuencia de los RNAi diseñados fue sometida a un BLAST contra la base de datos genómica de poroto, y sólo mostraron similitud con la secuencia a partir de la cual fueron diseñados, sugiriendo que son secuencias específicas de gen. Estas secuencias fueron clonadas como repeticiones invertidas en el vector pK7GWIWG2D(II),0, bajo el control del promotor p35S. El vector elegido expresa GFP bajo el promotor *rolD*, lo cual permite seleccionar las raíces transgénicas para los análisis fenotípicos mediante la observación bajo luz UV (Figura 45B). El análisis por transcripción reversa seguida de PCR cuantitativo (RT-qPCR) de los niveles de *NIPK* a partir de raíces transformadas con *NIPK* RNAi 1 y 2, mostró que en ambos casos se genera una reducción en los niveles de *NIPK* respecto a raíces *GUS*

RNAi (Figura 45C). Para raíces *NIPK* RNAi 1 el silenciamiento fue de aproximadamente 75 %, mientras que para *NIPK* RNAi 2 fue del 90 %. Para determinar si el silenciamiento es específico de gen, se cuantificaron por RT-qPCR los niveles del gen *Phvul.011G070500*, que es el mejor homólogo a *NIPK* con un tamaño proteico y dominios funcionales similares. Este análisis fue realizado mediante un BLASTN contra la base de datos genómica de poroto (Phytozome). Los resultados muestran que en raíces *NIPK* RNAi 2 se produce una fuerte reducción en los niveles de *Phvul.011G070500* respecto al control, mientras que en raíces *NIPK* RNAi 1 el efecto fue moderado, produciéndose un silenciamiento de sólo el 40 % (Figura 45C). Estos resultados eran esperables teniendo en cuenta la región a partir de la cual fueron diseñados los RNAi y sus tamaños.

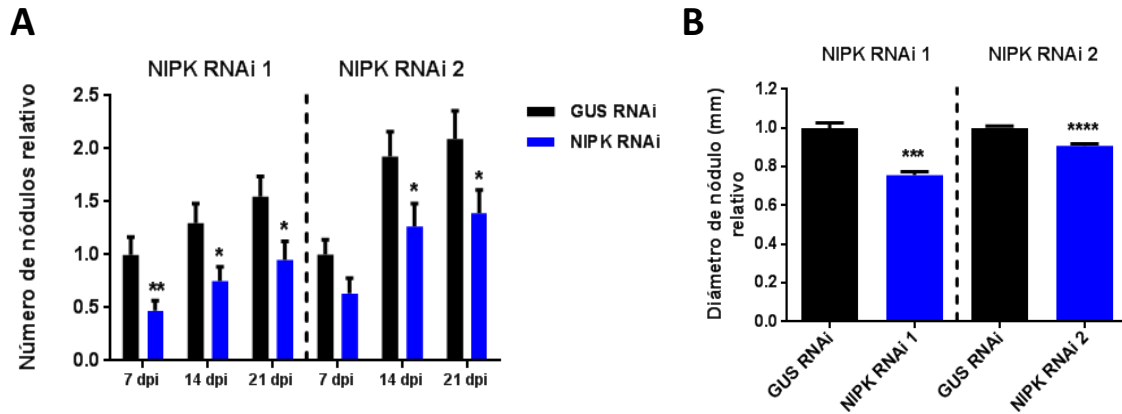


**Figura 45. Silenciamiento postranscripcional de *NIPK*.**

**A.** Se muestra la posición de los *primers* utilizados para la obtención de *NIPK* RNAi 1 (en verde) y *NIPK* RNAi 2 (en rojo) sobre el mRNA maduro de *NIPK*. **B.** Raíces transgénicas de poroto *NIPK* RNAi 1 observadas bajo la lupa de fluorescencia. **C.** Niveles de expresión relativa de *NIPK* (barras negras) y *Phvul.011G070500* (barras grises) en plantas transformadas con *NIPK* RNAi 1, *NIPK* RNAi 2 y *GUS* RNAi como control. Los resultados son la media de dos repeticiones técnicas y las barras de error representan el SEM. Los asteriscos indican diferencias significativas respecto al control en un t-test con  $p < 0,05$  (\*,  $p < 0,05$ ; \*\*,  $p < 0,01$ ). Los datos fueron normalizados respecto a los valores de expresión de *EF-1 $\alpha$*  y expresados en relación a *GUS* RNAi.

## 6. *NIPK* es requerido para la organogénesis del nódulo

Con el objetivo de determinar si *NIPK* es requerido para el establecimiento de la asociación simbiótica fijadora de nitrógeno, plantas compuestas transformadas con las construcciones *NIPK* RNAi 1 ó 2 o con la construcción *GUS* RNAi fueron inoculadas con *R. etli* SC15, y el número de nódulos por raíz fue cuantificado a 7, 14 y 21 días post infección (dpi). Las cinéticas de nodulación mostraron que las raíces con niveles disminuidos de *NIPK* forman una menor cantidad de nódulos respecto a las raíces control a todos los tiempos observados (Figura 46A). La expresión de las construcciones *NIPK* RNAi 1 y 2 produjo una reducción de aproximadamente el 60% en el número de nódulos respecto al control *GUS* RNAi a lo largo de la cinética (Figura 46A). A 21 dpi el tamaño de los nódulos formados fue cuantificado con el fin de determinar si el silenciamiento de *NIPK* afecta su desarrollo. Los nódulos formados a este tiempo mostraron una reducción significativa de su tamaño respecto al control (Figura 46B). La disminución del diámetro del nódulo en plantas *NIPK* RNAi fue de 24,39 % y 9,24 % para los RNAi 1 y 2, respectivamente (Figura 46B). Estos resultados sugieren que *NIPK* es requerido para la organogénesis y desarrollo del nódulo.

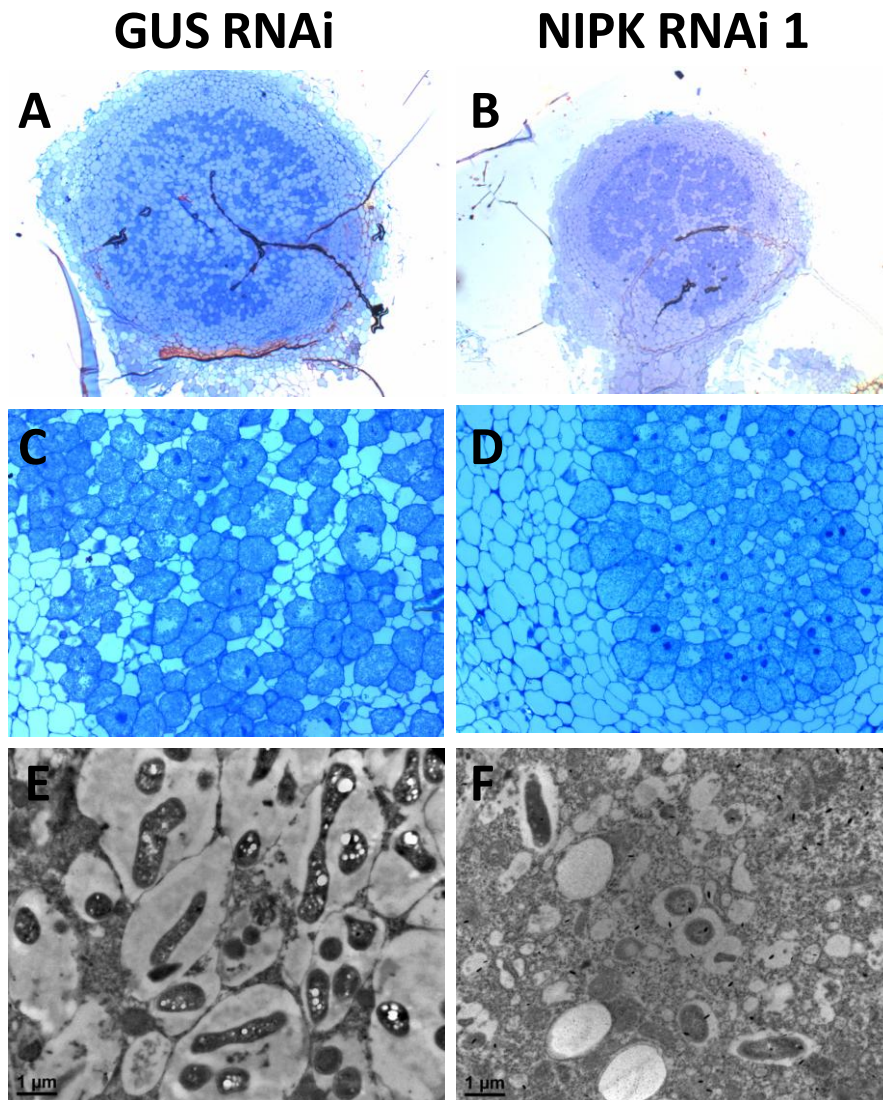


**Figura 46. El silenciamiento de NIPK afecta el número y tamaño de los nódulos.**

Plantas compuestas de poroto transformadas con *NIPK* RNAi 1, *NIPK* RNAi 2 o *GUS* RNAi fueron crecidas durante 7 días en ausencia de nitratos y luego fueron inoculadas con un cultivo de *R. etli* SC15. El número de nódulos por raíz fue cuantificado a 7, 14 y 21 dpi y luego normalizado por el valor de *GUS* RNAi a 7 dpi ( $n > 10$  para cada condición) (A). A 21 dpi los nódulos fueron fotografiados y se les midió el diámetro mediante el software Photoshop ( $n > 100$  para cada condición). Los valores fueron normalizados por el control (B). En (A) y (B) las barras de error representan el SEM y los asteriscos indican diferencias significativas respecto al control en un t-test con  $p < 0,05$  (\*,  $p < 0,05$ ; \*\*,  $p < 0,01$ ; \*\*\*,  $p < 0,001$ ; \*\*\*\*,  $p < 0,0001$ ). Se muestran experimentos representativos de tres experimentos independientes.

A pesar que los nódulos *NIPK* RNAi mostraron un menor tamaño, formaron lenticelas y mostraron un color rosado semejante al de las plantas control, indicando que eran capaces de realizar un correcto intercambio gaseoso y de llevar a cabo la síntesis de leghemoglobina, sugiriendo que los nódulos estaban ocupados por el simbiote. Para determinar si esta hipótesis es correcta, se realizaron cortes histológicos de nódulos de 14 dpi. Los cortes semifinos muestran que el área infectada por *R. etli* en los nódulos *NIPK* RNAi 1 es comparable a la de nódulos control (Figura 47A y B). El análisis de las células infectadas muestra que tampoco existen diferencias obvias en la formación de simbiosomas entre nódulos *NIPK* RNAi 1 y control (Figura 47C y D). Para determinar si existe algún defecto en la ultraestructura o grado de diferenciación de los simbiosomas, se realizaron cortes ultrafinos para estudiarlos por microscopía de transmisión electrónica (TEM). Las micrografías muestran que tanto las células

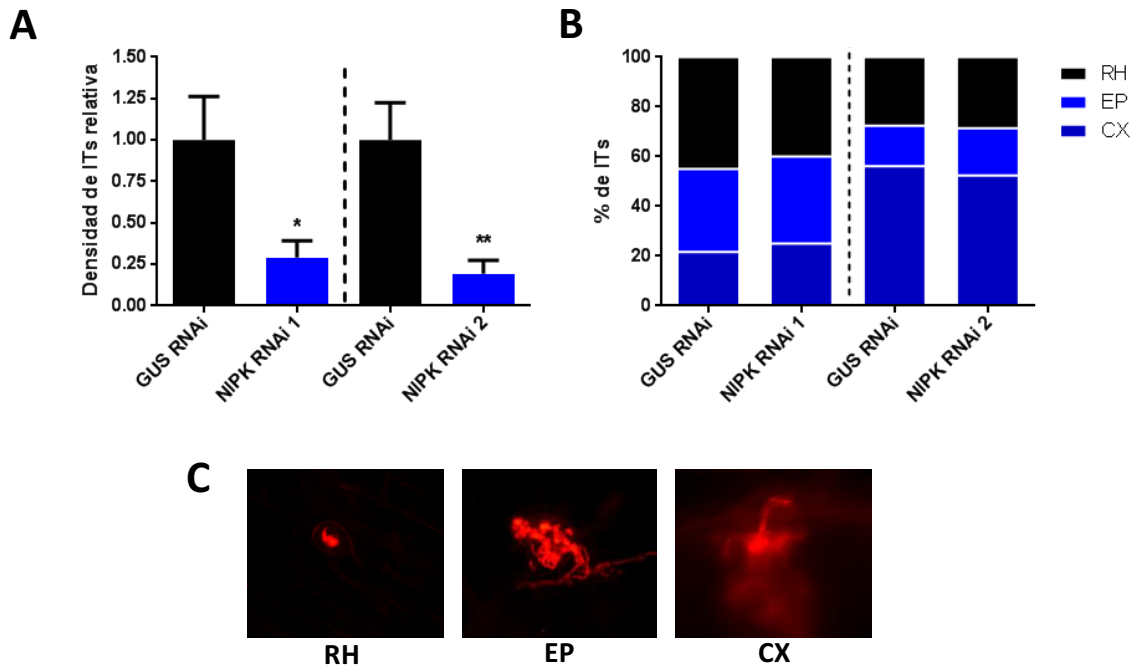
infectadas de nódulos provenientes de raíces *NIPK* RNAi 1 y *GUS* RNAi poseen simbiosomas diferenciados, indicando que la internalización y diferenciación de la bacteria no es afectada drásticamente por el silenciamiento de *NIPK* (Figura 47E y F).



**Figura 47. Estructura del nódulo e infección en plantas con raíces *NIPK* RNAi y control.** Nódulos *GUS* RNAi o *NIPK* RNAi 1 de 14 días inoculados con la cepa de *R. etli* CFNX5 que expresa dsRED fueron embebidos en parafina y teñidos con azul de toluidina. Posteriormente se realizaron cortes semifinos para microscopía óptica (A-D) o cortes ultrafinos para microscopía de transmisión electrónica (E-F).

### **7. *NIPK* es requerido para la formación de ITs**

Teniendo en cuenta que las raíces con niveles disminuidos de *NIPK* mostraron una menor cantidad de nódulos, nos preguntamos si la causa de este fenotipo sólo se debe a un defecto en la organogénesis y desarrollo del nódulo, o si también podría estar afectado el proceso de infección. Para determinar esto, plantas compuestas fueron inoculadas con una cepa de *R. etli* CFN5X que expresa de manera constitutiva la proteína fluorescente roja dsRED. De esta manera es posible seguir los eventos de infección mediante microscopía de fluorescencia. La cuantificación de la cantidad de hilos de infección por cm de raíz muestra que las raíces *NIPK* RNAi 1 y 2 forman una menor cantidad de hilos de infección que las raíces control *GUS* RNAi (Figura 48A). Los hilos de infección formados fueron clasificados en base a su grado de progresión como: hilos que se encuentran aun en el pelo radical, hilos que alcanzaron las células epidérmicas e hilos que alcanzaron el córtex de la raíz (Figura 48B y C). Los resultados muestran que no hay diferencias significativas en el grado de progresión de los hilos de infección entre plantas control y *NIPK* RNAi 1 ó 2 (Figura 48B). En resumen, estos resultados sugieren que *NIPK* es requerido para la iniciación de los hilos de infección, pero una vez formados, éstos son capaces de progresar normalmente hacia las células corticales donde se está desarrollando el nódulo.



**Figura 48. *NIPK* RNAi afecta la formación, pero no la progresión de hilos de infección.**

Plantas compuestas *GUS* RNAi, *NIPK* RNAi 1 o *NIPK* RNAi 2 fueron inoculadas con un cultivo de *R. etli* CFN5X que expresa de manera constitutiva la proteína fluorescente roja dsRED. Cuatro días luego de la inoculación se cuantificó la cantidad de hilos de infección formados por centímetro (cm) de raíz a partir de imágenes obtenidas mediante microscopía de fluorescencia. Los datos fueron normalizados respecto al control. Las barras de error representan el SEM y los asteriscos indican diferencias significativas respecto al control en un t-test con  $p < 0,05$  (\*,  $p < 0,05$ ; \*\*,  $p < 0,01$ ). Se muestran experimentos representativos de tres experimentos independientes. (A). Los hilos de infección formados fueron clasificados en tres grupos: hilos que se encuentran en el pelo radical (RH), hilos que alcanzaron las células epidérmicas (EP) y aquellos que alcanzaron el córtex (CX) (B). En (C) se muestran imágenes representativas de hilos de infección de cada grupo.

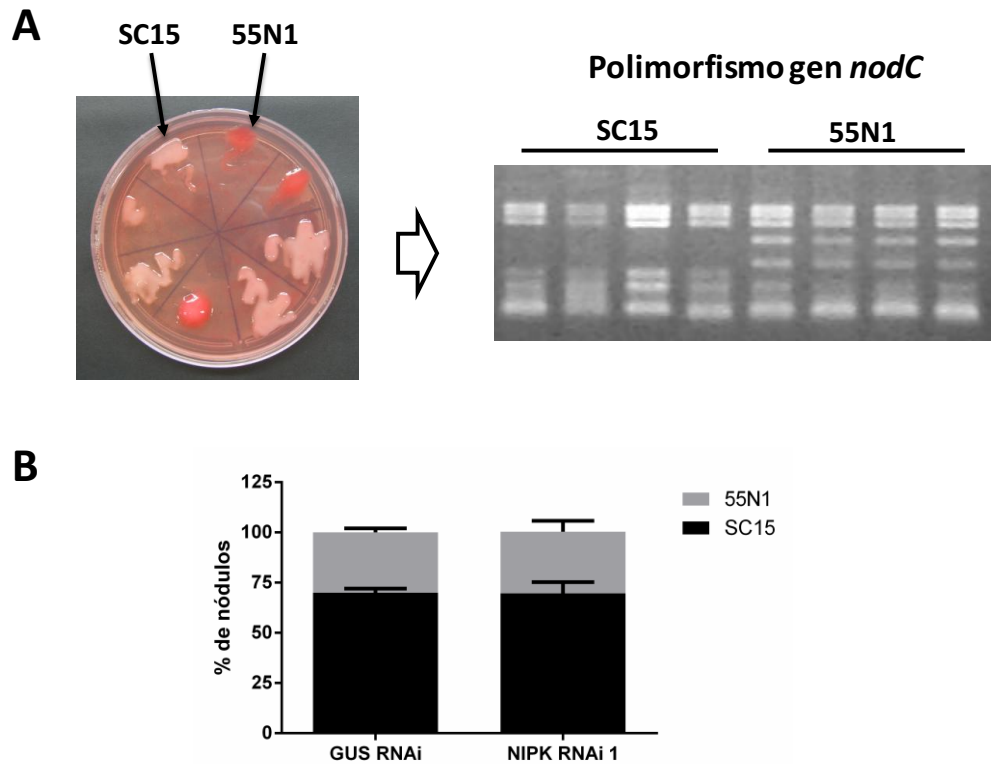
## 8. La especificidad de cepa no se ve afectada en raíces con niveles disminuidos de *NIPK*

Teniendo en cuenta que *NIPK* fue identificado como una proteína que interacciona con PvNF-YC1, el cual participa de los mecanismos moleculares que determinan la preferencia de cepa, y que el análisis fenotípico sugiere un rol en la formación del hilo de infección y en la organogénesis y posterior desarrollo de los nódulos, nos preguntamos si el fenotipo podría afectar la preferencia de las variedades Mesoamericanas por las cepas de *R. etli* tipo *nodC-α*. Para esto llevamos a cabo



### Capítulo III

experimentos de competencia en donde las raíces *GUS* RNAi y *NIPK* RNAi 1 del cultivar NAG12 fueron co-inoculadas con una mezcla equicelular de la cepa de *R. etli* tipo *nodC*- $\alpha$  SC15 y la cepa *nodC*- $\delta$  55N1. A los 21 dpi se colectaron los nódulos y se aislaron las bacterias de los mismos mediante su crecimiento en medio sólido. Luego, las mismas fueron identificadas como SC15 ó 55N1 en base a la coloración desarrollada por sus exopolisacáridos. La cepa SC15 toma un color rosa pálido, mientras que la cepa 55N1 toma un color rojo rubí. Para corroborar esta correlación, en cada experimento se realizó una caracterización molecular del polimorfismo del gen *nodC* mediante un análisis de ARFLP (Figura 49A). Los resultados muestran que los nódulos formados en raíces *NIPK* RNAi 1 son ocupados principalmente por la cepa más eficiente SC15, sin mostrar diferencias significativas con el control (Figura 49B). Estos resultados sugieren que *NIPK* no sería necesario para que opere la vía de transducción de señales que permite a los cultivares Mesoamericanos discriminar entre cepas que portan los alelos de *nodC* tipo - $\alpha$  o - $\delta$ .



**Figura 49. El silenciamiento de NIPK no afecta la preferencia del cultivar NAG12 por las cepas de *R. etli nodC-α* SC15.**

**A.** Se muestra una placa de YEM-agar rojo congo donde se plaquearon cepas SC15 y 55N1 aisladas de nódulos de 21 dpi. La cepa SC15 muestra un color rosa pálido y la cepa 55N1 un color rojo rubí. A la derecha se muestra un gel de agarosa con el resultado de la restricción con *HinfI* del producto de PCR del gen *nodC*. Las cepas SC15 y 55N1 muestran un patrón *nodC* tipo  $\alpha$  (*nodC-α*) y tipo  $\delta$  (*nodC-δ*), respectivamente. **B.** En el gráfico de barras se muestra el porcentaje de nódulos de raíces *GUS* RNAi y *NIPK* RNAi 1 que fueron ocupados por SC15 (barras negras) y por 55N1 (barras grises). Se muestra el promedio de dos experimentos independientes y el SEM (se utilizó un  $n > 70$  para cada condición).

## DISCUSIÓN

Mediante ensayos realizadas en levadura e *in planta*, hemos determinado que el factor de transcripción PvNF-YC1 interacciona físicamente con la pseudoquinasa NIPK, presumiblemente en el citoplasma de las células vegetales (Figura 47). Los análisis funcionales realizados mediante genética reversa sugieren que NIPK es un regulador positivo de la infección bacteriana y la organogénesis del nódulo, ya que plantas con

### Capítulo III

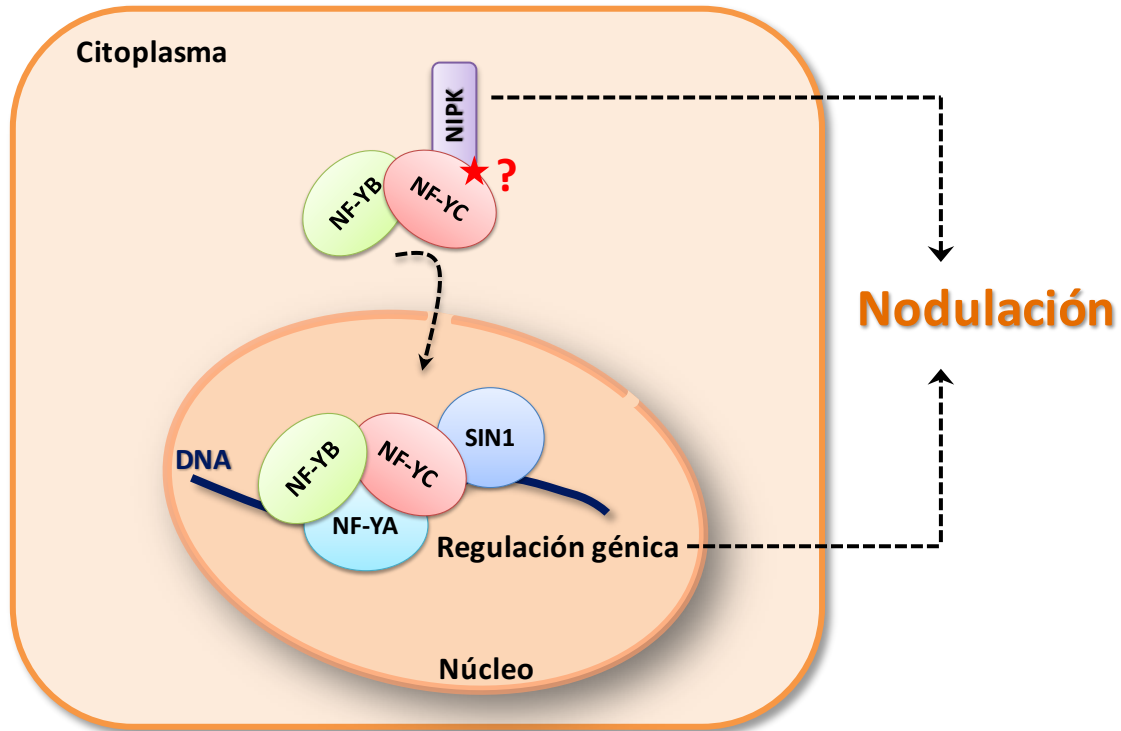
raíces transgénicas *NIPK* RNAi muestran una reducción en la densidad de hilos de infección, así como también, una disminución en el número y tamaño de los nódulos formados respecto al control. Este fenotipo es reminiscente del de plantas silenciadas en *NF-YC1*, las cuales también presentan defectos en la infección y organogénesis del nódulo que no son dependientes del tipo de cepa empleada en la inoculación, ya que plantas con raíces *NF-YC1* RNAi mostraron un número de nódulos reducido cuando fueron inoculadas con SC15 ó 55N1 (Zanetti et al., 2010). La similitud de los fenotipos y la interacción física de estas dos proteínas, sugieren que uno de los posibles roles de *NIPK* podría estar vinculado a la regulación positiva de la actividad de PvNF-YC1. Sin embargo, la disminución de los niveles de *NIPK* no afecta la nodulación preferencial por parte de la cepa tipo *nodC- $\alpha$*  SC15 respecto a la cepa menos eficiente *nodC- $\delta$*  55N1. Es decir que *NIPK* podría participar de los mecanismos moleculares generales que participan durante la infección rizobiana y el desarrollo de los nódulos, pero no de aquellos que son activados específicamente luego del reconocimiento de las cepas tipo *nodC- $\alpha$* , y que permiten a los cultivares Mesoamericanos nodular preferencialmente con estas cepas.

El análisis de la secuencia del dominio quinasa de *NIPK* sugiere que carecería de actividad fosfotransferasa, ya que no presenta dos de los tres aminoácidos reportados como fundamentales para la actividad catalítica de esta familia de proteínas, así como también otros aminoácidos fuertemente conservados (Hanks, 2003; Boudeau et al., 2006). Sin embargo, este criterio no es absoluto, dado que el gran número y diversidad de estas “pseudocinasas” -que en *A. thaliana* representan el 13% de las quinasas totales (Castells and Casacuberta, 2007)- hace difícil realizar predicciones generales respecto de su actividad (Boudeau et al., 2006). De hecho, proteínas como WNK1 (*with no lysine* (K)) y PRPK (*p53-related protein kinase*) que originalmente fueron clasificadas

### Capítulo III

como pseudoquinasas en base al análisis de su secuencia, posteriormente mostraron ser quinasas activas (Xu et al., 2000; Min et al., 2004; Abe et al., 2001). Igualmente, en los últimos años se ha puesto de manifiesto que diversas pseudoquinasas cumplen importantes roles biológicos a pesar de no poseer actividad catalítica, principalmente mediante la interacción y regulación de proteínas quinasas activas (Llompart et al., 2003; Boudeau et al., 2006; Chevalier et al., 2005; Halter et al., 2014; Kumar et al., 2016; Castells and Casacuberta, 2007). En consecuencia, independientemente de que NIPK sea o no una quinasa activa, podría estar regulando la actividad de NF-YC1 de manera postraduccional mediante su fosforilación (Figura 50). Este tipo de regulación ha sido reportada para NF-Y en células humanas, donde la subunidad NF-YA es fosforilada en dos serinas por la quinasa dependiente de ciclina CDK2. Una mutante de NF-YA que no puede ser fosforilada en estos residuos es capaz de formar el trímero con las subunidades B y C, pero se ve afectada su capacidad de unión al DNA, inhibiendo la activación transcripcional de los genes blanco *CDC2* y *CDK2* (Yun et al., 2003; Chae et al., 2004; Chan et al., 2010; Dolfini et al., 2012). También ha sido reportado que NF-YB y NF-YA pueden ser acetiladas por la proteína p300. Se ha demostrado que esta acetilación de NF-YA previene su ubiquitinación y consecuente degradación vía proteasoma, extendiendo su vida media (Li et al., 1998; Currie, 1998; Manni et al., 2008; Dolfini et al., 2012). Hipotéticamente, NIPK también podría estar regulando la acetilación o cualquier otro tipo de modificación postraduccional que ocurra en NF-YC1, ya que se ha reportado que diversas pseudoquinasas funcionan como andamios (*scaffolds*) para la formación de complejos multiproteicos (Boudeau et al., 2006). Desde ya, las posibles funciones de la interacción NIPK/NF-YC1 mencionadas son sólo especulaciones, pero sería interesante abordar ensayos funcionales para contrastar estas hipótesis. Lo que sí parece claro es que el dominio quinasa atípico de NIPK cumpliría

un rol relevante, ya que se encuentra conservado en todos los ortólogos putativos de las plantas vasculares estudiadas.



**Figura 50. Modelo de la interacción entre NIPK y NF-YC1.**

NIPK interacciona con NF-YC1 en el citoplasma, posiblemente promoviendo algún tipo de modificación postranscripcional en NF-YC1 (estrella roja). Durante la importación nuclear de NF-YC1 ocurriría la disociación del heterodímero. Una vez en el núcleo, NF-YC1 es capaz de formar el trímero simbiótico y asociarse a SIN1, generando los cambios transcripcionales requeridos para la interacción simbiótica.

Otra posible función de la interacción entre NF-YC1 y NIPK podría ser la de regular la localización subcelular de alguna de ellas. Dado que NIPK localiza en el citoplasma y la membrana plasmática, podría retener NF-YC1 en el citoplasma mediante su dimerización, impidiendo su importación al núcleo y la activación de los genes blancos requeridos para la nodulación. Sin embargo, esto posicionaría a NIPK con un rol negativo sobre la función de NF-YC1, lo cual no concuerda con los resultados funcionales realizados. Consistentemente, la sobreexpresión de NIPK no afectó la

### *Capítulo III*

localización subcelular de NF-YC1, al menos en *N. benthamiana*. De manera análoga, la función de la interacción NF-YC1/NIPK podría ser la translocación de este último al núcleo, pero los ensayos de BiFC no mostraron interacción nuclear, y la coexpresión de NF-YC1 junto con NIPK-GFP tampoco mostró una translocación de NIPK al núcleo (Figura 50).

En conclusión, el análisis funcional de la pseudokinasa putativa NIPK ha permitido revelar su función en la infección y organogénesis del nódulo, permitiendo ampliar el conocimiento sobre la vía de transducción de señales que involucra a NF-YC1. Este aporte plantea a su vez interesantes preguntas acerca de la posible función de la interacción que ocurre en el citoplasma entre un factor de transcripción y una proteína quinasa aparentemente inactiva.

# **CONCLUSIONES Y PERSPECTIVAS**

## *Conclusiones y perspectivas*

La simbiosis fijadora de nitrógeno permite a la planta obtener nitrógeno en formas asimilables independientemente de su disponibilidad en el suelo. Sin embargo, esta capacidad está restringida sólo a ciertas especies de angiospermas dispersas en cuatro órdenes relacionadas dentro del clado de Eurósidos I. Actualmente, la hipótesis más aceptada sugiere que la capacidad de establecer una asociación simbiótica fijadora de nitrógeno surgió hace 100 millones de años a partir de un antecesor común que adquirió una predisposición genética para formar este tipo de interacción (Matin et al., 2017). Posteriormente, la capacidad de formar nódulos fijadores de nitrógeno fue adquirida y perdida múltiples veces de manera independiente. Posiblemente, la clave de esta predisposición genética se basa en el reclutamiento de parte del programa genético requerido para el establecimiento de la simbiosis con micorrizas arbusculares (Kistner and Parniske, 2002). Esta interacción simbiótica intracelular con hongos del phylum Glomeromycota surgió hace 383-419 millones de años y se encuentra en más del 80 % de las plantas, permitiendo una mejor absorción de agua y diversos nutrientes tales como fósforo y zinc (Harrison, 2005). Se ha determinado que ambas simbiosis comparten componentes moleculares comunes de la planta que incluyen a SYMRK, CASTOR, POLUX, CCaMK, CYCLOPS y NUP85/133. Esta vía simbiótica común ha sido denominada *Common Symbiosis Pathway* (CSP).

El hecho que la simbiosis fijadora de nitrógeno haya surgido de manera independiente múltiples veces y la existencia de la vía simbiótica común, hace plausible la idea de transferir el programa genético requerido para la nodulación a otras especies vegetales de interés agronómico como los cereales (Oldroyd and Dixon, 2014). Diversos grupos de investigación enmarcados en el proyecto ENSA (*Engineering Nitrogen Symbiosis for Africa*), financiado por la fundación Bill y Melinda Gates, se encuentran abocados al estudio de estrategias que permitan identificar y transferir este programa genético



## *Conclusiones y perspectivas*

(<https://www.ensa.ac.uk/>). En este contexto, identificar y comprender las interrelaciones entre los componentes moleculares requeridos para el establecimiento de la simbiosis fijadora de nitrógeno es sumamente relevante, y los resultados de esta tesis aportaron nueva información referente a estos mecanismos. Se identificaron los cambios en la expresión génica en raíces de dos cultivares de *P. vulgaris* en respuesta a cepas de *R. etli* de distinto linaje evolutivo, que contribuyen a comprender la diversidad de respuestas moleculares asociadas a la nodulación, y cómo se correlacionan con su eficiencia simbiótica. A su vez, se ha identificado y caracterizado un nuevo gen requerido para la infección y organogénesis del nódulos en *P. vulgaris*, el cual codifica una proteína quinasa posiblemente inactiva que interacciona con el factor de transcripción NF-YC1. Será interesante investigar si esta interacción modifica a nivel postraduccional a NF-YC1 ya que este factor de transcripción es capaz de formar un complejo multiproteico con NF-YA1 y SIN1, los cuales también están implicados en la activación de programas genéticos necesarios para la infección y desarrollo de nódulos. Mediante la secuenciación masiva de las bibliotecas de DNA generadas a partir de muestras de CHIP, esperamos identificar y caracterizar las redes transcripcionales reguladas por estos factores de transcripción.

En último término, los resultados obtenidos constituyen un aporte a la comunidad científica abocada al estudio de la simbiosis fijadora de nitrógeno; un proceso biológico que atañe de manera directa aspectos críticos de nuestra sociedad, como lo son la producción de alimentos y el medio ambiente (Erisman et al., 2008).

# **MATERIALES Y MÉTODOS**

## 1. MATERIAL BIOLÓGICO

---

### 1.1. Semillas

Las semillas de *P. vulgaris* cultivar NAG12 (Mesoamericano) fueron provistas por Susana García Medina (INTA, Estación Experimental Agropecuaria Salta, Cerrillos, Argentina). Las semillas de *N. benthamiana* fueron provistas por la Dra. Carina Reyes (Laboratorio de Virus vegetales, IBBM).

### 1.2. Cepas de *E. coli*

---

Cepa	Aplicación	Referencia
DH5 $\alpha$	Utilizada para el mantenimiento y multiplicación de plásmidos	(Taylor et al., 1993)
TOP10	Cepa de alta eficiencia de transformación provista por Invitrogen para realizar transformaciones de las reacciones de pENTR y LR clonasa	Invitrogen
ccdB survival 2TR1	Cepa resistente al producto del gen <i>ccdB</i> utilizada para el mantenimiento y multiplicación de vectores de destino del sistema Gateway	Invitrogen

---

Las cepas de *E. coli* DH5 $\alpha$ , TOP10 y *ccdB survival 2 TR1* fueron crecidas en medio LB líquido a 37 °C y 250 rpm de agitación, o en cultivos sólidos en placas de Petri con medio LB-agar.

### 1.3 Cepas de *Agrobacterium*

Las cepas de *Agrobacterium* fueron crecidas en medio LB sólido o líquido a 28 °C y 250 rpm de agitación (si corresponde). Los medios fueron suplementado con los antibióticos correspondientes.

Especie	Cepa	Aplicación
<i>A. rhizogenes</i>	K599	Utilizada para generar raíces transgénicas en <i>P. vulgaris</i>
<i>A. tumefaciens</i>	GV3101	Utilizada para transformar de manera transiente hojas de <i>N. benthamiana</i>

#### 1.4. Cepas de *R. etli*

Cepa	Aplicación	Referencia
SC15 ( <i>NodC-α</i> )	Utilizada como cepa eficiente en ensayos de nodulación	Aguilar et al., 2004
CE3 ( <i>NodC-α</i> )	Utilizada como cepa eficiente en ensayos de nodulación	Zanetti et al., 2010
55N1 ( <i>NodC-δ</i> )	Utilizada como cepa poco eficiente en ensayos de nodulación	Aguilar et al., 2004
124N1 ( <i>NodC-δ</i> )	Utilizada como cepa poco eficiente en ensayos de nodulación	Zanetti et al., 2010
CFNX5 DsRed ( <i>NodC-α</i> )	Expresa de manera constitutiva la proteína fluorescente roja DsRed. Utilizada para el estudio de ITs	(Rípodas et al., 2013)

Las cepas de *R. etli* fueron crecidas en medio YEM líquido o sólido con Rojo Congo a 28 °C y 250 rpm de agitación (si corresponde). Los medios fueron suplementado con los antibióticos correspondientes.

#### 1.5. Medio de cultivo para *E. coli*, *A. rhizogenes* y *A. tumefaciens*

- **LB Luria-Bertani**

Extracto de levadura	5 g/L
NaCl	10 g/L
Triptona	10 g/L

El medio LB fue suplementado con el antibiótico adecuado para la selección de cada cepa. Para el medio LB-sólido se adicionaron 15 g/l de agar.

## 1.6. Medios de cultivo para rizobios

- **YEM (yeast extract mannitol)**

---

Extracto de levadura	0,4 g/L
Manitol	7 g/L
NaCl	0,1 g/L
MgSO <sub>4</sub> ·7H <sub>2</sub> O	0,2 g/L
K <sub>2</sub> HPO <sub>4</sub>	0,5 g/L

---

Para el medio de cultivo YEM sólido, el medio fue suplementado con 10 ml/l de una solución acuosa de Rojo Congo 2,5 g/l y 15 g/l de agar.

## 1.7. Medios de cultivo para levaduras

- **SD (Minimal Synthetic Defined Base)**

---

Base nitrogenada libre de aminoácidos	6,7 g/L
Glucosa	20 g/L

El medio SD se complementó con el *dropout*, que contiene todos los aminoácidos excepto los de selección de los vectores o de los marcadores de la interacción en el ensayo de doble híbrido. En esta tesis se utilizaron los siguientes *dropouts*:

-Trp-Leu: se empleó para la selección de los diploides transformados con pGBKT7 y pGAD-Rec.

-Trp-Leu-His, también denominado TDO (por *Triple Drop Out*): para selección de plásmidos e interacciones.

-Trp-Leu-Ade-His, también denominado QDO (por *Quadruple Drop Out*): para selección de plásmidos e interacciones.

## *Materiales y métodos*

El pH se ajustó a 5,8 y se esterilizó en autoclave por 15 min a 121 °C. Para preparar el medio sólido se agregaron 20 g/l de agar. Los medios sin histidina fueron suplementados con 10 mM de 3-AT (3-amino-1,2,4-triazole; Sigma Cat No. A-8056), que al ser un inhibidor de la síntesis de histidina previene el crecimiento de falsos positivos.

### **1.8. Antibióticos**

Se utilizaron las siguientes concentraciones de antibióticos en medios de cultivo sólidos o líquidos.

<b>Antibiótico</b>	<b><i>E. coli</i> (µg/ml)</b>	<b><i>A. tumefaciens</i> (µg/ml)</b>	<b><i>A. rhizogenes</i> (µg/ml)</b>	<b><i>R. etli</i> (µg/ml)</b>
Cloranfenicol	50	150	150	-
Kanamicina	50	50	200	-
Tetraciclina	-	-	-	5
Rifampicina	-	100	-	-
Espectinomicina	-	-	200	-
Gentamicina	-	15	-	-

### 1.9. Medio Fahraeus

Solución mineral nutritiva para riego de las plantas

	<b>Compuesto</b>	<b>Concentración 1000X</b>
<b>Macronutrientes</b>	CaCl <sub>2</sub>	0,78 M
	MgSO <sub>4</sub> .7H <sub>2</sub> O	0,49 M
	Citrato férrico	0,02 M
	KH <sub>2</sub> PO <sub>4</sub>	0,74 M
	Na <sub>2</sub> HPO <sub>4</sub>	0,84 M
<b>Micronutrientes</b>	KCl	50 mM
	H <sub>3</sub> BO <sub>3</sub>	25 mM
	MnSO <sub>4</sub> .H <sub>2</sub> O	5 mM
	CuSO <sub>4</sub> .5H <sub>2</sub> O	0,5 mM
	ZnSO <sub>4</sub> .7H <sub>2</sub> O	2 mM
	Na <sub>2</sub> MoSO <sub>4</sub> .2H <sub>2</sub> O	1 mM

Los macro y micronutrientes se prepararon como solución stock concentrada 1000 veces. Las soluciones de macronutrientes fueron esterilizadas por autoclave, mientras que los micronutrientes fueron esterilizados por filtrado a través de un filtro de 0,45 µm de poro. En el caso de Fahraeus suplementado con fuente de nitrógeno se adicionó KNO<sub>3</sub> hasta una concentración final de 8 mM.

### 1.10. Medio de preservación de microorganismos

La conservación a largo plazo de las cepas bacterianas se realizó por congelamiento de alícuotas de cultivos crecidos hasta fase logarítmica tardía que fueron luego suplementadas con glicerol estéril hasta alcanzar una concentración final de 10 % (v/v).

## *Materiales y métodos*

Las suspensiones de bacterias fueron congeladas a -80 °C y preservadas a esa temperatura.

### **1.11. Esterilización y germinación de las semillas**

Las semillas de poroto variedad NAG12 fueron esterilizadas con etanol 96° por 30 seg, seguido de una incubación con una solución de hipoclorito de sodio 10% (v/v) de 15 min con agitación suave. Luego de 6 lavados con agua destilada estéril, las semillas fueron colocadas en cajas de acrílico con papel absorbente humedecido con agua, y se incubaron durante 72 hs en cámara de cultivo a 26 °C y en oscuridad para su germinación.

### **1.12. Crecimiento de plántulas**

Las semillas germinadas fueron colocadas en potes plásticos conteniendo una mezcla 1:4 de arena perlita previamente esterilizada y fueron regadas con solución Fahraeus suplementada con KNO<sub>3</sub> 8 mM. Las plántulas se mantuvieron en una cámara de cultivo Sanyo MLR-350HT a 26 °C, 80 % de humedad y un ciclo luz/oscuridad de 16/8 horas durante 5 días.

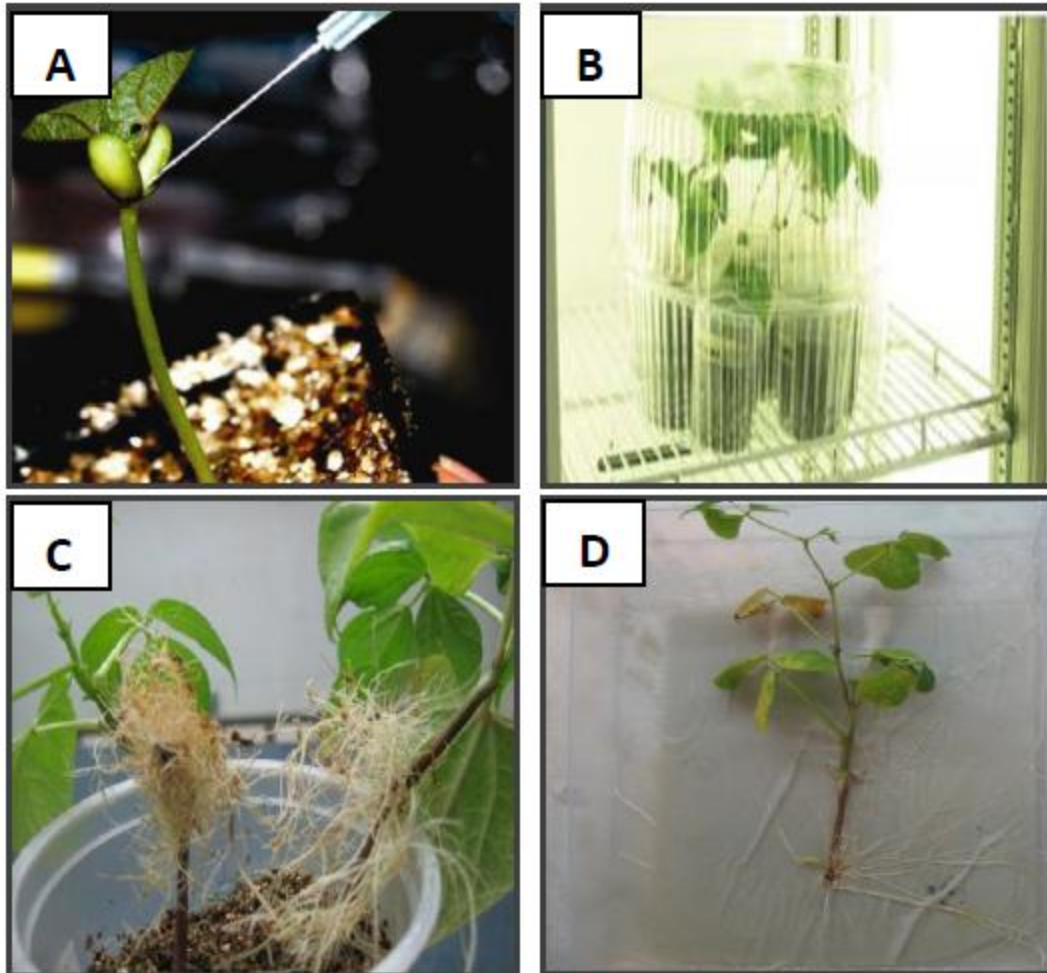
### **1.13. Crecimiento de *A. rhizogenes* K599**

Las cepas de *A. rhizogenes* K599 fueron crecidas por dos días en placas de petri con LB sólido suplementadas con el antibiótico correspondiente. Las células fueron colectadas con espátula y resuspendidas en 1 ml de agua destilada estéril hasta obtener una suspensión concentrada.



#### **1.14. Obtención de plantas compuestas**

Plantas crecidas por 5 días en potes con arena/perlita fueron transformadas siguiendo la técnica desarrollada por Estrada-Navarrete y colaboradores (Estrada-Navarrete et al., 2006). El procedimiento consiste en inocular 3 veces con jeringa Hamilton unos 30  $\mu$ l de una suspensión de *A. rhizogenes* K599 en el nodo del cotiledón (Figura 51A). Las plantas inoculadas fueron colocadas en recipientes plásticos cerrados y fueron regadas con abundante medio Fahraeus con  $\text{KNO}_3$  con el fin de generar una atmósfera con alta humedad que favorezca la infección bacteriana (Figura 51B). Las plantas fueron incubadas en cámara de cultivo a 26 °C y 80 % de humedad por un período de 15 días, momento en que las raíces transgénicas (*hairy roots*) tienen alrededor de 3 cm (Figura 51C). Por último, se corta el tallo 1 cm por debajo del cotiledón para remover la raíz salvaje y la parte aérea con las raíces transgénicas son colocadas en cajas con agar Fahraeus inclinado (Figura 51D).



**Figura 51. Generación de plantas compuestas.**

Para generar raíces transgénicas en plantas de *P. vulgaris* se inyectan 30  $\mu$ l de una suspensión de *A. rhizogenes* K599 en el nodo de los cotiledones utilizando una jeringa Hamilton (A). Para mantener la humedad saturante y favorecer el proceso de infección las plantas se colocan en recipientes plásticos sellados (B). En la zona de la inyección se genera un callo a partir del cual emergen las raíces transgénicas (C). Se corta el tallo 1 cm por debajo del cotiledón para remover la raíz salvaje y la planta se transfiere a cajas con agar Fahraeus inclinado (D).

### 1.15. Inoculación de plantas de poroto

Luego de un período de incubación de 7 días en ausencia de nitrato, las plantas compuestas fueron inoculadas con 5 ml de YEM o 5 ml de un cultivo de *R. etli* en fase exponencial con una  $DO_{600}$  de 0,8. Para obtener el cultivo se estrió la cepa de *R. etli* en placa de Petri con YEM sólido con rojo congo y se incubó durante 48 hs a 28 °C. A partir de la placa se inoculó un precultivo de 5 ml de YEM que se incubó por 48 hs a 28 °C y 250 rpm de agitación. A partir de estos precultivos, se inocularon erlenmeyers

## *Materiales y métodos*

conteniendo medio YEM (dilución 1/100) y los mismos fueron incubados en las mismas condiciones de crecimiento hasta alcanzar una OD<sub>600</sub> de 0,8 (fase exponencial tardía).

### **1.16. Generación de células de *E. coli* DH5α competentes químicas**

Células DH5α fueron plaqueadas por estriamiento en medio LB sólido a partir de un glicerol stock. Luego de 24 hs en estufa a 37 °C, se tomó una colonia aislada y se preparó un cultivo de LB líquido que fue incubado durante toda la noche a 37 °C y 250 rpm. A partir de este cultivo se realizó una dilución 1/100 y se incubó en las condiciones mencionadas hasta alcanzar una OD<sub>600</sub> de aproximadamente 0,8. El cultivo fue centrifugado a 2600 g durante 10 min y el precipitado fue resuspendido en 12,5 ml de la solución de transformación 1 (MOPS 10 mM, KCl 10 mM, pH 7,0). Luego se realizó otra centrifugación a 2600 g por 2 min y el precipitado fue resuspendido en 12,5 ml de solución de transformación 2 (MOPS 100 mM, KCl 10 mM, CaCl<sub>2</sub> 50 mM, pH 6,5). El homogenato de células fue incubado por 15 min en hielo y posteriormente se realizó una centrifugación de 2 min a 2600 g. El sobrenadante fue descartado y el precipitado de células se resuspendió en 2,5 ml de solución 2. El homogenato obtenido fue alicuotado, congelado rápidamente en nitrógeno líquido y posteriormente guardado a -80 °C.

### **1.17. Transformación de células de *E. coli* competentes químicas**

Alicuotas de células competentes químicas fueron mezcladas con 1-2 µl de plásmido e incubadas en hielo durante 30 min. Posteriormente se realizó un shock térmico a 42 °C durante 30 seg seguido de una incubación en hielo por 2 min. A continuación se agregó 1 ml de medio LB, y las células fueron incubadas a 37 °C durante una hora. El cultivo fue centrifugado a 5000 g durante 1 min, se descartó parte del sobrenadante y el

precipitado fue resuspendido en el medio residual. Finalmente, el homogenato fue esparcido en placas de Petri conteniendo medio LB agar suplementado con el antibiótico adecuado para la selección del plásmido de interés.

#### **1.18. Generación de células de *Agrobacterium rhizogenes* K599 y *A. tumefaciens* GV3101 electrocompetentes**

Células de *A. rhizogenes* K599 o *A. tumefaciens* GV3101 fueron plaqueadas por estriamiento en medio LB solido a partir de un glicerol stock. Luego de 2 días en estufa a 28 °C, se tomó una colonia aislada y se preparó un cultivo de LB líquido que fue incubado durante toda la noche a 28 °C y 250 rpm. A partir de este cultivo se realizó una dilución 1/100 en 200 ml de medio LB y se incubó a 28 °C y 250 rpm hasta alcanzar una OD<sub>600</sub> de aproximadamente 0,8. El cultivo fue enfriado en hielo durante 30 min y luego fue centrifugado a 6000 g y 4 °C durante 10 min. El sobrenadante fue descartado y se resuspendió el precipitado en 200 ml de glicerol 10 %. A continuación se volvió a centrifugar el homogenato, se descartó el sobrenadante y el precipitado se resuspendió en 4 ml de glicerol 10%. Este último paso se repitió en las mismas condiciones pero el precipitado se resuspendió en 2 ml de glicerol 10 %. Finalmente el homogenato fue alicuotado, congelado en nitrógeno líquido y guardado a -80 °C.

#### **1.19. Transformación de células de *Agrobacterium* electrocompetentes**

Alícuotas de células competentes químicas fueron mezcladas con 1 µl de plásmido y colocadas en cubetas de electroporación BioRad. Luego de una incubación por 5 min en hielo se realizó la transformación por electroporación mediante un pulso de 2,5 kV, 200U y 25 uFD. Inmediatamente se agregó 1 ml de medio LB y la suspensión bacteriana resultante fue incubada durante 1 h a 28 °C y 250 rpm. Luego el cultivo fue

centrifugado a 6000 g durante 1 min y resuspendido en 200 µl. Finalmente el homogenato fue esparcido en placas de Petri conteniendo medio LB agar suplementado con el/los antibiótico/s adecuado/s.

## **2. MANIPULACIÓN DE ÁCIDOS NUCLEICOS**

---

### **2.1. Minipreparación de DNA plasmídico para secuenciación**

A partir de una colonia bacteriana se inocularon 5 ml de medio LB con el antibiótico correspondiente al plásmido de interés. Luego el cultivo fue incubado a 37 °C con agitación durante aproximadamente 12 hs y se obtuvo un precipitado de células por centrifugación a 6000 g durante 10 min. A partir del precipitado se procedió a la minipreparación de DNA plasmídico mediante la utilización del *kit Wizard Plus SV Minipreps* (Promega), siguiendo las instrucciones del proveedor.

### **2.2. Minipreparación de DNA plasmídico de rutina**

A partir de una colonia bacteriana se inocularon 5 ml de medio LB con el antibiótico correspondiente al plásmido de interés. Luego el cultivo fue incubado a 37 °C con agitación durante aproximadamente 12 hs y se obtuvo un precipitado de células por centrifugación a 12000 g durante 30 seg. Luego de remover el sobrenadante el precipitado se resuspendió en 100 µl de Solución 1 (glucosa 50 mM, Tris-Cl 25 mM pH 8 y EDTA 10 mM pH 8). Se adicionaron 200 µl de Solución 2 (NaOH 0.2 N, SDS 1 % p/v) y se mezcló por inversión del tubo. A continuación fueron agregados 150 µl de Solución 3 (acetato de potasio 3 M pH 6). Se mezcló por inversión suave y luego se centrifugó a 12000 g por 5 min. El sobrenadante se pasó a un tubo nuevo y se precipitó el DNA con el agregado de 2 volúmenes de etanol. Se centrifugó a 12000 g por 5 min,

## *Materiales y métodos*

se descartó el sobrenadante y se dejó secar el precipitado. A continuación el precipitado fue lavado con etanol 70 % v/v, se centrifugó y descartó el sobrenadante como en el paso anterior. Por último se resuspendió el DNA en agua miliQ.

### **2.3. Reacción en Cadena de la Polimerasa (PCR)**

Las reacciones de PCR fueron realizadas en 20 µl finales con la siguiente concentración de reactivos: buffer de reacción 1X, MgSO<sub>4</sub> 1,5 mM (PBL-UNQ), dNTPs 2 mM, oligonucleótidos 0,25 µM, y enzima Taq Polimerasa 0,5 U (PBL-UNQ). Las reacciones de PCR fueron realizadas en un ciclador térmico Mastercycler gradient (Eppendorf). El tiempo de elongación del producto y la temperatura de apareamiento de los oligonucleótidos utilizados fue variable para cada reacción utilizada. El programa estándar consistió en una primera incubación a 94 °C por 5 min, seguido de 35 ciclos de 30 seg a 94 °C, 30 seg a 52 °C (o T<sub>m</sub> indicado), 1 min por kb a 72 °C y como paso final 10 min a 72 °C.

### **2.4. Electroforesis de DNA y RNA en geles de agarosa**

Los productos de las reacciones de PCR, de las minipreparaciones de plásmido y las extracciones de RNA fueron analizados en geles de agarosa en un rango de 0,8 - 2 % (p/v), dependiendo del tamaño de los fragmentos de DNA a separar. El buffer de corrida fue TBE 0,5 X (Tris base 45 mM, H<sub>3</sub>BO<sub>3</sub> 45 mM y EDTA 1 mM). A la mezcla agarosa/buffer se la suplementó con bromuro de etidio (BrEt) para visualizar el DNA/RNA por exposición a luz ultravioleta (UV). Para sembrar las muestras en las fosas del gel, se les agregó 1/6 volúmenes de *buffer* de siembra 6X para DNA (Tris-HCl 20 mM pH 8,0, glicerol 50 % (p/v), azul de bromofenol 0,5 % (p/v), EDTA 2 mM pH 8,0). Las electroforesis se llevaron a cabo en *buffer* TBE 0,5 X a voltaje constante

## *Materiales y métodos*

(usando 10 V por cada cm lineal de gel). Para visualizar y documentar los perfiles electroforéticos se utilizó un transiluminador de UV (UVP) y una cámara digital UVITEC Cambridge. Los geles de agarosa para evaluar la integridad del RNA fueron 1,2 % (p/v) en condiciones no desnaturalizantes, semejantes a lo descrito para evaluar las reacciones de PCR. La presencia de bandas discretas correspondientes al RNA ribosomal fueron indicadores de una buena integridad del RNA para ser usado como sustrato para reacciones enzimáticas.

### **2.5. Extracción de RNA**

La extracción de RNA vegetal se realizó a partir del macerado en nitrógeno líquido del tejido radical. A 200 mg de tejido se adicionó 1 ml de Trizol (Invitrogen). La mezcla fue incubada a temperatura ambiente por 5 min. Luego se agregaron 0,2 ml de cloroformo y se agitó vigorosamente. La mezcla se incubó durante 5 min a temperatura ambiente seguido de una centrifugación a 17000 g durante 10 min a 4 °C. La fase acuosa se transfirió a un nuevo tubo, y se agregaron 500 µl de isopropanol para precipitar el RNA. A continuación se centrifugó a 17000 g durante 10 min a 4 °C, se descartó el sobrenadante y el precipitado fue lavado 2 veces con etanol 75 % (v/v). Por último, el precipitado fue secado al aire y se resuspendió con 50 µl de agua miliQ seguido de una incubación de 10 min a 60 °C. El RNA disuelto fue cuantificado en un espectrofotómetro Nanodrop2000 (Thermo Scientific) utilizando volúmenes de 1 µl. La concentración de la muestra se calculó a partir de la  $A_{260}$  y se estimó su pureza a través de la relación  $A_{260}/A_{280}$  que en todos los casos fue cercana a 2. La integridad del mismo fue analizada mediante una electroforesis en gel de agarosa. La presencia de bandas discretas correspondientes al RNA ribosomal, fueron indicadores de una buena

integridad del RNA para ser usado como sustrato para reacciones enzimáticas. Las muestras de RNA fueron conservadas a  $-80\text{ }^{\circ}\text{C}$ .

## **2.6. Síntesis de cDNA**

Para la síntesis de la primera cadena del cDNA se partió de  $1\text{ }\mu\text{l}$  de RNA total, el cual fue digerido previamente con DNasa para eliminar DNA genómico que interfiera con las reacciones de RT-PCR. Para esto,  $1\text{ }\mu\text{g}$  de RNA se digirió con 1 U de la enzima RQ1 DNasa libre de RNasa (Promega) utilizando el buffer suministrado por el fabricante, en un volumen final de  $10\text{ }\mu\text{l}$ . Luego de incubar durante una hora a  $37\text{ }^{\circ}\text{C}$  se agregó  $1\text{ }\mu\text{l}$  de EDTA  $10\text{ mM}$  y se incubó durante  $10\text{ min}$  a  $65\text{ }^{\circ}\text{C}$  para inactivar la enzima. Para determinar el éxito de la reacción, una alícuota de la mezcla de reacción fue utilizada como molde para una PCR utilizando *primers* para *EF-1*. La ausencia de producto de amplificación fue tomada como criterio de RNA libre de DNA. A  $10\text{ }\mu\text{l}$  del RNA tratado con DNasa se le agregaron  $0,5\text{ }\mu\text{g}$  de oligo dT15 como cebador. La mezcla se incubó durante  $5\text{ min}$  a  $70\text{ }^{\circ}\text{C}$  para eliminar la formación de estructuras secundarias del RNA y luego se mantuvo en hielo. A continuación se agregó una mezcla conteniendo  $200\text{ U}/\mu\text{l}$  de transcriptasa reversa (MMLV RT, Promega), dNTPs  $0,5\text{ mM}$  (Promega) y buffer 1X provisto con la enzima en un volumen final de  $25\text{ }\mu\text{l}$ . La mezcla de reacción se incubó durante  $90\text{ min}$  a  $42\text{ }^{\circ}\text{C}$ , temperatura a la cual tiene lugar la síntesis de cDNA, y finalmente  $15\text{ min}$  a  $65\text{ }^{\circ}\text{C}$  para inactivar la enzima. Para determinar el éxito de la reacción, una alícuota de la mezcla de reacción fue utilizada como molde para una PCR utilizando *primers* para *EF-1 $\alpha$* . La presencia de un producto de amplificación de  $100\text{ pb}$  fue tomado como criterio de una síntesis de cDNA exitosa.



## **2.7. PCR cuantitativa en tiempo real (qRT-PCR)**

Las reacciones de qRT-PCR fueron llevadas a cabo en un termociclador Bio-Rad CFX96 Real Time System (BioRad) a partir de 5  $\mu$ l de una dilución 1/10 de las muestras de cDNA. La concentración de *primers* fue de 0,5  $\mu$ M en un volumen final de 20  $\mu$ l. Se utilizó la mezcla de reacción comercial iQ SYBR Green Supermix de BioRad. El protocolo de amplificación fue el siguiente: una primera etapa a 95 °C durante 5 min, seguido de 45 ciclos de 30 seg a 95 °C, 30 seg a la temperatura de *annealing* correspondiente a cada par de *primers* y 20 seg a 72 °C. Se realizaron dos réplicas técnicas, un control negativo sin cDNA para evaluar la presencia de alguna contaminación y una curva de calibración con diluciones seriadas del cDNA para cada par de *primers* utilizados. La posible dimerización de los *primers* o la amplificación de otros fragmentos fueron analizados a partir de los resultados de las curvas de *melting*. La obtención de un único producto de amplificación (de entre 100 y 200 pb) también se verificó por corrida electroforética en geles de agarosa. La normalización de los datos de expresión se realizó con el transcripto correspondiente al factor de elongación 1 $\alpha$  (EF1 $\alpha$ ) de *P. vulgaris*, según fuera descrito previamente por Peltzer-Meschini (2008). Los datos fueron analizados y graficados utilizando el programa GraphPad Prism V5.0 (GraphPad Prism Software).

## **2.8. Construcción de vectores mediante sistema GATEWAY**

Los vectores generados en este trabajo fueron construidos mediante el sistema de clonado GATEWAY. Brevemente, las secuencias de interés fueron clonadas en el vector pENTR/D-TOPO siguiendo las instrucciones del fabricante (Invitrogen) y posteriormente se clonaron en los vectores de destino correspondientes mediante

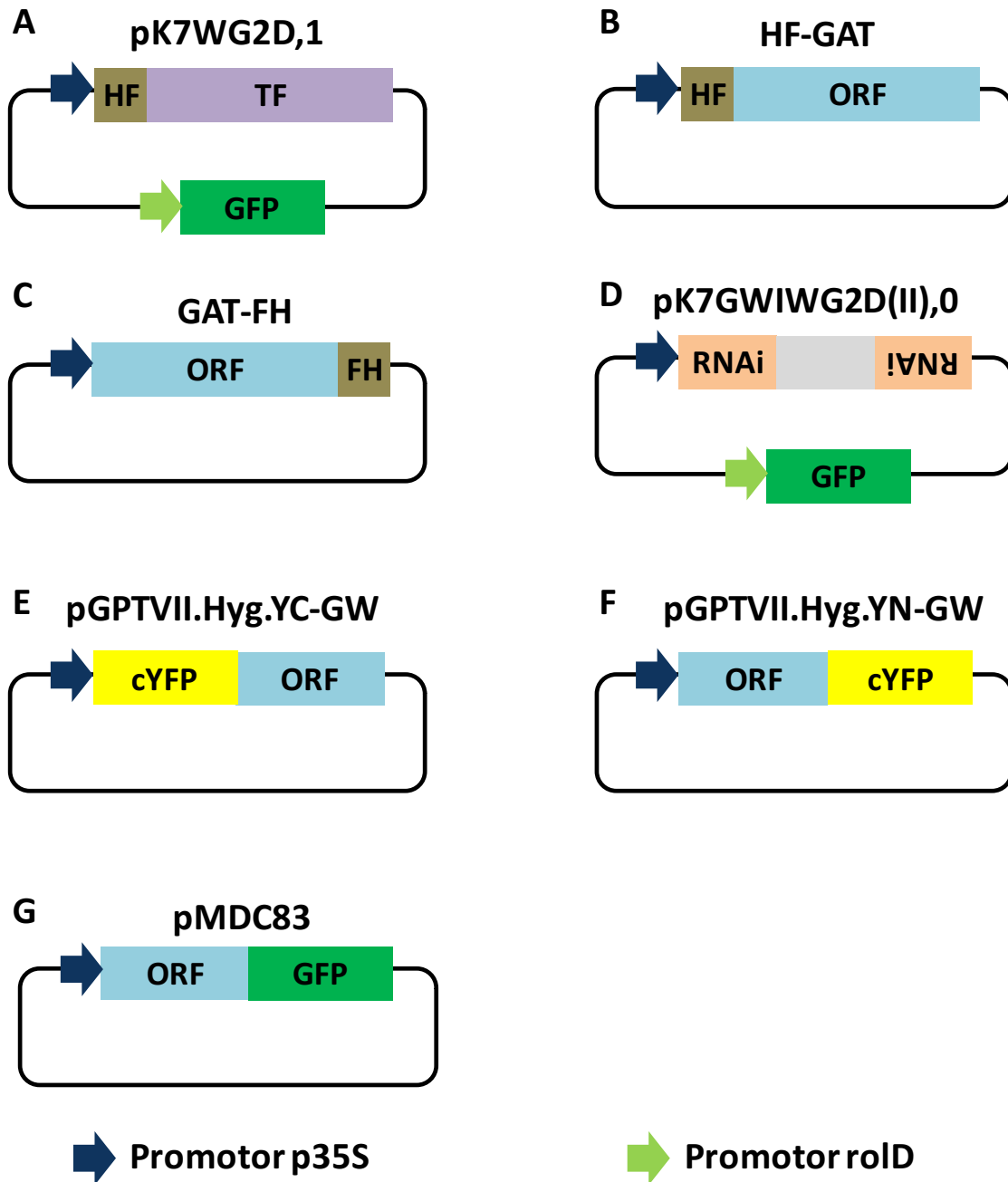
### *Materiales y métodos*

recombinación utilizando la enzima LR clonasa (Invitrogen), de acuerdo a las instrucciones provistas.

Para clonar las fusiones traduccionales HF-NF-YA1, HF-NF-YC1 y HF-SIN1 en el vector pK7WG2D,1, se usó como molde los vectores HF-GAT existentes en el laboratorio que contienen estos ORFs (Zanetti et al., 2010; Tesis doctoral de Carolina Rípodas). Los productos de PCR fueron clonados en el vector pENTR/D-TOPO, y luego este se utilizó para realizar la doble recombinación con el vector de destino pK7WG2D,1 (Figura 52A y B).

Para generar la construcción NIPK-HF se amplificó el ORF de NIPK con los primers NIPK ORF (Fw con ATG y Rv sin codón de STOP), y el producto de PCR fue clonado en el vector pENTR/D-TOPO. Por último, se realizó la recombinación con el vector de destino GAT-FH (Figura 52C).

Para el silenciamiento de *NIPK* mediante RNAi, se amplificaron por PCR un fragmento de 100 bp correspondiente a la región 3' UTR y un fragmento de 311 bp correspondiente al dominio quinasa, utilizando los *primers NIPK* RNAi 1 y 2 (Tabla 5) y cDNA de *P. vulgaris* como molde. Los fragmentos de PCR fueron clonados en el vector pENTR/D-TOPO, y posteriormente se recombinó con el vector pK7GWIWG2D(II) dando lugar a la formación de dos repeticiones invertidas del fragmento correspondiente en el vector de destino (Figura 52D). Las construcciones resultantes, *NIPK* RNAi 1 y 2, fueron analizadas en base a su patrón de restricción utilizando las enzimas *EcoRI* y *ApaI*. El plásmido control *GUS* RNAi fue previamente generado en el laboratorio (Blanco et al., 2009).



**Figura 52. Plásmidos utilizados.**

Esquemas de los plásmidos utilizados en este trabajo. En el plásmido pK7GW2D,1 se clonaron las fusiones traduccionales HF-NF-YA1, HF-NF-YC1 y HF-SIN1 para realizar los experimentos de ChIP (A). El plásmido HF-GAT (B) y GAT-FH (C) permiten fusionar el epítipo His-Flag (HF) al ORF de interés en el extremo N- y C-terminal, respectivamente. El plásmido pK7GWIWG2D(II),0 fue utilizado para silenciar NIPK mediante RNAi (D). El ORF de NIPK fue fusionado al extremo C-terminal de la YFP (cYFP) en su extremo N- (E) o C-terminal (F). El plásmido pMDC83 permite la fusión de la GFP al extremo C-terminal de un ORF (G).

## *Materiales y métodos*

Para los ensayos de BiFC, el ORF de *NIPK* fue amplificado por PCR a partir de cDNA mediante los *primers* NIPK ORF (Fw con ATG y Rv sin codón de STOP o Fw sin ATG y Rv con codón de STOP dependiendo si la fusión traduccional se realiza en el extremo C- o N-terminal, respectivamente) y clonado en el vector pENTR/D-TOPO como se describió anteriormente. El vector pENTR/D-TOPO con el inserto de interés se recombinó en los vectores pGPTVII.Hyg.YN-GW o pGPTVII.Hyg.YC-GW para obtener fusiones N- y C-terminales al fragmento C-terminal de la YFP (Figura 52E y F) (Hirsch et al., 2009). Las construcciones nYFP-NF-YC1, nYFP-NSP1 y cYFP-NSP2 fueron generadas por Battaglia et al., 2014.

Para el análisis de localización subcelular y coimmunopurificación (CoIP) se recombinó el vector pENTR/D-TOPO con el ORF de *NIPK* o la versión truncada  $\Delta$ N *NIPK*, con el vector de destino pMDC83. Este vector permite la expresión de proteínas como fusiones a GFP bajo el control del promotor p35S (Figura 52G).

## **3. CARACTERIZACIÓN FENOTÍPICA DE PLANTAS COMPUESTAS**

---

### **3.1. Cinéticas de nodulación**

Teniendo en cuenta que cada raíz de una planta compuesta es un evento de transformación independiente, se cuantificó el número de nódulos por raíz transgénica. Para esto, se observaron las raíces de las plantas en la lupa de fluorescencia previamente al momento de la inoculación, y se retiraron aquellas raíces que no expresaran GFP. Los vectores utilizados para realizar análisis funcionales permiten la expresión de GFP bajo el promotor rolD como gen reportero. La cantidad de nódulos por raíz fue determinada a los 7, 14 y 21 dpi, momento en que los nódulos alcanzan un estadio maduro. En todos

los experimentos se incluyeron plantas inoculadas con YEM como control negativo de la nodulación.

### **3.2. Determinación del tamaño de los nódulos**

Plantas compuestas crecidas en cajas de acrílico con nódulos de 21 dpi fueron utilizadas para estos experimentos. Los nódulos fueron fotografiados incorporando una regla como escala. El diámetro de los nódulos se midió a partir de las imágenes empleando herramientas del programa Adobe Photoshop 7.0.

### **3.3. Ensayos de coinoculación y análisis de polimorfismo del gen *nodC***

Para el ensayo de competencia de las cepas de *R. etli* SC15 y 55N1 se generaron plantas compuestas que fueron transferidas a potes conteniendo una mezcla perlita/arena y se regaron con medio Fahraeus sin nitratos. Luego de 7 días de crecimiento en cámara de cultivo, las plantas fueron inoculadas con 5 ml de una mezcla 1:1 de cultivos de SC15 y 55N1 con una  $DO_{600}$  de 0,8. Las plantas fueron incubadas 21 días en cámara de cultivo hasta que desarrollaron nódulos maduros. A continuación se colectaron los nódulos transgénicos expresando GFP como gen reportero y fueron esterilizados superficialmente con alcohol 96 % (v/v) durante 30 seg y una posterior incubación en agua oxigenada 30 vol durante 6 min. Para eliminar restos de agua oxigenada se lavaron los nódulos 6 veces con agua destilada estéril. Los nódulos fueron macerados individualmente en pocillos de placas de Elisa y posteriormente se agregaron 5  $\mu$ l de agua estéril para generar una suspensión bacteriana que fue esparcida en placas de Petri conteniendo medio YEM sólido con rojo congo. Las placas fueron incubadas durante 3 días a 28 °C, se tomaron muestras para determinar el polimorfismo del gen *nodC* y fueron guardadas a 4 °C durante una semana. Luego de esta incubación a 4 °C las cepas

## *Materiales y métodos*

SC15 y 55N1 pueden distinguirse en función de la coloración de sus exopolisacáridos. La cepa SC15 desarrolla un color rosa pálido, mientras que 55N1 desarrolla un color rojo rubí. Paralelamente, las muestras tomadas de las placas de Petri fueron utilizadas para inocular un cultivo líquido que fue incubado por 2 días a 28 °C y 250 rpm. El cultivo fue centrifugado a 15000 g, se retiró el sobrenadante, y el precipitado se resuspendió en 200 µl de NaCl 1 M. El homogenato fue centrifugado a 15000 g durante 2 min, se retiró el sobrenadante y se resuspendieron las células en 200 µl de agua. A continuación se repitió el paso anterior pero las células fueron resuspendidas en 200 µl de resina Chelex (Biorad). Las muestras fueron incubadas a 56 °C por 8 min seguido de una incubación a 99 °C por 10 min y centrifugadas a 15000 g por 2 min. El sobrenadante conteniendo el DNA fue utilizado como molde para una reacción de PCR para amplificar el gen *nodC* con los *primers* indicados en la tabla 5. El producto de PCR fue digerido con la enzima de restricción *HinfI* y posteriormente analizado por electroforesis en gel de agarosa.

### **3.4. Análisis de hilos de infección (ITs)**

El análisis de los hilos de infección fue realizado mediante microscopía de epifluorescencia con un microscopio invertido IX51 Olympus (Olympus Corporation). Para esto, plantas compuestas de poroto fueron crecidas en cajas de acrílico conteniendo agar Fahraeus inclinado sin nitrato durante 7 días y a continuación se inocularon con 5 ml de un cultivo de *R. etli* CFNX5 DsRed, que expresa de manera constitutiva la proteína fluorescente roja DsRed. Cuatro días luego de la infección se cuantificaron los hilos de infección por cm de raíz y se clasificaron en tres grupos de acuerdo a su progresión: ITs en pelo radical, ITs que alcanzan la epidermis e ITs que alcanzan el córtex de la raíz.

## **4. ANÁLISIS DE LA EXPRESIÓN DE PROTEÍNAS MEDIANTE *WESTERN BLOT***

---

### **4.1. Extracción de proteínas totales**

La extracción de proteínas fue realizada a partir de tejido de raíz de plantas compuestas de *P. vulgaris* u hojas de *N. benthamiana* pulverizadas en nitrógeno líquido. Se homogenizaron 100 mg de tejido con 200 µl de buffer de siembra 3 X (Tris-HCl 50 mM pH 6,8, SDS 2 % (p/v), glicerol 10 % (v/v), β-mercaptoetanol 0,2 % (v/v), azul de bromofenol 0,02 mg/ml) y para la desnaturalización de las proteínas las muestras fueron incubadas a 100 °C durante 5 min. Luego se precipitaron los restos celulares por centrifugación a 17000 g durante 10 min a 4 °C.

### **4.2. Separación de proteínas por SDS-PAGE**

La electroforesis de proteínas se realizó en geles de SDS-poliacrilamida al 12 % (p/v) de acuerdo a protocolos estándares (Sambrook, 2001) utilizando el Mini PROTEAN System 3 (Biorad). La electroforesis fue realizada en buffer Tris-Glicina (Tris-HCl 25 mM pH 8,8, glicina 190 mM, SDS 0,1 % (p/v)), con una corriente constante de 25 mA por gel durante aproximadamente 1 h.

### **4.3. Transferencia de las proteínas a membrana de nitrocelulosa**

Las proteínas fueron transferidas a una membrana de nitrocelulosa Hybond C (Amersham Pharmacia Biotech) utilizando el sistema de transferencia Mini Trans-Blot Cell (BioRad) y buffer de transferencia (Tris-HCl 25 mM, glicina 150 mM pH 8,0, metanol 20 %) con un voltaje constante de 100 V durante 60 min. La presencia de las

proteínas en las membranas de nitrocelulosa y la carga de proteínas en cada calle se evaluó por tinción de la membrana con una solución de Rojo Ponceau 0.5 % (p/v)

#### **4.4. Inmunodetección y revelado por quimioluminiscencia**

Las membranas fueron incubadas durante una noche a 4 °C con una solución bloqueante de leche descremada en polvo 5 % (p/v) en buffer PBS-T (las siglas de PBS-T provienen de Phosphate Buffered Saline with Tween 20; la composición del buffer 1 X es: Na<sub>2</sub>HPO<sub>4</sub> 3,2 mM, KH<sub>2</sub>PO<sub>4</sub> 0,5 mM, KCl 1,3 mM, NaCl 135 mM, Tween 20 0.05% (v/v), pH 7,4). Luego de descartar la solución de bloqueo y realizar 5 lavados con PBS-T, se agregó el anticuerpo primario diluido en TBS-T y se incubó a temperatura ambiente durante 60 min con agitación suave. Para la inmunodetección de proteínas fusionadas al epítopo FLAG o GFP se utilizaron los anticuerpos monoclonales anti-FLAG conjugado a la enzima peroxidasa (HRP) (1:500, Sigma-Aldrich) y anti-GFP conjugado a HRP (1:5000; Invitrogen), respectivamente. Luego de la incubación con el anticuerpo, se realizaron 5 lavados de la membrana con PBS-T 1X de 5 min cada uno. Para el revelado de las membranas se utilizó el sistema de revelado basado en la oxidación del luminol. Se mezclaron volúmenes iguales de una solución de luminol (en DMSO) 250 mM, ácido cumárico 90 mM y buffer Tris HCl 1M pH 8,5 con una solución de H<sub>2</sub>O<sub>2</sub> 30 vol en buffer Tris HCl 1M pH 8,5. Inmediatamente las membranas fueron expuestas a placas radiográficas durante aproximadamente 1 hora y fueron reveladas y fijadas siguiendo las instrucciones del proveedor (Amersham Pharmacia Biotech).



## **5. LOCALIZACIÓN SUBCELULAR E INTERACCIÓN PROTEÍNA-PROTEÍNA**

---

### **5.1. Ensayos de localización subcelular y complementación bimolecular por fluorescencia (BiFC) en *N. benthamiana***

Cepas de *A. tumefaciens* GV3101 fueron transformadas con las construcciones correspondientes para llevar a cabo los experimentos de localización subcelular y BiFC. En todos los ensayos fue utilizada una cepa de *A. tumefaciens* GV3101 que expresa la proteína supresora del silenciamiento P19, permitiendo altos niveles de expresión de proteínas heterólogas (Voinnet et al., 2003). Para realizar la agroinfiltración de hojas de *N. benthamiana* se partió de cultivos líquidos de las cepas correspondientes incubados toda la noche a 28 °C. Éstos fueron centrifugados a 2600 g y 4 °C durante 12 min, y el precipitado fue resuspendido en 5 ml de agromix (MgCl<sub>2</sub> 10 mM, MES/KOH 10 mM, pH 5.6 y acetosiringona 150 mM). El homogenato se incubó durante 2 h a temperatura ambiente y fue diluido hasta una DO<sub>600</sub> igual a 1. Por último se mezclaron las cepas necesarias para cada experimento en volúmenes iguales y se agroinfiltraron en hojas de *N. benthamiana* haciendo presión con una jeringa sin aguja como fue descrito por Voinnet et al., 2003. Muestras de hojas fueron analizadas mediante microscopía confocal 2 y 3 días luego de la agroinfiltración.

### **5.2. Ensayos de co-inmunoprecipitación**

Las cepas *A. tumefaciens* GV3101 fueron transformadas por electroporación con las construcciones correspondientes. La agroinfiltración de hojas de *N. benthamiana* se realizó como fue descrito en la sección anterior para los ensayos de BiFC. Tres días luego de la agroinfiltración se colectó y pulverizó el tejido de hoja en nitrógeno líquido.

## *Materiales y métodos*

A partir de 100 mg de este tejido se extrajeron las proteínas totales utilizando un buffer de extracción (Tris 10 mM pH 7,5, NaCl 150 mM, DTT 5 mM, glicerol 10 % (v/v), EDTA 1 mM, Triton X-100 0,1 % (v/v), CaCl<sub>2</sub> 5 mM) suplementado con polivinilpolipirrolidona (PVPP) 2 % (p/v) y un cóctel inhibidor de proteasas 1,65 % (v/v) (Sigma-Aldrich). Después de 30 min de incubación a 4 °C con agitación, el extracto se clarificó por centrifugación a 2400 g durante 15 min a 4 °C. El sobrenadante se mezcló con 40 µl de resina anti-FLAG conjugada a agarosa (Sigma-Aldrich) y se incubó en un agitador durante 1,5-2 hs a 4 °C. Después de una centrifugación a 17000 g, los coimmunoprecipitados se lavaron seis veces con 1 ml de buffer de extracción sin PVPP y posteriormente se incubaron en buffer de elución (péptido 3X FLAG Sigma-Aldrich con una concentración final de 200 ng/µl en buffer de extracción sin PVPP) durante 10 min a 4 °C. Las mezclas se centrifugaron y el sobrenadante se recuperó como el eluido de las coimmunoprecipitaciones. Como se describe en la sección 4 de “Materiales y Métodos”, las proteínas se separaron mediante SDS-PAGE y se detectaron por western blot usando un anticuerpo anti-FLAG (1:500; Sigma-Aldrich) o anti-GFP (1:5000; Invitrogen). Una alícuota del extracto crudo se sembró como control de entrada (*input*) en el inmunoblot.

### **5.3. Observaciones macroscópicas y microscópicas generales**

La observación del sistema radical para identificar las raíces transformadas que expresan la proteína fluorescente verde GFP se realizó en una lupa Leica MZFLIII (Leica Microsystems, Wetzlar, Germany). Las imágenes adquiridas en los campos visibles y de fluorescencia se digitalizaron con el programa Image-Pro Plus 5.1 (Inc. MD, U.S.A.). El seguimiento de la formación de hilos de infección (ITs) y la ocupación de nódulos por la cepa de *R. etli* marcada con DsRed se realizó con un microscopio

invertido Olympus IX51 utilizando luz blanca y UV con filtros apropiados (Olympus Corporation, Shinjuku Tokio, Japón). Las imágenes fueron capturadas usando una cámara Q-Color3 de alta resolución (Olympus Corporation). Para obtener secciones semifinas y ultrafinas de los nódulos de *P. vulgaris* se utilizó el servicio de la Facultad de Ciencias Veterinarias de la Universidad Nacional de La Plata, siguiendo el protocolo descrito previamente (Zanetti et al., 2010). Las secciones semifinas (de 1-2 mm) se tiñeron con azul de toluidina 0,04 % y se examinaron con un microscopio Olympus IX51 (Olympus Corporation). La microscopía confocal de raíces de *P. vulgaris*, hojas de *N. benthamiana* y la observación de ITs fue realizada con un microscopio confocal Leica (SP5) con objetivos 20X y 40X. Las muestras se excitaron con láser de Argón y los espectros de emisión utilizados para detectar las fluorescencias fueron los siguientes: GFP (498 nm – 550 nm), YFP (520-550) y DsRed (578 nm – 626 nm). Las imágenes fueron capturadas por la cámara del microscopio con el programa Leica Application Suite (Leica) (servicio provisto por la plataforma de microscopía avanzada PMA-FCEN, de la Universidad Nacional de La Plata).

#### **5.4. Ensayo de doble híbrido de levadura**

Las cepas de levadura diploides portando el ORF completo de *NIPK* en el vector pGADT7-Rec (Clontech) y el ORF completo de *NF-YC1* o sus versiones truncadas en el vector pGBKT7 fueron generadas por la Dra. Battaglia como fue descrito en Battaglia et al., 2014. Para realizar el ensayo de doble híbrido se utilizaron cultivos líquidos de las cepas diploides en el medio SD sin leucina y triptofano (Doble dropout: DDO) que fueron incubados a 28 °C durante 2 días. 5 µl de este cultivo fueron utilizados para sembrar una gota en placas de Petri conteniendo medio SD sólido careciendo de Leu y

Trp (DDO), Leu, Trp y Ade (TDO) o Leu, Trp, His y Ade (QDO). Las placas fueron incubadas durante 2 días a 28 °C y luego fueron fotografiadas.

### **5.5. Actividad $\beta$ -galactosidasa**

Los ensayos de medida de la actividad  $\beta$ -galactosidasa se realizaron usando orto-nitrofenil- $\beta$ -galactósido (ONPG) como sustrato, siguiendo las indicaciones provistas en el Yeast Protocols Handbook (Clontech). Cultivos de 5 ml en medio de selección SD líquido (SD-Trp-Leu), para mantener la selección de los plásmidos utilizados, se incubaron durante dos noches. El día del experimento, se preparó una solución de ONPG en buffer Z (Na<sub>2</sub>HPO<sub>4</sub> 0.06 M; NaH<sub>2</sub>PO<sub>4</sub> 0.04 M; KCl 0.01 M; MgSO<sub>4</sub> 0.001 M;  $\beta$ -mercaptoetanol 0,05 M) de 4 mg/ml. Se transfirieron 2 ml de cada cultivo a 10 ml de YPD y se incubaron a 30 °C durante 3-5 h con agitación (230-250 rpm) hasta que las células alcanzaron la fase semilogarítmica (OD<sub>600</sub>: 0,5-0,8). Se centrifugaron tres alícuotas de 1,5 ml de cultivo a 10000 g durante 30 seg y las células se resuspendieron en 300  $\mu$ l de buffer Z. Posteriormente se transfirieron 0,1 ml de las suspensiones celulares a nuevos tubos de microcentrífuga, los cuales se sometieron a tres ciclos de congelación/descongelación (nitrógeno líquido/baño a 37 °C) para lisar las células. Posteriormente se añadió 0,7 ml de buffer Z +  $\beta$ -mercaptoetanol 0,3 % (v/v) a cada reacción y a los blancos, e inmediatamente se añadieron 160  $\mu$ l de la solución de ONPG a cada tubo. Estos se colocaron en una incubadora a 30 °C y se comenzó a tomar el tiempo. A medida que las muestras fueron desarrollando el color amarillo, se añadieron 0,4 ml de Na<sub>2</sub>CO<sub>3</sub> 1M a la reacción y a las muestras blanco. Posteriormente se centrifugaron durante 10 min a 10000 g para sedimentar los restos celulares y finalmente el sobrenadante se transfirió a cubetas para medir su DO<sub>420</sub>. Una unidad de  $\beta$ -galactosidasa se define como la cantidad que hidroliza 1  $\mu$ mol de ONPG a o-

nitrofenol y D-galactosa por min por célula. A partir de esto la actividad  $\beta$ -galactosidasa fue calculada con la siguiente fórmula:

$$\text{Unidades } \beta\text{-galactosidasa} = 1,000 \times \text{OD}_{420} / (t \times 0.5 \times \text{OD}_{600})$$

t = tiempo transcurrido (en min) de incubación

OD 600/420 = A 600/420 de 1 ml de cultivo

## 6. ANÁLISIS BIOINFORMÁTICOS GENERALES

---

### 6.1. Análisis general de secuencias y diseño de *primers*

Las secuencias de los vectores utilizados, los marcos abiertos de lectura, los cromatogramas de la secuenciación de plásmidos, así como también el diseño de *primers*, fueron analizadas mediante el software libre UGENE (<http://ugene.net/>)

### 6.2. Bases de datos de secuencias genómicas

Se utilizó la base de datos genómica de *P. vulgaris* v1.0, *M. truncatula* 4.0v1 y *Glycine max* Wm82.a2.v1 disponibles públicamente (<http://phytozome.jgi.doe.gov/pz/portal.html>).

### 6.3. BLAST

La secuencia nucleotídica de *NIPK* fue utilizada como *query* en una búsqueda contra la base de datos genómica de *Phaseolus vulgaris* mediante un BLASTN.

Los ortólogos putativos de *NIPK* de *Medicago truncatula*, *Arabidopsis thaliana*, *Lotus japonicus*, *Glycine max*, *Amborella trichopoda*, *Solanum lycopersicum*, *Zea mays*, *Oryza sativa*, *Selaginella moellendorffii*, *Picea abies*, *Physcomitrella patens* y

*Chlamydomonas reinhardtii* fueron identificados mediante un tBLASTN utilizando la secuencia proteica de NIPK.

#### **6.4. Alineamientos múltiples de secuencias proteicas**

Los alineamientos múltiples de secuencias de proteínas fueron realizados con el programa Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). Posteriormente los alineamientos fueron coloreados utilizando el programa BoxShade ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)).

#### **6.5. Análisis filogenético**

Los análisis filogenéticos de NIPK y CYCP4;1 fueron realizados con el programa MEGA 6 (Tamura et al., 2013) a partir de los alineamientos múltiples obtenidos con Clustal Omega. Las relaciones entre las secuencias fueron inferidas utilizando el método *neighbor-joining* y la opción de *bootstrap* con 1000 repeticiones. Los árboles filogenéticos fueron diagramados en una escala definida por las distancias evolutivas con el programa MEGA5. Las distancias evolutivas se consideran como el número de sustituciones de aminoácidos por sitio.

#### **6.6. Identificación de dominios funcionales en NIPK**

Para la anotación funcional de NIPK, la secuencia de proteína fue utilizada para realizar una búsqueda en la base de datos UNIPROT mediante tBLASTN.

La búsqueda de dominios transmembrana fue realizada mediante el servidor TMHMM Server, v. 2.0.

## **7. RNA-SEQ**

---

### **7.1. Material biológico y construcción de las bibliotecas**

Las semillas del cultivar NAG12 y Alubia fueron germinadas según el protocolo descrito. A continuación se colocaron en cajas de acrílico con agar Fahraeus inclinado sin nitratos y se incubaron en cámaras de cultivo por un periodo de 7 días. Las plantas fueron inoculadas con 5 ml de YEM o de cultivo líquido de la cepa de *R. etli* correspondiente. Los cultivos de rizobios se crecieron según el protocolo descrito.

Como material biológico de partida para el RNA-seq se utilizaron raíces del cultivar NAG12 y Alubia 24 hpi con las cepas SC15, CE3, 55N1, 124N1 o YEM a modo de control. Para cada condición se realizaron 3 réplicas biológicas conteniendo tejido de 5 plantas cada una.

A partir del tejido obtenido, se realizó la extracción de RNA total de cada una de las muestras utilizando TRIZOL. El RNA obtenido fue analizado cuali y cuantitativamente mediante electroforesis capilar con el equipo 2100 Bioanalyzer (Agilent Technologies). Todas las muestras superaron el valor mínimo de RIN (*RNA integrity number*) sugerido por Illumina. Este parámetro estima la integridad del RNA. Para la construcción de las bibliotecas se partió de 3 µg de RNA y se siguió el protocolo TruSeq RNA Sample Preparation kit v2 de Illumina. El protocolo consta de una primera etapa en la cual el mRNA es aislado a partir de RNA total utilizando bolillas magnéticas conjugadas a oligo-dT y luego fragmentado utilizando cationes divalentes, generando fragmentos de aproximadamente 200 pb. EL RNA fragmentado es transcripto a cDNA utilizando hexanucleótidos con secuencias al azar y la transcriptasa reversa Superscript II (Invitrogen). La segunda etapa consiste en la reparación de extremos para generar extremos romos seguidos de una ligación de adenina en el extremo 3' y de una reacción

de ligación con oligonucleótidos adaptadores con timinas en sus extremos, generando fragmentos de cDNA de 200 pb aproximadamente. Finalmente, el protocolo requiere de una etapa de amplificación por PCR de los fragmentos de cDNA. Las etapas de purificación del cDNA, la ligación de adaptadores, y la amplificación de PCR requieren de un proceso de purificación en el que se utilizan bolillas AMPure XP beads (Beckman Coulter Genomics). El sistema de purificación de AMPure XP PCR consiste en una tecnología de fase sólida paramagnética para la automatización de la purificación de amplicones o DNA en ensayos de gran escala. Este sistema utiliza un buffer optimizado para unir selectivamente DNA de 100 pb o mayor tamaño a las bolillas paramagnéticas. El exceso de oligonucleótidos, nucleótidos, sales y enzimas se remueven en los pasos de lavados. El análisis de calidad de las bibliotecas se llevó a cabo con el equipo 2100 Bioanalyzer, usando el chip Agilent DNA-1000 (Agilent Technologies) de acuerdo a las indicaciones provistas en el manual. Finalmente, las bibliotecas generadas fueron enviadas al servicio de secuenciación de la Universidad de California (Davis), donde fueron *pooladas* y posteriormente secuenciadas mediante la plataforma HiSeq 4000 con lecturas simples de 90 pb.

## **7.2. Procesamiento y análisis de la información de secuenciación**

El análisis informático fue llevado a cabo a través de la interfase GALAXY (Blankenberg and Hillman-Jackson, 2014; Blankenberg et al., 2010). Los archivos fastq provistos por el servicio de secuenciación corresponden a un archivo que contiene las secuencias, pero además posee información de la calidad de las lecturas por cada base. La primera etapa del análisis requirió convertir el formato fastq de las secuencias al formato fastq Sanger con la opción Groomer de GALAXY permitiendo obtener un formato estándar fastq compatible con el software de análisis. Esta opción se encuentra



## *Materiales y métodos*

en NGS:QC Manipulation > FastQ Groomer. Luego se procedió a analizar la calidad de las lecturas con la opción NGS:QC and Manipulation > FASTQ QC > Fastqc para cada uno de los archivos fastq.

Esta herramienta calcula una serie de parámetros que permiten determinar si la calidad y confiabilidad de las secuencias provistas es aceptable. Los parámetros calculados son: valor medio de Q para cada base, índice de calidad por secuencia, el contenido de GC por secuencia, el porcentaje de cada nucleótido para cada posición en la secuencia, el contenido de bases no identificadas (N) por secuencia, la distribución de tamaños de las secuencias, cantidad de secuencias duplicadas, y la existencia de secuencias y Kmers sobrerrepresentadas. Todos estos parámetros superaron los estándares de calidad esperados para el tipo de secuenciación realizado.

El alineamiento de las lecturas de RNA-Seq con el genoma de referencia Andino de *P. vulgaris* (genoma Phaseolus vulgaris v1.0, fuente DOE-JGI, USDA-NIFA, ARRA), se realizó con el programa TopHat2 (Kim et al., 2013). Esta herramienta se encuentra en la opción NGS:RNaseq>TopHat2 y la misma utiliza el programa Bowtie2. Luego de seleccionar el genoma de referencia en la herramienta TopHat2, se seleccionó la opción Full Parameters y se modificaron los siguientes parámetros recomendables para genomas de plantas:

- > minimum intron length: 60
- > Maximum intron length: 6000
- > Allow indel search: YES
- > Minimum intron length that may be found during split-segment (default) search: 45
- > Maximum intron length that may be found during split-segment (default) search: 5000
- > Use own junction: YES
- > Use gene annotation model: YES> gene model gff3 archive

## *Materiales y métodos*

> Use coverage search: YES

Los archivos generados tienen una extensión BAM, que luego fueron analizados con el paquete de análisis Cufflinks que consta de tres herramientas de análisis: Cufflinks, cuffmerge y cuffdiff (Trapnell et al., 2013, 2012). La herramienta Cufflinks permitió ensamblar los transcritos para cada condición y se encuentra en la opción NGS\_RNAseq > Cufflinks. Para un mejor análisis estadístico se seleccionaron los siguientes parámetros:

> Reference anotation: YES

> Perform quartile normalization: YES

> Perform BIAS correction: YES

> Use multi-read correct: YES

> Use effective length correction: YES

Seguidamente, los transcritos ensamblados fueron combinados con la opción NGS:RNAseq> CuffMerge. Hasta esta etapa del análisis bioinformático, los archivos fastq de las réplicas biológicas fueron analizados independientemente. Sin embargo, en la herramienta CuffDiff, los archivos resultantes de TopHat2 para las réplicas biológicas fueron agrupados. Finalmente, los archivos con los alineamientos de TopHat2 y el archivo de todos los transcritos ensamblados obtenido luego del CuffMerge fueron analizados con Cuffdiff para detectar transcritos con expresión diferencial. Cuffdiff está ubicado en NGS\_RNAseq > CuffDiff. Los parámetros modificados fueron:

> Library dispersion method: Quartile

> Multi read correct: YES

> Perform Bias Correction: YES

> Use Muli read correct: YES

Cuffdiff genera varios archivos, entre ellos archivos que indican la abundancia de los transcritos a nivel del gen e isoformas. Esta herramienta permite estimar los valores de FPKM (del inglés Fragments Per Kilobase of transcript per Million mapped reads, Fragmentos por Kilobase de transcripto por millón de lecturas alineadas) para cada transcritos para cada una de las muestras con las anotaciones permitiendo encontrar diferencias significativas en la expresión de transcritos, splicing y el uso de promotores. Los archivos de Cuffdiff fueron descargados y analizados con CummeRbund en R versión 3.0.3 para Windows 64 bits (<http://cran.r-project.org/bin/windows/base/>). CummeRbund implementa numerosas funciones gráficas utilizadas para la visualización de datos. Mediante el uso de esta herramienta se realizaron los gráficos de box plot mostrando la media y dispersión de los valores de FPKM, así como también, los gráficos vulcano mostrando los genes diferenciales en las distintas comparaciones de a pares.

### **7.3. Análisis de *clustering* jerárquico y PCA**

El análisis de componentes principales y el de *clustering* jerárquico de la Figura 14 fue realizado con el software estadístico Past3. Para esto se utilizaron los valores del log<sub>2</sub> de los valores de expresión de todos los genes normalizados por el respectivo valor en YEM y se utilizó el método de correlación en ambos casos. Para el análisis de *clustering* jerárquico correspondiente a la Figura 16 fueron utilizados los valores del log<sub>2</sub> de los valores de expresión de los genes diferenciales normalizados por el respectivo valor en YEM. Los datos fueron analizados con el software Morpheus (<https://software.broadinstitute.org/morpheus/>) utilizando la herramienta *Hierarchical clustering* seleccionando los siguientes parámetros:

–Metric: one minus Pearson correlation

## *Materiales y métodos*

–Linkage method: Average

–Cluster: rows and columns

### **7.4. Análisis de enriquecimiento de GO (*gene ontology*)**

El análisis de enriquecimiento de GO fue realizado a partir de los genes que mostraron una expresión diferencial en muestras inoculadas con rizobios respecto al control. Los identificadores de estos genes fueron cargados en la herramienta *Singular Enrichment Analysis* (SEA) disponible en la plataforma AgriGO (<http://systemsbiology.cau.edu.cn/agriGOv2/>), con el genoma de *P. vulgaris* como referencia. Este análisis fue realizado por separado para los genes regulados positiva y negativamente para cultivar, y luego fueron comparados mediante la herramienta *Cross comparison of SEA* (SEACOMPARE).

### **7.5. Diagramas de Venn**

Para realizar los diagramas de Venn se identificaron los genes con expresión diferencial en respuesta a cada cepa respecto al control. Las listas con los identificadores de estos genes fueron cargadas en la herramienta web *Calculate and draw custom Venn diagrams*, disponible en el portal de la universidad de Gent (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). El programa genera el diagrama mostrando la superposición de elementos, así como también, tablas mostrando los elementos compartidos y únicos entre categorías.

## **8. CHIP-PCR Y CHIP-SEQ**

---

### **8.1 Fijación y colección de tejido**

El tejido de raíz fue colectado y sumergido completamente en 25 ml de Buffer MC (fosfato de sodio 10 mM pH 7,5; NaCl 50 mM; sacarosa 0,1 M) con 1 % de formaldehído para fijar las interacciones DNA-proteína. La incubación se realizó aplicando vacío por 15 min, se mezcló el tejido por inversión del tubo, y se aplicó vacío por 14 min más. Seguidamente se detuvo la fijación agregando 2,5 ml de glicina (1,25 M) y nuevamente se aplicó vacío por dos min. El tejido fue lavado 3 veces con buffer MC (25 ml por lavado), fue secado, se molió en nitrógeno líquido y finalmente fue conservado a -80 °C.

### **8.2 Aislamiento de núcleos y sonicación**

Para el aislamiento de núcleos se transfirieron aproximadamente 5 gr de tejido a un tubo Falcon de 50 ml conteniendo 30 ml de buffer 1 (sacarosa 0,4 M; Tris-HCl 10 mM, pH 8; MgCl<sub>2</sub> 10 mM; β-mercaptoetanol 5 mM; inhibidor de proteasas). Se mezcló con vórtex y se incubó en hielo durante 30 min y agitación suave. Luego se filtró la suspensión a través de una malla de 55 μm y se centrifugó a 2880 g a 4 °C durante 20 min. El sobrenadante fue descartado y el precipitado se resuspendió en 1 ml de Buffer 2 (sacarosa 0,25 M; Tris-HCl 10 mM, pH 8; MgCl<sub>2</sub> 2 mM; β-mercaptoetanol 5 mM; Triton 0,15%; inhibidor de proteasas). El homogenato fue centrifugado a 1133 g a 4 °C por 10 min. Nuevamente se removió el sobrenadante y se resuspendió en 300 μl de Buffer 2. Se agregaron 300 μl de Buffer 3 (sacarosa 1,7 M; Tris-HCl 10 mM, pH 8; MgCl<sub>2</sub> 10 mM; β-mercaptoetanol 5 mM; inhibidor de proteasas) a un nuevo tubo y se agregó la solución obtenida suavemente para evitar que se mezclen. Se centrifugó a

## *Materiales y métodos*

1133 g a 4 °C por 1 h. Se removió el sobrenadante y se resuspendió en 500 µl de Buffer de sonicación (Tris-HCl 10 mM, pH 8; EDTA 10 mM; SDS 1%; inhibidor de proteasas). La sonicación fue realizada mediante un sonicador de punta *Sonics Vibra Cell* para el caso de los experimentos de ChIP-PCR. Se sonicó al 50% de amplitud durante 15 seg (con el tubo en hielo) y se realizó una pausa de 45 seg para que se enfríe la muestra. El ciclo fue repetido 4 veces y luego se centrifugó a 1133 g por 10 min a 4 °C. El sobrenadante (*input*) fue utilizado para las reacciones de inmunoprecipitación. Los experimentos de ChIP destinados a la construcción de las bibliotecas para secuenciación masiva fueron realizados en el laboratorio del Dr. Federico Ariel (Laboratorio de Epigenética y ARNs no codificantes, Instituto de Agrobiotecnología del Litoral) utilizando un sonicador de baño.

### **8.3. Preparación de *Dynabeads* conjugadas a *anti-FLAG* o IgG**

50 µl de *Dynabeads* fueron resuspendidas y alicuotadas en un tubo de 1,5 ml y luego fueron sedimentadas utilizando un magneto por 3 min. Se descartó el sobrenadante y las bolillas se lavaron con 400 µl de Buffer B (Tris-HCl 16,7 mM, pH 8; EDTA 1,2 mM; NaCl 167 mM). Se realizaron 3 lavados. Por otro lado, se diluyeron 5 µl del anticuerpo anti-FLAG (1 µg/µl) o 1 µl de IgG (5 µg/µl) en 400 µl de Buffer B. Seguidamente se mezclaron el anticuerpo diluido con las bolillas y se incubaron 1 h a temperatura ambiente. Las bolillas fueron sedimentadas empleando un magneto durante 3 min. Se descartó el sobrenadante y se lavaron 2 veces las bolillas con 400 µl de Buffer B.

### **8.4. Inmunoprecipitación**

El sonificado se diluyó 1/10 en *ChIP dilution Buffer* (Tris-HCl 16,7 mM, pH 8; EDTA 1,2 mM; NaCl 167 mM; Triton 1,1%; inhibidor de proteínas) y fue mezclado con las

## *Materiales y métodos*

bolillas magnéticas conjugadas con anti-FLAG (ver Preparación de *Dynabeads* conjugadas a *anti-FLAG* o IgG). Se incubó la mezcla 2 horas en hielo con agitación. Luego se separaron las bolillas con 8 min de magneto, y seguidamente se removió el sobrenadante. Se hicieron 5 lavados con 1 ml de *ChIP dilution Buffer*. Las bolillas se resuspendieron en 70  $\mu$ l de *Elution Buffer* (Tris-HCl 50 mM, pH 8; SDS 10%; EDTA 10mM; DTT 50 mM) y se incubaron 5 min a 95 °C. Nuevamente se separaron las bolillas con 8 min de magneto y se guardó el sobrenadante conteniendo la proteína de fusión. Se repitió la elución con 80  $\mu$ l de *Elution Buffer* y finalmente se combinaron ambos volúmenes (150  $\mu$ l finales). La solución se centrifugó 2 min a 16200 g y se guardó el sobrenadante. Para los experimentos de ChIP-seq la elución se realizó durante el *reverse crosslinking* directamente.

### **8.5. Reverse crosslink**

Se mezclaron 150  $\mu$ l de TE con el sobrenadante obtenido en el punto anterior, se le agregaron 11,25  $\mu$ l de proteinasa K (20 mg/ml) y se incubó *overnight* a 37 °C. Luego se agregó una segunda alícuota de 11,25  $\mu$ l de proteinasa K y se incubó 8 h a 65 °C.

### **8.6. Precipitación del DNA**

Para precipitar el DNA se agregaron a la muestra 2,5 volúmenes de etanol, 1/10 volúmenes de acetato de sodio 3 M pH 5,4 y 1  $\mu$ l de glucógeno. Se dejó precipitando *overnight* a -20 °C. Seguidamente se centrifugó a 1133 g durante 30 min a 4 °C, se descartó el sobrenadante y se dejó secar el precipitado por 15 min. El precipitado fue resuspendido en 100  $\mu$ l de agua. Esta solución fue utilizada como molde para reacciones de PCR cuantitativa como se describió en la sección 2.7.

### **8.7. Análisis de datos**

En los experimentos de ChIP-PCR de la sección 3 de Resultados del Capítulo 2, los valores de Ct de las regiones promotoras de interés obtenidos por qPCR fueron normalizados por los valores de Ct obtenidos para una región genómica control utilizando los *primers* qNIPK. El grado de enriquecimiento fue calculado como el cociente entre los valores obtenidos para muestras HF-NF-YC1 y control (plantas transformadas con el vector vacío).

Para el caso de los experimentos de ChIP-PCR de la sección 4 de Resultados del Capítulo 2, los valores de Ct de las regiones promotoras de interés y la región genómica utilizada como control obtenidas en las muestras IP FLAG, fueron relativizadas a los valores de las muestras IP IgG.

## **9. PRIMERS UTILIZADOS**

---

### **9.1. Diseño**

Los primers de este trabajo se diseñaron con el programa Primer3 y luego fueron analizados en detalle de con la herramienta OligoAnalyzer 3.1 de IDT (<https://www.idtdna.com/calc/analyzer>).

### **9.2. Secuencias**

La tabla 5 muestra la secuencia de los primers utilizados en este trabajo:

<i>Primer</i>	<i>Secuencia</i>
His-FLAG Fw	CACCATGGGACATCACCATCATCACC
<i>NF-YAI</i> ORF Rv	TCAAACCTTTAAGGTTGCAGCAG
<i>NF-YCI</i> ORF Rv	CTAATGTTCTGGGGATGGTTGACG
<i>SINI</i> ORF Rv	TCACCTCCAAGCAGAAGAAGTGG



*Materiales y métodos*

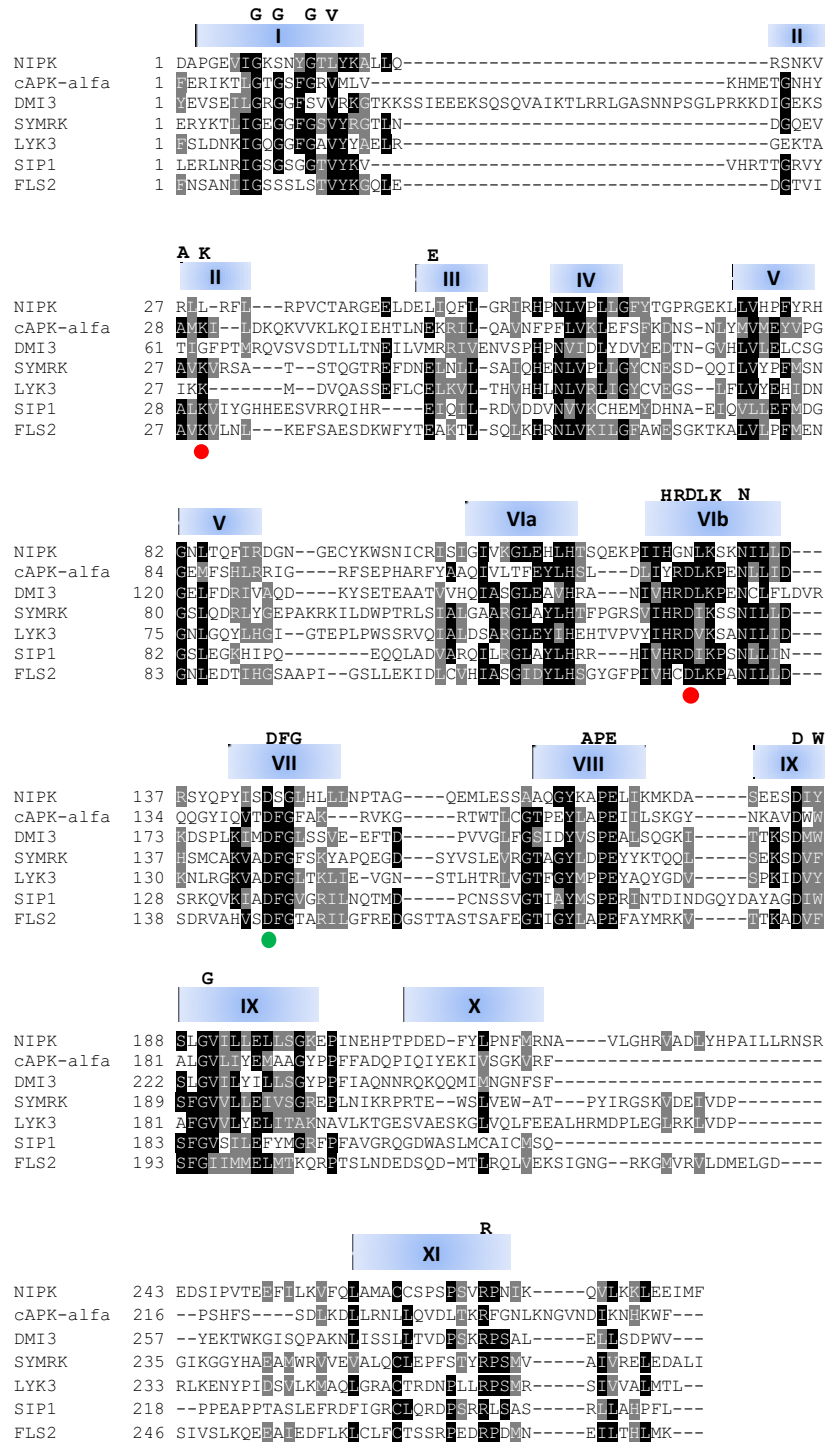
qp <i>ERN1</i> Fw	TGTTCTAGCCCTTGGGATGC
qp <i>ERN1</i> Rv	TTTTCAACCGCCCCTGTGTA
qp <i>ERN2</i> Fw	TTGTAGCATACCAATTCTTGC
qp <i>ERN2</i> Rv	CCAATTTCTCCTTATTTTGTTCGG
qp <i>CYCP</i> Fw	TGGAATGCCCAAACCCTTGT
qp <i>CYCP</i> Rv	TGTGGTGGCTTGTGTCCTTT
qp <i>WEE1</i> Fw	CATTAGTACGGCAAATCATTTC
qp <i>WEE1</i> Rv	TAAGCTAGTGGAAATACGGAGAG
qp <i>CYCC</i> Fw	ACCCAACATCGAAGTGAAATCC
qp <i>CYCC</i> Rv	AAAAGTGTGCAACTTCAGCTCG
qp <i>CDC2</i> Fw	AACGTTCTAATAGAAGAGTATTCCA
qp <i>CDC2</i> Rv	GACAAGTCACAATTATGACAACC
qp <i>CDC25</i> Fw1	ATGCGTGTGCTGATAAGTTG
qp <i>CDC25</i> Rv1	TTGACATCTGACACTTCAAGTA
qp <i>CDC25</i> Fw2	GTGGAAGTATTTTTCGTGGGAGT
qp <i>CDC25</i> Rv2	GTCCTAACAGTAAACGCAACACA
<i>NIPK</i> ORF Fw	<u>CACCATGTTGATACTAGTTCTTGGACTAGC</u>
<i>NIPK</i> ORF sin ATG Fw	<u>CACCTTGATACTAGTTCTTGGACTAGCTTC</u>
<i>NIPK</i> ORF Rv	CTAGAACATAATTTCTTCAAGTTTCTTGAGG
<i>NIPK</i> ORF sin STOP Rv	GAACATAATTTCTTCAAGTTTCTTGAGGAC
$\Delta$ <i>NIPK</i> Fw	<u>CACCGTGATTGGAAAATCCA</u> ACTATGGAAC
<i>NIPK</i> RNAi 1 Fw	<u>CACCGTGAGCTTGTGACATGGACAATG</u>
<i>NIPK</i> RNAi 1 Rv	GCTACAAATTCTTGGACAGCAGCAC
<i>NIPK</i> RNAi 2 Fw	<u>CACCGAAGCTTCTTGTT</u> CATCCCTTC
<i>NIPK</i> RNAi 2 Rv	TTACAAGGCACCTGAGCTCA
q <i>NIPK</i> Fw	<u>CACCGTGAGCTTGTGACATGGACAATG</u>
q <i>NIPK</i> Rv	GCTACAAATTCTTGGACAGCAGCAC
q <i>Phvul.011G070500</i> Fw	GTGGTCCGTTGTAGTTGCCA
q <i>Phvul.011G070500</i> Rv	TGAAAGGGCAAGAGGCAAAAC

## *Materiales y métodos*

<i>nodC</i> Fw	TGCTTGACACAACCAGCACTG
<i>nodC</i> Rv	GACAGCCAGTCGCTATTGGTCA
q <i>EF1-α</i> Fw	AACTACCACCGGCCACTTGAT
q <i>EF1-α</i> Rv	AGCACCCAGGCATACTTGAAT

**MATERIAL  
SUPPLEMENTARIO**

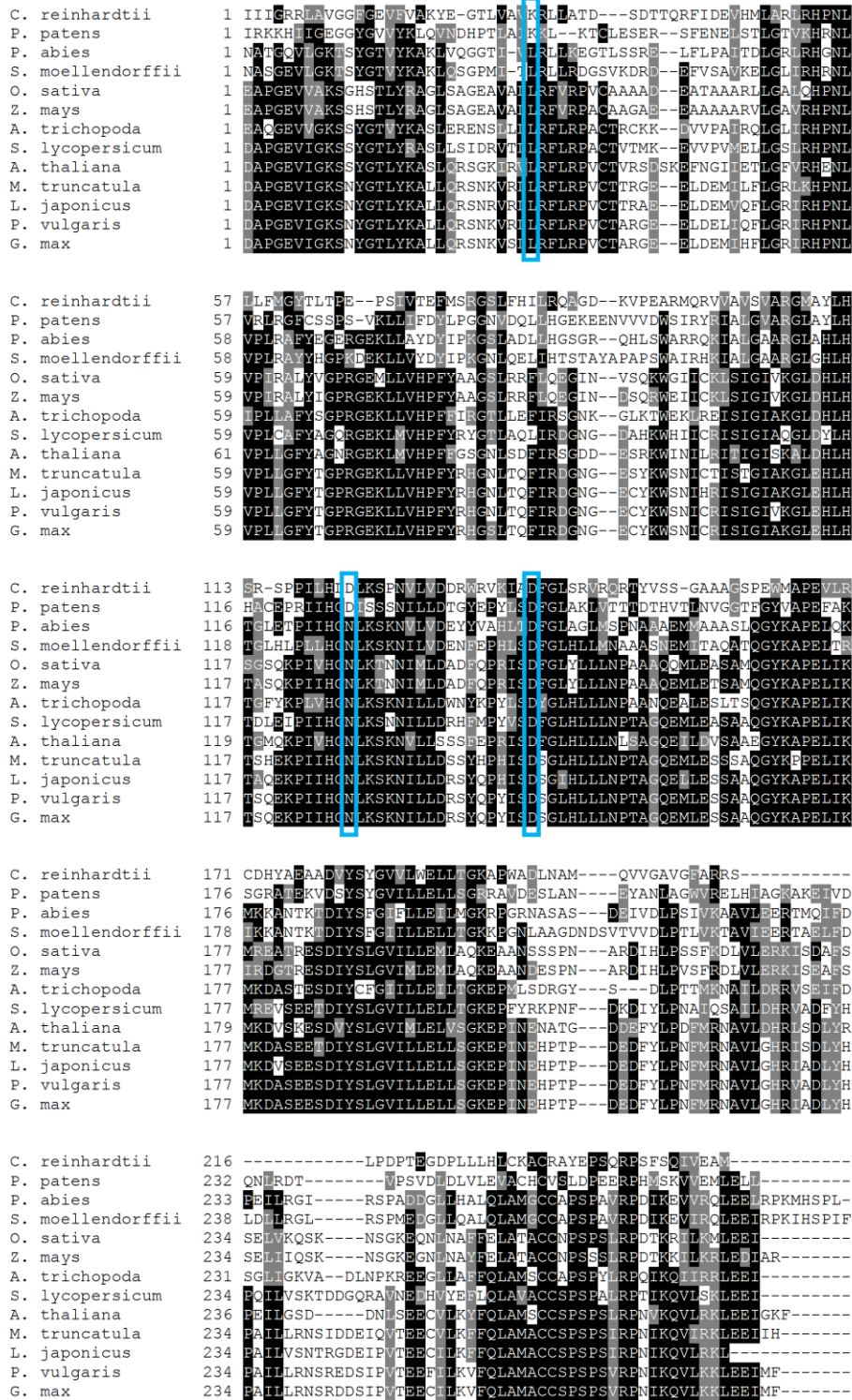
## Material suplementario



**Figura S1. Dominio quinasa de NIPK.**

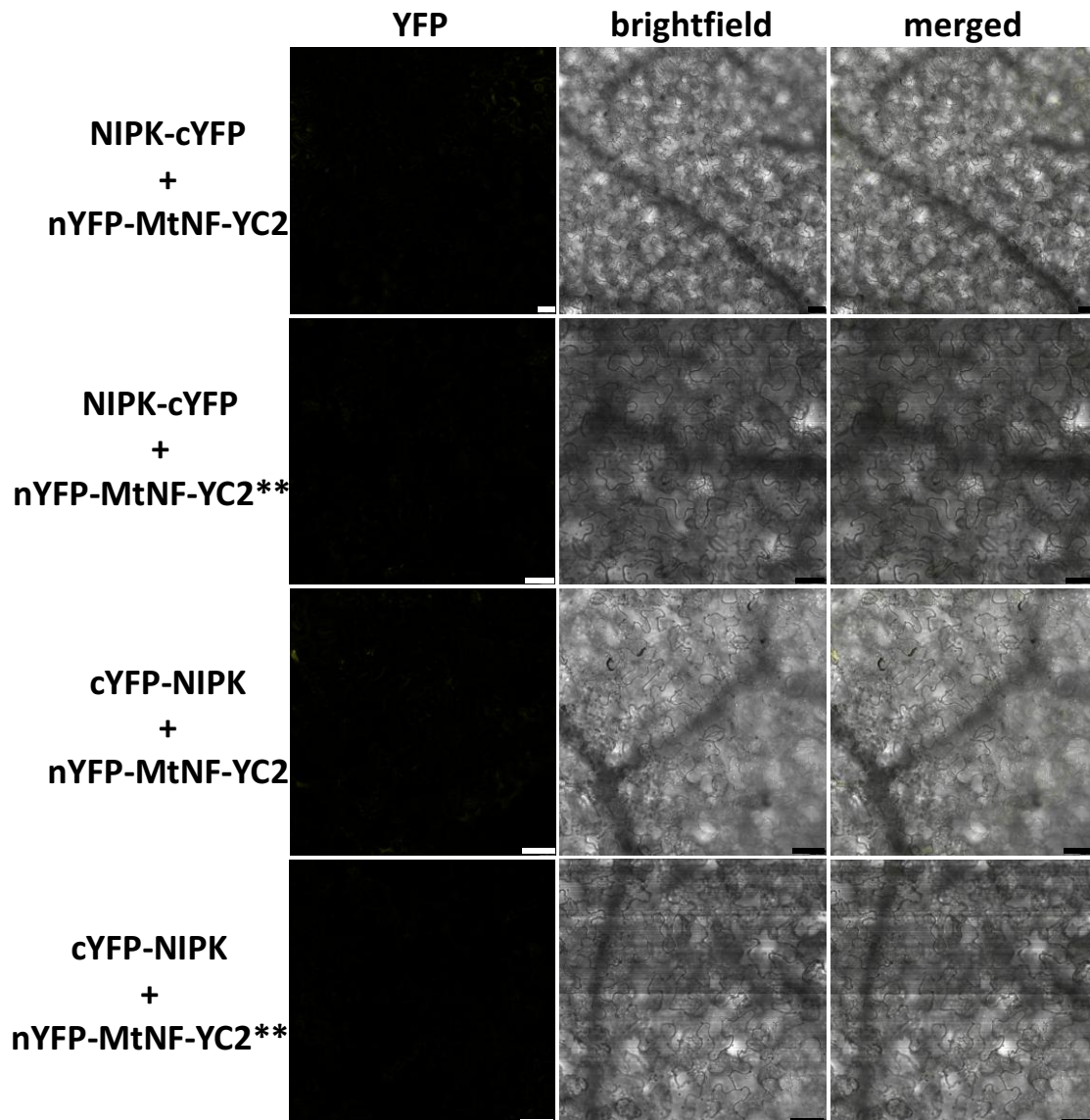
Alineamiento múltiple de la secuencia proteica del dominio quinasa de NIPK, cAPK-alfa, DMI3 (Gleason et al., 2006), SYMRK (Yoshida and Parniske 2004), LYK3 (Jayaraman et al., 2017), SIP2 (Chen et al., 2012) y FLS2 (Lu et al., 2009; Sun et al., 2013). Los rectángulos celestes indican la posición de los 11 motivos típicos del dominio quinasa, y sobre ellos, los residuos canónicos del motivo. Los círculos indican la posición de los tres aminoácidos requeridos para la actividad catalítica de las proteínas quinasa. El color verde y rojo de los círculos indica si el aminoácido se encuentra conservado o no en NIPK, respectivamente, en base a la secuencia considerada canónica.

Material suplementario



**Figura S2. Alineamiento múltiple de NIPK.**

Alineamiento múltiple de las secuencias proteicas de los dominios quinasa de PvNIPK y los ortólogos putativos en las especies vegetales *Medicago truncatula*, *Arabidopsis thaliana*, *Lotus japonicus*, *Glycine max*, *Amborella trichopoda*, *Solanum lycopersicum*, *Zea mays*, *Oryza sativa*, *Selaginella moellendorffii*, *Picea abies*, *Physcomitrella patens* and *Chlamydomonas reinhardtii*. El alineamiento fue generado mediante la herramienta ClustalOmega. Los recuadros celestes indican las posiciones donde deberían encontrarse los aminoácidos requeridos para la actividad catalítica.



**Figura S3. Ensayo de complementación bimolecular de fluorescencia.**

Hojas de *N. benthamiana* fueron agroinfiltradas con las construcciones indicadas en el margen izquierdo. En la figura se muestran imágenes representativas de microscopía confocal láser correspondientes a la señal de la YFP en los paneles izquierdos, al campo claro en los paneles del medio y la superposición de ambos canales en los paneles derechos (merged). Barra de escala: 50  $\mu$ m.

# **BIBLIOGRAFÍA**

- Abe, Y., Matsumoto, S., Wei, S., Nezu, K., Miyoshi, A., Kito, K., Ueda, N., Shigemoto, K., Hitsumoto, Y., Nikawa, J., and Enomoto, Y.** (2001). Cloning and Characterization of a p53-related Protein Kinase Expressed in Interleukin-2-activated Cytotoxic T-cells, Epithelial Tumor Cell Lines, and the Testes. *J. Biol. Chem.* **276**: 44003–44011.
- Aguilar, O., Riva, O., and Peltzer, E.** (2004). Analysis of *Rhizobium etli* and of its symbiosis with wild *Phaseolus vulgaris* supports coevolution in centers of host diversification. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 13548–13553.
- Aguilar, O.M., López, M. V, and Riccillo, P.M.** (2001). The diversity of rhizobia nodulating beans in Northwest Argentina as a source of more efficient inoculant strains. *J. Biotechnol.* **91**: 181–8.
- Aguilar, O.M., López, M.V., Riccillo, P.M., González, R.A., Pagano, M., Grasso, D.H., Pühler, A., Favelukes, G., and Lo, N.** (1998). Prevalence of the *Rhizobium etli* -Like Allele in Genes Coding for 16S rRNA among the Indigenous Rhizobial Populations Found Associated with Wild Beans from the Southern Andes in Argentina. *Appl Environ Microbiol.* **64**: 3520-3524
- Andrews S.** (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Ane, J., Kiss G., Riely B., Penmetsa R., Oldroyd G.E., et. al.** (2004). *Medicago truncatula* DMI1 Required for Bacterial and Fungal Symbioses in Legumes. *Science.* **303**: 1364-1367.
- Arrighi J., Barre A., Ben Amor B., Bersoult A., Soriano L., et. al.** (2006). The *Medicago truncatula* Lysine Motif-Receptor-Like Kinase Gene Family Includes NFP and New Nodule-Expressed Genes. *Plant Physiol.* **142**: 265-279.
- Battaglia, M., Rípodas, C., Clúa, J., Baudin, M., Aguilar, O.M., Niebel, A., Zanetti, M.E., and Blanco, F.A.** (2014). A Nuclear Factor Y Interacting Protein of the GRAS Family Is Required for Nodule Organogenesis, Infection Thread Progression, and Lateral Root Growth. *Plant Physiol.* **164**: 1430–1442.
- Baudin, M., Laloum, T., Lepage, A., Ripodas, C., Ariel, F., Frances, L., Crespi, M., Gamas, P.C., Blanco, F.A., Zanetti, M.E., de Carvalho-Niebel, F., and Niebel, A.** (2015). A phylogenetically conserved group of NF-Y transcription factors interact to control nodulation in legumes. *Plant Physiol.* **169**: 2761-2773.
- Bensmihen, S.** (2015). Hormonal Control of Lateral Root and Nodule Development in Legumes. *Plants.* **4**: 523–547.
- Bhawna, Bonthala, V.S., and Gajula, M.P.** (2016). PvTFDB: a *Phaseolus vulgaris* transcription factors database for expediting functional genomics in legumes. *Database (Oxford).* **2016**: 1–6.
- Bitocchi E., Bellucci E., Giardini A., Rau D., Rodriguez M., Biagetti E., Santilocchi R., Spagnoletti Zeuli P., Gioia T., Logozzo G., Attene G., Nanni L., and Papa R.** (2013). Molecular analysis of the parallel domestication of the common bean ( *Phaseolus vulgaris* ) in Mesoamerica and the Andes. *New Phytol.* **197**: 300–313.
- Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P.S., Logozzo, G., Stougaard, J., McClean, P., Attene, G., and Papa, R.** (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *Proc. Natl. Acad. Sci. U. S. A.* **109**: 788-796.
- Bitocchi, E., Rau, D., Bellucci, E., Rodriguez, M., Murgia, M.L., Santo, D., Nanni, L., Attene, G., and Papa, R.** (2017). Beans (*Phaseolus* spp.) as a Model for Understanding Crop Evolution. *Front. Plant Sci.* **8**:722.
- Blanco, F.A., Meschini, E.P., Zanetti, M.E., and Aguilar, O.M.** (2009). A small GTPase of the Rab family is required for root hair formation and preinfection stages of the common bean-*Rhizobium* symbiotic association. *Plant Cell* **21**: 2797–2810.



- Blankenberg, D. and Hillman-Jackson, J.** (2014). Analysis of Next-Generation Sequencing Data Using Galaxy. *Methods Mol Biol.* **1150**: 21–43.
- Blankenberg, D., Kuster, G. Von, Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J.** (2010). Galaxy: A Web-Based Genome Analysis Tool for Experimentalists. In *Current Protocols in Molecular Biology* (John Wiley & Sons, Inc.: Hoboken, NJ, USA), p. Unit 19.10.1-21.
- Boudeau, J., Miranda-Saavedra, D., Barton, G.J., and Alessi, D.R.** (2006). Emerging roles of pseudokinases. *Trends Cell Biol.* **16**: 443–52.
- Brewin, N.J. and Kardailsky, I.V.** (1997). Legume lectins and nodulation by Rhizobium. *Trends Plant Sci.* **2**: 92–98.
- Broughton, W.J., Hernandez, G., Blair, M., Beebe, S., Gepts, P., and Vanderleyden, J.** (2003). Beans (*Phaseolus* spp.)-model food legumes. *Plant Soil* **252**: 55–128.
- Burghardt, L.T., Guhlin, J., Chun, C.L., Liu, J., Sadowsky, M.J., Stupar, R.M., Young, N.D., and Tiffin, P.** (2017). Transcriptomic basis of genome by genome variation in a legume-rhizobia mutualism. *Mol. Ecol.* **26**: 6122–6135.
- Capoen, W., Sun, J., Wysham, D., Otegui, M.S., Venkateshwaran, M., Hirsch, S., Miwa, H., Downie, J.A., Morris, R.J., Ané, J.-M., and Oldroyd, G.E.D.** (2011). Nuclear membranes control symbiotic calcium signaling of legumes. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 14348–53.
- Cárdenas, L. and Quinto, C.** (2008). Reactive oxygen species (ROS) as early signals in root hair cells responding to rhizobial nodulation factors. *Plant Signal. Behav.* **3**: 1101–2.
- Castells, E. and Casacuberta, J.M.** (2007). Signalling through kinase-defective domains: The prevalence of atypical receptor-like kinases in plants. *J. Exp. Bot.* **58**: 3503–3511.
- Cerri, M.R., Frances, L., Laloum, T., Auriac, M.-C., Niebel, A., Oldroyd, G.E.D., Barker, D.G., Fournier, J., and de Carvalho-Niebel, F.** (2012). Medicago truncatula ERN Transcription Factors: Regulatory Interplay with NSP1/NSP2 GRAS Factors and Expression Dynamics throughout Rhizobial Infection. *Plant Physiol.* **160**: 2155-72.
- Chae, H.B., Yun, J., Bang, Y.J., and Shin, D.Y.** (2004). Cdk2-dependent phosphorylation of the NF-Y transcription factor is essential for the expression of the cell cycle-regulatory genes and cell cycle G1/S and G2/M transitions. *Oncogene* **23**: 4084–4088.
- Chan, Q.K.Y., Lam, H.M., Ng, C.F., Lee, A.Y.Y., Chan, E.S.Y., Ng, H.K., Ho, S.M., and Lau, K.M.** (2010). Activation of GPR30 inhibits the growth of prostate cancer cells through sustained activation of Erk1/2, c-jun/c-fos-dependent upregulation of p21, and induction of G2 cell-cycle arrest. *Cell Death Differ.* **17**: 1511–1523.
- Charpentier, M., Bredemeier, R., Wanner, G., Takeda, N., Schleiff, E., and Parniske, M.** (2008). Lotus japonicus CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis. *Plant Cell* **20**: 3467–79.
- Charpentier, M., Sun, J., Vaz Martins, T., Radhakrishnan, G. V, Findlay, K., Soumpourou, E., Thouin, J., Véry, A.-A., Sanders, D., Morris, R.J., and Oldroyd, G.E.D.** (2016). Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science* **352**: 1102–5.
- Chen, T., Zhu, H., Ke, D., Cai, K., Wang, C., Gou, H., Hong, Z., and Zhang, Z.** (2012). A MAP kinase kinase interacts with SymRK and regulates nodule organogenesis in Lotus japonicus. *Plant Cell* **24**: 823–38.
- Chevalier, D., Batoux, M., Fulton, L., Pfister, K., Yadav, R.K., Schellenberg, M., and Schneitz, K.** (2005). STRUBBELIG defines a receptor kinase-mediated signaling pathway regulating organ development in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 9074–9079.
- Combier, J.P., Frugier, F., de Billy, F., Boualem, A., El-Yahyaoui, F., Moreau, S., Vernié, T., Ott,**

- T., Gamas, P., Crespi, M., and Niebel, A.** (2006). MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev.* **20**: 3084–8.
- Cui, H., Tsuda, K., and Parker, J.E.** (2015). Effector-Triggered Immunity: From Pathogen Perception to Robust Defense. *Annu. Rev. Plant Biol.* **66**: 487–511.
- Currie, R.A.** (1998). NF-Y is associated with the histone acetyltransferases GCN5 and P/CAF. *J. Biol. Chem.* **273**: 1430–4.
- Dalla Via, V., Narduzzi, C., Aguilar, O.M., Zanetti, M.E., and Blanco, F.A.** (2015). Changes in the Common Bean Transcriptome in Response to Secreted and Surface Signal Molecules of *Rhizobium etli*. *Plant Physiol.* **169**: 1356–1370.
- Damiani, I., Pauly, N., Puppo, A., Brouquisse, R., and Boscari, A.** (2016). Reactive Oxygen Species and Nitric Oxide Control Early Steps of the Legume – Rhizobium Symbiotic Interaction. *Front. Plant Sci.* **7**: 1–8.
- Delgado-Salinas, A., Bibler, R., and Lavin, M.** (2006). Phylogeny of the Genus *Phaseolus* (Leguminosae): A Recent Diversification in an Ancient Landscape. *Syst. Bot.* **31**: 779–791.
- Denancé, N., Szurek, B., and Noël, L.D.** (2014). Emerging functions of nodulin-like proteins in non-nodulating plant species. *Plant Cell Physiol.* **55**: 469–474.
- Díaz, C.L., Melchers, L.S., Hooykaas, P.J.J., Lugtenberg, B.J.J., and Kijne, J.W.** (1989). Root lectin as a determinant of host–plant specificity in the Rhizobium–legume symbiosis. *Nature* **338**: 579–581.
- Dolfini, D., Gatta, R., and Mantovani, R.** (2012). NF-Y and the transcriptional activation of CCAAT promoters. *Crit. Rev. Biochem. Mol. Biol.* **47**: 29–49.
- Duhamel, M. and Vandenkoornhuysse, P.** (2013). Sustainable agriculture: possible trajectories from mutualistic symbiosis and plant neodomestication. *Trends Plant Sci.* **18**: 597–600.
- Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kaló, P., and Kiss, G.B.** (2002). A receptor kinase gene regulating symbiotic nodule development. *Nature* **417**: 962–966.
- Erismán, J.W., Galloway, J.N., Seitzinger, S., Bleeker, A., Dise, N.B., Petrescu, a M.R., Leach, A.M., and de Vries, W.** (2013). Consequences of human modification of the global nitrogen cycle. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **368**: 20130116.
- Erismán, J.W., Sutton, M.A., Galloway, J., Klimont, Z., and Winiwarter, W.** (2008). How a century of ammonia synthesis changed the world. *Nat. Geosci.* **1**: 636–639.
- Estrada-Navarrete, G., Alvarado-Affantranger, X., Olivares, J.E., Díaz-Camino, C., Santana, O., Murillo, E., Guillén, G., Sánchez-Guevara, N., Acosta, J., Quinto, C., Li, D., Gresshoff, P.M., Sánchez, F.** (2006). *Agrobacterium rhizogenes* Transformation of the *Phaseolus* spp.: A Tool for Functional Genomics. **19**: 1385–1393.
- Ferguson, B.J., Indrasumunar, A., Hayashi, S., Lin, M.-H., Lin, Y.-H., Reid, D.E., and Gresshoff, P.M.** (2010). Molecular analysis of legume nodule development and autoregulation. *J. Integr. Plant Biol.* **52**: 61–76.
- Fonouni-Farde, C., Diet, A., and Frugier, F.** (2016a). Root Development and Endosymbioses: DELLAs Lead the Orchestra. *Trends Plant Sci.* **21**: 898–900.
- Fonouni-Farde, C., Tan, S., Baudin, M., Brault, M., Wen, J., Mysore, K.S., Niebel, A., Frugier, F., and Diet, A.** (2016b). DELLA-mediated gibberellin signalling regulates Nod factor signalling and rhizobial infection. *Nat. Commun.* **7**: 12636.
- Fowler, D., Pyle, J. a, Raven, J. a, Sutton, M. a, and B, P.T.R.S.** (2013). The global nitrogen cycle in the twenty-first century. *Philos Trans R Soc Lond B Biol Sci.* **368**: 20130164.
- Foyer, C.H., Lam, H.M., Nguyen, H.T., Siddique, K.H., Varshney, R.K., Colmer, T.D., Cowling, W., Bramley, H., Mori, T.A., Hodgson, J.M., Cooper, J.W., Miller, A.J., Kunert, K., Vorster, J. Cullis, C. Ozga, J.A., Wahlqvist, M.L., Liang, Y., Shou, H., Shi, K., Yu, J., Fodor,**

- N., Kaiser, B.N., Wong, F.L., Valliyodan, B., Considine, M.J.** (2016). Neglecting legumes has compromised human health and sustainable food production. *Nat. Plants*. **2**: 1–10.
- Gage, D.J. and Margolin, W.** (2000). Hanging by a thread: Invasion of legume plants by rhizobia. *Curr. Opin. Microbiol.* **3**: 613–617.
- Galloway, J.N., Leach, A.M., Bleeker, A., and Erisman, J.W.** (2013). A chronology of human understanding of the nitrogen cycle. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **368**: 20130120.
- Galloway, J.N., Townsend, A.R., Erisman, J.W., Bekunda, M., Cai, Z., Freney, J.R., Martinelli, L. a, Seitzinger, S.P., and Sutton, M. a** (2008). Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science* **320**: 889–892.
- Gans, J., Wolinsky, M., and Dunbar, J.** (2005). Computational Improvements Reveal Great Bacterial Diversity and High Metal Toxicity in Soil. *Science* **309**: 1387–1390.
- Gepts, P., Osborn, T.C., Rashka, K., and Bliss, F.A.** (1986). Phaseolin-protein Variability in Wild Forms and Landraces of the Common Bean (*Phaseolus vulgaris*): Evidence for Multiple Centers of Domestication. *Econ. Bot.* **40**: 451–468.
- Gerland, P. et al.** (2014). World population stabilization unlikely this century. *Science* **346**: 234–7.
- Geurts, R., Lillo, A., and Bisseling, T.** (2012). Exploiting an ancient signalling machinery to enjoy a nitrogen fixing symbiosis. *Curr. Opin. Plant Biol.* **15**: 438–443.
- Gleason, C., Chaudhuri, S., Yang, T., Mun, A., Muñoz, A., Poovaiah, B.W., and Oldroyd, G.E.D.** (2006a). Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature* **441**: 1149–1152.
- Gleason, C., Chaudhuri, S., Yang, T., Muñoz, A., Poovaiah, B.W., and Oldroyd, G.E.D.** (2006b). Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature* **441**: 1149–1152.
- Gnesutta, N., Kumimoto, R.W., Swain, S., and Chiara, M.** (2017). CONSTANS imparts DNA sequence-specificity to the histone-fold NF-YB / NF-YC dimer. *Plant Cell* **29**: 1516-1532.
- Gonzalez-Rizzo, S., Crespi, M., and Frugier, F.** (2006). The *Medicago truncatula* CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with *Sinorhizobium meliloti*. *Plant Cell* **18**: 2680–93.
- Graham, P.H. and Vance, C.P.** (2003). Update on Legume Utilization Legumes : Importance and Constraints to Greater Use. *Plant Physiol.* **131**: 872–877.
- Groth, M., Takeda, N., Perry, J., Uchida, H., Dräxl, S., Brachmann, A., Sato, S., Tabata, S., Kawaguchi, M., Wang, T.L., and Parniske, M.** (2010). NENA, a *Lotus japonicus* homolog of Sec13, is required for rhizodermal infection by arbuscular mycorrhiza fungi and rhizobia but dispensable for cortical endosymbiotic development. *Plant Cell* **22**: 2509–26.
- Gruber, N. and Galloway, J.N.** (2008). An Earth-system perspective of the global nitrogen cycle. *Nature* **451**: 293–296.
- Guinel, F.C.** (2015). Ethylene, a Hormone at the Center-Stage of Nodulation. *Front. Plant Sci.* **6**: 1121.
- Halter, T., Imkampe, J., Mazzotta, S., Wierzba, M., Postel, S., Bücherl, C., Kiefer, C., Stahl, M., Chinchilla, D., Wang, X., Nürnberger, T., Zipfel, C., Clouse, S., Borst, J.W., Boeren, S., de Vries, S.C., Tax, F., Kemmerling, B.** (2014). The leucine-rich repeat receptor kinase BIR2 is a negative regulator of BAK1 in plant immunity. *Curr. Biol.* **24**: 134–143.
- Hanks, S.K.** (2003). Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. *Genome Biol.* **4**: 111.
- Hanks, S.K., Quinn, A.M., and Hunter, T.** (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**: 42–52.

- Harrison, M.J.** (2005). Signaling in the Arbuscular Mycorrhizal Symbiosis. *Annu. Rev. Microbiol.* **59**: 19–42.
- Hayashi, S., Gresshoff, P.M., and Ferguson, B.J.** (2014). Mechanistic action of gibberellins in legume nodulation. *J. Integr. Plant Biol.* **56**:971–978.
- Heath, K.D., Burke, P. V., and Stinchcombe, J.R.** (2012). Coevolutionary genetic variation in the legume-rhizobium transcriptome. *Mol. Ecol.* **21**: 4735–4747.
- Heckmann, A.B., Lombardo, F., Miwa, H., Perry, J. a, Bunnewell, S., Parniske, M., Wang, T.L., and Downie, J.A.** (2006). Lotus japonicus nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume. *Plant Physiol.* **142**: 1739–50.
- Hirsch, A.M., Bhuvaneshwari, T. V., Torrey, J.G., and Bisseling, T.** (1989). Early nodulin genes are induced in alfalfa root outgrowths elicited by auxin transport inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **86**: 1244-1248.
- Hirsch, S., Kim, J., Munoz, A., Heckmann, A.B., Downie, J.A., and Oldroyd, G.E.D.** (2009). GRAS Proteins Form a DNA Binding Complex to Induce Gene Expression during Nodulation Signaling in *Medicago truncatula*. *Plant Cell* **21**: 545–557.
- Hirsch, S. and Oldroyd, G.E.D.** (2009). GRAS-domain transcription factors that regulate plant development. *Plant Signal Behav.* **4**: 698–700.
- Horváth, B., Yeun, L.H., Domonkos, A., Halász, G., Gobbato, E., Ayaydin, F., Miró, K., Hirsch, S., Sun, J., Tadege, M., Ratet, P., Mysore, K.S., Ané, J.M., Oldroyd, G.E., Kaló, P.** (2011). *Medicago truncatula* IPD3 Is a Member of the Common Symbiotic Signaling Pathway Required for Rhizobial and Mycorrhizal Symbioses. *Mol. Plant-Microbe Interact.* **24**: 1345–1358.
- Jayaraman, D., Richards, A.L., Westphall, M.S., Coon, J.J., and Ané, J.-M.** (2017). Identification of the phosphorylation targets of symbiotic receptor-like kinases using a high-throughput multiplexed assay for kinase specificity. *Plant J.* **90**: 1196–1207.
- Jin, Y., Liu, H., Luo, D., Yu, N., Dong, W., Wang, C., Zhang, X., Dai, H., Yang, J., and Wang, E.** (2016). DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. *Nat. Commun.* **7**: 1–14.
- Jones, K.M., Kobayashi, H., Davies, B.W., Taga, M.E., and Walker, G.C.** (2007). How rhizobial symbionts invade plants: the Sinorhizobium–*Medicago* model. *Nat. Rev. Microbiol.* **5**: 619.
- Kaffman, A., Herskowitz, I., Tjian, R., and O’Shea, E.K.** (1994). Phosphorylation of the transcription factor PHO4 by a cyclin-CDK complex, PHO80-PHO85. *Science* **263**: 1153–6.
- Kahle, J., Baake, M., Doenecke, D., and Albig, W.** (2005). Subunits of the Heterotrimeric Transcription Factor NF-Y Are Imported into the Nucleus by Distinct Pathways Involving Importin and Importin 13. *Mol. Cell. Biol.* **25**: 5339–5354.
- Kaló, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R.M., Hirsch, S., Jakab, J., Sims, S., Long, S.R., Rogers, J., Kiss, G.B., Downie, J.A., Oldroyd, G.E.** (2005). Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science* **308**: 1786–9.
- Kanamori, N., Madsen, L.H., Radutoiu, S., Frantescu, M., Quistgaard, E.M., Miwa, H., Downie, J.A., James, E.K., Felle, H.H., Haaning, L.L., Jensen, T.H., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., Stougaard, J.** (2006). A nucleoporin is required for induction of Ca<sup>2+</sup> spiking in legume nodule development and essential for rhizobial and fungal symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 359–364.
- Keller, J., Imperial, J., Ruiz-Argüeso, T., Privet, K., Lima, O., Michon-Coudouel, S., Biget, M., Salmon, A., Ainouche, A., and Cabello-Hurtado, F.** (2018). RNA sequencing and analysis

- of three *Lupinus nodulosus* provide new insights into specific host-symbiont relationships with compatible and incompatible *Bradyrhizobium* strains. *Plant Sci.* **266**: 102–116.
- Kevei, Z., Lougnon, G., Mergaert, P., Horváth, G.V., Kereszt, A., Jayaraman, D., Zaman, N., Marcel, F., Regulski, K., Kiss, G.B., Kondorosi, A., Endre, G., Kondorosi, E., Ané, JM.** (2007). 3-hydroxy-3-methylglutaryl coenzyme a reductase 1 interacts with NORK and is crucial for nodulation in *Medicago truncatula*. *Plant Cell* **19**: 3974–89.
- Kim, D., Perteza, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L.** (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**: R36.
- Kistner, C. and Parniske, M.** (2002). Evolution of signal transduction in intracellular symbiosis. *Trends Plant Sci.* **7**: 511–518.
- Kouchi, H., Shimomura, K., Hata, S., Hirota, A., Wu, G.J., Kumagai, H., Tajima, S., Suganuma, N., Suzuki, A., Aoki, T., Hayashi, M., Yokoyama, T., Ohyama, T., Asamizu, E., Kuwata, C., Shibata, D., Tabata, S.** (2004). Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*. *DNA Res.* **11**: 263–274.
- Kumar, D., Kumar, R., Baek, D., Hyun, T.-K., Chung, W.S., Yun, D.-J., and Kim, J.-Y.** (2016). An *Arabidopsis thaliana* RECEPTOR DEAD KINASE1 (RDK1) gene functions as a positive regulator in plant responses to ABA. *Mol. Plant* **10**: 223–243.
- Laloum, T., Baudin, M., Frances, L., Lepage, A., Billault-Penneteau, B., Cerri, M.R., Ariel, F., Jardinaud, M.F., Gamas, P., de Carvalho-Niebel, F., and Niebel, A.** (2014). Two CCAAT-box-binding transcription factors redundantly regulate early steps of the legume-rhizobia endosymbiosis. *Plant J.* **79**: 757–768.
- Laloum, T., De Mita, S., Gamas, P., Baudin, M., and Niebel, A.** (2013). CCAAT-box binding transcription factors in plants: Y so many? *Trends Plant Sci.* **18**: 157–66.
- Laporte, P., Lepage, A., Fournier, J., Catrice, O., Moreau, S., Jardinaud, M.-F., Mun, J.-H., Larrainzar, E., Cook, D.R., Gamas, P., and Niebel, A.** (2013). The CCAAT box-binding transcription factor NF-YA1 controls rhizobial infection. *J. Exp. Bot.* **61**: 481-494.
- Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.P., Ané, J.M., Lauber, E., Bisseling, T., Dénarié, J., Rosenberg, C., Debelle, F.** (2004). A Putative Ca<sup>2+</sup> and Calmodulin-Dependent Protein Kinase Required for Bacterial and Fungal Symbioses. *Science* **303**: 1361–1364.
- Li, Q., Herrler, M., Landsberger, N., Kaludov, N., Ogryzko, V. V., Nakatani, Y., and Wolffe, A.P.** (1998). *Xenopus* NF-Y pre-sets chromatin to potentiate p300 and acetylation-responsive transcription from the *Xenopus hsp70* promoter *in vivo*. *EMBO J.* **17**: 6300–6315.
- Li, Y., Tian, C.F., Chen, W.F., Wang, L., Sui, X.H., and Chen, W.X.** (2013). High-Resolution Transcriptomic Analyses of *Sinorhizobium* sp. NGR234 Bacteroids in Determinate Nodules of *Vigna unguiculata* and Indeterminate Nodules of *Leucaena leucocephala*. *PLoS One* **8**: e70531.
- Libault, M., Farmer, A., Brechenmacher, L., Drnevich, J., Langley, R.J., Bilgin, D.D., Radwan, O., Neece, D.J., Clough, S.J., May, G.D., and Stacey, G.** (2010a). Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection. *Plant Physiol.* **152**: 541–52.
- Libault, M., Farmer, A., Joshi, T., Takahashi, K., Langley, R.J., Franklin, L.D., He, J., Xu, D., May, G., and Stacey, G.** (2010b). An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. *Plant J.* **63**: 86–99.
- Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and Geurts, R.** (2003). LysM Domain Receptor Kinases Regulating Rhizobial Nod Factor-Induced Infection. *Science*

302: 630–633.

- Llompарт, B., Castells, E., Ríó, A., Roca, R., Ferrando, A., Stiefel, V., Puigdomènech, P., and Casacuberta, J.M.** (2003). The Direct Activation of MIK, a Germinal Center Kinase (GCK)-like Kinase, by MARK, a Maize Atypical Receptor Kinase, Suggests a New Mechanism for Signaling through Kinase-dead Receptors. *J. Biol. Chem.* **278**: 48105–48111.
- Lohar, D.P., Haridas, S., Gantt, J.S., and VandenBosch, K.A.** (2007). A transient decrease in reactive oxygen species in roots leads to root hair deformation in the legume-rhizobia symbiosis. *New Phytol.* **173**: 39–49.
- Lohar, D.P., Sharopova, N., Endre, G., Peñuela, S., Samac, D., Town, C., Silverstein, K. a T., and VandenBosch, K. a** (2006). Transcript analysis of early nodulation events in *Medicago truncatula*. *Plant Physiol.* **140**: 221–234.
- Lu, D., Wu, S., Gao, X., Zhang, Y., Shan, L., and He, P.** (2010). A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 496–501.
- Madsen, E.B., Madsen, L.H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N., and Stougaard, J.** (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* **425**: 637–40.
- Majer, C. and Hochholdinger, F.** (2011). Defining the boundaries: structure and function of LOB domain proteins. *Trends Plant Sci.* **16**: 47–52.
- Mallory, A.C. and Vaucheret, H.** (2006). Functions of microRNAs and related small RNAs in plants. *Nat. Genet.* **38**: 31–36.
- Mamidi, S., Rossi, M., Moghaddam, S.M., Annam, D., Lee, R., Papa, R., and McClean, P.E.** (2013). Demographic factors shaped diversity in the two gene pools of wild common bean *Phaseolus vulgaris* L. *Heredity (Edinb.)* **110**: 267–276.
- Manni, I., Caretti, G., Artuso, S., Gurtner, A., Emiliozzi, V., Sacchi, A., Mantovani, R., and Piaggio, G.** (2008). Posttranslational regulation of NF-YA modulates NF-Y transcriptional activity. *Mol. Biol. Cell* **19**: 5203–13.
- Mantovani, R.** (1999). The molecular biology of the CCAAT-binding factor NF-Y. *Gene* **239**: 15–27.
- Marsh, J.F., Rakocevic, A., Mitra, R.M., Brocard, L., Sun, J., Eschstruth, A., Long, S.R., Schultze, M., Ratet, P., and Oldroyd, G.E.D.** (2007). *Medicago truncatula* NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. *Plant Physiol.* **144**: 324–35.
- Martin, F.M., Uroz, S., and Barker, D.G.** (2017). Ancestral alliances: Plant mutualistic symbioses with fungi and bacteria. *Science* **356**: eaad4501.
- Mathesius, U., Schlaman, H.R.M., Spaink, H.P., of Sautter, C., Rolfe, B.G., and Djordjevic, M.A.** (1998). Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivatives of chitin oligosaccharides. *Plant J.* **14**: 23–34.
- Mazziotta, L., Reynoso, M.A., Aguilar, O.M., Blanco, F.A., and Zanetti, M.E.** (2013). Transcriptional and functional variation of NF-YC1 in genetically diverse accessions of *Phaseolus vulgaris* during the symbiotic association with *Rhizobium etli*. *Plant Biol.* **15**: 808–818.
- McDowell, J.M. and Woffenden, B.J.** (2003). Plant disease resistance genes: Recent insights and potential applications. *Trends Biotechnol.* **21**: 178–183.
- Meschini, E.P., Blanco, F.A., Zanetti, M.E., Beker, M.P., Küster, H., Pühler, A., and Aguilar, O.M.** (2008). Host Genes Involved in Nodulation Preference in Common Bean (*Phaseolus vulgaris*)–*Rhizobium etli* Symbiosis Revealed by Suppressive Subtractive Hybridization. **21**:

459–468.

- Messinese, E., Mun, J.-H., Yeun, L.H., Jayaraman, D., Rougé, P., Barre, A., Lougnon, G., Schornack, S., Bono, J.-J., Cook, D.R., and Ané, J.-M. (2007). A Novel Nuclear Protein Interacts With the Symbiotic DMI3 Calcium- and Calmodulin-Dependent Protein Kinase of *Medicago truncatula*. *Mol. Plant-Microbe Interact.* **20**: 912–921.
- Middleton, P.H., Jakab, J., Penmetsa, R.V., Starker, C.G., Doll, J., Kaló, P., Prabhu, R., Marsh, J.F., Mitra, R.M., Kereszt, A., Dudas, B., VandenBosch, K., Long, S.R., Cook, D.R., Kiss, G.B., Oldroyd, G.E. (2007). An ERF transcription factor in *Medicago truncatula* that is essential for Nod factor signal transduction. *Plant Cell* **19**: 1221–34.
- Min, X., Lee, B.-H., Cobb, M.H., and Goldsmith, E.J. (2004). Crystal Structure of the Kinase Domain of WNK1, a Kinase that Causes a Hereditary Form of Hypertension. *Structure* **12**: 1303–1311.
- Mitra, R.M., Shaw, S.L., and Long, S.R. (2004). Six nonnodulating plant mutants defective for Nod factor-induced transcriptional changes associated with the legume-rhizobia symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 10217–22.
- Montiel, J., Arthikala, M.K., Cárdenas, L., and Quinto, C. (2016). Legume NADPH oxidases have crucial roles at different stages of nodulation. *Int. J. Mol. Sci.* **17**: 1–12.
- Murray, J.D., Karas, B.J., Sato, S., Tabata, S., Amyot, L., and Szczyglowski, K. (2007). A Cytokinin Perception Mutant Colonized by Rhizobium in the Absence of Nodule Organogenesis. *Science* **315**: 101–104.
- Nardini, M., Gnesutta, N., Donati, G., Gatta, R., Forni, C., Fossati, A., Vonrhein, C., Moras, D., Romier, C., Bolognesi, M., and Mantovani, R. (2013). Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. *Cell* **152**: 132–43.
- Nour, S.M., Drouin, P., Amarger, N., Laguerre, G., Sanjuan, J., and Macheret, V. (2001). Classification of rhizobia based on nodC and nifH gene analysis reveals a close phylogenetic relationship among Phaseolus vulgaris symbionts. *Microbiology* **147**: 981–993.
- 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., Hernandez, G., and Vance, C.P. (2014). An RNA-Seq based gene expression atlas of the common bean. *BMC Genomics* **15**: 866.
- Oldroyd, G.E. and Dixon, R. (2014). Biotechnological solutions to the nitrogen problem. *Curr. Opin. Biotechnol.* **26**: 19–24.
- Oldroyd, G.E.D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* **11**: 252–263.
- Oldroyd, G.E.D. and Downie, J.A. (2008). Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annu. Rev. Plant Biol.* **59**: 519–46.
- Oldroyd, G.E.D., Murray, J.D., Poole, P.S., and Downie, J.A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* **45**: 119–44.
- Ovchinnikova, E. et al. (2011). IPD3 Controls the Formation of Nitrogen-Fixing Symbiosomes in Pea and *Medicago* Spp. *Mol. Plant-Microbe Interact.* **24**: 1333–1344.
- Perret, X., Staehelin, C., and Broughton, W.J. (2000). Molecular Basis of Symbiotic Promiscuity. *Microbiol. Mol. Biol. Rev.* **64**: 180–201.
- Petroni, K., Kumimoto, R.W., Gnesutta, N., Calvenzani, V., Fornari, M., Tonelli, C., Holt, B.F., and Mantovani, R. (2012). The promiscuous life of plant NUCLEAR FACTOR Y transcription factors. *Plant Cell* **24**: 4777–92.
- Puppo, A., Pauly, N., Boscari, A., Mandon, K., and Brouquisse, R. (2013). Hydrogen Peroxide and Nitric Oxide: Key Regulators of the Legume—*Rhizobium* and Mycorrhizal Symbioses.

- Antioxid. Redox Signal. **18**: 2202–2219.
- Radutoiu, S., Madsen, L.H., Madsen, E.B., Felle, H.H., Umehara, Y., Grønlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., and Stougaard, J.** (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425**: 585–592.
- Ramu, S.K., Peng, H.M., and Cook, D.R.** (2002). Nod factor induction of reactive oxygen species production is correlated with expression of the early nodulin gene *rip1* in *Medicago truncatula*. *Mol. Plant-Microbe Interact.* **15**: 522–528.
- Reay, D.S., Davidson, E. a., Smith, K. a., Smith, P., Melillo, J.M., Dentener, F., and Crutzen, P.J.** (2012). Global agriculture and nitrous oxide emissions. *Nat. Clim. Chang.* **2**: 410–416.
- Reid, D.E., Liu, H., Kelly, S., Kawaharada, Y., Mun, T., Andersen, S.U., Desbrosses, G.J., and Stougaard, J.** (2018). Dynamics of *Lotus japonicus* ethylene production in response to compatible Nod factor. *Plant Physiol.* **176**: 1764-1772.
- Ren, H. and Gray, W.M.** (2015). SAUR Proteins as Effectors of Hormonal and Environmental Signals in Plant Growth. *Mol. Plant* **8**: 1153–1164.
- Rendón-Anaya, M. et al.** (2017). Genomic history of the origin and domestication of common bean unveils its closest sister species. *Genome Biol.* **18**: 60.
- Reynoso, M.A., Blanco, F.A., Bailey-Serres, J., Crespi, M., and Zanetti, M.E.** (2012). Selective recruitment of mRNAs and miRNAs to polyribosomes in response to rhizobia infection in *Medicago truncatula*. *Plant J.* **73**: 289–301.
- Reynoso, M.A., Blanco, F.A., and Zanetti, M.E.** (2013). Insights into post-transcriptional regulation during legume-rhizobia symbiosis.
- Riera, A., Barbon, M., Noguchi, Y., Reuter, L.M., Schneider, S., and Speck, C.** (2017). From structure to mechanism—understanding initiation of DNA replication. *Genes Dev.* **31**: 1073–1088.
- Rípodas, C., Via, V.D., Aguilar, O.M., Zanetti, M.E., and Blanco, F.A.** (2013). Knock-down of a member of the isoflavone reductase gene family impairs plant growth and nodulation in *Phaseolus vulgaris*. *Plant Physiol. Biochem.* **68**: 81–89.
- Romier, C., Cocchiarella, F., Mantovani, R., and Moras, D.** (2003). The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. *J. Biol. Chem.* **278**: 1336–1345.
- Roux, B. et al.** (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J.* **77**: 817–837.
- Rubio, M.C., James, E.K., Clemente, M.R., Bucciarelli, B., Fedorova, M., Vance, C.P., and Becana, M.** (2004). Localization of superoxide dismutases and hydrogen peroxide in legume root nodules. *Mol. Plant. Microbe. Interact.* **17**: 1294–305.
- Saito, K., Yoshikawa, M., Yano, K., Miwa, H., Uchida, H., Asamizu, E., Sato, S., Tabata, S., Imaizumi-Anraku, H., Umehara, Y., Kouchi, H., Murooka, Y., Szczyglowski, K., Downie, J.A., Parniske, M., Hayashi, M., Kawaguchi, M.** (2007). NUCLEOPORIN85 Is Required for Calcium Spiking, Fungal and Bacterial Symbioses, and Seed Production in *Lotus japonicus*. *Plant Cell* **19**: 610-624.
- Santos, R., Herouart, D., Sigaud, S., Touati, D., and Puppo, a** (2001). Oxidative burst in alfalfa-Sinorhizobium meliloti symbiotic interaction. *Mol. Plant-Microbe Interact.* **14**: 86–89.
- Sauter, M.** (2015). Phytosulfokine peptide signalling. *J. Exp. Bot.* **66**: 5161–5169.
- Schauser, L., Roussis, A., Stiller, J., and Stougaard, J.** (1999). A plant regulator controlling development of symbiotic root nodules. *Nature* **402**: 191–195.
- Schmutz, J., McClean, P.E., Mamidi, S., Wu, G.A., Cannon, S.B., Grimwood, J., Jenkins, J., Shu, S., Song, Q., Chavarro, C., Torres-Torres, M., Geffroy, V., Moghaddam, S.M., Gao D.,**



- Abernathy B., Barry, K., Blair M., et al.** (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* **46**: 707–13.
- Simpson, C., Thomas, C., Findlay, K., Bayer, E., and Maule, A.J.** (2009). An Arabidopsis GPI-anchor plasmodesmal neck protein with callose binding activity and potential to regulate cell-to-cell trafficking. *Plant Cell* **21**: 581–594.
- Singh, S., Katzer, K., Lambert, J., Cerri, M., and Parniske, M.** (2014). CYCLOPS, A DNA-Binding Transcriptional Activator, Orchestrates Symbiotic Root Nodule Development. *Cell Host Microbe* **15**: 139–152.
- Smets, B., Ghillebert, R., De Snijder, P., Binda, M., Swinnen, E., De Virgilio, C., and Winderickx, J.** (2010). Life in the midst of scarcity: adaptations to nutrient availability in *Saccharomyces cerevisiae*. *Curr. Genet.* **56**: 1–32.
- Smil, V.** (1999). Nitrogen in crop production: An account of global flows. *Global Biogeochem. Cycles* **13**: 647–662.
- Smit, P., Raedts, J., Portyanko, V., Debellé, F., Gough, C., Bisseling, T., and Geurts, R.** (2005). NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. *Science* **308**: 1789–91.
- Soyano, T., Kouchi, H., Hirota, A., and Hayashi, M.** (2013). Nodule inception directly targets NF-Y subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. *PLoS Genet.* **9**: e1003352.
- Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Szczyglowski, K., and Parniske, M.** (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* **417**: 959–962.
- Sun, X., Jones, W.T., and Rikkerink, E.H. a** (2012). GRAS proteins: the versatile roles of intrinsically disordered proteins in plant signalling. *Biochem. J.* **442**: 1–12.
- Sun, Y., Li, L., Macho, A.P., Han, Z., Hu, Z., Zipfel, C., Zhou, J.-M., and Chai, J.** (2013). Structural Basis for flg22-Induced Activation of the Arabidopsis FLS2-BAK1 Immune Complex. *Science* **342**: 624–628.
- Suzuki, T., Ito, M., and Kawaguchi, M.** (2013). Genetic basis of cytokinin and auxin functions during root nodule development. *Front. Plant Sci.* **4**: 42.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., and Kumar, S.** (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**: 2725–9.
- Taylor, R.G., Walker, D.C., and McInnes, R.R.** (1993). *E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res.* **21**: 1677–8.
- Tian, C., Wan, P., Sun, S., Li, J., and Chen, M.** (2004). Genome-wide analysis of the GRAS gene family in rice and Arabidopsis. *Plant Mol Biol.* **54**: 519–532.
- Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., and Su, Z.** (2017). AgriGO v2.0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **45**: 122–129.
- Tilman** (2002). Agriculture sustainability and intensive production practices. *Nature* **418**: 671–677.
- Tirichine, L., Sandal, N., Madsen, L.H., Radutoiu, S., Albrechtsen, A.S., Sato, S., Asamizu, E., Tabata, S., and Stougaard, J.** (2007). A Gain-of-Function Mutation in a Cytokinin Receptor Triggers Spontaneous Root Nodule Organogenesis. *Science* **315**: 104–107.
- Torres Acosta, J.A., de Almeida Engler, J., Raes, J., Magyar, Z., De Groodt, R., Inz, D., and De Veylder, L.** (2004). Molecular characterization of Arabidopsis PHO80-like proteins, a novel class of CDKA;1-interacting cyclins. *Cell. Mol. Life Sci.* **61**: 1485–1497.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L.,**

- Rinn, J.L., and Pachter, L.** (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**: 562–578.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L.** (2013). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**: 562–578.
- Trapnell, C., Williams, B. a, Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L.** (2011). Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. *Nat. Biotechnol.* **28**: 511–515.
- Via, V.D., Zanetti, M.E., and Blanco, F.** (2016). How legumes recognize rhizobia. *Plant Signal. Behav.* **11**: e1120396.
- Vitousek, P.M., Cassman, K., Cleveland, C., Crews, T., Field, C.B., Grimm, N.B., Howarth, R.W., Marino, R., Martinelli, L., Rastetter, E.B., and Sprent, J.I.** (2002). Towards an ecological understanding of biological nitrogen fixation. In *The Nitrogen Cycle at Regional to Global Scales* (Springer Netherlands: Dordrecht), pp. 1–45.
- Vlasova, A., Capella-Gutiérrez, S., Rendón-Anaya, M., Hernández-Oñate, M., Minoche, A.E., Erb, I., Câmara, F., Prieto-Barja, P., Corvelo, A., et al.** (2016). Genome and transcriptome analysis of the Mesoamerican common bean and the role of gene duplications in establishing tissue and temporal specialization of genes. *Genome Biol.* **17**: 32.
- Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D.** (2003). Retracted: An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* **33**: 949–956.
- Wang, G., Kong, H., Sun, Y., Zhang, X., Zhang, W., Altman, N., dePamphilis, C.W., and Ma, H.** (2004). Genome-Wide Analysis of the Cyclin Family in Arabidopsis and Comparative Phylogenetic Analysis of Plant Cyclin-Like Proteins. *Plant Physiol.* **135**: 1084–1099.
- Werner, G.D. a, Cornwell, W.K., Sprent, J.I., Kattge, J., and Kiers, E.T.** (2014). A single evolutionary innovation drives the deep evolution of symbiotic N<sub>2</sub>-fixation in angiosperms. *Nat. Commun.* **5**: 4087.
- Xiao, T.T., Schilderink, S., Moling, S., Deinum, E.E., Kondorosi, E., Franssen, H., Kulikova, O., Niebel, A., and Bisseling, T.** (2014). Fate map of Medicago truncatula root nodules. *Development* **141**: 3517–28.
- Xu, B., English, J.M., Wilsbacher, J.L., Stippec, S., Goldsmith, E.J., and Cobb, M.H.** (2000). WNK1, a novel mammalian serine/threonine protein kinase lacking the catalytic lysine in subdomain II. *J. Biol. Chem.* **275**: 16795–801.
- Yano, K. et al.** (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 20540–20545.
- Yoshida, S. and Parniske, M.** (2005). Regulation of Plant Symbiosis Receptor Kinase through Serine and Threonine Phosphorylation. *J. Biol. Chem.* **280**: 9203–9209.
- Yun, J., Chae, H.D., Choi, T.S., Kim, E.H., Bang, Y.J., Chung, J., Choi, K.S., Mantovani, R., and Shin, D.Y.** (2003). Cdk2-dependent phosphorylation of the NF-Y transcription factor and its involvement in the p53-p21 signaling pathway. *J. Biol. Chem.* **278**: 36966–36972.
- Zanetti, M.E., Blanco, F.A., Beker, M.P., Battaglia, M., and Aguilar, O.M.** (2010). A C subunit of the plant nuclear factor NF-Y required for rhizobial infection and nodule development affects partner selection in the common bean-Rhizobium etli symbiosis. *Plant Cell* **22**: 4142–4157.
- Zanetti, M.E., Rípodas, C., and Niebel, A.** (2017). Plant NF-Y transcription factors: key players in plant-microbe interactions, root development and adaptation to stress. *Biochim. Biophys. Acta* **1860**: 645–654.

## *Bibliografia*

**Zumft, W.G. and Mortenson, L.E.** (1975). The nitrogen-fixing complex of bacteria. *Biochim. Biophys. Acta* **416**: 1–52.