



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

Tesis de doctorado en Ciencias Biológicas

**Identificación de actores moleculares involucrados en
la interacción entre β -rizobios y leguminosas
hospederas**

Cecilia Rodríguez Esperón

Orientador:

Dr. Raúl Platero

Co-Orientadores:

Dr. Andrés Iriarte y Dra. Elena Fabiano

Tribunal

Dr. Omar Borsani

Dra. Silvia Batista

Dr. Jorge Monza

Laboratorio de Microbiología Ambiental
Departamento de Bioquímica y Genómica Microbianas
Instituto de Investigaciones Biológicas Clemente Estable
Montevideo, 2024

Agradecimientos

Quiero agradecer a todo BIOGEM del IIBCE, por haberme dado un lugar en el cual crecí como persona y como profesional.

Especialmente a mi orientador Rufo, del cual aprendí mucho y al cual admiro muchísimo, y me mostró siempre el como seguir y no bajar los brazos. A mis co-orientadores Andrés y Elena, por estas siempre a disposición.

Al tribunal por haberse dedicado a la lectura de la tesis, y por sus sugerencias.

Y finalmente a todos mis compañeros del laboratorio y amigos de la vida por estar siempre, y apoyarme durante este proceso increíble, que fue realizar un Doctorado.

Tabla de contenidos

Índice de figuras y tablas.....	10
Resumen	15
INTRODUCCIÓN	18
1 El nitrógeno y su importancia biológica	18
1.1 Fijación biológica del nitrógeno.....	20
1.1.1 Complejo enzimático nitrogenasa	21
1.2 Los rizobios	24
1.3 Las leguminosas	28
1.4 Características del género <i>Cupriavidus</i>	29
1.4.1 Homeostasis de metales.....	29
1.4.2 Degradación de compuestos aromáticos.....	32
1.4.3 Producción de PHB.....	33
2 Simbiosis rizobio–leguminosa.....	34
2.1 Genes simbióticos.....	38
2.2 Especificidad mediada por flavonoides.....	41
2.3 Especificidad mediada por la percepción de factores Nod	45
2.4 Especificidad mediada por exopolisacáridos bacterianos.....	46
2.5 Especificidad mediada por la inmunidad innata del huésped	47
3 Antecedentes del grupo de trabajo	48
3.1 Modelo de Estudio: La cepa UYPR 2.512 de <i>Cupriavidus necator</i>	48
3.2 Hipótesis de trabajo	48
3.3 Objetivo general y específicos	49
CAPÍTULO I	50
Descifrando el genoma de la cepa UYPR2.512 de <i>Cupriavidus</i> sp.....	50
Objetivo general	50

Objetivos específicos	50
Resumen	50
1 Genómica. Tecnologías de secuenciación y NGS.....	51
1.1 Secuenciación por semiconducción mediante Ion Torrent	54
1.2 Secuenciación por síntesis (SBS) mediante Illumina	56
1.3 Secuenciación SMRT-seq de PacBio	57
1.4 Estructura de los genomas de rizobios	58
1.4.1 Ubicación y contenido de las islas simbióticas	60
1.4.2 Genes simbióticos en β -rizobios	61
1.5 Genes relacionados con la interacción planta-microorganismo.....	63
1.5.1 Quimiotaxis y movilidad del flagelo	63
1.5.2 Sistemas de secreción	63
2 Materiales y métodos	67
2.1 Clasificación de la cepa de UYPR2.512 de <i>Cupriavidus</i> sp., en base a la información genómica disponible.....	67
2.2 Obtención del ADN genómico y secuenciación del genoma de UYPR2.512 69	
2.2.1 Ensamblaje y anotación del genoma de UYPR2.512 de <i>C. necator</i>	70
2.2.1.1 Búsqueda de genes característicos del género <i>Cupriavidus</i>	70
2.2.1.2 Búsqueda de genes relacionados con la interacción planta-bacteria ...	71
3 Resultados	71
3.1 Clasificación taxonómica de la cepa UYPR2.512	71
3.2 Descripción del nuevo genoma	74
3.2.1 Mejoras de la secuencia del genoma respecto a la versión anterior.	78
3.3 Anotación funcional del genoma de UYPR2.512	79
3.4 Genes involucrados en funciones de replicación y mantenimiento de los replicones.....	79
3.4.1 Características genómicas de la cepa UYPR2.512 relacionadas a la especie <i>necator</i>	80

3.4.1.1	Sistemas involucrados en la tolerancia a metales	80
3.4.1.2	Degradación de compuestos aromáticos.....	85
3.4.1.3	Síntesis y degradación de PHB	90
3.4.2	Genes relacionados con la interacción planta-microorganismo.....	91
3.4.2.1	Quimiotaxis y movilidad del flagelo.....	92
3.4.2.2	Genes involucrados en la interacción simbiótica y fijación biológica del nitrógeno	94
3.4.2.3	Sistemas de secreción.....	96
3.4.2.4	Modulación de fitohormonas.....	104
4	Discusión.....	104
4.1	Asignación de especie a la cepa UYPR2.512 de <i>Cupriavidus</i> sp.....	104
4.2	Análisis genómico	105
4.3	Características genómicas.....	106
4.3.1	Tolerancia a metales pesados	106
4.3.2	Degradación de compuestos aromáticos.....	107
4.3.3	Metabolismo de PHB	108
4.4	Características genómicas involucradas en la interacción de UYPR2.512 con sus hospederos.....	109
4.4.1	Quimiotaxis y movilidad bacteriana.....	109
4.4.2	Simbiosis y fijación biológica del nitrógeno.....	110
4.4.3	Sistemas de secreción	111
4.4.4	Síntesis de fitohormonas y su regulación	115
	CAPITULO II	116
	Análisis transcriptómico de la respuesta de la cepa UYPR2.512 de <i>Cupriavidus necator</i> a la presencia de luteolina	116
	Objetivo general	116
	Objetivos específicos	116
	Resumen.....	116
1	Introducción	117

1.1	Transcriptómica. RNA-Seq como aproximación funcional	117
1.2	Análisis de datos de secuencias obtenidas mediante RNA-Seq	118
2	Materiales y métodos	121
2.1	Cepas bacterianas, medios y condiciones de cultivo.....	121
2.2	Obtención de las cepas de UYPR2.512 de <i>C. necator</i> portando pCBM01 o pCZ388	122
2.3	Evaluación de la expresión del pnodB19424-lacZ en UYPR2.512	122
2.4	Extracción y purificación de ARN total (ARNtot) de UYPR2.512	123
2.5	Análisis transcriptómico a la presencia luteolina mediante RNA-Seq.....	124
2.5.1	Secuenciación mediante Ion Torrent	124
2.5.1.1	Enriquecimiento de ARN mensajero (ARNm) o depleción de ARN ribosómico (ARNr).....	124
2.5.1.2	Construcción de librerías de ADN copia (ADNc) a partir de los ARNm obtenidos.....	125
2.5.1.3	Fragmentación del ARNm.....	125
2.5.1.4	Amplificación de ADNc.....	125
2.5.1.5	Secuenciación de las librerías	126
2.5.2	Secuenciación mediante Illumina.....	126
2.5.3	Análisis de las secuencias obtenidas (Ion Torrent / Illumina)	126
2.5.3.1	Análisis de la expresión diferencial de genes	127
2.5.3.2	Análisis de categorías funcionales.....	127
2.5.4	Validación de RNA-Seq mediante q-PCR.....	128
2.5.4.1	Extracción de ARN total y síntesis de ADN copia.....	128
2.5.4.2	Especificidad de los cebadores para la q-PCR.....	128
2.5.4.3	Selección de los genes a evaluar la expresión diferencial	130
3	Resultados	131
3.1	Evaluación de la expresión del gen pnodB19424-lacZ en UYPR2.512	131
3.2	Obtención del ARNtot bacteriano.....	133
3.3	Construcción de librerías de RNA-Seq para Ion Torrent	134
3.3.1	Enriquecimiento del ARNm o depleción de los ARNr	134

3.3.2	Fragmentación del ARNm.....	135
3.3.3	Obtención del ADNc.....	136
3.3.4	Análisis transcriptómico de UYPR2.512 en presencia de luteolina	138
3.3.4.1	Primera aproximación: Secuenciación en Ion Torrent.....	138
3.3.5	Análisis de la expresión diferencial de genes	142
3.3.6	Análisis de enriquecimiento funcional	138
3.3.7	Genes sobre-expresados en presencia de luteolina.....	143
3.3.7.1	Genes pertenecientes al pCne512e (pSym).....	143
3.3.7.2	Genes pertenecientes al chr2	146
3.3.7.3	Genes pertenecientes al chr1	148
3.3.7.4	Genes sub-expresados en presencia de luteolina.....	149
3.4	Validación de RNA-Seq mediante q-PCR.....	154
3.5	Comparativa del análisis transcriptómico de la cepa UYPR2.512 de <i>C. necator</i> en presencia de luteolina mediante NGS	155
4	Discusión.....	156
CAPÍTULO III		166
Identificación de las señales involucradas en el diálogo molecular entre la cepa UYPR2.512 <i>Cupriavidus necator</i> y <i>Mimosa pudica</i>		166
1	Introducción	167
1.1	Metabolómica.....	167
1.2	Técnicas de separación cromatográficas aplicada a la metabolómica	168
1.2.1	Cromatografía líquida de alta presión (HPLC).....	168
1.2.2	Cromatografía de gases (GC).....	169
1.3	Métodos de detección para Cromatografía líquida de alta presión (HPLC) y Cromatografía de gases (GC).....	169
1.4	Métodos de identificación para Cromatografía líquida de alta presión (HPLC-MS) y Cromatografía de gases (GC-MS).....	171
1.5	La metabolómica aplicada al estudio de la simbiosis rizobio-leguminosa .	172
1.5.1	Comunicación en la rizósfera.....	172

2	Materiales y métodos.....	174
2.1	Obtención de exudados radiculares.....	174
2.2	Evaluación de la expresión del gen pnodB19424-lacZ	174
2.3	Búsqueda de moléculas involucradas en la interacción entre UYPR2.512-M. pudica	175
2.3.1	Estrategia I. Búsqueda de flavonoides en exudados radiculares mediante técnicas de HPLC-DAD	175
2.3.1.1	Colecta y obtención de exudados radiculares	175
2.3.1.2	Búsqueda de compuestos flavonoides	175
2.3.1.3	Análisis de los cromatogramas obtenidos	176
2.3.2	Estrategia II-A y II-B. Búsqueda de flavonoides mediante HPLC-ESI-MS/MS	176
2.3.2.1	Extracción de moléculas presentes en los exudados radiculares	176
2.3.2.2	Obtención y fraccionamiento de extractos celulares	177
2.3.2.3	Análisis cromatográficos cualitativo de metabolitos.....	178
2.3.3	Estrategia III. Búsqueda de metabolitos en exudados radiculares mediante GC/MS	179
2.3.3.1	Obtención de exudados radiculares	179
2.3.3.2	Preparación de extractos a partir de los exudados radiculares colectados	179
2.3.3.3	Identificación cualitativa de metabolitos.....	180
2.3.3.4	Estudio del efecto inductor de exudados radiculares de M. pudica en URPR2.512	181
2.4	Estudio del efecto inductor de flavonoides y otros compuestos fenólicos en URPR2.512.....	181
3	Resultados	181
3.1	Los exudados radiculares de M. pudica inducen la expresión del pnodB19424-lacZ en UYPR2.512.....	181
3.2	Búsqueda de moléculas involucradas en la interacción entre UYPR2.512-M. pudica	182

3.2.1	Estrategia I. Búsqueda de flavonoides mediante HPLC-DAD	182
3.2.2	Estrategia II-A. Espectros obtenidos mediante HPLC-UV y HPLC-DAD-ELSD	186
3.2.3	Estrategia II-B. Espectros obtenidos mediante HPLC-ESI-MS/MS	188
3.3	Caracterización de los picos observados.....	189
3.4	Estrategia III. Identificación de metabolitos presentes en exudados radiculares mediante GC/MS	193
3.4.1	La fracción metanólica presenta metabolitos inductores del pnodB19424-lacZ en UYPR2.512	198
3.4.2	Flavonoides y otras moléculas son inductores de los genes nod en cultivos de UYPR2.512	199
4	Discusión.....	200
4.1	Respuesta de los genes nod en UYPR2.512 a exudados radiculares, flavonoides y otros compuestos.....	200
4.2	Compuestos presentes en los exudados de M. pudica	203
	CONCLUSIONES.....	207
	ANEXOS	209
1	Medios de cultivos	209
1.1	Medio de cultivo líquido mínimo M9.....	209
1.2	Medio de cultivo líquido Luria – Bertani (LB)	210
1.3	Medio de cultivo líquido Howelson 10X	210
2	Tablas suplementarias	211
2.1	Capítulo I.....	211
2.2	Capítulo II.....	213
	BIBLIOGRAFÍA	224

INDICE DE FIGURAS

INTRODUCCIÓN

Figura 1. Ilustración del Ciclo del nitrógeno de un ecosistema terrestre.	19
Figura 2. Ilustración del complejo enzimático de la nitrogenasa.....	22
Figura 3. Ilustración de los principales sistemas de transportadores de metales en bacterias.....	30
Figura 4. Ilustración del proceso de infección durante la simbiosis rizobio-leguminosa y morfogénesis de un nódulo indeterminado.	35
Figura 5. Ilustración de la estructura básica de los lipo-chito-oligosacáridos (LCOs), y función de los genes <i>nod</i>	39
Figura 6. Localización de flavonoides en las raíces de <i>Medicago truncatula</i> , mediante microscopía de fluorescencia UV.....	42
Figura 7. Ilustración de algunos de los esqueletos de flavonoides más estudiados y subclases, dependiendo de la modificaciones estructurales y sustituyentes.	43

CAPITULO I

Figura 8. Esquema del proceso de secuenciación por semiconducción en Ion Torrent.	55
Figura 9. Esquema del sistema de secuenciación por síntesis utilizado por Illumina....	56
Figura 10. Secuenciación de molécula única en tiempo real (SMRT).	58
Figura 11. Organización y contexto genómico de los genes <i>nod</i> , <i>nod</i> y <i>noe</i> de diferentes especies de rizobios.....	61
Figura 12. Comparación de islas simbióticas de diferentes β -rizobios.	62
Figura 13. Ilustración resumida de los principales sistemas de secreción bacterianos.	67
Figura 14. Árbol filogenómico del género <i>Cupriavidus</i> y cepas seleccionadas pertenecientes a la familia <i>Burkholderiaceae</i>	72
Figura 15. Relaciones filogenómicas entre cepas y especies de <i>Cupriavidus</i> que contienen rizobios.	73

Figura 16. Representación circular de los 4 replicones cerrados del genoma de UYPR2.512.	77
Figura 17. Representación circular de los <i>contigs</i> ensamblados.	78
Figura 18. Ilustración de las vías periféricas y centrales de degradación de compuestos aromáticos en UYPR2.512.	89
Figura 19. Ilustración de la isla simbiótica de UYPR2.512.	94
Figura 20. Estudio comparativo del operón <i>nod</i> entre UYPR2.512 <i>C. necator</i> y LMG19424 de <i>C. taiwanensis</i>	96
Figura 21. Sistema de secreción del tipo I (SSTI) en UYPR2.512.	97
Figura 22. Sistema de secreción del tipo II (SSTII) en UYPR2.512.	98
Figura 23. Sistema de secreción del tipo III (SSTIII) en UYPR2.512.	99
Figura 24. Sistema de secreción del tipo IV (SSTIV) en UYPR2.512.	100
Figura 25. Sistema de secreción del tipo VI (SSTVI) en UYPR2.512.	101
Figura 26. Imagen tomada mediante microscopía electrónica de transmisión (MET) de un bacteroide de la cepa UYPR2.512 en el interior de un nódulo de <i>M. pudica</i>	108

CAPITULO II

Figura 27. Respuesta de la región promotora del gen <i>nodB</i> a concentraciones decrecientes de luteolina en cultivos de UYPR2.512.	132
Figura 28. Respuesta de la región promotora del gen <i>nodB</i> en presencia de luteolina en cultivos de UYPR2.512.	133
Figura 29. Extracción y purificación del ARN _{tot} de UYPR2.512.	134
Figura 30. Análisis cualitativo mediante Bioanalizador del ARN _{tot} obtenido a partir del cultivo de UYPR2.512 en presencia y en ausencia de luteolina.	135
Figura 31. Cromatogramas de ARNm depletado y purificado de UYPR2.512.	136
Figura 32. Librería de ADN _c amplificada de UYPR2.512.	137
Figura 33. Análisis de componentes principales (PCA).	142
Figura 34. Diagrama de volcán mostrando los genes diferencialmente expresados (DEG) en presencia de luteolina.	143
Figura 35. Comparativa de los niveles de expresión de los genes expresados diferencialmente.	137

Figura 36. Análisis de enriquecimiento funcional de las categorías COG en los genes sobre-expresados y en sub-expresados.	140
Figura 37. Estudio de la expresión de los genes <i>nod</i> en presencia de luteolina.	145
Figura 38. Análisis comparativo de DEG obtenidos mediante la secuenciación por Ion Torrent vs Illumina.	154
Figura 39. Estudio de la expresión de los genes <i>nodS</i> y <i>hdcP</i> mediante RNA-Seq y q-PCR.	155
Figura 40. Modelo para la respuesta transcriptómica de UYPR2.512 a la presencia de luteolina.	165

CAPITULO III

Figura 41. Respuesta de la región promotora del gen <i>nodB</i> a la presencia de exudados radiculares frescos en cultivos de UYPR2.512.	182
Figura 42. Perfil cromatográfico de exudados radiculares de plántulas de <i>M. pudica</i> inoculadas con UYPR2.512.	183
Figura 43. Perfil cromatográfico de exudados radiculares de plántulas de <i>M. pudica</i> inoculadas con UYPR2.512.	184
Figura 44. Perfil cromatográfico de exudados radiculares de plántulas de <i>M. pudica</i>	184
Figura 45. Perfiles cromatográficos de flavonoides utilizados como estándares.	185
Figura 46. Perfiles cromatográficos obtenidos mediante HPLC-DAD-ELSD.	187
Figura 47. Cromatogramas de los extractos de exudados radiculares obtenidos mediante HPLC-ESI-MS/MS.	188
Figura 48. Perfiles obtenidos mediante HPLC-ESI-MS/MS de los extractos de exudados radiculares de <i>M. pudica</i>	189
Figura 49. Red molecular generada a partir de los espectros obtenidos mediante HPLC-ESI-MS/MS.	191
Figura 50. Clúster de interés con nodos anotados como correspondientes a flavonoides según la librería del GNPS.	192
Figura 51. Cromatogramas de los exudados analizados mediante GC-MS.	195
Figura 52. Respuesta de la región promotora del gen <i>nodB</i> a la presencia de exudados radiculares en cultivos de UYPR2.512.	198

Figura 53. Respuesta de la región promotora del gen <i>nodB</i> en presencia de flavonoides y otros polifenoles en cultivos de UYPR2.512.	199
--	-----

INDICE DE TABLAS

INTRODUCCIÓN

Tabla 1. Agrupamiento taxonómico de rizobios simbióticos y no simbióticos relacionados filogenéticamente, pertenecientes a las subclases de α - y β -proteobacterias.....	26
Tabla 2. Resumen de plataformas de secuenciación y sus características más relevantes	53

CAPITULO I

Tabla 3. Cepas del género <i>Cupriavidus</i> utilizadas para el análisis ANI.....	69
Tabla 4. ANI entre los genomas de distintas cepas del género <i>Cupriavidus</i>	74
Tabla 5. Características generales del genoma ensamblado de UYPR2.512	76
Tabla 6. Características principales del nuevo genoma y valores comparativos con la versión anterior.	79
Tabla 7. Presencia de genes relacionados a la replicación, conjugación y mantenimiento en el crómido y plásmidos en UYPR2.512.....	80
Tabla 8. Características y ubicación genómica de genes relacionados con los principales mecanismos implicados en la homeostasis de metales en UYPR2.512	83
Tabla 9. Características y ubicación genómica de genes relacionados con la homeostasis del hierro en UYPR2.512.	85
Tabla 10. Características, función y ubicación genómica de genes relacionados con la degradación de compuestos aromáticos en UYPR2.512.	86
Tabla 11. Características y ubicación genómica de genes relacionados con el metabolismo de PHB en UYPR2.512.	91
Tabla 12. Características y ubicación genómica de genes/proteínas relacionadas con la quimiotaxis y en la movilidad flagelar en UYPR2.512.....	93
Tabla 13. Características y ubicación genómica de genes/proteínas relacionadas con la simbiosis y con la fijación biológica del nitrógeno en UYPR2.512.....	95
Tabla 14. Características y ubicación genómica de genes relacionados con los sistemas de secreción de moléculas en UYPR2.512.....	102

CAPITULO II

Tabla 15. Cepas y plásmidos utilizados.	121
Tabla 16. Secuencias nucleotídicas de cebadores de genes normalizadores, sobre-expresados y sub- expresados, evaluados para la validación del <i>RNA-Seq</i> , mediante q-PCR.	130
Tabla 17. Genes sobre-expresados en presencia de luteolina.	139
Tabla 18. Genes sub-expresados en presencia de luteolina.	140
Tabla 19. Características generales de las secuencias obtenidas.	141
Tabla 20. Distribución de genes sobre-expresados y sub-expresados	138
Tabla 21. Análisis de enriquecimiento funcional de las vías KO en genes sub-expresados	141
Tabla 22. Análisis de enriquecimiento funcional de las vías KO en genes sobre-expresados.....	143
Tabla 23. Genes sobre-expresados en presencia de luteolina.	149
Tabla 24. Genes sub-expresados en presencia de luteolina.	153

CAPITULO III

Tabla 25. Cantidad de metabolitos identificados mediante GC-MS para cada una de las de las fracciones metanólicas analizadas	193
Tabla 26. Metabolitos detectados en exudados radiculares de <i>M. pudica</i> en ausencia y/o en presencia de UYPR2.512.....	196

Resumen

Los rizobios son bacterias del suelo pertenecientes a las subdivisiones alfa y beta de proteobacterias, las cuales se caracterizan por asociarse simbióticamente con plantas de la familia *Fabaceae*. Esta asociación requiere un complejo diálogo molecular entre los simbioses, el cual implica mecanismos de señalización y mutuo reconocimiento, culminando con la formación de nódulos en las raíces de las plantas, donde las bacterias llevarán a cabo la fijación biológica del nitrógeno. Este diálogo ocurre en la rizósfera, donde la planta, a través de sus raíces libera una gran diversidad de moléculas entre las que se encuentran carbohidratos, ácidos orgánicos y compuestos aromáticos. En dicho contexto, los flavonoides como la luteolina son moléculas clave durante los primeros pasos de la simbiosis rizobio-leguminosa, las cuales al ser secretadas por la planta son capaces de inducir la transcripción de los genes *nod* bacterianos, responsables de la síntesis y secreción de moléculas lipo-quito-oligosacáridos (LCOs) claves para el reconocimiento de la planta huésped. Actualmente, el conocimiento acerca de este proceso proviene mayoritariamente de modelos de estudios, donde se utilizan los alfa-rizobios, desconociéndose en gran medida lo que ocurre durante la interacción entre beta-rizobios y sus leguminosas hospederas. En este trabajo se utilizó como modelo de estudio un beta-rizobio, *Cupriavidus necator* UYPR2.512 y su par simbiote, *Mimosa pudica*. Primeramente, con el objetivo de conocer las respuestas moleculares de UYPR2.512 en los primeros pasos de la interacción simbiótica, se utilizó la luteolina como inductor de los genes *nod*. Con este fin, se decidió (i) re-secuenciar y mejorar la anotación del genoma de UYPR2.512, (ii) estudiar el efecto de la luteolina en la expresión de los genes *nod* bacterianos, (iii) describir las capacidades de interaccionar con el hospedero de UYPR2.512 a nivel genómico, (iv) estudiar la respuesta de la bacteria a la presencia de luteolina, utilizando como herramienta el RNA-Seq y (v) caracterizar las moléculas involucradas en la señalización de la interacción simbiótica, mediante una aproximación metabolómica.

La secuenciación de UYPR2.512 mostró la existencia de un genoma compuesto de 9,8Mb, incluyendo 4 replicones circulares: dos cromosomas y dos plásmidos mayores a

300 Kb, uno de ellos conteniendo los genes simbióticos *nod*, *nif* y *fix*. Además, se identificaron genes implicados en la interacción con el hospedero (genes simbióticos, genes relacionados con sistemas de secreción, degradación de compuestos aromáticos, homeostasis de metales, entre otros).

Por otro lado, se evaluó la respuesta de los genes *nod* bacterianos a la presencia de luteolina en UYPR2.512 portando el plásmido pCBM01 (*pnodB*_{19424-lacZ}), indicando que la luteolina es una molécula señal fuerte y altamente específica en este beta-rizobio.

Mediante el RNA-Seq se observó un total de 145 genes expresados diferencialmente, con un número similar de genes sub-expresados y sobre-expresados, y una distribución particular entre los replicones. La mayoría de los genes sub-expresados se asignaron al cromosoma principal (chr1), mientras que los genes sobre-expresados se distribuyeron entre el resto de las moléculas (chr2, pCne512e o pSym, pCne512b). Los genes sobre-expresados incluyeron el operón *nod*, y genes implicados en la biosíntesis de exopolisacáridos y flagelos, así como también genes implicados en el metabolismo del cobre y otros metales. Entre los genes reprimidos se identificaron genes implicados en el metabolismo global del nitrógeno y del carbono. Estos resultados obtenidos sugieren que, en respuesta a la luteolina, UYPR2.512 no solo produce NFs, sino que también es capaz de remodelar su superficie y alterar su metabolismo con el fin de prepararse para la próxima interacción simbiótica con la planta huésped.

Mediante varias técnicas de metabolómica no dirigida se intentó caracterizar las moléculas presentes en exudados radiculares de plántulas de *Mimosa pudica*. A partir de LC-ESI-MS/MS en exudados radiculares de plántulas de *Mimosa pudica*, se lograron detectar 13 compuestos presentes únicamente en presencia del simbionte bacteriano, indicando la especificidad de dicha interacción. En esta línea, debido a que estos compuestos no fueron encontrados en las bases de datos disponibles hasta el momento, y sumado a la poca información sobre este modelo, se sugiere que estas moléculas expresadas diferencialmente podrán ser compuestos potencialmente desconocidos.

Se comprobó que otros flavonoides como crisina, apigenina, ácido cumárico y la miricetina, además de la luteolina, pueden activar los genes *nod* en UYPR2.512, lo que confirma su papel como señales importantes en los primeros pasos de la interacción con sus hospederos. Aunque no se encontraron flavonoides específicos en los exudados de

la planta *Mimosa pudica*, se identificaron otras moléculas que podrían estar involucradas en la interacción, como aminoácidos, ácidos aromáticos, carboxilatos y polialcoholes. La variación en la composición química de los exudados radiculares de la planta en respuesta a la presencia de UYPR2.512 refleja la complejidad de esta interacción. El uso combinado de las ciencias ómicas diversas, como la genómica, transcriptómica y metabolómica, nos permitió proponer un modelo de simbiosis entre los beta-rizobios y las leguminosas, confirmando la especificidad molecular de esta relación y abriendo nuevas perspectivas en la comprensión de las simbiosis entre microorganismos y plantas.

INTRODUCCIÓN

1 El nitrógeno y su importancia biológica

El nitrógeno es un componente esencial de la mayoría de las biomoléculas, siendo un elemento principal para la síntesis de ácidos nucleicos (ADN, ARN), aminoácidos, fosfolípidos, vitaminas, nucleótidos trifosfato (ATP, GTP, CTP, UTP) y otros compuestos orgánicos nitrogenados [1]. Además, es uno de los componentes de la clorofila, pigmento importante en la fotosíntesis [2].

Si bien el nitrógeno es uno de los elementos más abundante en nuestro planeta, se encuentra mayoritariamente presente en la atmósfera como dinitrógeno molecular (N_2), un gas muy poco biodisponible para los organismos, debido a la incapacidad de la mayoría de éstos de romper el triple enlace de la molécula de dinitrógeno [3].

En el caso de las plantas, solo las formas reactivas de nitrógeno como ser la oxidada (por ejemplo, NO_x) o reducida (por ejemplo, NH_3 o aminas), pueden ser asimiladas y utilizadas como fuente de nitrógeno [4], [5].

Las relaciones entre los estados de oxidación del nitrógeno, que en conjunto conforman el ciclo del N, son consecuencia de la actividad metabólica de microorganismos, plantas y animales. Además de la actividad biológica, también hay factores abióticos que influyen en el ciclo del nitrógeno. Los rayos cósmicos también participan en el ciclo del nitrógeno, facilitando la formación de formas oxidadas de N. Los óxidos de nitrógeno generados, ya sea por procesos bióticos como abióticos se combinan con la humedad atmosférica para producir ácido nítrico (HNO_3), que puede precipitar como lluvia ácida y transportar compuestos de nitrógeno hacia el suelo. En el suelo, las plantas pueden absorber estos compuestos de nitrógeno y utilizarlos para su crecimiento y desarrollo (Fig. 1).

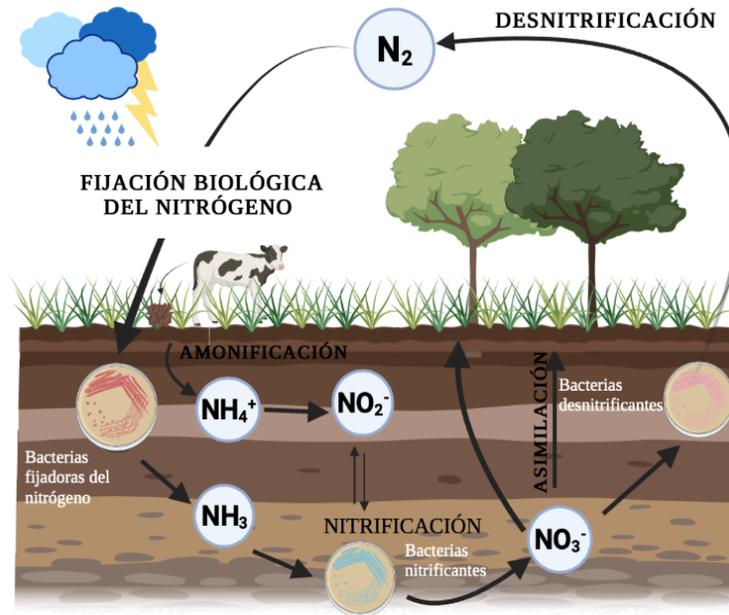


Figura 1. Ilustración del Ciclo del nitrógeno de un ecosistema terrestre.

Los principales procesos involucrados en el ciclo del nitrógeno incluyen: la fijación biológica del nitrógeno, la amonificación (N_2 a NH_4^+), la nitrificación (NH_4^+ a NO_2^- y/o a NO_3^-), la absorción/asimilación (incorporación de NO_3^- o NH_3 a las plantas) y la desnitrificación (NO_3^- a N_2). En el ciclo del nitrógeno intervienen reacciones químicas redox que involucran diferentes estados redox del nitrógeno, las cuales son llevadas a cabo por animales, plantas y microorganismos, que permiten mantener un equilibrio del ecosistema terrestre. Tomada y modificada de [6].

Particularmente en el suelo, el N se encuentra como inorgánico (nitrógeno mineral, ~ 2%) y orgánico (~ 98%). Las formas inorgánicas incluyen amoníaco (NH_3), amonio (NH_4^+), nitrito (NO_2^-) y nitrato (NO_3^-), mientras que las formas orgánicas se encuentran en la materia orgánica viva (biota del suelo y restos frescos de animales y plantas) y materia orgánica no viva, incluida compuestos humificados y no humificados.

El NO_3^- y el NH_4^+ son las principales formas inorgánicas de nitrógeno absorbido/asimilado por las plantas. El NH_4^+ producido a partir de los restos de seres vivos puede ser absorbido por las plantas, transformarse en NO_3^- , mediante el proceso de nitrificación, o perderse como NH_3 por volatilización. Parte del NO_3^- es reducido a NH_4^+ mediante la amonificación y en condiciones limitantes de O_2 , el NO_3^- puede ser respirado por algunos microorganismos. La respiración de NO_3^- tiene como consecuencia la pérdida de nitrógeno de los ecosistemas (desnitrificación). Las pérdidas de nitrógeno mediante la erosión del suelo, la desnitrificación, la lixiviación, la volatilización química y, quizás lo

más importante, la eliminación de los residuos de cultivos que contienen nitrógeno afecta negativamente la salud del suelo y el medio ambiente [7].

El nitrógeno suele ser limitante en los suelos de cultivo, afectando su crecimiento y productividad. Esto ha llevado a la utilización de fertilizantes en la producción agrícola [8].

Por lo tanto, la utilización de fertilizantes nitrogenados conlleva problemas ambientales y económicos, contaminando las aguas y las capas freáticas, generando gastos derivados del uso de compuestos orgánicos no renovables como el petróleo para la producción de fertilizantes [9].

Contrariamente al proceso oxidativo, el principal mecanismo de ingreso de nitrógeno a los ecosistemas ocurre por un proceso reductivo, denominado fijación biológica de nitrógeno (FBN), el cual es realizado exclusivamente por microorganismos procariotas [10].

1.1 Fijación biológica del nitrógeno

La FBN es un proceso natural de transformación del nitrógeno atmosférico (N_2) en una forma simple soluble no tóxica (NH_4^+) y es realizado por un grupo reducido de procariotas, pertenecientes a los dominios Arquea y Bacteria denominados diazótrofos (diazó = N, trofos = comida).

En el dominio bacteria, la habilidad de fijar nitrógeno, se distribuyen en 6 phyla; Chlorobi (bacterias verdes del azufre), Cyanobacteria, Firmicutes (Gram positivas de bajo contenido de G+C), Actinobacteria (Gram positivas con alto contenido de G+C), Spirochaetes y Proteobacterias, siendo esta última la división bacteriana con más abundancia en diazótrofos, y a la que pertenecen los rizobios [11]. Los diazótrofos presentan metabolismos muy diversos, incluyendo bacterias heterótrofas aerobias, anaerobias y anaerobias facultativas, bacterias quimioautótrofas y microorganismos fotosintéticos.

En las tierras cultivadas, la FBN aporta un 60% del N fijado, mientras que un 30% se debe a procesos industriales de Haber-Bosch y un 10% proviene de procesos atmosféricos no biológicos [12].

La FBN no requiere combustibles fósiles, el nitrógeno fijado es menos susceptible a la desnitrificación, lixiviación y volatilización, con respecto al agregado a los suelos en forma de fertilizantes de síntesis química, debido a que es directamente metabolizado por las plantas [13]. Esto hace que surja gran interés de optimizar la FBN para producir alimentos en forma sustentable.

1.1.1 Complejo enzimático nitrogenasa

La FBN es catalizada por el complejo enzimático nitrogenasa. Este complejo está formado por dos componentes: el componente I (dinitrogenasa o proteína Fe-Mo, hierro molibdeno) y por el componente II (dinitrogenasa reductasa o proteína Fe). Si bien todos los diazótrofos poseen la dinitrogenasa clásica, que posee ambos componentes, existen algunos microorganismos que poseen además una o dos nitrogenasas alternativas, donde el Mo del sitio catalítico es sustituido por vanadio o hierro (Fe-V o Fe-Fe nitrogenasas) [4].

De las diferentes dinitrogenasas reportadas, la Fe-Mo-nitrogenasa de *Azotobacter vinelandii* es la mejor caracterizada [14]. En este modelo, en el componente I se encuentra el sitio de unión al nitrógeno y es donde ocurre la reducción de N_2 a NH_4^+ , mientras que el componente II cede electrones al componente I, en una reacción acoplada a la hidrólisis de Mg-ATP (Fig. 2). Dado que para la reducción de una molécula de N_2 a NH_3 son necesarios 8 equivalentes de reducción, la reacción consume en condiciones óptimas, un mínimo de 16 moléculas de Mg-ATP por cada molécula de N_2 fijada.

El componente I, consta de cuatro heterodímeros $\alpha\beta$. Cada dímero $\alpha\beta$ tiene asociado un cofactor molibdeno-hierro (FeMo-co), el cual provee el sitio activo de la nitrogenasa para la unión y reducción del sustrato, y un grupo de hierro o denominado clúster P, involucrado en la transferencia de electrones desde el componente II al componente I. El clúster P consiste en dos clústeres $[Fe_4-S_4]$, que comparten un átomo de S, formando una estructura simétrica con respecto a éste. El FeMo-co consiste en una estructura elongada que contiene siete átomos de Fe, un átomo de Mo, y nueve de S. El componente II (también llamado nitrogenasa reductasa) de la nitrogenasa consta de dos subunidades idénticas que coordinan un único clúster $[Fe_4-S_4]$ y un sitio de unión a ATP

en cada subunidad. Esta proteína actúa como una reductasa dependiente de ATP, que obtiene electrones desde una ferredoxina o flavodoxina reducida y los transfiere al componente I, en un proceso que consume dos moléculas de ATP por cada electrón transferido.

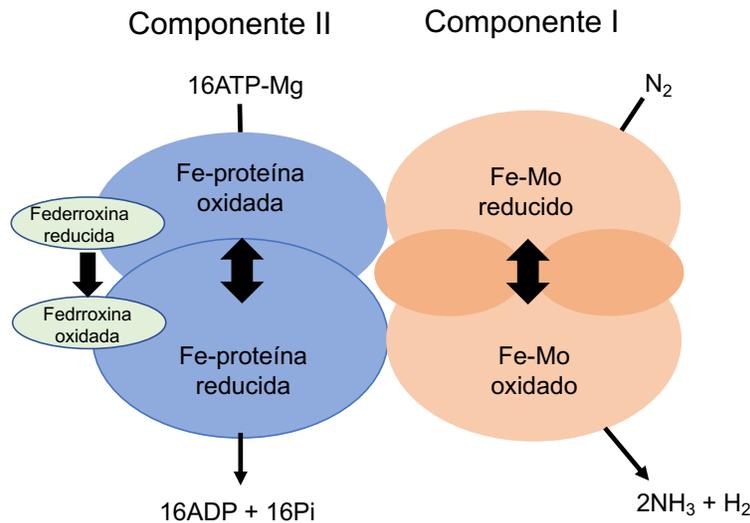


Figura 2. Ilustración del complejo enzimático de la nitrogenasa.

El complejo enzimático está formado por el componente II o la Fe proteína y por el componente I o Fe-Mo proteína. El flujo de electrones comienza desde el agente reductor (por ejemplo, ferredoxina o flavodoxina) hacia el componente II y luego al componente I. La unión e hidrólisis del ATP al componente II son necesarias para la reducción del N_2 por el componente I. Tomada y modificada de [6].

Los dos componentes están codificados por el gen *nifJ* (codifica para los transportadores de electrones), los genes *nifD* y *nifK* que codifican para las cadenas que componen la Fe-Mo dinitrogenasa y el gen *nifH* que codifica para la nitrogenasa reductasa o Fe proteína Fe [1], [15].

Además, para el completo ensamblaje del complejo de la nitrogenasa se requieren los productos de otros genes *nif*, involucrados en la síntesis del FeMo-co que incluye los genes *nifB*, *nifQ*, *nifE*, *nifN*, *nifX*, *nifU*, *nifS*, *nifV* y *nifY*, *nifS* y *nifU* para el ensamblaje de grupos Fe_4S_4 y *nifW* y *nifZ* para la maduración de los componentes de la nitrogenasa [15].

A su vez están presentes los genes *fix*, los cuales tienen un rol importante en la regulación de la FBN. Los genes *fixGHIS*, codifican para un complejo enzimático reductor de electrones y que a su vez ayuda a mantener la estabilidad de la nitrogenasa, *fixNOQP* codifican una citocromo oxidasa terminal del tipo Cbb3, la cual tiene alta afinidad por el oxígeno y permite que se mantenga baja la concentración de O_2 , lo cual es necesario para la actividad de la nitrogenasa. Los genes *fixL*, *fixJ*, *fixK*, codifican para proteínas reguladoras, las cuales además de regular la expresión de *nifA* (regulador global de la FBN), regulan la actividad de otros genes *fix*. La proteína sensora de oxígeno FixL, es capaz de autofosforilarse y transferir el fosfato a FixJ, regulando la expresión de los genes *nifA* y *fixK*, cuyos productos están implicados a su vez en la inducción de la expresión de los genes *nif* y *fix* [16].

1.1.2 Regulación de la fijación biológica del nitrógeno

La nitrogenasa se inactiva en un ambiente aeróbico, ya que la enzima es extremadamente sensible al oxígeno [17]. Los microorganismos fijadores de nitrógeno han desarrollado varias estrategias para evitar el efecto inhibitorio o tóxico del oxígeno. Por ejemplo, muchas bacterias diazotróficas fijan N_2 solo en condiciones anaeróbicas o microaeróbicas. En los quimiótrofos y fotótrofos aeróbicos que necesitan acceder al oxígeno o producirlo como parte de su metabolismo, estas bacterias logran un buen compromiso entre la eficiencia del uso de oxígeno como aceptor de electrones y la inactivación de la nitrogenasa [17], [18].

Con el fin de adaptar el proceso de FBN a los requerimientos fisiológicos, las bacterias diazótroficas han desarrollado varios mecanismos de sensado de señales ambientales, tales como la disponibilidad de N_2 (principalmente NH_4^+ y glutamina) o a la presencia de O_2 , como se comentó anteriormente. La regulación de la FBN ocurre mayoritariamente por medio de la regulación transcripcional y en algunos casos postransduccional, de la nitrogenasa, la cual puede inactivarse de forma irreversible [19].

La proteína NifA es el principal regulador global de la fijación de nitrógeno y es un activador de la transcripción de los genes *nifHDKE* en la mayoría de los organismos diazotróficos [20]. Los mecanismos de regulación de NifA se relacionan principalmente

con la disponibilidad de oxígeno y de amonio en el ambiente, siendo *Azotobacter vinelandii*, uno de los modelos más estudiados. Cuando los niveles de oxígeno son altos, NifL se une a NifA e inhibe su actividad, previniendo así la transcripción de los genes de la FBN. Cuando los niveles de oxígeno disminuyen, NifL ya no se une a NifA, lo que permite que NifA se active y promueva la transcripción de los genes necesarios para la FBN. En condiciones de exceso de nitrógeno, el metabolismo de este resulta en bajas concentraciones de 2-oxoglutarato (2-OG) y altas concentraciones de la proteína GlnK, que responde a la glutamina. Mediante la formación de un complejo GlnK, NifA y NifL, se suprime la actividad de NifA, lo que conlleva a la inactivación de la expresión de la nitrogenasa. Cuando el nitrógeno es limitado, el metabolismo del nitrógeno se ve afectado, provocando la acumulación de 2-OG y el consumo excesivo de GlnK. Esto hace que se interrumpa la formación del complejo ternario y se active NifA. La proteína NifA en su forma activa promueve la expresión de la nitrogenasa [21].

1.2 Los rizobios

La palabra rizobio deriva del griego “rizo” (raíz) y “bios” (vida) y es un nombre genérico adjudicado a ciertas bacterias, conocidas por su capacidad de asociarse simbióticamente con plantas pertenecientes a la familia *Fabaceas* (también denominada *Leguminosae*) y excepcionalmente, con ciertas especies no-leguminosas del género *Parasponia* [22].

En esta interacción, la planta hospedera obtiene nutrientes nitrogenados de los rizobios, brindándole a la bacteria un ambiente favorable para fijar nitrógeno. Esta simbiosis permite a las plantas leguminosas crecer sin fertilizantes nitrogenados y sin empobrecer los suelos.

En las raíces de las plantas (y excepcionalmente en los tallos de algunas leguminosas) los rizobios inducen la formación de órganos especializados denominados nódulos, donde llevarán a cabo la FBN. Además de establecer asociaciones simbióticas con leguminosas, los rizobios son microorganismos saprófitos que habitan en el suelo, condición denominada "vida libre". Suelen localizarse en la superficie de las raíces (rizoplano), en el suelo que rodea inmediatamente la raíz (rizósfera) y en menor cantidad en suelos no rizosféricos.

A pesar de que tanto los rizobios como las plantas leguminosas son filogenéticamente diversas, ninguna cepa bacteriana es capaz de asociarse simbióticamente con todas las especies de leguminosas y viceversa. Contrariamente, el establecimiento de la simbiosis rizobio-leguminosa es un proceso altamente específico, y en general cada rizobio es capaz de entablar relaciones simbióticas con una o unas pocas especies vegetales. Si bien existen rizobios con un amplio rango de hospederos, como por ejemplo *Sinorhizobium fredii* NGR234 capaz de nodular 112 géneros de leguminosas, estos son la excepción y no la regla [23].

Algunas leguminosas como *Glycine max* (soja) y *Phaseolus vulgaris* (poroto) pueden ser noduladas por más de una especie de rizobios [24]. El establecimiento de esta interacción no siempre se traduce en una simbiosis efectiva que beneficie a la planta, ya que una sola planta puede ser colonizada por múltiples bacterias que difieren en la efectividad de la fijación de nitrógeno [25]. De hecho, diferentes cepas de la misma especie de rizobios pueden diferir en su efectividad [26], incluso en asociación con la misma especie de leguminosa hospedera [27]. En consecuencia la eficiencia de fijación de nitrógeno difiere significativamente entre diferentes combinaciones de plantas y bacterias [28].

Si bien se consideran rizobios aquellas bacterias capaces de establecer asociaciones simbióticas con las plantas, resulta común que los rizobios se relacionen filogenéticamente con bacterias no simbióticas en diferentes niveles taxonómicos. Por ejemplo, un grupo de bacterias no simbióticas del género *Rhizobium* aisladas de la endósfera de la raíz del maíz [29], y bacterias del *Neorhizobium* provenientes de un suelo agrícola de tierras secas [30], presentan alta similitud con especies simbióticas de dichos géneros. De la misma forma el género *Cupriavidus* incluye cepas fijadoras de nitrógeno ubicadas filogenéticamente cerca de otras no fijadoras, reflejando que la capacidad de ser rizobio se adquirió en diferentes momentos durante la evolución [31].

En la actualidad, el phylum Proteobacteria es la división más grande y diversa dentro del dominio Bacteria. Mediante la utilización de estudios filogenéticos utilizando un marcador molecular como es el gen 16S del ARN ribosomal, las Proteobacterias se agrupan en 5 linajes designados como subdivisiones α , β , γ , δ , ϵ . Particularmente, los rizobios no son un grupo filogenéticamente discreto, y se encuentran mayoritariamente distribuidos en

las subclases de las α -proteobacteria (α -rizobios) y β -proteobacteria (β -rizobios). El reporte de que la γ -proteobacteria *Pseudomonas* sp. forma nódulos fijadores en *Robinia pseudoacacia*, no ha sido confirmado [24].

Las bacterias perteneciente a los α - y β -rizobios, se distribuyen en alrededor de 180 especies incluidas en 20 géneros bacterianos [32] (Tabla 1). Los α -rizobios son el grupo que ha sido mejor caracterizado, con una amplia distribución geográfica y mayor número de plantas hospederas, en comparación con los β -rizobios, los cuales fueron descritos posteriormente y tienen aparentemente una distribución más restringida.

Tabla 1. Agrupamiento taxonómico de rizobios simbióticos y no simbióticos relacionados filogenéticamente, pertenecientes a las subclases de α - y β -proteobacterias.

Se incluyen el Orden, Familia, Género y algunos ejemplos de sus plantas hospederas. Tomado y adaptado de [32].

Clase	Orden	Familia	Género	Planta hospedera (representantes)
<i>Alphaproteobacteria</i> (α -rizobia)	<i>Hyphomicrobiales</i>	<i>Bradyrhizobiaceae</i>	<i>Bradyrhizobium</i>	<i>Glycine max</i> , <i>Arachis hypogaea</i> , <i>Vigna unguiculate</i> , <i>Aeschynomene indica</i> , <i>Vigna</i> sp., <i>Phaseolus lunatus</i> , <i>Erythrophleum fordii</i> , <i>Kummerowia stipulacea</i> , etc.
		<i>Brucellaceae</i>	<i>Ochrobactrum</i>	<i>Cytisus scoparius</i> , <i>Lupinus honoratus</i> , etc.
		<i>Hyphomicrobiaceae</i>	<i>Devosia</i>	<i>Neptunia natans</i>
		<i>Methyllobacteriaceae</i>	<i>Methyllobacterium</i> , <i>Microvirga</i>	<i>Crotalaria juncea</i> , <i>Sesbania aculeata</i> , <i>Lotononis</i> sp., <i>Vicia alpestris</i> , etc
		<i>Phyllobacteriaceae</i>	<i>Aminobacter</i> , <i>Mesorhizobium</i> , <i>Phyllobacterium</i>	<i>Cicer</i> sp., <i>Astragalus</i> , <i>Amorpha fruticose</i> , <i>Caragana</i> , <i>Alhagi</i> , <i>Lotus</i> sp., <i>Sophora</i> sp., <i>Robinia</i> sp., <i>Acacia</i> , <i>Lathyrus numidicus</i> , etc.
		<i>Rhizobiaceae</i>	<i>Agrobacterium</i> , <i>Allorhizobium</i> , <i>Rhizobium</i> , <i>Shinaella</i> , <i>Ensifer</i> , <i>Sinorhizobium</i> , <i>Neorhizobium</i> , <i>Pararhizobium</i>	<i>Vicia</i> , <i>Pisum sativum</i> , <i>Trifolium</i> , <i>Vicia</i> , <i>Lens</i> , <i>Lathyrus</i> , <i>Phaseolus</i> , <i>Sophora</i> , <i>Kummerowia stipulacea</i> , <i>Medicago</i> , <i>Melilotus</i> , <i>Trigonella</i> , <i>Sesbania cannabina</i> , <i>Caragana</i> , <i>Galgae</i> , <i>Astragalus</i> , etc.
		<i>Xanthobacteraceae</i>	<i>Azorhizobium</i>	<i>Sesbania</i> sp.
<i>Betaproteobacteria</i> (β -rizobia)	<i>Burkholderiales</i>	<i>Burkholderiaceae</i>	<i>Cupriavidus</i> , <i>Trinickia</i> , <i>Paraburkholderia</i>	<i>Mimosa</i> , <i>Phaseolus vulgaris</i> , <i>Podalyria calyprate</i> , <i>Hypocalyptus</i> sp., <i>Virgilia oroboides</i> , <i>Robinia pseudoacacia</i> , etc.

Los α -rizobios se distribuyen en 17 géneros de siete familias, que incluyen *Bradyrhizobiaceae*, *Brucellaceae*, *Hyphomicrobiaceae*, *Methylobacteriaceae*, *Phyllobacteriaceae*, *Rhizobiaceae* y *Xanthobacteraceae*, y son miembros del orden *Hyphomicrobiales* denominado recientemente [33]. El género *Rhizobium*, que alberga 112 especies, es el género que contiene más especies de rizobios [34], [35].

Los α -rizobios se encuentran generalmente asociados a leguminosas cultivadas por el ser humano, como ser la soja (*Glycine max*), el maní (*Arachis hypogaea*), el garbanzo (*Cicer arietinum*), la arveja (*Pisum sativum*), el poroto común (*Phaseolus vulgaris*), la alfalfa (*Medicago sativa*) y el poroto mungo (*Vigna radiata*), entre otros.

Por otro lado, los β -rizobios descritos hasta ahora pertenecen a tres géneros bacterianos, *Cupriavidus*, *Paraburkholderia* y *Trinickia* [36], de la familia *Burkholderiaceae* dentro del orden *Burkholderiales*. Los β -rizobios se han aislado principalmente de nódulos de las raíces de leguminosas tropicales pertenecientes al clado *Mimosoideae*, tales como *Mimosa* spp. [37], [38], [39], [40], [41], y *Parapiptadenia rigida* [42], pero también al *Papilionoideae* como *Phaseolus vulgaris* [43], y *Rhynchosia ferulifolia* [44]. Los primeros reportes de la existencia de β -rizobios fueron en el año 2001, donde se enfocaron en el estudio de dos cepas del género *Burkholderia tuberum* (cepa STM678 y cepa STM815), aisladas de *Aspalathus carnosa* (*Papilionoideae*) en África del Sur y de *Machaerium lunatum* (*Papilionoideae*) en la Guyana Francesa [45]. Posteriormente, estas cepas fueron, denominadas como *Burkholderia tuberum* y *Burkholderia phymatum* y finalmente incluidas dentro del género *Paraburkholderia* [46], [47]. También en 2001, se describió la especie *Ralstonia taiwanensis*, aislada a partir de nódulos de dos especies de *Mimosa* (*Mimosa pudica* y *Mimosa diplotricha*) en Taiwan [48], la cual fue posteriormente renombrada como *Cupriavidus taiwanensis* [48], [49]. Actualmente, se sabe que las bacterias pertenecientes a los géneros *Paraburkholderia* y *Cupriavidus* muestran una extrema diversidad de hábitats. Cepas de estos géneros se han aislado de suelo [50], de plantas de tratamiento de aguas residuales [51], de rizósfera [52], de insectos [53] y de infecciones en humanos [54] reflejando su notable capacidad de adaptación.

1.3 Las leguminosas

Las leguminosas, plantas pertenecientes a la familia de las *Fabaceae*, se componen de aproximadamente 19.500 especies dentro de 751 géneros, presentes como hierbas, arbustos, enredaderas o árboles en hábitats principalmente terrestres [55], [56].

Dicha familia se ha dividido tradicionalmente en tres subfamilias, *Caesalpinioideae*, *Mimosoideae* y *Papilionoideae* [57]. Sin embargo, esta clasificación taxonómica, se encuentra en revisión porque no refleja el conocimiento actual sobre sus relaciones filogenéticas. Actualmente, el grupo de trabajo sobre la filogenia de leguminosas *The Legume Phylogeny Working Group* (LPWG) ha propuesto una reclasificación en seis subfamilias: *Cercidoideae*, *Detarioideae*, *Duparquetioideae*, *Dialioideae*, *Caesalpinioideae* y *Faboideae*. Entre las modificaciones más relevantes que se realizaron, se definieron cuatro subfamilias nuevas a partir de *Caesalpinioideae*, las restantes *Caesalpinioideae* se fusionaron con *Mimosoideae*, manteniéndose esta última como un clado dentro de la subfamilia *Caesalpinioideae* [24], [56]. Sin embargo, al no haberse realizado una filogenia consensuada, la clasificación realizada por LPWG continúa en discusión y no está del todo aceptada. En base a esto, en el presente trabajo mantendremos la clasificación clásica.

Las leguminosas presentan distribución cosmopolita, con capacidad de crecer en condiciones climáticas muy diversas. Su distribución abarca regiones tropicales, subtropicales y templadas. Se encuentran en una variedad de condiciones edáficas, así como en una gran diversidad de ambientes terrestres y acuáticos [58]. Aproximadamente un 70% de las leguminosas son capaces de establecer asociaciones simbióticas con los rizobios y fijar N_2 , siendo la característica más estudiada desde el punto de vista ecológico. En base a ello, su utilización como fuente de nitrógeno, en contraposición al uso de fertilizantes nitrogenados, es cada vez más valorado en sistemas de rotación con otros cultivos, ofreciendo una práctica de agricultura más sustentable y amigable con el medio ambiente [59]. Particularmente, las leguminosas pertenecientes al clado *Mimosoideae*, son utilizadas en la recuperación de suelos pobres, ayudando a la fertilización de estos, así como cultivos de soporte en sistemas de producción alimentaria [60].

El género *Mimosa* abarca unas 500 especies incluyendo plantas herbáceas, arbustos y árboles, capaces de crecer en hábitats ampliamente diversos, como bosques tropicales y subtropicales, sabanas, matorrales, desiertos, praderas, y humedales, así como en diversos suelos (suelos con pocos nutrientes, pH ácido o suelos contaminados con metales pesados). Su principal centro de diversificación se encuentra en las regiones del Cerrado y la Caatinga en el centro de Brasil [61], en donde existen más de 200 especies endémicas, y su distribución abarca desde la zona subtropical de Sudamérica (Paraguay, Argentina, Uruguay y el sur de Brasil) a México y sur de Estados Unidos [62].

Las especies del género *Mimosa* son conocidas por su capacidad de asociarse simbióticamente con rizobios pertenecientes a las α - y β -proteobacterias [63], postulándose que la preferencia por uno u otro endosimbionte depende de la ubicación geográfica, del ambiente y de su taxonomía. Actualmente, se ha reportado que las plantas de *Mimosa* spp. son preferentemente noduladas por *Paraburkholderia* en Brasil, *Ensifer* en México y *Cupriavidus* en Uruguay [64], [65].

1.4 Características del género *Cupriavidus*

El género *Cupriavidus* está compuesto por bacterias gram-negativas, móviles debido a la presencia de flagelos, que utilizan una amplia variedad de fuentes de carbono, lo que les proporciona una versatilidad metabólica [66]. Han sido aisladas de diferentes ambientes, como muestras clínicas [67], de suelos [68], de lodos [69], y de humedales para el tratamiento de aguas residuales [70]. En general son de baja patogenicidad. Sin embargo, uno de los casos más estudiados es el del patógeno humano *C. galliardii* [71]. A continuación, se describirán algunas de las características metabólicas de los *Cupriavidus*.

1.4.1 Homeostasis de metales

Las bacterias del género *Cupriavidus* son conocidas por su alta resistencia a metales pesados. Esto las convierte en candidatas ideales para su uso como inoculantes en áreas que requieren biorremediación [72]. Las bacterias utilizan varios mecanismos para mantener la homeostasis de los metales en su interior, como sistemas transportadores que regulan la entrada y salida de metales a través de la membrana celular, bombas de

eflujo para eliminar metales tóxicos del citoplasma, unión a proteínas de almacenamiento para evitar su toxicidad y sistemas quelantes como los sideróforos para transportar metales al interior celular [73]. Estos mecanismos incluyen ATPasas tipo P, proteínas de resistencia, nodulación y división celular (RND, de sus siglas en inglés *Resistance Nodulation and Cell Division Proteins*) y sistemas que facilitan la difusión de metales (CDF, de sus siglas en inglés *Cation Diffusion Faciliator*) (Fig.3). De esta manera, las bacterias mantienen niveles adecuados de metales y evitan su toxicidad.

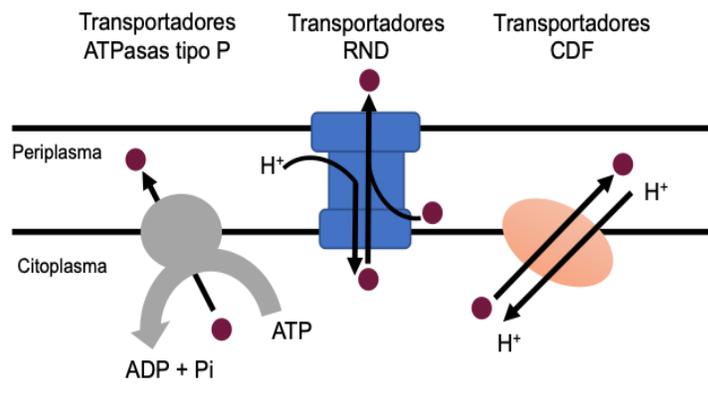


Figura 3. Ilustración de los principales sistemas de transportadores de metales en bacterias.

Se indican : (I) los sistemas de transportadores de ATPasas tipo P, encargados de la exportación del metal a través de una bomba al periplasma con gasto de energía que se obtiene de la hidrólisis de ATP, (II) *RND*, encargados de expulsar el metal hacia el exterior celular, mediante una proteína de fusión a membrana y un factor adicional para traspasar la membrana externa y (III) *CDF*, median el transporte del metal a partir de un gradiente quimiosmótico desde el citoplasma hacia el periplasma.

Dentro del género *Cupriavidus*, la especie *C. necator* se caracteriza por su capacidad para tolerar y emplear el ión cobre. Las bacterias de esta especie son especialistas en captar, transportar y almacenar este metal. El nombre "*Cupriavidus*" proviene del latín *cuprum*, que significa cobre, y "*avidus*", que significa ávido o amante. La cepa tipo, *C. necator* N1, fue descrita como un microorganismo del suelo cuyo crecimiento se ve estimulado por la presencia de cobre [74]. El cobre es un micronutriente esencial que actúa como cofactor en numerosas enzimas clave del metabolismo. Las cuproproteínas, que contienen cobre, son fundamentales para la transferencia y transporte de electrones en los citocromos, implicados en procesos vitales como la respiración y la protección celular contra el estrés oxidativo [75]. Sin embargo, el cobre también puede generar

especies reactivas de oxígeno (ROS, de sus siglas en inglés *Reactive Oxygen Species*) que causan daño oxidativo en lípidos, ácidos nucleicos y proteínas [76]. En concentraciones elevadas, el cobre puede ser tóxico para la mayoría de las células. Por esta razón, las bacterias han desarrollado diversos mecanismos para mantener la homeostasis del cobre, incluyendo regulación de la captación y expulsión del ion metálico, y su oxidación en el espacio periplásmico. En el género *Cupriavidus*, se han identificado tres sistemas para la homeostasis de cobre que involucran transportadores *RND*, codificado por *cusCBA*, ATPasas de eflujo de tipo P_{IBI}, como ser CopF y la familia codificada por el clúster *copABCD*, y transportadores del tipo *CDF* [77]. En dicho contexto, la cepa CH34 de *C. metallidurans* es una cepa modelo conocida por su resistencia a metales pesados, capaz de tolerar concentraciones milimolares de mercurio, cadmio, cobre, cromo, níquel, zinc y plomo [78]. A su vez, alberga numerosos determinantes de resistencia a metales pesados que incluyen diversos mecanismos, como la capacidad de formar complejos, sistemas de expulsión, reducción y precipitación reductora, que permiten la desintoxicación celular y la supervivencia en ambientes altamente contaminados con metales tóxicos [79], [80].

Otro elemento esencial para los microorganismos es el hierro, y su importancia se debe a que participa en un gran número de reacciones enzimáticas, así como forma parte de macromoléculas (citocromos, ferroproteínas). En el suelo, el hierro se encuentra poco biodisponible, ya que se encuentra formando óxidos poco solubles. Si bien es un elemento abundante, su baja solubilidad y su unión a proteínas han supuesto un desafío para los microorganismos que habitualmente se enfrentan a condiciones limitantes del metal. Como consecuencia, las bacterias han desarrollado varios sistemas de adquisición y utilización del metal, a partir de diferentes fuentes según los ambientes que colonizan [81], [82], [83].

En condiciones limitantes de hierro, las bacterias son capaces de obtener el metal en la forma férrica (Fe^{+3}), mediante la utilización de compuestos de bajo peso molecular denominados sideróforos, que son excretados al medio extracelular y actúan como quelantes del ion férrico con alta afinidad [84]. Una vez en el exterior, los sideróforos unen y solubilizan el hierro, generando un complejo ferri-sideróforo que puede ser internalizado. Particularmente, en el caso de bacterias gram-negativas, este complejo es

reconocido por receptores de membrana externa y transportado hacia el periplasma en un proceso dependiente del complejo de proteínas TonB-ExbB-ExbD.

Se pueden distinguir cuatro tipos de sideróforos en función de la naturaleza química del grupo funcional implicado en la quelación de hierro. Estos son los de tipo catecol, fenolato, hidroxamato y carboxilato, aunque también son comunes sideróforos que contienen más de un grupo funcional [85]. Particularmente, en bacterias de la especie *C. necator* se ha reportado la síntesis de sideróforos tipo carboxilato, como ser en la cepa JMP 134 de *C. necator* [86].

1.4.2 Degradación de compuestos aromáticos

Los hidrocarburos aromáticos se encuentran ampliamente distribuidos en la naturaleza. La fracción más prevalente se halla en la lignina de las plantas superiores y en sus productos de descomposición, contribuyendo a la formación de materia orgánica persistente en los suelos. Por otra parte, el uso extensivo de hidrocarburos aromáticos naturales y compuestos xenobióticos en procesos industriales, junto con prácticas inadecuadas de gestión de residuos, ha situado a estos compuestos como contaminantes orgánicos estables y persistentes.

La degradación de los compuestos aromáticos es llevada a cabo casi exclusivamente por microorganismos, los cuales que han desarrollado diversas estrategias para obtener beneficios energéticos y fuentes carbonadas [86]. A su vez, los compuestos aromáticos pueden ser potencialmente tóxicos para los microorganismos, incluso para aquellos que pueden degradarlos. Una estrategia para sobrevivir en presencia de compuestos aromáticos potencialmente tóxicos para la célula, consiste en desarrollar mecanismos de tolerancia a estos, de forma que puedan mantener su viabilidad [87].

Existe una gran variedad de compuestos aromáticos y la estrategia para su degradación se basa en transformar primeramente estos compuestos mediante vías periféricas, que confluyen dando lugar a un menor número de compuestos centrales [88].

En el caso del catabolismo aerobio bacteriano, los compuestos aromáticos son activados mediante reacciones de oxigenación (por las oxigenasas de hierro no hemo de *Rieske*, las flavoproteínas monooxigenasas y las oxigenasas multicomponentes de dihierro

soluble) y/o CoA ligasas, incorporando oxígeno molecular al anillo aromático para permitir su posterior apertura [86].

Luego, oxidaciones sucesivas mediante hidroxilasas, monooxigenasas o dioxigenasas permiten la conversión de hidrocarburos aromáticos periféricos a intermediarios dihidroxilados que pueden converger en alguna ruta catabólica central, como por ejemplo el ciclo de los ácidos tricarboxílicos (TCA, de sus siglas en inglés *Tricarboxylic Acid Cycle*) [88].

La cepa JMP134 de *C. pinatubonensis* es una de las más estudiadas, debido a su capacidad de degradar compuestos aromáticos. Su genoma presenta 11 de las 12 rutas centrales de degradación de compuestos aromáticos reportadas en proteobacterias. Entre ellas, posee una de las principales vías catabólicas que converge en la formación de β -cetoadipato, la cual posee dos ramas principales: la rama del catecol (generado a partir de la degradación de benzoato y algunos monómeros de lignina) y la rama del protocatecuato (generado a partir 4-hidroxibenzoato, y diversos monómeros de lignina) [89].

1.4.3 Producción de PHB

El PHB (poli(3-hidroxibutirato)) es un poliéster termoplástico biodegradable que pertenece a la familia de los polihidroxialcanoatos (PHA), compuesto por hidroxiacidos de entre tres a cinco átomos de carbono. En las últimas décadas, ha generado gran interés en aplicaciones industriales y campos de investigación debido a su biodegradabilidad y biocompatibilidad [90].

En la naturaleza, el PHB se encuentra como un cuerpo de inclusión citosólica en varios microorganismos, como bacterias y arqueas. Estos microorganismos utilizan el PHB como almacenamiento de carbono y energía, y también como fuente de poder reductor [91], [92], especialmente en situaciones de estrés. Esto les brinda ventajas para sobrevivir en ambientes cambiantes y competitivos. *C. necator* es una bacteria ampliamente estudiada para la producción de PHB debido a su capacidad para acumular grandes cantidades de este poliéster, llegando a representar hasta el 90 % de su peso seco celular [89].

La ruta clásica de biosíntesis del PHB involucra tres reacciones secuenciales y se expresa a partir de diferentes fuentes de carbono y ácidos grasos que generan acetil-CoA como metabolito intermedio. En la primera reacción, dos moléculas de acetil-CoA se combinan para formar acetoacetil-CoA, reacción catalizada por la enzima β -cetotiolasa/3-cetotiolasa/acetil-CoA acetiltransferasa (*phaA*). Luego, en una segunda reacción catalizada por la acetoacetil-CoA reductasa/hidroxibutiril-CoA deshidrogenasa (EC 1.1.1.36) (dependiente de NADPH) (*phaB*), el acetoacetil-CoA se reduce a 3-hidroxibutiril-CoA. Finalmente, la enzima PHB sintasa/poli(3-hidroxibutirato) sintetasa (*phaC*) polimeriza dos monómeros de 3-hidroxibutiril-CoA para formar el poli(3-hidroxibutirato) [93].

2 Simbiosis rizobio–leguminosa

El modelo de asociación simbiótico rizobio-leguminosa es uno de los más estudiados debido a que varias leguminosas son fuentes de alimentación tanto para humanos como para el ganado. Además, durante esta asociación ocurren los procesos de FBN más eficaces (20–300 Kg hectárea⁻¹año⁻¹) [94]. El aislamiento de rizobios a partir de nódulos colectados en el campo y el uso de plantas trampa, ha permitido seleccionar cepas fijadoras de nitrógeno, con el fin de utilizarlas como inoculantes que permitan el desarrollo de leguminosas en suelos empobrecidos [95].

La interacción simbiótica entre α -rizobios y leguminosas hospederas ha sido un modelo de estudio durante más de 100 años [96]. Esta interacción simbiótica tiene la característica de que las bacterias (rizobios) en su estado diferenciado (bacteroides), dentro de las células vegetales de los nódulos de las leguminosas, llevan a cabo el proceso de la FBN. El establecimiento de esta relación mutualista implica un complejo diálogo molecular entre la planta huésped y el microorganismo fijador de nitrógeno [28]. Las etapas de formación del nódulo se resumen en la figura 4 y se describen a continuación.

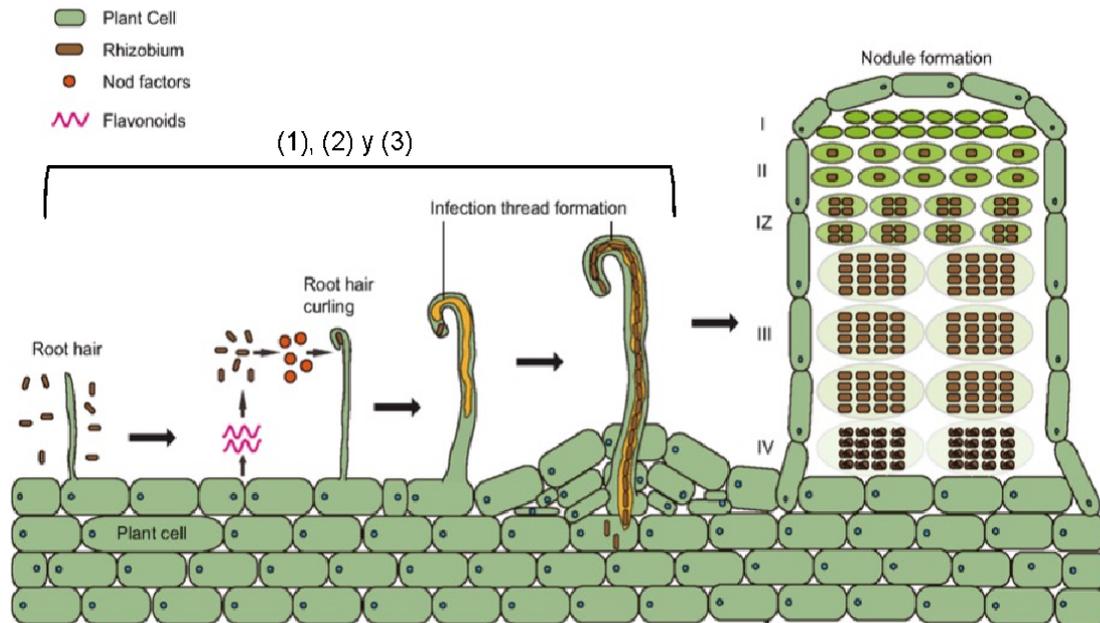


Figura 4. Resumen del proceso de infección durante la simbiosis rizobio-leguminosa y morfogénesis de un nódulo indeterminado.

Ilustración del proceso de infección durante la simbiosis rizobio-leguminosa y morfogénesis de un nódulo indeterminado. Se indican: (1) el reconocimiento y adhesión de las bacterias al pelo radicular, (2) la curvatura del pelo radicular e infección, (3) el transporte de las bacterias por el hilo de infección hacia las células del primordio y seguidamente hacia las células corticales de la planta, endocitosis y liberación de las bacterias. A la derecha se muestran las cinco zonas de un nódulo maduro indeterminado, en las cuales las bacterias se diferenciarán a bacteroide, que incluye zona I: zona meristemática, zona II: zona de infección, zona IZ: interzona; zona III: zona de fijación de nitrógeno y zona IV: zona de senescencia. Tomado y adaptado de [28].

1) Reconocimiento y adhesión de las bacterias a la planta. El reconocimiento está mediado por la liberación de compuestos presentes en los exudados radiculares de las plantas, entre los que se encuentran los flavonoides, los cuales atraen químicamente a las bacterias presentes en la rizósfera hacia la región apical de los pelos radiculares. En presencia de flavonoides se induce la expresión de los genes *nod* bacterianos, responsables de la síntesis y secreción de los factores de nodulación o factores Nod (NFs, de sus siglas del inglés *Nod Factors*). Los NFs son a su vez reconocidos por receptores específicos de las células vegetales, desencadenando en la planta los procesos que permitirán la infección, como ser la deformación de los pelos radiculares, la formación del hilo de infección, la formación del nódulo y la división de las células

corticales. Este primer paso de reconocimiento le da especificidad a la interacción simbiótica, ya que cada rizobio es capaz de responder a uno o unos pocos flavonoides, y a su vez cada planta hospedera reconoce solo algún NF. En dicho escenario, los rizobios para poder ingresar a la planta, previamente deben adherirse a la raíz. Se han reportado proteínas bacterianas como las adhesinas (ricadhesinas y *RAPs* (proteínas adherentes de rizobios), celulasas y glucanansas) las cuales estarían involucradas en la adhesión [97], así como también en la formación de biopelículas [98]. También se ha observado que ciertos componentes de la superficie celular de las bacterias tienen un rol importante en esta primera etapa de la interacción, como ser los polisacáridos extracelulares (EPS) y lipopolisacáridos (LPS) de las bacterias [99] y diversas lectinas del lado de la planta [100].

2) Curvatura del pelo radicular e infección. Al ser reconocidos los NFs bacterianos por los receptores en las plantas, ocurren cambios morfológicos y fisiológicos en los pelos radiculares de las plantas. Dentro de las respuestas inducidas por los NFs se han identificado cambios en los niveles del influjo y eflujo de iones calcio, cloro, potasio y protones, siendo estos últimos los que provocan cambios en el pH intracelular de las células vegetales, que se vuelve más alcalino. Esto provoca la activación de las proteínas reguladas por el calcio, las cuales serán responsables de hidrolizar y reorganizar el citoesqueleto de las células de los pelos radiculares. Como consecuencia de que ocurre un hinchamiento en la zona apical del pelo, provocando un enroscamiento de este, quedando las bacterias que se encuentran localizadas en esa zona (atrapadas en un “bolsillo”). En ese sitio, las bacterias inducen lesiones locales mediante hidrólisis de la pared celular de la planta, y penetran a la célula vegetal mediante la invaginación de la membrana. La planta hospedera responde depositando material de pared celular sobre la lesión, formándose un tubo, denominado “hilo de infección”, que crece en sentido longitudinal hacia el interior, por el que ingresan las bacterias hasta las células del córtex de la raíz de la planta [101].

3) Endocitosis y liberación de las bacterias. Cuando el hilo de infección alcanza las células del primordio del nódulo, las bacterias son liberadas al citoplasma de estas células, mediante un proceso similar al de endocitosis. En dicho proceso los rizobios no

quedan en contacto con el citoplasma de las células infectadas, sino porque quedan envueltos en una membrana de origen vegetal, denominada “membrana peribacteroidea”, distinta en su composición a la membrana plasmática. La membrana peribacteroidea limita las vesículas, denominadas simbiosomas, donde las bacterias luego se diferenciarán a bacteroides.

4) Diferenciación de bacteria a bacteroide y formación del nódulo. Dentro de los simbiosomas, las bacterias se diferencian a bacteroides, que incluye aumento de tamaño y forma celular, y pérdida de la capacidad de dividirse y adquieren la capacidad de fijar nitrógeno.

Los nódulos de leguminosas se pueden agrupar en dos tipos principales: indeterminados y determinados. Los nódulos indeterminados se originan a partir de divisiones celulares en la corteza interna y poseen un meristema apical persistente. Presentan una forma cilíndrica, con un gradiente de desarrollo desde el ápice hasta la base del nódulo, en el que se pueden individualizar cinco zonas (Fig. 4). Los nódulos determinados son el resultado de divisiones celulares en la corteza media o externa de la raíz, carecen de un meristema persistente y tienen forma esférica. Las divisiones celulares de un nódulo determinado cesan al principio de las etapas del desarrollo y el nódulo maduro se desarrolla a través del agrandamiento celular, por lo cual las células infectadas se desarrollan más o menos sincrónicamente con la etapa de FBN. En ambos tipos de nódulos, las células vegetales experimentan una endorreduplicación del genoma, lo que lleva a la poliploidización y al agrandamiento celular.

Paralelamente al desarrollo de las células del nódulo, se produce la diferenciación de los bacteroides fijadores de nitrógeno. Dependiendo del hospedador y del tipo de nódulo, la diferenciación puede ser terminal o reversible. La diferenciación terminal se caracteriza por la endorreduplicación del genoma, la elongación celular, el aumento de la permeabilidad de la membrana y la pérdida de la capacidad reproductiva. En la diferenciación reversible, los bacteroides conservan el tamaño celular y el contenido de ADN de manera similar a las bacterias en “vida libre” [102], [103]. En comparación con las bacterias de “vida libre”, los bacteroides muestran cambios drásticos en el

transcriptoma, la estructura de la superficie celular y en las actividades metabólicas para que se adapten al entorno intracelular y a fijar nitrógeno [104].

El poder reductor que utiliza el bacteroide para la FBN proviene de la planta, predominantemente en forma de ácidos dicarboxílicos, que a su vez, aportan los esqueletos carbonados que servirán como vehículo para el nitrógeno fijado. A su vez el nitrógeno fijado sale del bacteroide hacia el espacio peribacteroidal a través de la membrana por difusión, en este espacio el pH ácido promueve la protonación a NH_4^+ . Este NH_4^+ es exportado al citosol de la célula nodular mediante transportadores de cationes monovalentes dispuestos en la membrana peribacteroidea, es asimilado por el sistema GS/GOGAT para formar glutamato, que pasa al sistema vascular de la planta. La FBN tiene un alto requerimiento energético en forma de ATP, por lo que el metabolismo respiratorio de los bacteroides debe ser muy activo. Sin embargo, el O_2 necesario para cumplir tal demanda energética es perjudicial para la nitrogenasa, ya que el complejo es irreversiblemente inactivado en presencia del mismo. Por ello, los diazótrofos han desarrollado distintos mecanismos de protección. En el caso de los rizobios, la presencia en el espacio peribacteroidal de una alta concentración de leghemoglobina, una hemoproteína, permite, por un lado, actuar como proteína de transporte, asegurando la provisión suficiente de O_2 para el metabolismo de los bacteroides. Al mismo tiempo, regula el nivel de O_2 para evitar que afecte la actividad del complejo nitrogenasa [105].

2.1 Genes simbióticos

En rizobios los genes simbióticos mejor caracterizados, son los involucrados en el establecimiento de la simbiosis (*nod*, *nol* y *noe*) [34]. Los genes *nod* se clasifican en tres categorías, cada una de las cuales aporta un grado de especificidad en la interacción simbiótica. La primera incluye a los genes *nod* más conservados en los rizobios (*nodABC*), los cuales codifican para proteínas involucradas en la biosíntesis del esqueleto y ensamblaje de los factores Nod (NFs, de sus siglas en inglés *Nod Factors*), y la unión del ácido graso. Los NFs, son lipo-quitooligosacáridos (LCOs), que consisten en un tetra o pentasacárido de N-acetil- β -D-glucosamina, que presenta en el C₂ de su extremo no reductor una cadena de ácido graso insaturado. Los NFs se distinguen en el

largo y la saturación de la cadena del ácido graso, y a su vez pueden presentar otros sustituyentes tanto en su extremo no reductor como en el reductor. La cadena de ácido graso generalmente tiene entre 16 a 18 átomos de carbono y el número de repeticiones central (n) varía entre 2 o 3. La posición de los sustituyentes varía de R1 a R7, a las cuales se la adicionan distintos sustituyentes por ejemplo de azúcar (arabinosa, fucosa) acetilo, metilo, carbamoilo, grupos sulfurilo o glicosilo (Fig. 5).

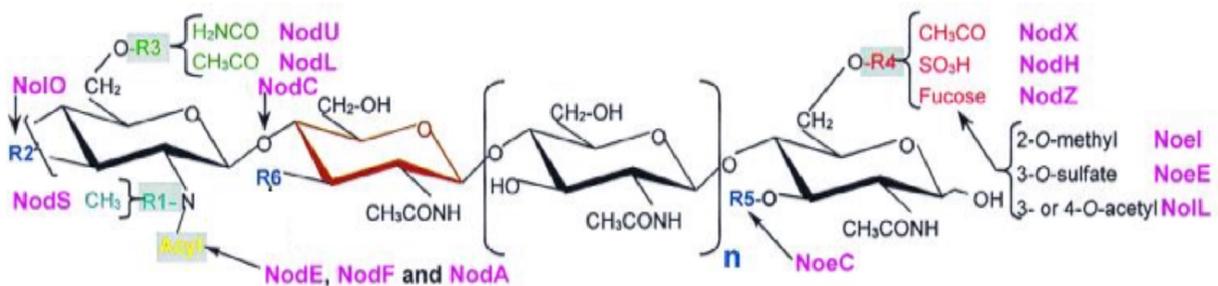


Figura 5. Ilustración de la estructura básica de los lipochito-oligosacáridos (LCOs), y función de los genes *nod*.

En la mayoría de los NFs, n varía entre 2 o 3, y la posición de los sustituyentes varía desde R1 a R7 (indicado con los diferentes colores), así como también los NFs sintetizados a partir de cada decoración. R1: Me (metilo) y H (hidrógeno), R2: OH (hidroxilo, H y Ca (carbamoil)), R3: H, Ca, y Ac, R4: H, Fu (fucosa), MeFu (metil fucosa), AcMeFu (acetilmetilfucosa), R5: H y Ara (arabinosilo) y R6: OH, Ac y Fu. Tomado de [106].

Particularmente, *nodA*, *nodB* y *nodC*, codifican para una quitin oligosacárido aciltransferasa, para una polisacárido desacetilasa y para una quitín oligosacárido sintasa, respectivamente.

El segundo grupo, incluye los genes *nod* reguladores (*nodD*), controlan la transcripción del operón *nod*, generalmente a través de su interacción con flavonoides (ver más adelante en 2.2). El número de copias de los genes reguladores varía entre las distintas especies de rizobio. Por ejemplo, hay tres genes *nodD* en *S. meliloti* (*nodD1*, *nodD2* y *nodD3*), que responden a distintos estímulos, ampliando el repertorio de moléculas señal involucradas para el reconocimiento simbiótico [106]. Tal es el caso de los genes *nodD1* y *nodD2* de la cepa CIAT 899 de *R. tropici* que responden al flavonoide luteolina y a la presencia de NaCl, respectivamente [107]. Por otro lado, en *L. japonicum* se demostró que los genes *nodD* de *M. loti* están involucrados en diferentes etapas del proceso,

mientras que la proteína NodD2 participa principalmente en la inducción de genes *nod* en la rizósfera y dentro de los nódulos, NodD1 induce los genes *nod* principalmente dentro de los hilos de infección del pelo radicular [108].

El tercer grupo, son los genes *nod* específicos (genes *nod*, *nol* y *noe*), con presencia variable en distintas especies de rizobios, codifican para enzimas responsables del agregado de los diferentes sustituyentes en los extremos de los NFs, determinando el rango de hospedero [106]. Por ejemplo, los genes *nodH*, *nodQ*, *nodS* y *nodU*, están involucrados en decoraciones de los NFs, mediante O-sulfatación (sulfotransferasa y adenililsulfato quinasa), N-metilación (metiltransferasa) y carboamiliación (carbamoiltransferasa), respectivamente [109], [110], [111]. Mientras los genes *nolO*, *noeE*, *noel*, *noeC* y *nolK*, también están involucrados en la decoración de los NF mediante, la 3-O-carbamoilación, sulfatación, 2-O-metilación y el agregado de azúcares, respectivamente [112].

Los genes *nodJ* y *nodI*, codifican para una proteína transportadora de eflujo tipo ABC (permeasa) y para una proteína transportadora de unión al ATP involucrada en el transporte de los NF desde el citoplasma bacteriano al espacio extracelular [113].

Si bien los genes *nod* tienen un rol central en muchas interacciones simbióticas, no son esenciales en todas las simbiosis. Se han descrito cepas de *Bradyrhizobium* que no poseen genes *nod* y que son capaces de nodular y fijar nitrógeno en simbiosis con especies de *Aeschynomene*, indicando la existencia de mecanismos de interacción simbiótica independiente de los NFs [114].

En todos los genomas de rizobios pertenecientes al género *Cupriavidus*, los genes de nodulación se transcriben de manera divergente del gen regulador *nodD*, ya que se encuentran en la otra hebra. En este contexto, el producto de *nodD* se uniría a las cajas *nod* presentes en las regiones operadoras del operón *nod*, lo cual activaría la expresión de los genes *nod* del operón. En la cepa LMG 19424 de *C. taiwanensis*, se observó que existen dos cajas *nod*, una delante del gen *nodB*, primer gen del operón *nod*, y la otra precediendo al gen *noeM*, por fuera del operón *nod*, aunque estaría involucrado en la decoración de los NFs producidos por esta bacteria [115].

2.2 Especificidad mediada por flavonoides

En la simbiosis rizobio-leguminosa, la inducción de la expresión de los genes *nod* está mediada por las proteínas NodD, reguladores transcripcionales de tipo LysR [110], las cuales son activadas por flavonoides sintetizados y secretados por la planta. La unión del complejo flavonoide-NodD, a la caja *nod* en el ADN bacteriano, activa la expresión de los genes *nod*, implicados en la síntesis y secreción NFs, que son a su vez reconocidos específicamente por la planta hospedera. La especificidad de este reconocimiento constituye un punto de control temprano de la simbiosis [116].

Los flavonoides son metabolitos secundarios de bajo peso molecular, que presentan diversas funciones biológicas que incluyen la protección de la planta contra la radiación ultravioleta evitando el daño fotooxidativo [117], entre otros [118]. A su vez, en los tejidos vegetales modulan los niveles de especies reactivas del oxígeno, mediante la inhibición de enzimas que generan ROS [119], o mediante la formación de complejos quelantes de iones metálicos capaces de catalizar la producción de ROS [120]. Los flavonoides están presentes generalmente en todos los órganos de las plantas y su composición depende de varios factores como ser el estado hídrico de la planta, disponibilidad de nutrientes, tipo de suelo y la edad [121].

Estudios de microscopía de fluorescencia han demostrado que los flavonoides se acumulan en determinadas células en la punta de las raíces, desde donde son exudados hacia el suelo [101] (Fig. 6).

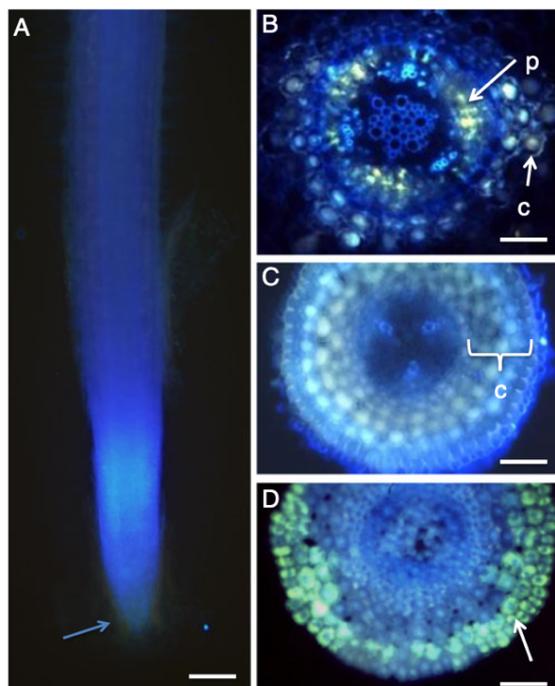


Figura 6. Localización de flavonoides en las raíces de *Medicago truncatula*, mediante microscopía de fluorescencia UV.

Se indican en (A) Acumulación de flavonoides con autofluorescencia de color azul y de color naranja , en las células de la punta de la cubierta y en la punta de la raíz (flecha azul), (B) Acumulación de flavonoides en una raíz madura de trébol blanco *Trifolium repens* L. (trébol blanco), en células del periciclo, (p) y corteza (c), (C) Acumulación de flavonoides en una sección de raíz joven de trébol blanco, en la células de la corteza, como (c), (D) Acumulación de flavonoides en una sección de la punta de la raíz del trébol blanco, en los núcleos de las células meristemáticas (azul claro), en el citoplasma de las células epidérmicas y corticales externas (se indican con una flecha). Las barras de aumento son de 500 μm en a y de 100 μm en b, c y d. Tomada de [101].

Por otro lado, los flavonoides se encuentran ampliamente distribuidos en nuestra dieta diaria y actualmente se consideran un componente indispensable en una variedad de aplicaciones nutracéuticas, farmacéuticas, medicinales y cosméticas [122]. Esto se atribuye a sus propiedades antimicrobianas, antialérgicas, antioxidantes, antiinflamatorias, antiviral, antifúngica, antimutagénicas y anticancerígenas [123].

Desde el punto de vista estructural, los flavonoides están compuestos por un esqueleto común de difenilpropano conocido como aglicona ($\text{C}_6\text{-C}_3\text{-C}_6$), en el que un anillo de benceno (A) está unido a un anillo de pirona (C), a través de una cadena de tres carbonos y a su vez en la posición 2 o 3 se une a un anillo de fenilo (B) como sustituto [122]. Las

distintas clases de flavonoides varían en su nivel de oxidación y en el patrón de sustitución de los distintos anillos, por lo cual conlleva a diferentes tipos de subclases, como se puede observar en la figura 7. La mayoría de los flavonoides se pueden subclasificar en varios grupos, que incluyen: flavonas (B), flavonoles (C), flavanonas (D), flavanoles (E), antocianidinas (F) e isoflavonas (G), y más recientemente las auronidinas [124].

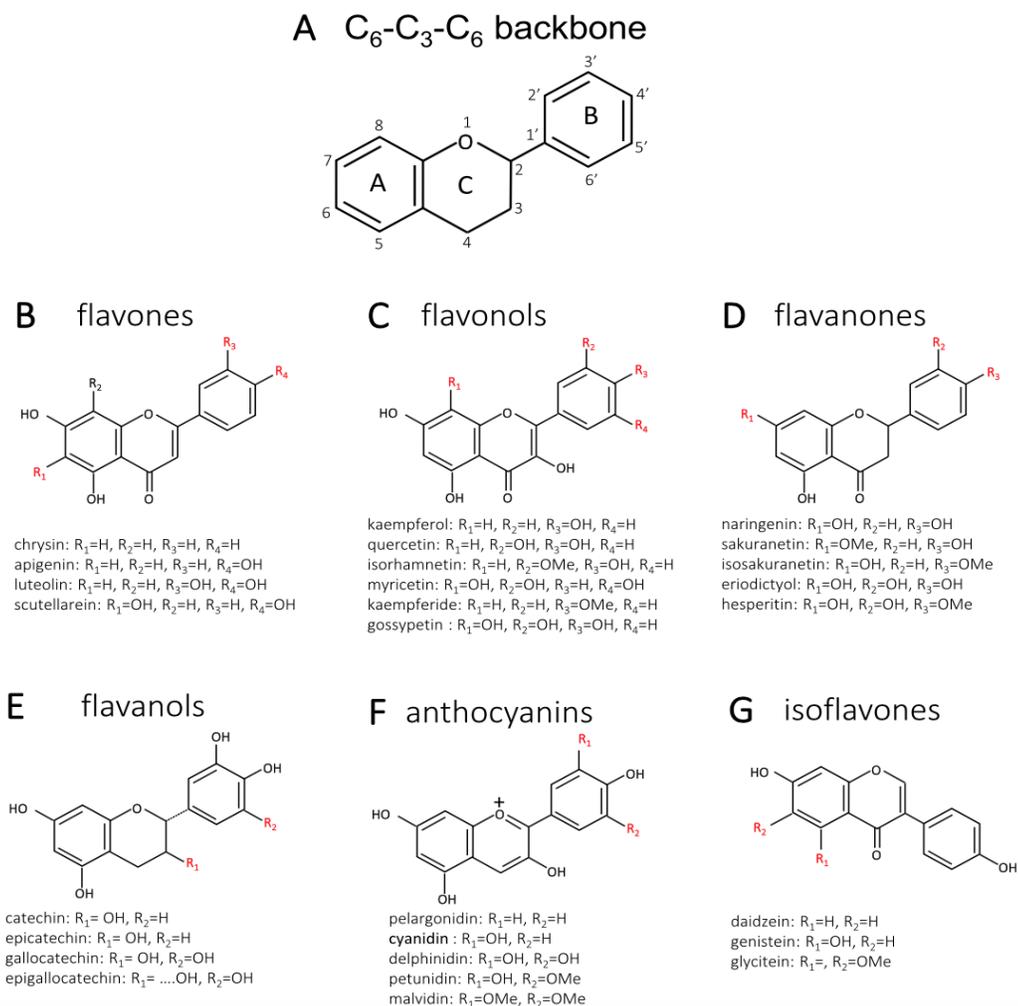


Figura 7. Ilustración de algunos de los esqueletos de flavonoides más estudiados y subclases, dependiendo de la modificaciones estructurales y sustituyentes.

La unidad básica es del tipo C₆-C₃-C₆ (o aglicona) y se define por dos anillos aromáticos, unidos mediante un anillo heterocíclico de pirona, anillo C. Se indican las clases flavonas (B), flavonoles (C), flavanonas (D), flavanoles (E), antocianidinas (F) e isoflavonas (G), y las subclases dependiendo del grupo sustituyente. Tomado y adaptado de [125].

Los flavonoides derivan de la fenilalanina y la malonil-CoA [126], y las vías de biosíntesis están bien caracterizadas [127]. Son sintetizados en el citosol, desde donde son secuestrados y transportados en vesículas hacia la vacuola para su almacenamiento o hacia otros destinos, donde pueden funcionar como moléculas bioactivas. El secuestro vacuolar de flavonoides involucra el tráfico de vesículas, transportadores de membrana del tipo ABC, transportadores del tipo extrusión de compuestos tóxicos y multidroga (*MATE*, de sus siglas en inglés *Multidrug And Toxic Extrusion compound*), y enzimas como el glutatión S-transferasa (GST) [128], [129].

Los pasos finales en la biosíntesis de flavonoides implican modificaciones como ser la hidroxilación, glicosilación, metilación y acilación, mejorando la actividad biológica, la solubilidad y el almacenamiento de estos compuestos [130]. La hidroxilación es la modificación más frecuente (78%) seguida por la glicosilación (56%), metilación (25%) y la fenilacilación (5%) [131], estando la mayoría de ello en la naturaleza como glucósidos [125]. Suelen acumularse en forma de mono-, di- o tri- glucósidos, en los que el O-glucósido y el C-glucósido son los más comunes. Sin embargo, se han encontrado unido a otros azúcares como ser galactosa, xilosa, ramnosa, arabinosa y el ácido glucurónico [132].

Distintos tipos de flavonoides son capaces de inducir la expresión de los genes *nod* en algunas especies de rizobios, e inhibir su expresión en otros. Otros estimulan el crecimiento bacteriano, facilitando la colonización radicular e incluso actuar como mecanismo de defensa frente a microorganismos patógenos [133].

La luteolina fue el primer flavonoide identificado producido por leguminosas, capaz de activar la expresión de los genes *nod* en cepas de rizobios además de actuar como quimioatrayente en *P. putida* [134], [135]. Los flavonoides genisteína y daidzeína (isoflavonas), producidos por raíces de soja, inducen la producción de NFs en *B. japonicum*, pero son inhibitorios de su síntesis en *S. meliloti* y *B. elkanii* [135], [136]. La naringenina es inductor de los genes *nod* en la cepa ORS285 de *Bradyrhizobium* sp. Sin embargo, la naringenina O-metilada, no induce la expresión de los genes *nod*, aunque estimula el crecimiento de la cepa ORS2185. Se postula que la estimulación del crecimiento bacteriano y los cambios metabólicos inducidos por naringenina, serían parte

de un mecanismo que facilitaría la colonización e infección de plantas que la exudan [137].

La genisteína y daidzeína, también pueden actuar como fitoalexinas, desencadenando un mecanismo de defensa contra infecciones de la raíz, mediante la inhibición del crecimiento de microorganismos patógenos [138].

Por otro lado, los estudios que emplean el β -rizobio *C. taiwanensis* LMG19424, han reportado que la biosíntesis y la exportación de NFs se inducen en respuesta a los flavonoides luteolina y apigenina [139]. Particularmente, en el caso de las *Burkholderia* no se han reportado estudios donde flavonoides estén relacionados con la simbiosis.

Además de los flavonoides, otras moléculas como ser azúcares, aminoácidos, ácidos dicarboxílicos (succinato, malato, fumarato), amonios cuaternarios (betaínas), ácidos aldónicos, y diferentes compuestos aromáticos (shikimato, quinato, protocatecuato, vainillina, acetosiringona, catecol), son quimioatrayentes de diversas bacterias y se ha demostrado que están implicados en el establecimiento de la simbiosis en el género *Rhizobium* [140], [141].

2.3 Especificidad mediada por la percepción de factores Nod

Los NFs son reconocidos por los receptores del factor Nod en plantas (por ejemplo, NFR1 y NFR5 en *Lotus japonicus*). Estas proteínas receptoras tienen un dominio LysM extracelular y un dominio quinasa intracelular [142], [143], [144]. La unión directa de NFs a los dominios LysM extracelulares del complejo receptor conduce a la activación de las vías de señalización de nodulación a través del dominio quinasa [145]. Como consecuencia, se induce la expresión de los genes que están involucrados en la regulación del nivel de calcio intracelular, acompañado de cambios en el citoesqueleto del pelo radicular. Estos cambios conducen a un aumento de las divisiones celulares en la raíz, lo que finalmente concluye con la formación de hilos de infección para la formación del nódulo. La señal provista por los NFs, estimula la expresión de los genes de síntesis de flavonoides y se propuso que estos interfieren con el transporte de auxinas, hormona de la planta que promueve la división celular [146].

A nivel de la planta, se observó que la transferencia de los genes codificantes para los receptores NFR1 y NFR5 de *L. japonicus* a *M. truncatula* permite la nodulación de los simbiontes de *L. japonicus*, *M. loti* [147].

2.4 Especificidad mediada por exopolisacáridos bacterianos

Además de los NFs, también los polisacáridos extracelulares de la superficie bacteriana y los secretados, tales como los exopolisacáridos (EPS), los lipopolisacáridos (LPS) y los polisacáridos capsulares (KPS), son importantes para establecer relaciones simbióticas en una gran variedad de rizobios [140], [141]. La producción de EPS es común en rizobios y su composición es variable entre las diferentes especies [148], pudiendo incluso inhibir o promover la simbiosis según su composición molecular [149].

En el caso de los EPS, se ha demostrado que median la colonización de los pelos radiculares y la formación de los hilos de infección, y que son reconocidos por una proteína localizada en la superficie celular que contiene tres dominios LysM extracelulares y un dominio quinasa intracelular tipo 3 (EPR) [150]. En dicho escenario, se observó que tanto mutantes de *Lotus* en el receptor EPR3, como la inoculación de las plantas salvajes con mutantes de *M. loti* defectuosas en la producción de EPS, provoca una reducción de la cantidad de hilos de infección y una acumulación inusual de los mismos en las células del córtex, afectando la eficiencia del proceso de infección [151]. En la simbiosis entre *S. meliloti*, *M. sativa* y *M. truncatula*, se demostró que un tipo de EPS (succinoglicano) es necesario para el inicio y elongación de los hilos de infección, y que un aumento en su producción mejora la capacidad de nodulación [152], [153]. Sin embargo, el papel simbiótico de EPS resulta más complejo en otro tipo de interacciones, como ser en la interacción entre *M. loti* y *Lotus* spp. [154]. En este caso, se ha visto que la producción de EPS en *M. loti* es necesaria para modular las respuestas de defensa de las plantas y permitir la infección bacteriana. Sin embargo, los mutantes que no producen EPS, igualmente conservan la capacidad de formar nódulos, mientras que aquellas que forman un EPS truncado no son capaces de progresar en la simbiosis, ya que desencadenan respuestas de defensa de la planta que resultan en un bloqueo de la infección [154].

2.5 Especificidad mediada por la inmunidad innata del huésped

La inmunidad innata se refiere a su capacidad para detectar y responder rápidamente a la presencia de patógenos o microorganismos invasores, como hongos, bacterias o virus, activando mecanismos de defensa preexistentes. En dicho contexto, las plantas hospederas han desarrollado mecanismos de reconocimiento para distinguir entre patógenos y simbioses, y responder diferencialmente [155].

Dentro de los mecanismos de inmunidad innata en las plantas incluyen la presencia de receptores de reconocimiento de patrones (PRRs, de sus siglas en inglés *Pattern Recognition Receptors*), los cuales son capaces de reconocer patrones moleculares asociados a microorganismos (MAMPs, de sus siglas en inglés *Microbe-Associated Molecular Patterns*) y a patrones moleculares asociados a patógenos (PAMPs, de sus siglas en inglés *Pathogen-Associated Molecular Patterns*), como ser componentes de la pared bacteriana (LPS), proteínas citosólicas, y moléculas efectoras, desencadenando la activación de los mecanismos tempranos de defensa de la planta. Algunos ejemplos conocidos de PRR genes en las plantas incluyen los receptores de quinasa asociados a flagelina (FLS2) que reconocen la flagelina bacteriana, el receptor de quinasa del ácido abscísico (BAK1) que reconoce diversos PAMPs [156], y los receptores tipo Nod (NLRs) que detectan efectores secretados por patógenos [157].

En dicho contexto, algunas bacterias utilizan los sistemas de secreción tipo II, III y IV (SSTII, SSTIII y SSTIV) para secretar ciertas proteínas efectoras en las células de la planta huésped, las cuales pueden activar o no, su respuesta de defensa [158]. A su vez, diferentes géneros de α - y β -rizobios, utilizan este sistema de efectores con el fin de promover la infección e incluso en ciertas situaciones son el mecanismo necesario para la nodulación, independiente de la síntesis de NFs [159].

Otro mecanismo de las plantas involucra a los genes de resistencia (genes R), los cuales en su mayoría codifican para receptores de superficie celular o intracelulares, los cuales pueden activar o elevar una respuesta de defensa específica. A modo de ejemplo, las leguminosas pueden aprovechar el reconocimiento mediado por los genes R para evitar ser noduladas por cepas deficientes en la fijación de nitrógeno e interactuar

selectivamente con cepas eficientes, como se observó con la cepa Rm41 de *S. meliloti* [160].

3 Antecedentes del grupo de trabajo

3.1 Modelo de Estudio: La cepa UYPR 2.512 de *Cupriavidus necator*

La cepa UYPR2.512 de *C. necator* se aisló de un nódulo de raíz presente en la leguminosa arbórea *Parapiptadenia rigida* en Uruguay. Esta cepa demostró ser capaz de formar nódulos Fix⁺ y promover el crecimiento tanto de su huésped original como de la planta modelo *M. pudica* [42]. Los análisis filogenéticos indicaron que UYPR2.512 pertenece al género *Cupriavidus*, aunque en un clado separado de *C. taiwanensis* [42], mientras que análisis posteriores revelaron su estrecha relación con la cepa UFLA-0271 de *Cupriavidus necator*, lo que sugiere que UYPR2.512 pertenece a la especie *necator* [161], [162]. En 2015, el genoma de UYPR2.512 de *Cupriavidus* sp. fue secuenciado y anotado, revelando una longitud de 7.858.949 pb, divididos en 369 *contigs*. Este borrador del genoma contiene 7.411 genes que codifican proteínas y 76 genes que codifican RNA [163]. Hasta ese momento, el único genoma secuenciado y publicado de un β -rizobio perteneciente al género *Cupriavidus* era el de la cepa LMG 19424T de *C. taiwanensis* [139].

Al igual que se ha visto en otras cepas de los genomas del género *Cupriavidus*, en el genoma de UYPR2.512 los operones *nod* y *nif* están fuertemente compactados. Los genes *nod* se organizan en un solo operón precedido por una caja de *nod* y se transcriben de manera divergente a la única copia del gen *nodD* presente en estos genomas [162], [163], [164], [165], [166]. En este contexto, el genoma de la cepa LMG19424 de *C. taiwanensis* contiene la isla simbiótica más compacta (35 kb) (que contiene los genes *nod*, *nif* y *fix*) identificada hasta la actualidad [139].

3.2 Hipótesis de trabajo

Nuestra hipótesis es que los β -rizobios utilizan mecanismos similares a los observados en α -rizobios para el establecimiento de relaciones simbióticas con plantas hospedadoras, sin embargo, también cuentan con estrategias exclusivas. La descripción de estos

mecanismos permitirá desarrollar estrategias de aprovechamiento de este recurso microbiano abundante en nuestro país, así como aumentar nuestro conocimiento sobre las interacciones benéficas entre plantas y microorganismos.

3.3 Objetivo general y específicos

El objetivo general de este trabajo es establecer los mecanismos moleculares involucrados en los primeros pasos de la interacción entre el β -rizobio UYPR2.512 y la leguminosa hospedera *Mimosa pudica*.

Para esto nos planteamos los siguientes objetivos particulares:

- 1) Definir la especie a la que pertenece la cepa UYPR2.512 y describir las principales características de su genoma.
- 2) Identificar los cambios a nivel de la expresión génica provocados por la luteolina en UYPR2.512.
- 3) Evidenciar los metabolitos presentes en los exudados de *Mimosa pudica* en presencia o ausencia de la bacteria simbiote UYPR2.512.

Con el fin de contestar los objetivos planteados se planteó la utilización de tres abordajes metodológicos complementarios comprendidos en las ciencias ómicas, como la genómica, la transcriptómica y la metabolómica, las cuales se describirán en los siguientes capítulos II, III y IV.

Durante la realización de esta tesis de doctorado, se realizaron 2 publicaciones:

-M. Rodríguez-Esperón *et al.*, “Genomics and transcriptomics insights into luteolin effects on the beta-rhizobial strain *Cupriavidus necator* UYPR2.512,” *Environ Microbiol*, vol. 24, Nov. 2021, doi: 10.1111/1462-2920.15845.

-C. Rodríguez-Esperón *et al.*, “Nodulation in the absence of nod genes induction: alternative mechanisms involved in the symbiotic interaction between *Cupriavidus* sp. UYMMa02A and *Mimosa pudica*,” *Environmental Sustainability*, vol. 6, no. 3, pp. 383–401, Jul. 2023, doi: 10.1007/s42398-023-00286-5.

CAPÍTULO I

Descifrando el genoma de la cepa UYPR2.512 de *Cupriavidus* sp.

Objetivo general

Definir la especie a la que pertenece la cepa UYPR2.512 y describir las principales características de su genoma.

Objetivos específicos

1. Definir la especie a la cual pertenece la cepa UYPR2.512.
2. Identificar los genes y principales vías metabólicas en el genoma relacionados al género *Cupriavidus* y a la interacción con plantas.

Resumen

La cepa UYPR2.512 de *Cupriavidus* sp, es un β -rizobio, capaz de formar asociaciones simbióticas efectivas con plantas leguminosas. Fue aislada de Angico (*Parapatadenia rígida*) y se determinó que pertenece al género *Cupriavidus*. Estudios genómicos y filogenéticos demostraron su relación cercana con especies como *C. necator*, *C. oxalaticus* y *C. taiwanensis*. En el año 2015 se presentó una versión borrador de su genoma, reflejando una versión incompleta del mismo. Este capítulo se centró en el análisis *in-silico* del genoma de la cepa UYPR2.512. El análisis filogenómico y estimación de distancias nucleotídicas para definir especies (ANI), sugirió que pertenece a la especie *necator*. A su vez, con el fin de mejorar la calidad del genoma, se llevó a cabo una nueva secuenciación mediante tecnología de lecturas largas (PacBio). A partir del genoma, se infirieron algunas funciones que podrían estar expresándose en la bacteria. Se estudiaron particularmente los genes característicos de la especie *necator*, así como

los genes posiblemente involucrados en la interacción entre UYPR2.512 y sus hospederos. La cepa UYPR2.512 presentó un genoma de un tamaño de 9,6 Mb, compuesto por cuatro replicones cerrados. Presenta una isla simbiótica conservada de unos 36 Kb, ubicada en el plásmido simbiótico (pSym), la cual incluye a los genes *nod*, *nif* y *fix*. En relación a las características genómicas relacionadas al género *Cupriavidus*, UYPR2.512 presentó más de 50 genes involucrados en la exportación y a la tolerancia de diversos metales divalentes (Cu, Cd, Co, Ni, Zn), lo que sugiere una adaptación a ambientes potencialmente desafiantes. A su vez, se observaron genes involucrados en vías periféricas y centrales de degradación de una gran variedad de compuestos aromáticos, así como también en el metabolismo de PHB, reforzando su versatilidad metabólica. En cuanto a la interacción con sus hospederos, la cepa UYPR2.512 mostró la presencia de componentes de maquinaria de sistemas de secreción (del SSI al SSTVI), genes de quimiotaxis y movilidad bacteriana, así como genes relacionados con la simbiosis y la fijación biológica del nitrógeno. Estos hallazgos genómicos aportan valiosa información para comprender mejor los mecanismos subyacentes a la simbiosis de UYPR2.512 con sus hospederos y su capacidad para prosperar en diferentes ambientes.

1 Genómica. Tecnologías de secuenciación y NGS

Las primeras tecnologías de secuenciación, desarrolladas por Sanger y Maxam-Gilbert en 1977, marcaron el inicio del campo de la secuenciación de ADN [167], [168]. Su descubrimiento permitió estudiar el código genético de los seres vivos, inspirando a los investigadores a buscar tecnologías más rápidas y eficientes. Sin embargo, a partir del 2005, una nueva era de tecnologías de secuenciación emergió con la aparición de la tecnología 454 de Roche [169]. Estas nuevas tecnologías ofrecían análisis masivos con un rendimiento mucho mayor (tamaño de lecturas en Gb, *gigabyte*), en menos tiempo y a un costo considerablemente inferior al de las primeras tecnologías de secuenciación. Esto abrió nuevas perspectivas para la exploración y el análisis de genomas. Estas técnicas se denominaron Tecnologías de secuenciación de próxima generación (NGS, por sus siglas en inglés *Next Generation Sequence*) o Tecnologías de secuenciación de alto rendimiento [170]. Las tecnologías NGS se dividieron en dos tipos: de primera y de

segunda generación, en función de su orden cronológico de aparición y se caracterizan por la preparación de bibliotecas amplificadas previo a la secuenciación [171].

No obstante, las tecnologías NGS siguen mejorando y tanto el número como la diversidad de secuenciadores ha aumentado de manera significativa en los últimos años. Entre los avances recientes surgen las tecnologías de secuenciación de tercera generación (TGS, de sus siglas en inglés *Third generation sequence*) [172]. A diferencia de las de segunda generación, se clasificaron como Tecnología de secuenciación de una sola molécula [173], ya que pueden secuenciar una sola molécula sin la necesidad de crear bibliotecas de amplificación. Además, son capaces de generar lecturas más largas a costos mucho más bajos y en menos tiempo.

Las diferentes plataformas de secuenciación basadas en distintos instrumentos, ofrecen diferencias en la cantidad de datos generados por corrida, la longitud promedio de las lecturas, el tipo de lectura (extremos pareados o PE *-paired-end-* o extremo simple o SE *-simple-end-*), los tipos de error más frecuentes (indel, inserción o delección, o sustitución o *mismatch*) y el porcentaje de error [174]. A continuación, se listan las principales características de las diferentes plataformas; de primera generación: ABI, Sanger; de segunda generación: 454, Illumina, Solid e Ion Torrent; y de tercera generación: PacBio y Oxford Nanopore (Tabla 2).

Tabla 2. Resumen de plataformas de secuenciación y sus características más relevantes

Resumen de plataformas de secuenciación y sus características más relevantes.

Se describen las principales características de las plataformas de secuenciación de primera, segunda y tercera generación. NA: no disponible, SE: *single-end*, PE: *paired-end*, M: millón, Gb: *Gigabytes*, Tb: *Terabytes*. Tomada y adaptada de [174].

Platform	Instrument	Reads per run	Avg Read length (pb)	Read Type	Error Type	Error Rate (%)	Data Generated per run (Gb)	Year
First Generation								
ABI Sanger	3730xl	96	400 – 900*	SE	NA	0.3	0.00069 to 0.0021	2002
Second Generation								
454	GS20	200	100	SE, PE	indel	1	0.02	2005
454	GS FLX	400	250	SE, PE	indel	1	0.1	2007
454	GS FLX Titanium	1 M	450	SE, PE	indel	1	0.45	2009
454	GS FLX Titanium+	1 M	700	SE, PE	indel	1	0.7	2011
454	GS Junior	100	400	SE, PE	indel	1	0.04	2010
454	GS Junior+	100	700	SE, PE	indel	1	0.07	2014
Illumina	MiniSeq	25M (maximum)	150	SE, PE	mismatch	1	7.5 (maximum)	2013
Illumina	MiSeq	25M (maximum)	300	SE, PE	mismatch	0.1	15 (maximum)	2011
Illumina	NextSeq	400M (maximum)	150	SE, PE	mismatch	1	120 (maximum)	2014
Illumina	HiSeq	5B (maximum)	150	SE, PE	mismatch	0.1	1.5Tb (maximum)	2012
Illumina	HiSeq X	6B (maximum)	150	SE, PE	mismatch	0.1	1.8Tb (maximum)	2014
SOLiD	5500 W	3B	75	SE	mismatch	~0.1	160	2011
SOLiD	5500xl W	6B	75	SE	mismatch	~0.1	320	2013
Ion Torrent	PGM 314 chip v2	400.000-550.000	400	SE	indel	1	0.06 to 0.1	2011
Ion Torrent	PGM 316 chip v2	2M - 3M	200	SE	indel	1	0.6 to 1	2011
Ion Torrent	PGM 318 chip v2	4M - 5.5M	400	SE	indel	1	1.2 to 2	2013
Ion Torrent	Ion Proton	60M - 80M	200	SE	indel	1	10	2012
Ion Torrent	Ion S5/S5XL 520	3M - 5M	400	SE	indel	1	1.2 to 2	2015
Ion Torrent	Ion S5/S5XL 530	15M-20M	400	SE	indel	1	03 to 05	2015
Ion Torrent	Ion S5/S5XL 540	60M - 80M	400	SE	indel	1	NA	2015
Third Generation								
PacBio	RS C1	432	1300	SE	indel	15	0.54	2011
PacBio	RS C2	432	2500	SE	indel	15	0.5 to 1	2012
PacBio	RS C2 XL	432	4300	SE	indel	15	0.5 to 1	2012
PacBio	RS II C2 XL	564	4600	SE	indel	15	0.5 to 1	2013
PacBio	RS II P5 C3	528	8500	SE	indel	13	0.5 to 1	2014
PacBio	RS II P6 C4	660	13500	SE	indel	12	0.5 to 1	2014
PacBio	Sequel	350	10000	SE	NA	NA	7	2016
Oxford Nanopore	MinION Mk	100	9545	1D,2D	indel/mismatch	12	1.5	2015
Oxford Nanopore	PromethION	NA	9846	1D,2D	NA	NA	2Tb to 4Tb	2016
*depending on run module; NA: Not available; SE: Single End; PE: Paired End; M: Million; B: Billion; Gb: Gigabytes; Tb: Terabytes								

Las técnicas NGS comparten la capacidad de secuenciar una gran cantidad de fragmentos de ADN de forma paralela en un corto lapso, con un alto rendimiento a partir de múltiples muestras y a un menor costo [170].

A su vez, comparten un abordaje metodológico semejante que se puede resumir en cinco pasos: 1) segmentación del ADN en fragmentos, 2) marcaje del ADN por medio de cebadores o adaptadores que indican el punto de partida para la síntesis de ADN, 3) amplificación de los fragmentos de ADN marcados con adaptadores por métodos basados en reacción en cadena de la polimerasa (PCR), 4) secuenciación y 5) reconstrucción de la secuencia completa por medio de secuencias de referencia y exportación a ficheros de almacenamiento de datos [175].

Algunas aplicaciones incluyen la secuenciación de genomas y la anotación de genes, la secuenciación de ARN (cuantificación de transcritos o RNA-seq) y amplicones de PCR (por ej. secuenciación masiva de genes que codifican para subunidad 16S del ARNr bacteriano). Son utilizadas para identificar marcadores epigenéticos, identificar interacción ADN-proteína, determinar estructura de cromatina (ChIP-Seq, Methyl-seq, DNase-seq), entre otras aplicaciones más utilizadas. Los métodos de secuenciación de NGS se diferencian principalmente por la forma en que se preparan las muestras de ADN o ARN y las opciones de análisis de datos utilizadas.

A continuación, se detallan las metodologías de secuenciación Ion Torrent, Illumina y SMRT de PacBio, ya que fueron las utilizadas en este trabajo.

1.1 Secuenciación por semiconducción mediante Ion Torrent

Esta tecnología involucra la utilización de un chip de secuenciación, formado por micropocillos, los cuales contienen microesferas con grupos o *clúster* de ADNc (previamente fragmentados, purificados y amplificados), ligados a las mismas mediante adaptadores (Fig. 8). Los nucleótidos se van agregando de forma secuencial a los pocillos, en el caso de que el nucleótido sea incorporado, se libera un protón, lo cual provoca un cambio en el pH del pocillo, el cual es detectado como un cambio de voltaje en cada clúster. Este cambio es detectado por un sensor o transistor de efecto de campo sensible a iones (ISFET, de sus siglas en inglés *Ion-Sensitive Field-Effect Transistor*)

acoplado a un circuito integrado, ubicado en el fondo de cada micropocillo y convertido en una señal de voltaje que es proporcional al número de nucleótidos incorporados. Para la identificación del tipo de nucleótido agregado, inicialmente se adiciona solo un tipo de base en el pocillo y el sensor ISFET detectará o no el voltaje dependiendo de la complementariedad de la base. Si esta no es complementaria, se agrega otro nucleótido hasta que se detecte voltaje. Este tipo de secuenciación difiere de los demás tipos en que no se usan nucleótidos modificados químicamente, la detección no se realiza por métodos ópticos, sino por detección de cambios en el pH, siendo una tecnología de secuenciación más simple y rápida (< 3 h/corrida). Por otro lado, presenta la limitación de la longitud de los fragmentos de ADN a ser secuenciados, la cual alcanza una longitud de lectura de hasta 400 pb y con un rendimiento que puede alcanzar los 10 Gb por corrida.

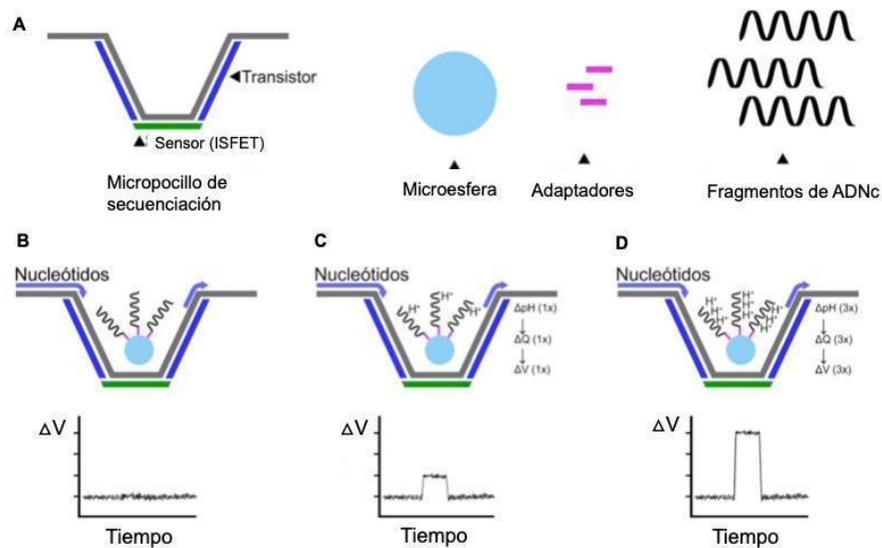


Figura 8. Esquema del proceso de secuenciación por semiconducción en Ion Torrent.

Se indican en: (A) La tecnología de secuenciación utiliza transistores de campo sensible a iones (ISFET) en los micropocillos de secuenciación, donde cada microesfera, mediante la adición de adaptadores a un extremo, tiene fragmentos de igual tamaño unido. La secuenciación ocurre mediante el agregado de forma secuencial y controlada de cada nucleótido a los micropocillos, al registrarse un cambio de voltaje (ΔV), debido al cambio de pH (ΔpH). (B) Si en un micropocillo no se adicionan nucleótidos a la cadena naciente, no se registra un ΔV (B), a diferencia de aquellos micropocillos donde un nucleótido es adicionado (C). Cuando en un micropocillo se adiciona más de un nucleótido debido a repeticiones continuas de la misma

base en la secuencia, el ΔV será proporcional al número de nucleótidos adicionados (D). Luego de haber registrado la presencia o ausencia del ΔV en un ciclo, se lavan las moléculas del nucleótido que no fueron adicionados y el ciclo puede repetirse mediante la inyección de un nucleótido diferente. Tomado y adaptado de [175].

1.2 Secuenciación por síntesis (SBS) mediante Illumina

Esta tecnología de secuenciación por síntesis (SBS, de sus siglas en inglés *Sequencing by Synthesis*), desarrollada por la compañía Solexa, que actualmente pertenece a Illumina, utiliza una celdilla o superficie de secuenciación la cual tiene ligado, mediante adaptadores, fragmentos de ADNc, los cuales fueron previamente fragmentados y purificados (Fig. 9). La técnica utiliza clúster de ADN amplificados y nucleótidos modificados con fluoróforos. Cada ciclo de secuenciación identifica los nucleótidos incorporados en cada grupo. Después de capturar imágenes, se remueven los fluoróforos y se continúa el proceso de secuenciación. Esta secuenciación permite la obtención de salida de datos que van desde 300 kb (kilobases) hasta múltiples Tb (terabases) en una sola corrida, dependiendo del instrumento [176].

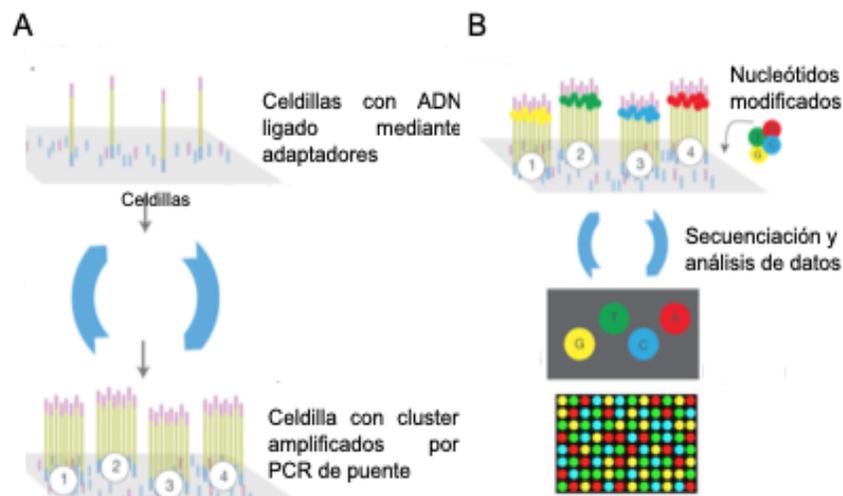


Figura 9. Esquema del sistema de secuenciación por síntesis utilizado por Illumina.

Se indican en: (A) La tecnología utiliza celdillas que contienen fragmentos de ADN ligados mediante adaptadores, los cuales son amplificados por PCR de puente, generando grupos o *clústeres*. (B) La adición de nucleótidos fluorescentes, modificados con un terminador limita cada ciclo de síntesis a la adición de un solo nucleótido por fragmento a secuenciar. Luego de que el nucleótido se agregó a la cadena nascente, se capturan cuatro imágenes para identificar al nucleótido que se agregó en cada clúster. Una vez que se

resuelven los nucleótidos adicionados en un ciclo de secuenciación en cada clúster, el fluoróforo, así como el bloqueo del sitio 3' se escinden químicamente y se procede a un nuevo ciclo de secuenciación. Tomado y adaptado de [176].

1.3 Secuenciación *SMRT-seq* de PacBio

La secuenciación en tiempo real de una sola molécula (SMRT, de sus siglas en inglés *Single Molecule Real Time Sequencing*), desarrollada por *Pacific BioSciences* (PacBio), es una tecnología de secuenciación definida como de tercera generación (TGS) [177]. Esto es debido a que captura la información de la secuencia durante el proceso de replicación de la molécula de ADN objetivo y no requiere una pausa entre los pasos de lectura (Fig. 10). La secuenciación ofrece lecturas más largas (entre 40 y 60.000 nt) y en menor tiempo que los métodos SGS, pero se ve obstaculizada por un rendimiento más bajo, una tasa de error más alta y un costo por nucleótido más alto. A su vez, a diferencia de la secuenciación por SBS (Illumina) o por Ion Torrent, no requiere la generación de clústeres.

Se utiliza una plataforma de secuenciación, la cual está equipada con un dispositivo llamado Chip de célula única (SMRTbell, de sus siglas en inglés *Single Molecule Real-Time Cell*), que contiene micro pozos donde se capturan moléculas de simple cadena de ADN. Este es un ADN circular monocatenario cerrado que se crea ligando adaptadores de horquilla a ambos extremos (de secuencia conocida) del fragmento de ADN a secuenciar (ADNc), con lo cual, al desnaturalizarse, dicho fragmento queda circularizado convirtiéndose en una molécula simple cadena. La SMRTbell es colocada en un chip llamado "celda SMRT" (Fig. 10), el cual difundirá en una unidad de secuenciación, denominada guía de ondas de modo cero (ZMW, de sus siglas del inglés *Zero Mode Waveguides*). Esta unidad de secuenciación contará con la ADN polimerasa, la cual reconoce los adaptadores y un conjunto de 4 nucleótidos marcados con fluoróforos en los extremos trifosfatos. El fluoróforo se escinde tras la formación del enlace fosfodiéster durante la síntesis del DNA y se registrará la emisión de fluorescencia del nucleótido que se está adicionando. Los procesos de síntesis en todos los ZMW de una celda SMRT se registran mediante una "película" de pulsos de luz, y los pulsos correspondientes a cada

ZMW se pueden interpretar como una secuencia de bases llamada lectura continua larga (CLR, de sus siglas en inglés *Long Continuous Reading*).

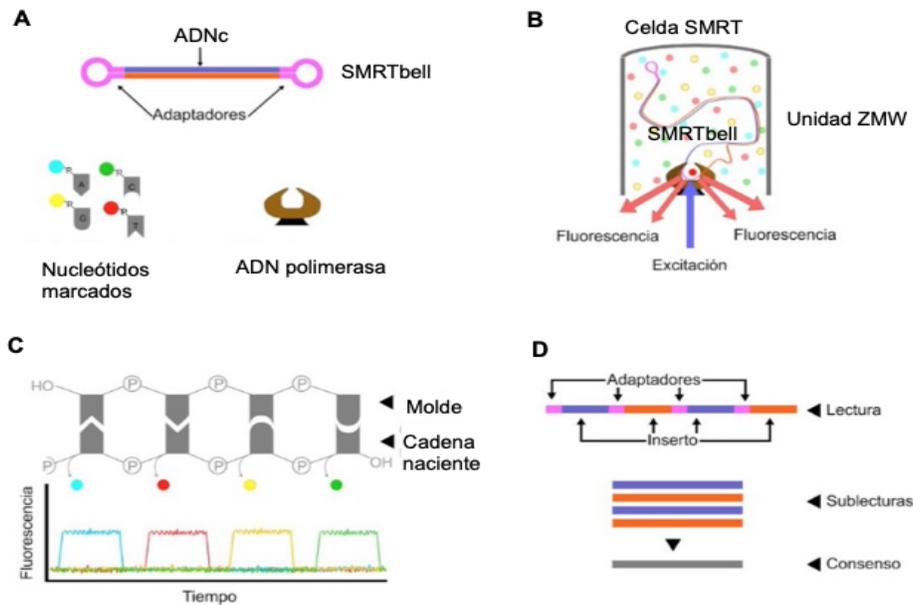


Figura 10. Secuenciación de molécula única en tiempo real (SMRT).

Se indican en : (A) La tecnología de secuenciación SMRT utiliza nucleótidos marcados con un fluoróforo diferente en el extremo trifosfatado, adaptadores que circularizan al fragmento de ADNc a secuenciar y una ADN polimerasa, (B) Un SMRTbell difunde en una unidad de secuenciación ZMW y el adaptador se une a la polimerasa inmovilizada en la parte inferior, (C) La secuencia se obtiene al excitar con láser a los fluoróforos que marcan a cada nucleótido, y debido a que la ADN polimerasa inmoviliza al nucleótido modificado mientras este se incorpora a la cadena naciente, el fluoróforo de este nucleótido emitirá fluorescencia durante más tiempo que los nucleótidos que difunden libremente en la solución, (D) Luego de la remoción de los adaptadores las sublecturas se alinean para obtener una secuencia consenso. Tomada y adaptada de [178].

1.4 Estructura de los genomas de rizobios

Los genomas bacterianos se encuentran mayoritariamente constituidos por un cromosoma principal, que incluye la mayoría de los genes necesarios para el mantenimiento de las funciones celulares básicas (genes esenciales o genes *housekeeping*, del inglés).

Además del cromosoma principal, se ha demostrado que muchas bacterias poseen otras moléculas de ADN con la capacidad de replicarse, denominadas replicones secundarios. Los replicones secundarios incluyen a los plásmidos, megaplásmidos y crómidos, los

cuales se clasifican de acuerdo con su tamaño y a su contenido génico. Las moléculas de ADN de tamaño menor a 350 Kb se denominan plásmidos, mientras que los replicones de tamaño mayor a 350 kb se denominan megaplásmidos, y si además contienen genes esenciales, se denominan cromosomas secundarios o crómidos [179].

En general los crómidos tienen una composición de nucleótidos y un uso de codones muy similares al de los cromosomas con los que están asociados [180]. Tanto crómidos como megaplásmidos poseen sistemas de replicación y partición similares a los encontrados en plásmidos. Estos replicones presentan bajo número de copias y contienen sistemas de replicación y mantenimiento que aseguran su partición a las células hijas tras la división celular. Estos loci incluyen los genes *repA*, *repB* y *repC*, los cuales codifican para proteínas involucradas en la replicación y en un origen de replicación (*oriV*) [181]. A su vez, la presencia de los genes *parB* y *parA*, los cuales codifican para una proteína que se une a un sitio similar a un centrómero *parS*, y para una ATPasa (la cual separa y dirige las moléculas del plásmido hacia cada polo celular), se ha reportado [182]. Además, la presencia de sistemas de toxina-antitoxina que incluyen los sistemas Phd (proteína asociada a la prevención de la muerte del huésped) - Doc (proteína de muerte al curar) y VapC (proteína asociada a la virulencia C) - (antitoxina para VapC), contribuyen al mantenimiento y estabilidad de los plásmidos, y lo realizan mediante un mecanismo general en el cual la unión de la toxina lábil a la antitoxina estable, evita su pérdida o degradación [183].

Aproximadamente un 10% de las bacterias presentan genomas multipartitos, formados por cromosomas y replicones secundarios, siendo esta característica particularmente común en miembros de las familias *Rhizobiaceae* y *Burkholderiaceae* [184]. En estas bacterias el tamaño de los genomas suele ser casi el doble o más del tamaño medio del genoma bacteriano, lo cual podría reflejar la adaptación a entornos complejos y altamente competitivos tales como el suelo y/o la rizosfera [179]. En dicho contexto, la presencia de replicones secundarios le confiere funciones accesorias, específicas o asociadas al nicho, como por ejemplo la presencia de genes relacionados con la resistencia a metales [185], la degradación de compuestos aromáticos [186], relacionados a la simbiosis y a la FBN [187], [188].

1.4.1 Ubicación y contenido de las islas simbióticas

Los genes bacterianos necesarios para establecer una relación simbiótica con leguminosas hospederas se conocen como genes simbióticos e incluyen tanto a los genes necesarios para el reconocimiento entre los simbioses (genes *nod*, *nol*, *noe*) como aquellos implicados en la fijación biológica del nitrógeno (genes *nif*, *fix*, *fdx*) [189]. En algunos α -rizobios como *S. meliloti* 1021 y *R. leguminosarum* bv. *viciae* 3841, los genes simbióticos se encuentran en megaplásmidos, denominados “plásmidos simbióticos” pSymA (1,35 Mpb) y pRL10 (0,49 Mpb) respectivamente. Mientras que, en otros rizobios, como ser *Bradyrhizobium*, *Mesorhizobium* y *Azorhizobium*, los genes simbióticos se encuentran asociados a su cromosoma, y se encuentran agrupados en regiones específicas denominadas “islas simbióticas”, formando parte de elementos conjugativos integradores (ICE, de sus siglas en inglés *Integrative Conjugative Elements*) [190].

A pesar de que es común encontrar a los genes simbióticos agrupados formando estas “islas”, existe una variación considerable en el número y contenido génico entre los diversos rizobios (Fig. 11) [31]. En particular considerando los genes *nod*, *nol* y *noe*, esto se refleja en diferentes sustituyentes y modificaciones de los NFs, lo cual afecta la especificidad de la interacción con plantas huéspedes [106]. Con respecto a los genes *nif*, *A. caulinodans* tienen 15 genes, *S. meliloti* y *R. leguminosarum* bv. *viciae* tienen un conjunto restringido de 9 y 8 genes *nif*, respectivamente [191]. En base a estas diferencias, y que aún en la ausencia de alguno de esos genes, las bacterias siguen siendo capaces de ejercer su función como rizobio, hasta la actualidad no existiría un conjunto mínimo de genes que lo definan como tal.

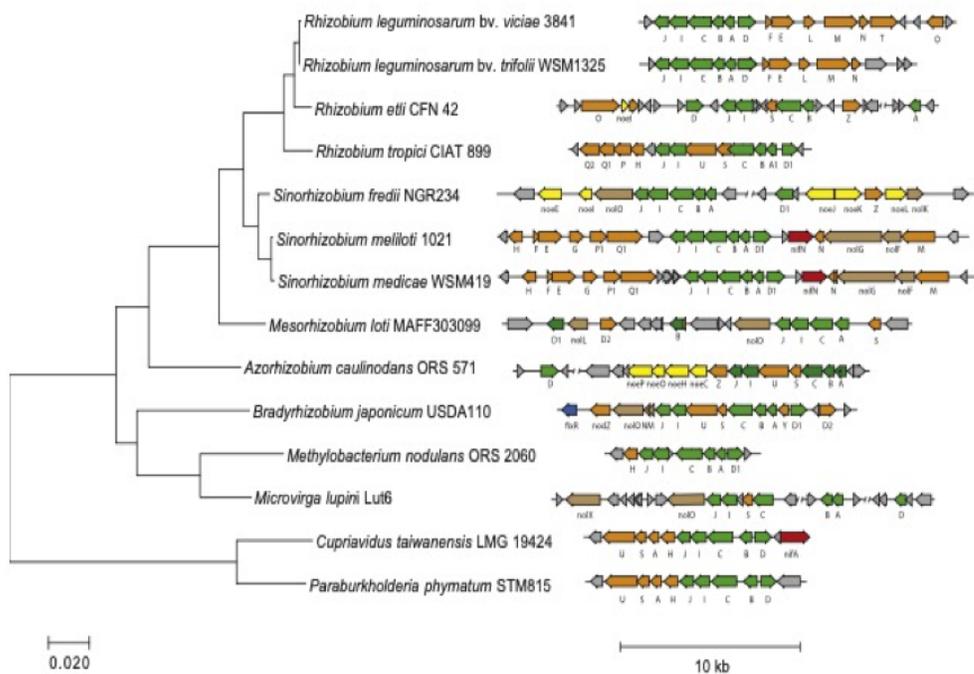


Figura 11. Organización y contexto genómico de los genes *nod*, *nol* y *noe* de diferentes especies de rizobios.

A la izquierda se muestra un árbol filogenético construido por máxima verosimilitud en base a los genes simbióticos y a la derecha los órdenes y agrupamientos de los genes involucrados en la simbiosis. Los genes *nod* centrales están coloreados en verde, los genes *nod* accesorios en naranja, los genes *noe* en amarillo y los genes *nol* en marrón. Tomado de [192].

1.4.2 Genes simbióticos en β -rizobios

En todos los β -rizobios que se han analizado hasta la fecha, los genes simbióticos se encuentran agrupados en una isla simbiótica compacta y conservada (Fig. 12), ubicada en megaplásmidos. Particularmente, la cepa LMG19424 de *C. taiwanensis* presenta un plásmido simbiótico de 560 kb (pRalta), que incluye una de las “islas simbióticas” más compactas conocidas (35 kb), compuesta por los genes *nodQUSAHJICBD*, *nifAENXQWVBZHDK*, y *fixXCBAU* [164]. De las 583 secuencias codificantes presentes en el plásmido pRalta, 207 son secuencias de elementos transponibles (*TE*, de sus siglas en inglés *Sequences of Transposable Elements*), siendo una característica que también se observa en muchos de los plásmidos pSym y en los ICE de los α -rizobios, lo que es

consistente con la movilidad de estas regiones y su herencia a través de la transferencia lateral o horizontal de genes [164].

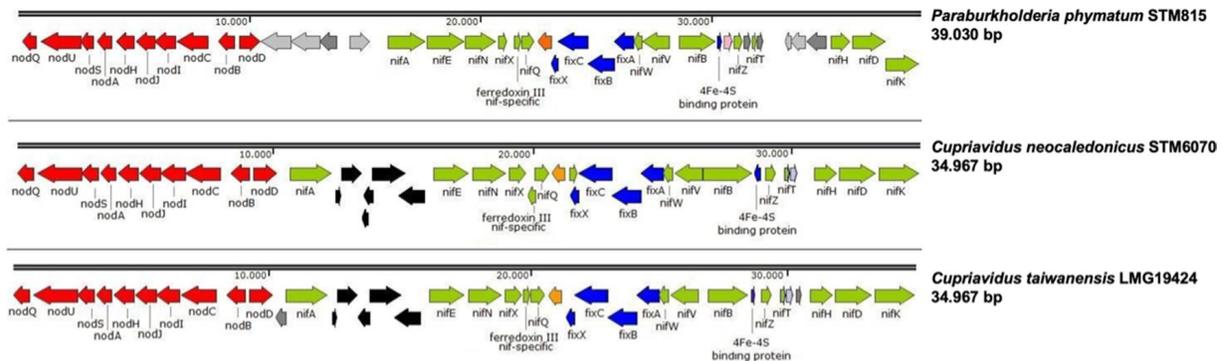


Figura 12. Comparación de islas simbióticas de diferentes β -rizobios.

El tamaño de cada isla simbiótica se indica debajo de los nombres de las cepas. Los genes pertenecientes a los operones *nod/nif/fix* se indican en rojo/verde/azul respectivamente, y los genes relacionados con transposones se colorearon en negro. Otros colores representan genes que no pertenecen a los operones mencionados anteriormente. Tomado de [193].

Si bien en los α -rizobios están presentes los genes que conforman el operón *nif*, en β -rizobios no se han encontrado homólogos a los genes *fixNOQP* y *fixGHIS*, los cuales codifican para un citocromo oxidasa terminal que funciona a bajas concentraciones de oxígeno y para *fixL*, el cual codifica para una proteína sensora de oxígeno, en los genomas de β -rizobios [194], [195], [196]. Sin embargo, en el género *Cupriavidus*, se encontró que el cromosoma de LMG de *C. taiwanensis* contiene un operón, *ccoNOQP*, que codifica para una oxidasa de tipo *cbb3* con una alta afinidad por el oxígeno que podría cumplir la misma función de *fixNOQP* [194]. Contrariamente, no se detectó la presencia de ninguna oxidasa de tipo *Cbb3* en los genomas de rizobios del género *Parabrakholderia* [195]. Sin embargo, el análisis del transcriptoma de nódulos de *Phaseolus vulgaris* (frijol común) infectados por *Parabrakholderia phymatum* STM815, reveló una fuerte inducción de los genes que codifican para un citocromo O ubiquinol oxidasa, sugiriendo que este sistema sería importante para la respiración de los rizobios dentro del nódulo. A su vez, el análisis transcriptómico de una mutante en *cyoB* (parte del complejo citocromo O ubiquinol oxidasa), mostró una menor expresión de la nitrogenasa en comparación con la cepa salvaje, apoyando la importancia de *cyoB* en el establecimiento de una simbiosis efectiva [197].

1.5 Genes relacionados con la interacción planta-microorganismo

1.5.1 Quimiotaxis y movilidad del flagelo

La quimiotaxis es un mecanismo mediante el cual las bacterias detecta sustancias a diferentes concentraciones denominados quimioattractantes, y en respuesta ajustan su movimiento, dirigiéndose hacia los que les resultan favorables o alejarse de los que les resultan perjudiciales. Estos compuestos pueden ser moléculas liberadas por otros organismos (fuentes de carbono y nitrógeno), aceptores finales de electrones, iones, neurotransmisores, moléculas de señalización de quórum, así como hormonas vegetales [198].

En proteobacterias, la quimiotaxis se lleva a cabo a través de los quimiorreceptores sensores denominados proteínas de quimiotaxis aceptoras de metilo (MCP, de sus siglas en inglés *Methyl Chemotaxis Proteins*). Estos quimiorreceptores están formados por un dominio periplásmico como ligando y un dominio de señalización que funciona como un sitio de unión citoplásmico y por las proteínas CheA y CheW [199]. Estas proteínas forman un complejo ternario en el que la unión de una señal al quimiorreceptor modula la actividad de CheA, lo que resulta en la fosforilación de CheY. Luego CheY-P se une al motor flagelar y cambia su dirección de rotación, lo que afecta el comportamiento de natación de la bacteria para navegar en un gradiente químico. La sensibilidad del sistema se ajusta mediante la acción de las enzimas CheR y CheB, esenciales para la quimiotaxis, y la eliminación o inactivación de estas enzimas interrumpe este proceso [200].

La quimiotaxis, la movilidad y la adherencia a diferentes superficies son mediadas por estructuras celulares de la bacteria, como los flagelos y los pilis (*T-pilus*), y en proteobacterias se ha reportado que los genes *che*, *mot* y *fli* codifican para proteínas que participan en estos procesos [201].

1.5.2 Sistemas de secreción

Los sistemas de secreción (SS) bacterianos son complejos proteicos que se encuentran en la membrana celular de las bacterias y desempeñan un papel crucial en la secreción

de moléculas, de modo que puedan interactuar con el medio extracelular o con otras células.

Hasta la fecha, se han descrito 11 familias diferentes de SS (SSTI-SSXI) en bacterias [202]. Los efectores a ser translocados suelen ser proteínas, enzimas, complejos nucleoproteicos, ADN solo o pequeñas moléculas, dependiendo del sistema. Estos sustratos pueden ser translocados al medio extracelular o inyectados en el citoplasma de una célula receptora, que puede ser eucariota o procariota. A continuación se describirán de forma resumida los sistemas de secreción bacterianos más estudiados [203] (Fig. 13).

El SSTI es un complejo heterotrímico, ejemplificado por el sistema de hemolisina, que involucra tres componentes. La proteína HlyB, una proteína de múltiples dominios transmembrana, ubicada en la membrana interna que lleva un dominio de unión y de hidrólisis de ATP en su lado citosólico. La proteína de membrana externa TolC, y una proteína de fusión de membrana periplásmica HlyD, la cual interactúa tanto con HlyB, y con TolC de modo de generar un canal continuo que se extiende desde el citosol hasta el medio extracelular, evitando por completo el periplasma [204]. Los sustratos a ser transportados por este sistema no llevan péptidos señales dependientes del sistema Sec, sino que en su lugar tienen una señal de direccionamiento no escindible en el extremo c-terminal que los dirige hacia HlyB. Proteínas, enzimas, toxinas y factores de virulencia pueden ser transportados por este sistema [203].

El SSTII está codificado por 13 genes centrales que conforman la vía secretora general (GSP), generalmente formando el operón, *gsp* [205]. El sistema se puede subdividir en tres subcomplejos: (1) un complejo de la membrana externa, que incluye el canal secretor (conocido como secretina) formado por GspD y GspS; (2) un complejo de la membrana interna llamado plataforma de ensamblaje, que incluye GspC, GspF, GspL y GspM, GspN, una ATPasa hexamérica citoplasmática GspE; (3) el pseudopilus, que está formado por la pseudopilina principal GspG y las pseudopilinas menores GspH, GspI, GspJ y GspK [203]. A su vez, la prepilina peptidasa GspO, necesaria para el procesamiento proteolítico de las pseudopilinas GspGHIJK a una forma madura lista para

el ensamblaje del pseudopilus. Otras dos proteínas GspA y GspB, facilitan la correcta localización de GspD en la membrana externa. Estas proteínas colaboran en la transferencia de GspD desde la membrana interna hasta la membrana externa, asegurando su correcta posición y función en el complejo de la membrana externa del sistema [206].

A su vez, el SSTII incluye otras dos vías de exportación de proteínas, las cuales se denominan dependientes de Sec (transporta sustratos no plegados) e independientes de Sec o conocida como vía de translocación de arginina gemela (TAT, de sus siglas en inglés *Twin Arginine Translocation*), la cual transporta sustratos plegados a través de la membrana interna. La vía Sec consiste en un canal de translocación de proteínas formado por el contacto de dos complejos heterotriméricos SecYEG y SecDFYajC, los cuales estarían involucrados en la translocación de las moléculas desde el citoplasma hacia el periplasma. Por otro lado, el multicomplejo TatABC, está formado por la proteína TatA que forma el canal de transporte mientras que las proteínas TatBC actúan como un complejo receptor que reconoce el péptido señal de la proteína a translocar [207].

Cabe destacar que el SSTII, a pesar de transportar moléculas efectoras hacia el espacio extracelular, forma parte de una superfamilia mucho más amplia de sistemas que comparten componentes homólogos, que incluyen el *pilus* de tipo IVa y IVb [208], el *pilus* de adherencia fuerte (Tad) ([209], el *pilus* de hemaglutinina sensible a manosa [210] y el *pilus* de competencia [211].

El SSTIII es una nanomáquina compleja compuesta por proteínas que forman una estructura conocida como inyectorio que crea poros en la membrana del huésped. A través de estos poros, con la ayuda de una estructura filamentosa "tipo aguja" ubicada en la membrana interna, las moléculas efectoras pueden ingresar al citosol de la célula huésped [212]. Su principal mecanismo de secreción ha evolucionado a partir del SSTIII responsable del autoensamblaje de los flagelos bacterianos [213], llegando a secretar una diversidad mucho mayor de proteínas sustrato. Uno de los primeros sistemas más estudiados ha sido el de especies patógenas del género *Yersinia* (genes *ysc*) [214] a partir de los cuales se han encontrado homólogos en otras bacterias. Además de ser importantes en la virulencia en patógenos bacterianos, también son utilizados por

algunas cepas de rizobios en interacciones simbióticas con sus leguminosas hospederas [215].

Los SSTIV son sistemas de transporte de macromoléculas que pueden ser utilizados por las bacterias para enviar material genético o proteínas a células huésped. Se dividen en dos subfamilias principales: sistemas conjugativos y sistemas patogénicos, formados por varias proteínas con diferentes funciones [216]. Los sistemas están compuestos por proteínas nombradas de VirB1 a VirB11, que se pueden dividir en tres categorías: proteínas citoplasmáticas asociadas a la membrana interna con actividad ATPasa (VirB4, VirB11, VirD4), proteínas que forman el canal o *core* (VirB6, VirB8, VirB10, las más conservadas) y proteínas que se encuentran en la membrana externa formando la estructura conocida como *pilus* o *T-pilus* (VirB2, VirB3, VirB7 y VirB9) [217].

Los SSTV difieren de otros sistemas de secreción en el sentido de que se secretan a sí mismos o se autotransportan a través de la membrana externa, a expensas de la hidrólisis de ATP. Se clasifican en subtipos del Va al Ve, y comparten la presencia de una o dos proteínas. Las proteínas o "dominios pasajeros" son transportados exclusivamente a través de la membrana externa, después de la exportación dependiente de Sec o Tat al periplasma [218]. En los subtipos Va, Vc, Vd y Ve, el dominio pasajero está fusionado al dominio de la membrana. Particularmente, el subtipo Vb, presenta la vía de secreción de dos componentes (TPS, de sus siglas en inglés *Two Protein Secretion*), compuesta por la proteína pasajero genéricamente llamada TpsA, la cual está separada de su transportador TpsB [219]. En dicho contexto, se ha reportado que distintos sistemas TPS secretan dominios pasajeros o efectores, con actividad enzimática (lipasas, proteasas), adhesinas involucradas en la adhesión a superficies del huésped o receptores, y colonización mediante formación de biopelículas [220].

El SSTVI ha sido estudiado recientemente y gran parte de la información disponible se ha obtenido a través de análisis de homología de secuencias *in silico*. Fue descubierto inicialmente en bacterias patógenas, como *E. coli*, así como en otras proteobacterias, incluyendo *P. aeruginosa*, *V. cholerae* y *B. cenocepacia* y presenta una alta similitud con los mecanismos de acción de los fagos [221]. El sistema está compuesto por 13 proteínas

estructurales centrales codificadas por los genes *tss*, y por genes accesorios denominados *tags* [222]. Su estructura básica consta de un complejo transmembrana (TssMLJ), una placa base citoplasmática (TssEFGK), un doble tubo, donde la parte exterior está formada por una vaina contráctil (TssBC) y la parte interior por la proteína Hcp (proteína coregulada con hemolisina) (TssD). Esta proteína termina en el dispositivo de inyección formado por la proteína VgrG y una pequeña proteína que tiene un dominio característico PAAR (dominio prolina-alanina-alanina-arginina), el cual facilita la secreción de efectores, como proteínas con actividad lipolítica (lipasas), DNasas y peptidoglicanasas, desde la membrana externa hacia la célula diana.[223]. Una ATPasa, codificada por *tssH* (ClpV), desarma la estructura del SSTVI para su reciclaje.

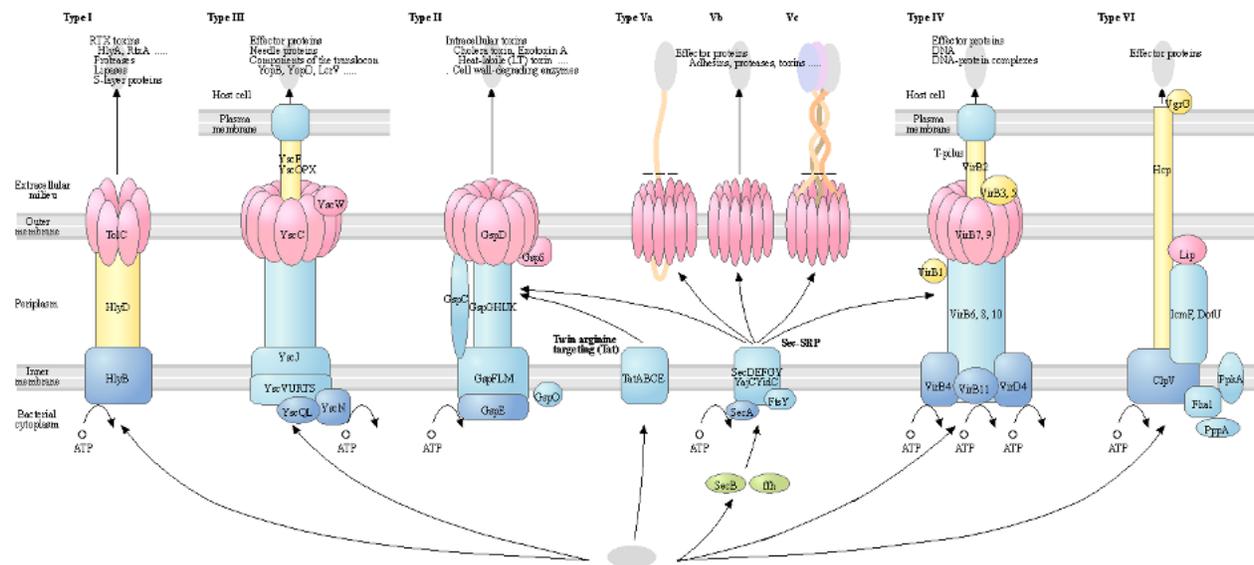


Figura 13. Ilustración resumida de los principales sistemas de secreción bacterianos.

Se indican los mecanismos utilizados por los diferentes sistemas de secreción bacterianos, las proteínas involucradas y su ubicación celular y moléculas efectoras a ser transportadas. Imagen obtenida de la anotación realizada con el servicio web KAAS y la base de datos KEGG. Tomado y adaptado de [224].

2 Materiales y métodos

2.1 Clasificación de la cepa de UYPR2.512 de *Cupriavidus* sp., en base a la información genómica disponible

Se descargaron todos los genomas del género *Cupriavidus* disponibles en el NCBI (al 2018) así como algunos genomas de miembros representativos de la subfamilia beta

proteobacterias. En total se incluyeron 71 (Tabla suplementaria S1 del Anexo) genomas que comprenden 40 especies. Las secuencias codificantes de los 71 genomas fueron traducidas a aminoácidos usando la herramienta *transeq* incluida en el paquete de EMBOS [225]. Los genes ortólogos entre los distintos genomas fueron identificados aplicando un umbral mínimo de identidad del 40% usando el algoritmo OrthoMCL v1.4, incluida en el programa *get_homologues.pl* v23102017 [226]. Se seleccionaron 22 grupos de proteínas ortólogas, las cuales se alinearon utilizando el método *G-INS-i* con el programa *MAFFT* v7.182 [227]. Las regiones de baja homología se eliminaron utilizando el programa *Gblocks* [228]. Para cada uno de los grupos se construyó un árbol filogenético empleando el método de máxima similitud (*Maximum Likelihood*) con el programa *RAxML* v8.2.3, usando la opción *PROTGAMMAAUTO* [229]. Para cada árbol obtenido, se estimó la significancia de los nodos mediante la realización de 1000 *bootstrap*. Finalmente, el árbol de especies se obtuvo con *ASTRAL-III* [230]. Los miembros del género *Cupriavidus* potencialmente noduladores se identificaron por la presencia de los genes *nod*, empleando la herramienta *BLASTp*.

Luego se seleccionó al menor grupo monofilético que contuviera a todas las cepas nodulantes dentro del género de los *Cupriavidus*, realizándose una nueva reconstrucción filogenética con la misma metodología descrita previamente, pero esta vez empleando 1.155 grupos de ortólogos identificados entre estos genomas.

De forma complementaria se realizó el análisis de la identidad media de nucleótidos (ANI, de sus siglas en inglés *Average nucleotide identity*) entre los genomas de UYPR5.512 y 18 genomas de cepas y especies más cercanas filogenéticamente (Tabla 4). Dicho análisis se realizó empleando el script *ani.rb* disponible en enveomics.blogspot.com desarrollado por L.M. Rodríguez. Estos análisis fueron realizados en colaboración con el Laboratorio de Biología Computacional, del Departamento de Desarrollo Biotecnológico (Instituto de Higiene, Facultad de Medicina), por la estudiante Lucía Balastrazzi.

Tabla 3. Cepas del género *Cupriavidus* utilizadas para el análisis ANI.

Nombre de la cepa	Identificador del genoma en el banco de datos del NCBI
<i>C. necator</i> N1 ^T	GCA_000219215.1
<i>C. necator</i> A5-1	GCA_000744095.1
<i>C. necator</i> PHE3-6	GCA_001853325.1
<i>C. necator</i> H16	GCA_000009285.1
<i>C. necator</i> NBRC 102504	GCA_001598755.1
<i>C. pinatubonensis</i> JMP134	GCA_000203875.1
<i>Cupriavidus</i> sp. UYMMa02A	GCA_001725945.1
<i>Cupriavidus</i> sp. amp6	GCA_000426345.1
<i>Cupriavidus</i> sp. NH9	GCA_002011925.2
<i>Cupriavidus</i> sp. UYPR2.512	GCA_000379565.1
<i>Cupriavidus</i> sp. SK4	GCA_000611125.1
<i>Cupriavidus</i> sp. GA3-3	GCA_000389805.1
<i>Cupriavidus</i> sp. IDO	GCA_000812465.1
<i>C. taiwanensis</i> LMG19424	GCA_000069785.1
<i>C. taiwanensis</i> STM6018	GCA_000472465.1
<i>C. taiwanensis</i> STM6070	GCA_000372525.1
<i>C. necator</i> NBRC 102504	GCA_001598755.1
<i>C. alkaliphilus</i>	GCA_900094595.1
<i>C. nantogensis</i>	GCA_001598055.1

2.2 Obtención del ADN genómico y secuenciación del genoma de UYPR2.512

La cepa UYPR2.512 de *C. necator* se cultivó en medio LB (Medios de cultivos y soluciones de Anexos) durante 18 h a 30 °C, con un agitador rotatorio a 200 rpm. Los cultivos se centrifugaron a 5,000 x g durante 10 min y el ADN se aisló y purificó mediante la utilización del kit de aislamiento y purificación comercial Genomic-tip 100/G (QIAGEN, Alemania), de acuerdo con las recomendaciones del fabricante. Se eligió este método dado que se minimiza la agitación durante la purificación, lo que permite obtener

fragmentos de ADN de tamaños de entre 50 y 150 Kb, lo cual es necesario para obtener lecturas de buena calidad al secuenciar con la tecnología SMRT.

La calidad y pureza del ADN genómico obtenido fue verificada mediante la realización de geles de agarosa 1%. El ADN genómico purificado fue secuenciado por MacroGen (MacroGen Inc, Corea) con la tecnología SMRT cell 8Pac V3 (Pacific Biosciences) (versión 2) y utilizando el kit de unión de polimerasa de ADN P6 (PacBio) y el reactivo de secuenciación de ADN 4.0 v2.

2.2.1 Ensamblaje y anotación del genoma de UYPR2.512 de *C. necator*

El ensamblaje del genoma se realizó mediante el software Canu [231] y la anotación se realizó en el servidor web RAST (*Rapid Annotations Subsystems Technology*) [232]. El análisis se complementó con la anotación mediante el servidor online KAAS (*KEGG Automatic Annotation Server*). Este servidor proporciona anotaciones funcionales de genes mediante comparaciones BLAST (búsqueda de nucleótidos) contra la base de datos curada *Kyoto Encyclopedia of Genes and Genomes* (KEGG). El resultado contiene asignaciones de KO (*KEGG Orthology*) y vías metabólicas KEGG generadas automáticamente [233]. A su vez, utilizando la herramienta BLASTn, BLASTx y BLASTp del NCBI (*National Center for Biotechnology Information*), se realizó una búsqueda por homología de secuencia, de genes y/o proteínas claves que codifican y/o intervienen para otras actividades de interés que no fueron anotadas por el servicio web RAST o KAAS.

Para realizar la representación circular del genoma bacteriano ensamblado, se utilizó el software Circos [234]. La secuencia completa del genoma de la cepa UYPR2.512 de *C. necator* se encuentra disponible en el NCBI con ID de BioProyecto PRJNA720661, y número de acceso CP074381–CP074390.

2.2.1.1 Búsqueda de genes característicos del género *Cupriavidus*

Se buscó la presencia de los genes *crtR*, *ciA*, *copZ*, *cutA*, *cutE*, *corC*, *copS*, *chasE2*, *cusR/cusS*, *czsS/silS*, relacionados con la homeostasis del cobre. Así como también de los genes *bdhA*, *aaS*, *phaABC*, *phaF*, *EH*, y *HBP*, implicados en la síntesis y degradación

del PHB y los genes *ben*, *tmoABCDEF*, *pob*, *qui*, *bph*, *cat* y *pca*, involucrados en las principales rutas de degradación de diversos compuestos aromáticos [86].

2.2.1.2 Búsqueda de genes relacionados con la interacción planta-bacteria

Se buscó la presencia de genes involucrados en las etapas de reconocimiento y la colonización radicular, como ser la adhesión bacteriana y quimiotaxis (genes *che*), y relacionados a la síntesis y movilidad flagelar (*MCP*, *aer*, *flh*, *fli*, *mot*, *flg*). A su vez, se buscaron genes codificantes para los sistemas de secreción: SSTI (*tolC* y *hly*), SSTII (*gsp*, *sec* y *tat*), SSTIII (*ysc*), SSTIV (genes *pil*), SSTV (*Tps*), SSTVI (*Vgr*, *PAAR*, *clp*, *hcp*, *imp*, *lip*, *lcm*), entre los principales. También, se buscaron los genes implicados en la nodulación y en la fijación biológica del nitrógeno (*nod*, *fix*, *nif*, *noe*, *nol*), en la síntesis y degradación de fitohormonas (*aldh* y *iac*) y en la síntesis, adquisición y regulación de sideróforos (*pvs*, *exb* y *fur*). La organización genómica de algunos genes de interés se comparó con la organización de homólogos visualizado mediante el servidor RAST, presentes en otras β -rizobios y en bacterias del mismo género como ser la cepa tipo LMG 19424 de *C. taiwanensis*. La anotación de los genes y los estudios de homología fue realizada mediante la misma metodología que en el punto anterior.

3 Resultados

3.1 Clasificación taxonómica de la cepa UYPR2.512

El análisis de 71 genomas representantes de la familia *Burkholderiaceae* nos permitió identificar 22 proteínas ortólogas. En base a estos resultados, se construyó un árbol de especies que evidencia cuatro ramas conteniendo cepas de rizobios dentro del género *Cupriavidus* (Fig. 14). Estas ramas se encuentran a su vez en un grupo monofilético, separado de otro grupo de *Cupriavidus* que no tienen representantes rizobianos.

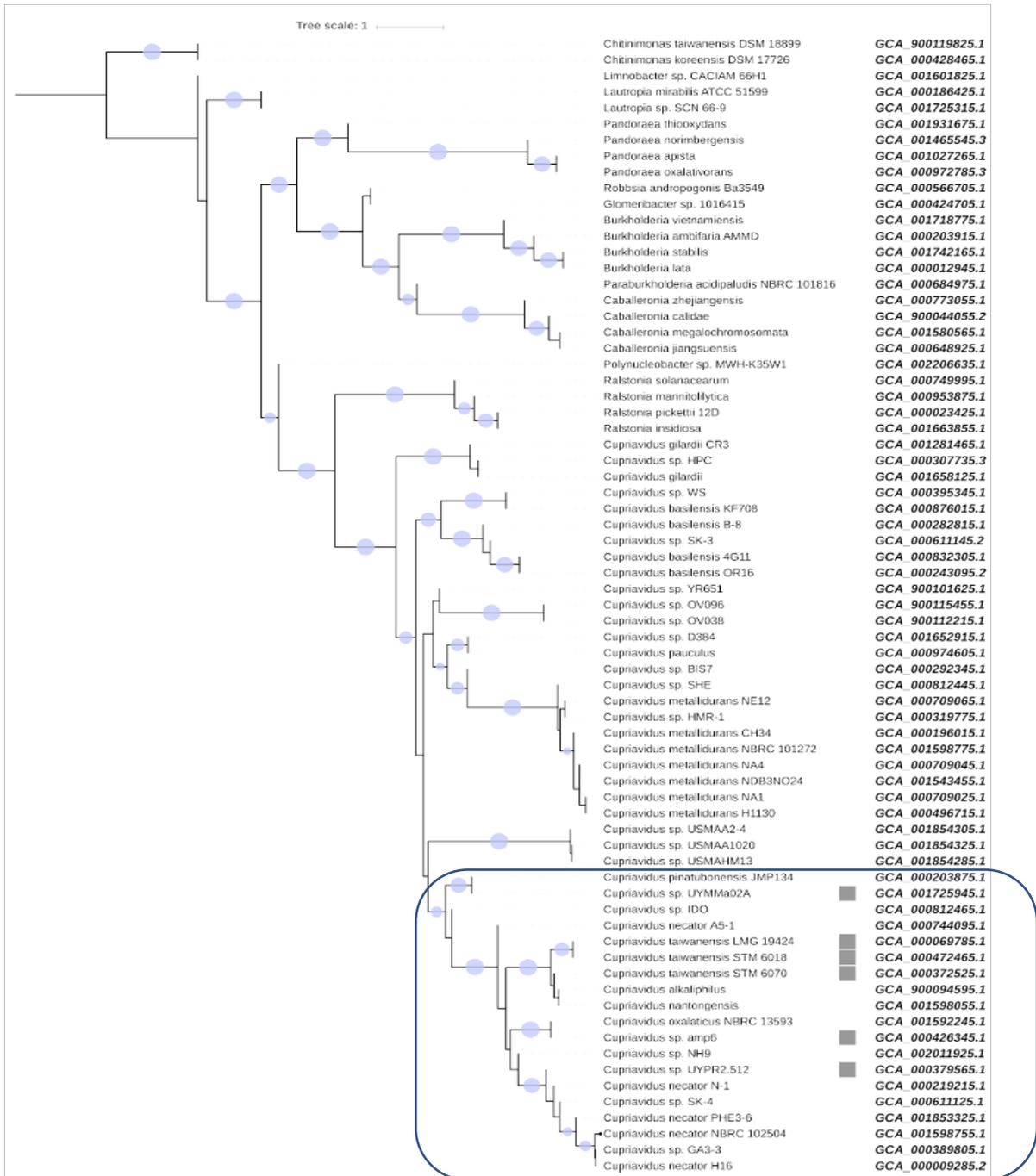


Figura 14. Árbol filogenómico del género *Cupriavidus* y cepas seleccionadas pertenecientes a la familia *Burkholderiaceae*.

Árbol consenso construido en base a 22 genes ortólogos. El recuadro azul indica una rama con representantes de *Cupriavidus* noduladores y no noduladores. Se indican con un cuadrado gris las cepas capaces de nodular. A la derecha se indican los identificadores de acceso según el Catálogo de genomas ensamblados (GCA, de sus siglas en inglés *Catalog of Assembled Genomes*) de cada uno de los genomas

utilizados provenientes de la base de datos del *NCBI*. Los círculos de color violeta indican las ramas que recibieron un soporte de *bootstrap* igual o mayores a 0.75.

Para mejorar la resolución de la filogenia, se seleccionaron las cepas de *Cupriavidus* que contenían los genes *nod* y se realizó una nueva búsqueda de ortólogos. Se obtuvieron en este caso un total 1.155 genes con los que construyó un nuevo árbol (Fig. 15). El análisis mostró que UYPR2.512 se encuentra en un clado junto con las cepas de *C. necator*, revelando por primera vez la existencia de cuatro clados dentro de los *Cupriavidus* que contienen rizobios.

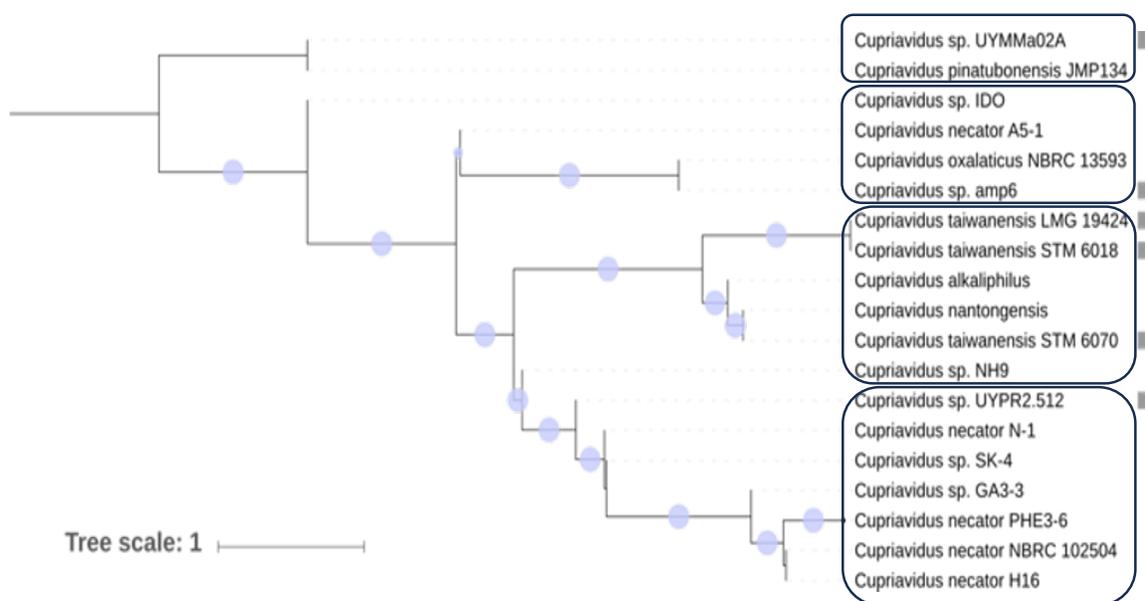


Figura 15. Relaciones filogenómicas entre cepas y especies de *Cupriavidus* que contienen rizobios. Árbol construido en base a secuencias proteicas de 1,155 genes ortólogos. Los recuadros grises indican los 4 clados que incluyen especies nodulantes y los cuadrados de color gris indican a las cepas capaces de establecer relaciones simbióticas. Los círculos de color violeta indican las ramas que recibieron un soporte de *bootstrap* igual o mayores a 0.75.

De forma complementaria al análisis filogenómico, se realizó el análisis de la identidad media de nucleótidos (ANI) entre los distintos genomas (Tabla 4). Tomando en cuenta que un valor de ANI mayor entre 95-96% indicaría que dos cepas pertenecen a la misma especie [235], la cepa UYPR2.512 mostró un valor de 96,26% con la cepa tipo de *C. necator*, por lo cual se sigue que pertenece a la especie *Cupriavidus necator* (Figura 15). A su vez, los resultados mediante ANI, sugieren que UYPR2.512 es la cepa más divergente dentro de las *necator*, ya que el resto presentan valores más cercanos y de

hecho se confirma al ser la más basal, según el estudio filogenómico (Fig. 15). Con respecto a las otras especies, por ejemplo en el caso de la cepa 6070 de *C. taiwanensis*, mostró un valor ANI semejante a otras dos cepas no pertenecientes a las *taiwanensis* (*C. alkaliphilus* y *C. oxalaticus*), en comparación a las cepas *taiwanensis* dentro del mismo clado. Lo cual sugirió, que la cepa 6070 de *C. taiwanensis*, estaría mal anotada.

Tabla 4. ANI entre los genomas de distintas cepas del género *Cupriavidus*.

Los valores de ANI en color negro indican valores de similitud menores al 94%, en color naranja de entre 94 y el 95%, mientras que los valores en color rojo indican similitudes mayores al 95%.

Strain	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 <i>C. pinatubonensis</i> JMP134 GCA_000203875.1	X	89.00	83.86	83.79	83.87	83.84	83.96	83.90	84.09	84.07	83.91	83.73	84.04	83.87	84.01	83.72	83.86	83.76	83.79
2 <i>Cupriavidus</i> sp. UYMMA02A GCA_001725945.1	89.00	X	84.56	84.43	84.35	84.31	84.29	84.38	84.62	84.57	84.40	84.33	85.28	84.43	84.58	84.29	84.49	84.34	84.44
3 <i>Cupriavidus</i> sp. IDO GCA_000812465.1	83.86	84.55	X	86.78	85.53	85.55	85.48	85.57	85.00	85.72	85.17	86.43	86.15	86.64	86.57	86.04	86.19	86.26	86.25
4 <i>C. necator</i> A5-1 GCA_000744095.1	83.79	84.42	86.76	X	88.44	88.39	88.28	88.60	88.72	88.88	88.41	89.32	89.80	90.17	90.14	89.79	89.99	89.88	89.90
5 <i>C. taiwanensis</i> LMG 19424 GCA_000069785.1	83.88	84.35	85.54	88.46	X	99.03	94.62	93.89	94.05	88.71	88.40	89.86	89.71	89.65	89.81	89.76	90.01	89.82	89.99
6 <i>C. taiwanensis</i> STM 6018 GCA_000472465.1	83.85	84.31	85.55	88.41	99.02	X	94.71	93.89	94.06	88.71	88.40	89.86	89.64	89.59	89.78	89.74	90.00	89.75	89.93
7 <i>C. taiwanensis</i> STM 6070 GCA_000372525.1	83.96	84.30	85.48	88.29	94.62	94.71	X	94.62	94.79	88.83	88.33	89.85	89.57	89.47	89.82	89.67	89.99	89.76	89.94
8 <i>C. alkaliphilus</i> GCA_900094595.1	83.90	84.38	85.57	88.61	93.89	93.89	94.62	X	95.12	88.84	88.41	90.36	90.01	90.06	90.02	90.11	90.22	90.01	90.07
9 <i>C. nantongensis</i> GCA_001598055.1	84.08	84.61	85.58	88.71	94.03	94.04	94.79	95.12	X	88.82	88.32	90.45	90.14	90.06	90.11	90.36	90.39	90.23	90.31
10 <i>C. oxalaticus</i> NBRC 13593 GCA_001592245.1	84.07	84.57	85.72	88.87	88.70	88.71	88.83	88.85	88.82	X	94.42	88.46	88.68	88.75	88.83	88.74	88.79	88.82	88.87
11 <i>Cupriavidus</i> sp. amp6 GCA_000426345.1	83.91	84.40	85.17	88.41	88.40	88.40	88.33	88.41	88.29	94.42	X	87.79	88.11	88.28	88.40	88.20	88.43	88.35	88.51
12 <i>Cupriavidus</i> sp. NH9 GCA_002011925.1	83.73	84.36	86.42	89.29	89.86	89.89	89.84	90.36	90.52	88.46	87.85	X	92.01	92.03	92.36	92.36	92.55	92.22	92.35
13 <i>Cupriavidus</i> sp. UYPR2.512 GCA_000379565.1	84.04	85.29	86.15	89.80	89.70	89.64	89.57	90.02	90.13	88.68	88.11	91.99	X	95.24	95.37	94.79	95.11	94.95	95.12
14 <i>C. necator</i> N-1 GCA_000219215.1	83.87	84.43	86.64	90.18	89.66	89.60	89.47	90.06	90.06	88.75	88.28	92.03	95.26	X	96.54	95.77	96.10	95.71	95.92
15 <i>Cupriavidus</i> sp. SK-4 GCA_000611125.1	84.01	84.58	86.57	90.14	89.81	89.78	89.82	90.02	90.11	88.83	88.40	92.35	95.37	96.54	X	96.67	96.86	96.68	96.72
16 <i>C. necator</i> PHE3-6 GCA_001853325.1	83.72	84.29	86.04	89.78	89.75	89.74	89.67	90.11	90.36	88.74	88.20	92.35	94.79	95.76	96.67	X	99.58	99.20	99.42
17 <i>Cupriavidus</i> sp. GA3-3 GCA_000389805.1	83.87	84.49	86.19	89.99	90.02	90.00	89.98	90.22	90.39	88.79	88.43	92.54	95.11	96.10	96.86	99.58	X	99.51	99.53
18 <i>C. necator</i> H16 GCA_000009285.2	83.75	84.33	86.25	89.87	89.82	89.74	89.75	90.01	90.23	88.82	88.36	92.19	94.95	95.71	96.68	99.20	99.51	X	99.99
19 <i>C. necator</i> NBRC 102504 GCA_001598755.1	83.79	84.44	86.25	89.90	89.98	89.93	89.94	90.07	90.31	88.87	88.51	92.35	95.12	95.91	96.72	99.42	99.53	99.99	X

3.2 Descripción del nuevo genoma

Mediante la re-secuenciación del genoma de UYPR2.512 (versión 2) utilizando la tecnología PacBio, se obtuvieron 18.759 lecturas, totalizando 9.605.692 pb secuenciados, lo que representa una cobertura del genoma completo 37,5 veces (37.5X). El ensamblado del genoma generó 10 *contigs* (tig00000001, tig00003541, tig00000004, tig00000105, tig00000006, tig00000107, tig0000103, tig0000106, tig000000104 y tig000003540), con un tamaño total del genoma de 9,6 Mb. De los 10 *contigs* obtenidos,

cuatro correspondieron a moléculas cerradas (tig00000001, tig00003541, tig00000004 y tig00000107), 3 tienen tamaños semejantes a megaplásmidos, pero no lograron cerrarse (tig00000105, tig00000006, tig0000103) y los otros 3 estaban formados por una única lectura o secuencias con muy baja cobertura (tig000000104, tig00000106, tig00003540) (Tabla 4).

En función de su tamaño y siguiendo la clasificación propuesta en la bibliografía [179], [180], [192] el genoma de UYPR2.512 estaría compuesto por un replicón principal tig00000001 (con un tamaño de 4,1 Mb, clasificado como cromosoma o chr1), y por un segundo replicón tig00003541 (con un tamaño de 2,8 Mb, clasificado como cromosoma 2 o chr2). Se decidió llamar cromosoma 2 a este contig debido a que codifica para genes de mantenimiento conservados como ser *vapB*, *vapC*, *doc* y *phd*, y varios ARNr y ARNt (Tabla 4). Los restantes *contigs* se nombraron de acuerdo a la nomenclatura utilizada para designar plásmidos de rizobios [236] partiendo de pCne512a a pCne512e, de acuerdo al tamaño de menor a mayor, o nombrados como no asignados (NA, de sus siglas en inglés *Non-Assigned*) si estaban compuestos por secuencias repetidas o por una sola lectura. El replicón denominado pCne512e (tig00000004), con un tamaño de 771 Kb se definió como el plásmido simbiótico (pSym), ya que contenía los genes *nod*, *nif* y *fix*, mientras que el replicón pCne512b (tig00000107) de 388 Kb sería el segundo plásmido cerrado. Los tres *contigs* restantes fueron nombrados como pCne512d, de 516 Kb, pCne512c, de 487 Kb y pCne512a de 381 Kb.

La anotación del genoma mediante el servidor web *RAST* [232] identificó 9.808 secuencias codificantes (CDS), de los cuales 69 corresponden a ARN de transferencia (58 en el chr1, 10 en el chr2 y uno en el pCne512e) y 15 a distintas subunidades de ARN ribosomal (cinco juegos completos de subunidades 5S, 16S y 23S, tres en el chr1 y dos en el chr2). En dicho contexto, más del 70% de los CDS se distribuyeron entre chr1 y chr2 (Tabla 6). El contenido de GC del genoma se estimó en 64,59%, valor similar al descrito en la primera versión del genoma (65,25%) [163] y dentro del rango informado para miembros de la familia *Burkholderiaceae* [237]. El contenido GC del chr1 y chr2, se estimaron en 66,0% y en 66,1% respectivamente, siendo mayor al contenido GC

estimado en los plásmidos, como también se ha observado en otras especies del género *Cupriavidus* [164].

Tabla 5. Características generales del genoma ensamblado de UYPR2.512

Se indican el nombre asignado a cada contig, la longitud de la secuencia (pb), el número de lecturas, el total de secuencias codificantes (CDS) y el contenido GC (%). En negrita se indican las moléculas circulares.

Contig	Nombre	Longitud (pb)	Lecturas	Número de CDC	(% CDS)	% GC
tig00000001	Chr1	4.137956	7.261	4.246	43,3%	66,0
tig00003541	Chr2	2.841.770	6.541	2.858	29,1%	66,1%
tig00000004	pCne512e (pSym)	770.893	1.767	863	8,8%	59,4%
tig00000105	pCne512d	516.130	815	471	4,8%	60,5%
tig00000006	pCne512c	486.770	859	508	5,2%	61,9%
tig00000107	pCne512b	387.719	795	422	4,3%	60,7%
tig00000103	pCne512a	3.809.33	674	353	3,6%	62,0%
tig00000106	NA	32.064	33	21	0,2%	64,2%
tig00000104	NA	24.173	13	19	0,2%	64,1%
tig00003540	NA	27.284	1	47	0,5%	56,5%
Total		9.605.692	18.759	9.808		64,6%

En la figura 16, se muestra una representación circular de chr1, chr2, pCne512e (pSym) y pCne512b y en la figura 17 se representan todos los *contigs* en una única molécula circular.

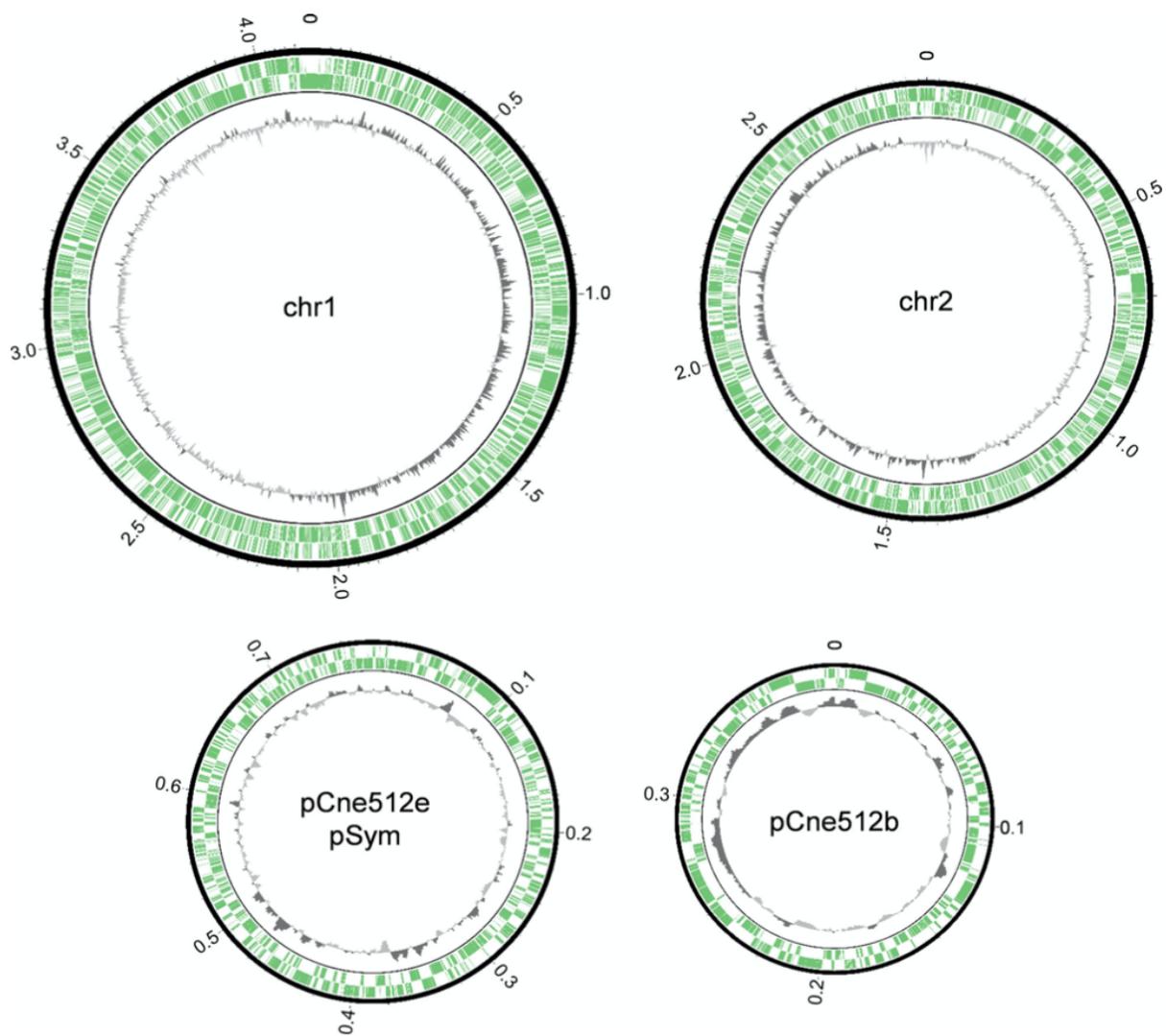


Figura 16. Representación circular de los 4 replicones cerrados del genoma de UYPR2.512. De afuera hacia adentro: escala de tamaño en Mb, *contigs* ensamblados, secuencias de codificación anotadas en orientación sentido y antisentido, y sesgo de GC para el cromosoma 1 (chr1), el cromosoma 2 (chr2) y los dos plásmidos cerrados pCne512e (pSym) y pCne512b.

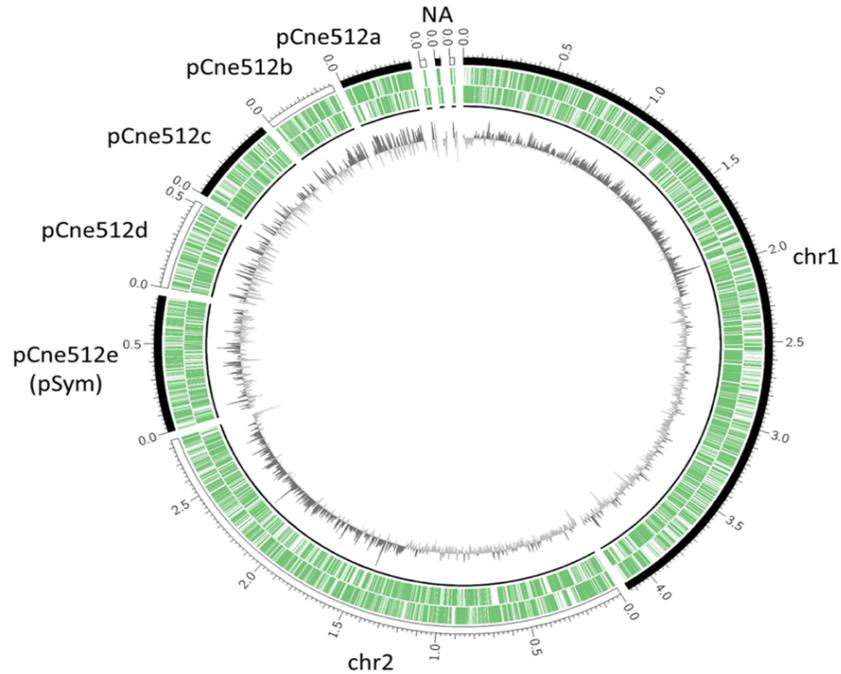


Figura 17. Representación circular de los *contigs* ensamblados.

De afuera hacia dentro se indican los nombres de los *contigs* ensamblados, las secuencias codificantes anotadas en orientación sentido y anti-sentido, y el sesgo de bases GC.

3.2.1 Mejoras de la secuencia del genoma respecto a la versión anterior.

La nueva versión del genoma reduce el número de *contigs* de 369 a 10 y agrega alrededor de 2 Mb de secuencia, la cual codifica para 2.146 genes que no habían sido descritos en la versión anterior (versión 1), llevando el tamaño del genoma de UYPR2.512 a 9,6 Mb (Tabla 6). Entre ellos, se describieron siete nuevos ARNt: Arg-ACG, Ala-TGC, Ile-GAT, Thr-GGT, Gly-TCC, Tyr-GTA y Arg-CCG. La mayoría de los nuevos genes descritos se ubicaron en los plásmidos, lo que llevó a un aumento en su número de genes al 37,2 % en el pCne512b, y al 74,7 % en el pCne512d. Además, en esta nueva versión del genoma, se consiguió anotar la secuencia completa de todas las subunidades de ARNr de UYPR2.512.

Tabla 6. Características principales del nuevo genoma y valores comparativos con la versión anterior. Se indica los valores de longitud (pb) y número de genes para cada replicón, asignados a cada contig. En negrita se indican los replicones en los cuales se observó un mayor aumento en el número de genes.

Contig	versión_2 del genoma			Cobertura de la versión_1				Aumento en la versión_2			
	Nombre	Longitud	#genes	Longitud	Genes	Longitud	Genes	Longitud	Genes	Longitud	Genes
tig00000001	Chr1	4.137.956	4.447	3.915.61	94,6%	4.102	92,2%	222.345	5,4%	345	7,8%
tig00003541	Chr2	2.841.770	2.906	2.743.21	96,5%	2.762	95,0%	98.552	3,5%	144	5,0%
tig00000004	pCne512e (pSym)	770.893	866	416.905	54,1%	430	49,7%	353.988	45,9%	436	50,3%
tig00000105	pCne512d	516.130	471	134.480	26,1%	119	25,3%	381.650	73,9%	352	74,7%
tig00000006	pCne512c	486.770	508	144.086	29,6%	138	27,2%	342.684	70,4%	370	72,8%
tig00000107	pCne512b	387.719	422	227.239	58,6%	265	62,8%	160.480	41,4%	157	37,2%
tig00000103	pCne512a	380.993	353	88.212	23,2%	98	27,8%	292.781	76,8%	255	72,2%
tig00000106	NA	32.064	21	-	0%	-	0%	32.064	100%	21	100%
tig00000104	NA	24.173	19	-	0%	-	0%	24.173	100%	19	100%
tig00003540	NA	27.824	47	-	0%	-	0%	27.824	100%	47	100%

3.3 Anotación funcional del genoma de UYPR2.512

3.4 Genes involucrados en funciones de replicación y mantenimiento de los replicones

Con el fin de establecer si las moléculas capaces de replicarse eran efectivamente replicones se buscaron los genes implicados en el mantenimiento y la replicación de plásmidos. Mediante el servicio web RAST, se identificaron homólogos a los genes *repB*, *parA* y *parB*, los cuales están involucrados en la replicación y partición de plásmidos [238] (Tabla 7). La presencia de los genes *parA* (ID = 6430), *parB* (ID = 6429) y *repB* (ID = 6431) en el contig tig00000105, los genes *parA* (ID = 5374), *parB* (ID = 5373) y *repB* (ID = 5375) en el contig tig00000006 y los genes *parA* (ID = 5618) y *repB* (ID = 6519 al 5620) en el contig tig00000103, sugiere que estos replicones son plásmidos y por lo tanto, se denominaron pCne512d, pCne512c y pCne512a, respectivamente. Además, se identificaron los sistemas toxina-antitoxina Phd/Doc (proteína asociada a la prevención de la muerte del huésped) y VapC/VapB (proteína C asociada a la virulencia) involucrados en el mantenimiento de plásmidos [239] en los replicones pCne512e, pCne512b y pCne512d. A su vez, la presencia de homólogos al gen *trhI* (que codifica para una ADN helicasa/relaxasa) y el gen *traD* (que codifica para una proteína de

acoplamiento tipo IV) en los replicones pCne512e y pCne512b, sugeriría que estos dos plásmidos son movilizables, mientras que la presencia adicional de genes que codifican un SSTIV en pCne512b sugiere que se trata de un plásmido conjugativo [240].

Tabla 7. Presencia de genes relacionados a la replicación, conjugación y mantenimiento en el crómido y plásmidos en UYPR2.512.

Replicón	Genes relacionados con la partición y la replicación (ID)	Genes relacionados con la conjugación (ID)	Sistemas de mantenimiento y estabilización (ID)
Chr2	<i>parA</i> (8266) <i>parB</i> (8267) <i>repB</i> (8264)	<i>Flp</i> (8438) <i>rcpA</i> (8441) <i>tadA</i> (8443) <i>tadB</i> (8444)	<i>vapB</i> (9112) <i>vapC</i> (9113) <i>doc</i> (8115) <i>phd</i> (8116)
pCne512e (pSym)	<i>parA</i> (4577) <i>parB</i> (4578) <i>repB</i> (4576)	<i>traD</i> (4568) <i>trhI</i> (4574)	<i>doc</i> (5063) <i>phd</i> (5062)
pCne512b	<i>parA</i> (6741) <i>parB</i> (6742) <i>repB</i> (6740)	<i>traD</i> (6745) <i>trhI</i> (6649) <i>trbB</i> (6748) <i>trbE</i> (6751)	<i>doc</i> (6508) <i>phd</i> (6509)
pCne512d	<i>parA</i> (6430) <i>parB</i> (6429) <i>repB</i> (6431)	ND	<i>vapB</i> (6249) <i>vapC</i> (6250) <i>doc</i> (6099) <i>phd</i> (6098)
pCne512c	<i>parA</i> (5374) <i>parB</i> (5373) <i>repB</i> (5375)	ND	ND
pCne512a	<i>parA</i> (5618) <i>repB</i> (5619-5620)	ND	ND

3.4.1 Características genómicas de la cepa UYPR2.512 relacionadas a la especie *necator*

3.4.1.1 Sistemas involucrados en la tolerancia a metales

En el genoma de UYPR2.512 se identificaron varios genes que codifican para posibles proteínas pertenecientes a las superfamilias que incluyeron: (1) familias de facilitadores de difusión de cationes (CDF), (2) bombas de eflujo impulsadas por proteínas de la superfamilia de resistencia, nodulación y división celular (RND), (3) transportador de iones cromato (CHR) y (4) ATPasas del tipo P (Tabla 8).

Formando parte de los sistemas de transporte CDF, se identificó el gen *czcD* (ID = 4107, 4843, 507) con cuatro copias, ubicadas en el chr1, el cual codifica para una proteína de membrana interna, la cual está involucrada en el transporte activo de cobalto, zinc y

cadmio (Co/Zn/Cd), hacia el periplasma. A su vez, *czcD* actúa como represor del sistema CZC, el cual incluye el operón *czcABC*. El sistema completo CZC está formado por CzcA ubicada en la membrana interna, la proteína de anclaje periplásmico CzcB y la proteína de membrana externa CzcC [185]. A sí mismo, se observaron varias copias adicionales del operón *czcABC*, cuatro copias *czcA* (ID = 4854 y 4855, 7542, 8123, 8787), cinco copias de *czcB* (ID = 4856, 1685, 7543, 8124, 8788) y cinco copias de *czcC* (ID = 4853, 7544, 8125, 8789), ubicadas en los cromosomas. Sin embargo, en este caso las proteínas codificadas por dicho operón formaban parte de un sistema de transporte RND, también involucrado en la resistencia a Co/Zn/Cd. La utilización de estos sistemas de resistencia a metales y su ubicación genómica, ha sido observada en otras cepas de *Cupriavidus*, como ser CH34 de *C. metallidurans* [241].

Como integrantes de los sistemas CHR se identificaron los genes *chrBACF* (ID = 9271 al 9274) formando parte de un posible operón, el cual codifica para posibles proteínas implicadas en la resistencia a compuestos con cromo, organización conservada en genomas del género *Cupriavidus* [241]. Los genes *chrA*, *chrB*, *chrC*, y *chrF* codifican para una bomba de eflujo, para un posible regulador transcripcional, para una superperóxido reductasa y para un factor de represión transcripcional (Tabla 9) [242].

Parte de los sistemas de transporte mediante ATPasas tipo P se encontraron varios genes involucrados en la resistencia a Cu/Zn/Cd, plomo y mercurio, distribuidos tanto en el chr1 como en el chr2. Se identificaron 4 copias del gene *ciA* (ID = 49, 485, 493, 7220), los cuales codifican para proteínas involucradas a la resistencia de metales divalentes. Abajo de la copia *ciA* (ID = 49), se identificó un posible regulador transcripcional *crtR* (ID = 48) de respuesta a Cu^{+1} , y junto a este gen se identificó el gen *copZ* (ID = 50) quien codifica para una chaperona de cobre relacionada con la regulación del metal. Abajo de otras de las copias *ciA* (ID = 7220) se observaron diversos genes posiblemente implicados en su regulación; *copS* y *copR* los cuales codifican para un sistema de dos componentes involucrado en la transducción de señales, en el que CopS (ID = 7223) es la proteína sensora y CopR codifica para un potencial factor de transcripción regulador de la respuesta a cobre (ID = 7224).

A su vez, implicados en la resistencia al cobre, se identificó un posible operón *copABCD* (ID = 7885 al 7888), en el cual CopA es una oxidasa multicobre periplasmática, que al formar un complejo junto con CopB ubicada en la membrana externa, constituyen el mínimo sistema para evitar niveles tóxicos de cobre. Por otro lado, CopD es una proteína de membrana que está involucrada en el transporte activo de cobre desde el periplasma al citoplasma de la célula y la proteína CopC que es un regulador de la homeostasis del metal dentro de la célula, organización que ha sido observada en el género *Cupriavidus* [185]. Aguas arriba al operón *cop*, se observaron los genes *cusR/cusS* (ID = 7889, 7890), que codifican para reguladores transcripcionales de dichos genes. Seguido a *cusR/cusS*, se identificó *chasE2* (ID = 7891), el cual codifica para una proteína con un dominio sensor extracelular conservado en receptores transmembrana involucrados en vías de traducción de señales de dos componentes 10.1128/JB.185.1.285-294.2003. A su vez, anotados como genes que codifican para proteínas involucradas en la tolerancia a cationes divalentes, se encontraron los genes *cutA* (ID = 4011) y *corC* (ID = 780), y *cutE* (ID = 799), que también estarían asociados a la homeostasis de cobre.

Involucrados en la resistencia de metales se encontraron posibles reguladores *hmrR/trmeR* (ID = 248, 9268) de la familia *merR* de reguladores transcripcionales bacterianos en respuesta al sensado de metales [243], los cuales activan la transcripción de genes aguas abajo, muchos de los cuales codifican para bombas de eflujo [244]. También, se identificó *zraR* (ID = 9049, 9050), el cual interactúa con sigma 54 y regula la expresión de genes relacionados con la resistencia a zinc. Finalmente, el gen anotado con la ID = 7486, el cual codifica para un mercurio reductasa, fue el único encontrado implicado en la resistencia a mercurio.

Tabla 8. Características y ubicación genómica de genes relacionados con los principales mecanismos implicados en la homeostasis de metales en UYPR2.512

Se muestran los diferentes sistemas de resistencia/expulsión de metales pesados que incluyen proteínas pertenecientes a las familias de facilitadores de difusión de cationes (CDF), bombas de eflujo impulsadas por proteínas de la superfamilia de resistencia, nodulación y división celular (RND), transportador de iones cromato (CHR) y ATPasas del tipo P.

Sistema de transporte	Genes/proteínas relacionadas a la homeostasis de metales (ID)	Sustrato para transportar	Función asociada	Replicón
CDF	<i>czcD</i> (4107, 4843, 507)	Co/Zn/Cd	Proteína de unión a membrana, represor del sistema <i>czc</i>	Chr1
RND	<i>czcA/cusA</i> (4854 y 4855) <i>czcB/cusB</i> (4856, 1685) <i>czcC</i> (4853)	Co/Zn/Cd	<i>CzcA</i> : transportador de membrana, <i>CusA</i> : Bomba de eflujo de cationes <i>CzcB</i> : proteína de anclaje periplásmico de unión a metales <i>CzcC</i> : Proteína de membrana externa del sistema de eflujo	Chr1
	<i>czcABC/cusA</i> (7542 al 7544)			Chr2
	<i>czcABC</i> (8123 al 8125) <i>czcABC</i> (8787 al 8789)			Chr2
Regulador	<i>czrR/cusR</i> (ID = 7889), <i>cusS</i> (ID = 7890)	Co/Zn/Cd	<i>czrR/cusR</i> : Regulador de respuesta a metales pesados, de unión al ADN	Chr2
			<i>cusS</i> : Histidina quinasa sensora de metales pesados	Chr2
Regulador	<i>hmrR/trmeR</i> (ID = 248, 9268)	Co/Zn/Cd	Regulador transcripcional, familia MerR	Chr2
Regulador	<i>zraR</i> (ID = 9049, 9050)	Zn	Proteína reguladora de respuesta del sistema de dos componentes dependiente de sigma-54 y zinc.	Chr2
-	no asignado (7486)	Hg	Mercurio reductasa (oxidoreductasa dependiente de FAD y NAD(P)-disulfuro. Familia PF00070)	Chr2
CHR	<i>chrBACF</i> (9271 al 9274), <i>chrB</i> (6560)	Cr	<i>ChrB</i> : regulador transcripcional <i>ChrA</i> : proteína de membrana (bomba de eflujo) <i>ChrC</i> : superóxido dismutasa <i>ChrF</i> : represor transcripcional	Chr2, pCne512b Chr2 Chr2
ATPasa de tipo P	<i>ciA</i> (485, 49, 493) <i>ciA</i> (7220)	Cu	ATPasa transportadora de plomo, cadmio, zinc y mercurio (EC 3.6.3.3) (EC 3.6.3.5) /Proteína translocadora de cobre (EC 3.6.3.4)	Chr1 Chr2
Regulador	<i>crtR</i> (48)	Cu	Regulador transcripcional sensible a Cu+1	Chr1
-	<i>copZ</i> (50)	Cu	Chaperona	Chr1
-	<i>cutA</i> (4011)	met. diva.	Proteína periplásmica	Chr1

Sistema de transporte	Genes/proteínas relacionadas a la homeostasis de metales (ID)	Sustrato para transportar	Función asociada	Replicón
-	<i>cutE/CorC</i> (799 y 800)	metales divalentes	<i>cutE</i> : Apolipoproteína N-aciltransferasa <i>corC</i> : Proteína de eflujo de magnesio y cobalto	Chr1
-	<i>copDCBA</i> (7885 al 7888)	Cu	<i>copD</i> : Transportador de cobre <i>copC</i> : Regulador de la proteína CopD <i>copB</i> : forma un complejo con CopA, implicado en la resistencia al cobre <i>copA</i> : Multicobre oxidasa	Chr2
-	<i>copSR</i> (7223 y 7224)	Cu	Proteína sensora de cobre	Chr2
-	<i>chasE2</i> (7891)	metales divalentes	Sistema sensor de dos componentes	Chr2

A pesar de la importancia biológica del hierro, puede ser tóxico si se acumula en exceso o si se encuentra en formas reactivas que generan especies ROS [245].

Varios estudios han demostrado que desempeña un papel fundamental en diferentes procesos biológicos. El hierro tiene un rol directo en la competencia por los rizobios durante la colonización radicular [246] y es un cofactor muy importante de la nitrogenasa [247]. Además, uno de los mecanismos utilizados por las bacterias para la captación de este metal es la síntesis, excreción y la posterior recaptación de sideróforos [247]. El genoma de la cepa UYPR2.512 presenta 10 genes relacionados con la adquisición y síntesis de sideróforos (Tabla 9). Entre ellos se encuentran genes *pvsEDCBA* (ID = 237 al 241) implicados en la biosíntesis de un sideróforo del tipo catecol similar a la vibriobactina [245]. Se identificaron los genes *exbB* (ID = 894), *tonB* (ID = 9182) y *exbD* (ID = 893) que codifican para un complejo de proteínas responsable de energizar el transporte activo de sideróforos (y otras moléculas) a través de la membrana externa. También se identificó un gen (ID = 1794) que codifica para un receptor de membrana externa dependiente de TonB, que actúa como receptora del complejo hierro-sideróforo (ID = 1794). También relacionado al transporte de este metal, se identificaron tres genes que codifican para proteínas de un sistema de transporte del tipo ABC (ID = 8365 al 8367), ubicados en el segundo cromosoma. Corriente abajo de este sistema se identificó un homólogo al gen *fur* (ID = 8372), que codifica para el regulador transcripcional conservado en bacterias FUR, implicado en la homeostasis de hierro. Además, se

identificó una segunda copia del gen *fur* (ID = 3612) en el chr1. Sin embargo, esta copia codificaría en realidad para un regulador del tipo PerR, miembro de la familia Fur, implicado en la respuesta al estrés oxidativo [248].

Tabla 9. Características y ubicación genómica de genes relacionados con la homeostasis del hierro en UYPR2.512.

Anotación Rast/Manual	Genes/proteínas (ID)	Función asociada	Replicón
Adquisición de hierro y metabolismo. Síntesis de vibrioferrina	<i>pvsEDCBA</i> (237 al 241)	PvsE: Vibrioferrina descarboxilasa PvsD: Proteína formadora de enlace vibrioferrina-amida /Sideróforo sintetasa PvsC: Transporte a través de la membrana de la Vibrioferrina PvsB: Proteína formadora de enlace vibrioferrina-amida PvsA: Vibrioferrina ligasa/carboxilasa	Chr1
Sistemas de transporte Ton/Tol	<i>exbB</i> (894)	Proteína transportadora/sistema de transporte de sideróforos férricos	Chr1
Sistemas de transporte Ton/Tol	sin nombre asignado (9182)	Receptor de sideróforo dependiente de TonB	Chr1
Estrés oxidativo	<i>fur</i> (3612)	Regulador transcripcional	Chr1
Adquisición de hierro y metabolismo	<i>fur</i> (8372)	Proteína reguladora de la adquisición del ión férrico	Chr2

3.4.1.2 Degradación de compuestos aromáticos

Se identificaron genes implicados en las vías periféricas y centrales de degradación de compuestos aromáticos, en particular aquellos implicados en las vías del catabolismo del catecol (genes *cat*) y del protocatecuato (genes *pca*) (Tabla 10 y Fig. 18).

Los genes *benABCD* (ID = 2269 al 2272), codifican para el complejo enzimático formado por la benceno 1,2-dioxigenasa y la 1,2-dihidroxiclohexa-3,5-dieno-1-carboxilato deshidrogenasa, involucrado en la degradación aerobia de benzoato, cuyo producto es dirigido hacia la vía del catecol. A su vez, se identificaron los genes *badIHR* (ID = 7392 al 7394), involucrados en la degradación anaerobia de benzoato. También se

identificaron los genes *benK* (ID = 2029) y *benE2* (ID = 3104) posiblemente implicados en el transporte del benzoato.

Por otro lado, se identificaron genes involucrados en la degradación del tolueno, incluyendo los genes *tmoABCDEF* (ID = 4298 al 4303) que codifican para el complejo enzimático tolueno monooxigenasa (EC:1.14.13.2361.14.13.-), codificado en el pCne512e (pSym). Una segunda copia de los genes *tmoD*, *tmoE* y *tmoF* (ID = 6012 al 6015) se identificó en el pCne512.

También, se observaron genes involucrados en el transporte y degradación de hidroxibenzoato (ID = 2170, 7993, 8535, 9241 y ID = 7986), respectivamente. El gen *pobA* codifica para la enzima p-hidroxibenzoato 3-monooxigenasa (EC:1.14.13.2), que cataliza la conversión del 4-hidroxibenzoato en protocatecuato (3,4 hidroxibenzoato).

A su vez, se encontró un grupo de genes, *bphHJ2D* (ID = 8857, 8858, 8859) y *bphC* (ID = 8892), los cuales codifican para enzimas involucradas en la degradación de bifenilo, para ser convertido en catecol.

Tabla 10. Características, función y ubicación genómica de genes relacionados con la degradación de compuestos aromáticos en UYPR2.512.

Compuesto (Anotación Rast)	Genes/proteínas (ID)	Función asociada	Replicón
Tolueno	<i>tmoABCDEF</i> (4298 al 4303)	Tolueno-4-monooxigenasa, subunidades TmoA, TmoB, TmoC, TmoD, TmoE, TmoF	pCne512e (pSym)
	<i>tmoDEF</i> (6012 al 6015)	Tolueno-4-monooxigenasa, subunidades TmoD, TmoE, TmoF	pCne512d
Quinato	<i>quiB</i> (3643, y 7562)	3-deshidroquinato deshidratasa II (EC 4.2.1.10)	Chr1
Benzoato	<i>benK</i> (2029) <i>benK</i> (7451, 7452, 7611, 9060)	Transportador de benzoato de la familia de transportadores tipoMFS	Chr1 Chr2
	<i>benDCBA</i> (2269 al 2272)	<i>benD</i> :1,2-dihidroxiclohexa-3,5-dieno-1-carboxilato deshidrogenasa (EC 1.3.1.25) <i>benC</i> : Benzoato 1,2-dioxigenasa, componente de ferredoxina reductasa <i>benB</i> :Subunidad β de benzoato 1,2-dioxigenasa (EC1.14.12.10) <i>benA</i> : Subunidad α de benzoato 1,2-dioxigenasa (EC 1.14.12.10)	Chr1
	<i>benE2</i> (3104)	Transportador de benzoato	Chr1

Compuesto (Anotación Rast)	Genes/proteínas (ID)	Función asociada	Replicón
	<i>ohbAB</i> (1973 al 1974) <i>ohbAB</i> (4551 y 4550)	O-halobenzoato 1,2-dioxigenasa α -ISP proteína	Chr1 pCne512e (pSym)
Benzoato (Degradación anaeróbica)	<i>badI</i> (7392 al 7394)	<i>badI</i> :2-cetociclohexanocarboxil-CoA hidrolasa (EC:3.1.2.-) <i>badH</i> : 2-hidroxiciclohexanocarboxil-CoA dehidrogenasa (EC:1.1.1.-) <i>badR</i> :Regulador transcripcional, familia MarR	Chr2
p-hidroxibenzoato	HT (2170) HT (7993 ,8535, 9241) <i>pobA</i> (7986)	Transportador de 4-hidroxibenzoato p-hydroxybenzoato 3-monooxygenasa (EC:1.14.13.2)	Chr1 Chr2 Chr2
Bifenilo	<i>bphHJ2</i> (8857 y 8858), <i>bphH</i> (9248, 9540) <i>bphJ2</i> (8805) <i>bphD</i> (8859,8759, 8804) <i>bphD</i> (8861) <i>bphC</i> (8892)	<i>bphH</i> :2-ceto-4-pentenoato hidratasa (EC 4.2.1.80) <i>bphJ2</i> : Acetaldehído deshidrogenasa, acetilante (EC 1.2.1.10) <i>bphD</i> :4-hidroxi-2-oxovalerato aldolasa (EC 4.1.3.39) 2-hidroxi-6-oxo-6-fenilhexa-2,4-dienoato hidrolasa (EC 3.7.1.-) 2,3-dihidroxibifenil 1,2-dioxigenasa (EC 1.13.11.39)	Chr2 Chr2
Rama del Catecol	<i>catA</i> (2273) <i>catBCD</i> (2276 al 2278)	<i>catA</i> : Catecol 1,2-dioxigenasa (EC 1.13.11.1) <i>catB</i> :Muconato cicloisomerasa EC 5.5.1.1) <i>catC</i> :Muconolactona isomerasa (EC 5.3.3.4) <i>catD</i> : β -ceto adipato enol-lactona hidrolasa (EC 3.1.1.24)	Chr1
-	<i>pcaQ</i> (8224)	Pca: Proteína regulatoria del regulón <i>pcaR</i> . Rama del protocatecuato	Chr2
Rama del protocatecuato	<i>pcaHGB</i> (7989 al 7991)	<i>pcaH</i> : Protocatecuato 3,4-dioxigenasa cadena β (EC 1.13.11.3) <i>pcaG</i> : Protocatecuato 3,4-dioxigenasa cadena alfa (EC 1.13.11.3)	Chr2
Rama del catecol/rama del protocatecuato	<i>pcaCD</i> (7988)	<i>pcaB</i> : 3-carboxil-cis-cis muconato cicloisomerasa (EC 5.5.1.2) <i>pcaCD</i> : β -ceto adipato enol-lactona hidrolasa (EC 3.1.1.24) / 4-carboximuconolactona descarboxilasa (EC 4.1.1.44)	
-	<i>pobR</i> (7987)	Posible Regulador transcripcional PobR	Chr2
Rama del catecol/rama del protocatecuato (Vía del ceto adipato)	<i>pcaIJ</i> (8453, 8454)	<i>pcaI</i> : 3-oxoadipato CoA-transferasa subunidad A (EC2.8.3.6) <i>pcaJ</i> : 3-oxoadipato CoA-transferasa subunidad B (EC2.8.3.6)	Chr2

A su vez, se identificaron dos copias del gen *quiB* (ID = 3643, 7562), que codifica para la 3-deshidroquinato deshidratasa II (EC 4.2.1.10), involucrada en la conversión de quinato, un compuesto vegetal abundante, en precursores del catabolismo del protocatecuato [89].

Con respecto a las vías del metabolismo central se identificaron los genes *catABCD* (ID = 2273, 2276, 2777, 2278) que codifican para las reacciones involucradas desde la rotura del anillo del catecol, hasta la formación de β -cetoadipato enol-lactona (Tabla 10). Junto a estos genes se identificó el gen *catR* (ID = 2275) que codifica para un regulador transcripcional de la familia LysR, un arreglo que se ha observado en otras bacterias del género *Cupriavidus* [249]. A su vez, se observaron los genes *pcaCD* (ID= 7988), *pcaHGB* (ID= 7989 al 7991), *pcaIJF* (ID = 8453 al 8455), involucrados en la conversión de protocatecuato a β -cetoadipato. A su vez, se observó el gen *pcaQ* (ID = 8224), el cual codifica para un regulador transcripcional de la familia LysR que podría estar regulando a los genes *pcaCD*, *pcaHGB* y *pcaIJF*.

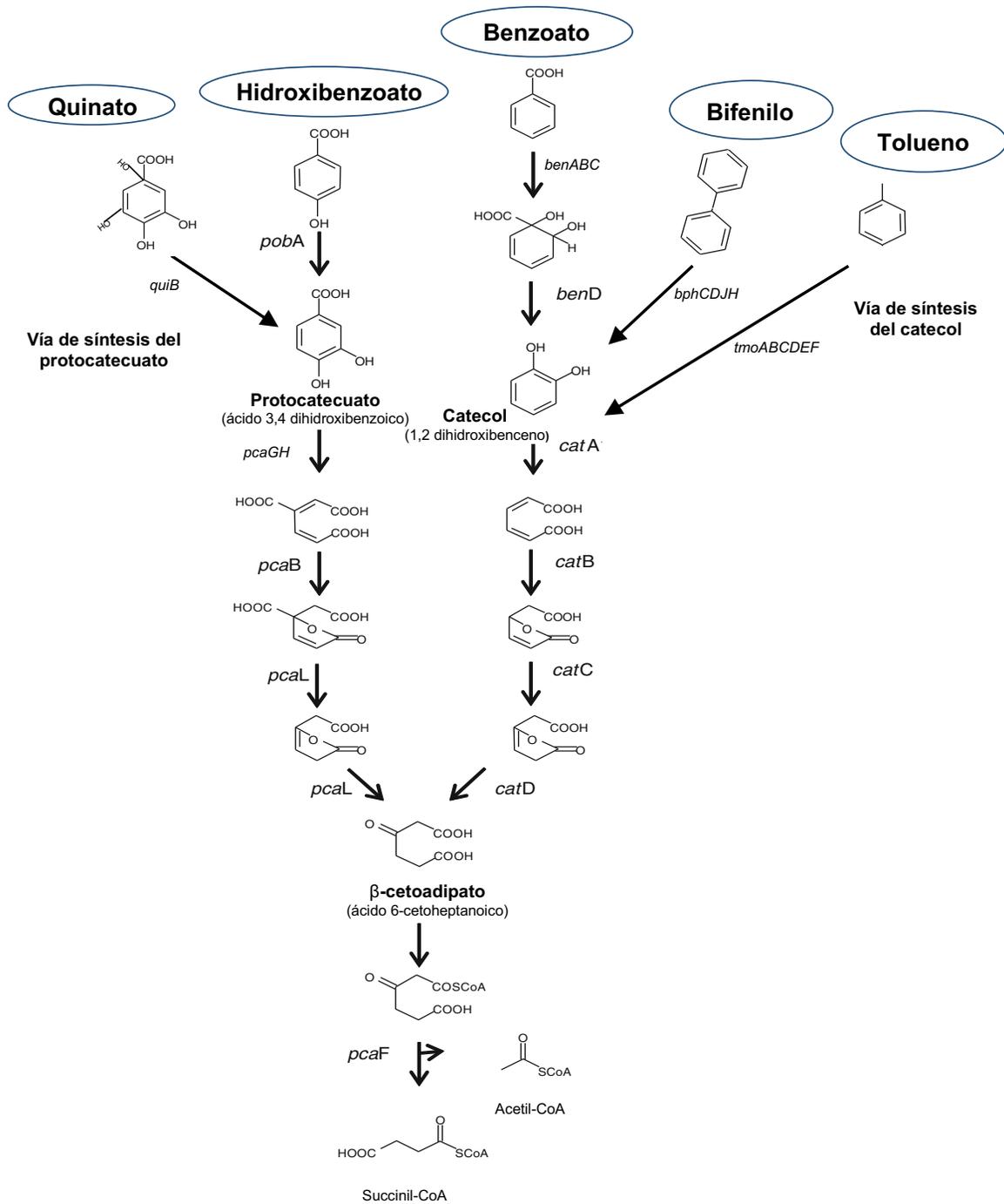


Figura 18. Ilustración de las vías periféricas y centrales de degradación de compuestos aromáticos en UYPR2.512.

Se muestran los compuestos de degradación, los genes implicados, y moléculas precursoras involucradas en la vía de síntesis del catecol y del protocatecuato, que convergen en la vía del β-cetoadipato, así como también otros precursores que tendrán como destino el ciclo de TCA.

3.4.1.3 Síntesis y degradación de PHB

Los análisis realizados permitieron identificar los genes involucrados tanto en la ruta de síntesis, como en la degradación de polihidroxicanoatos, así como en su regulación (Tabla 11). La presencia de los genes *phaA* (ID = 1658 al 1659), *phaB* (ID = 1652, 2604) y *phaC* (ID = 1649) sugieren la existencia de la ruta clásica de síntesis de PHB. A su vez, se observó la presencia de los genes *phaF* (ID = 2288) y *phaR* (ID=1653), que codifican para proteínas involucradas en la síntesis de los gránulos de polihidroxicanoatos y en la regulación de la síntesis de PHB, respectivamente [250].

Además de estos, se identificaron 25 copias de un gen que codifica para la proteína multifuncional Enoyl-CoA hidratasa (EC 4.2.1.17) y 8 copias del gen *bdhA* el cual además de codificar para la Enoyl-CoA hidratasa, codifica para la 3-hydroxyacyl-CoA dehidrogenasa (EC 1.1.1.35). La Enoyl-CoA hidratasa cataliza la conversión de enoil-CoA-3-hidroxitiril-CoA en 3-hidroxitirato-CoA, que luego se polimeriza para formar PHB. Cabe aclarar que estas enzimas además de estar involucradas en el metabolismo de PHB, catalizan el segundo paso del proceso de β -oxidación de ácidos grasos.

En la tabla 11, se listan los genes que codifican para las enzimas involucradas únicamente en la síntesis de PHB.

Se observó la presencia de varios genes involucrados en la degradación de PHB, que incluyeron dos copias del gen que codifica para la D- β -hidroxitirato deshidrogenasa (DHBD, ID=1541, ID=5926), dos copias del gen que codifica para la hidroxitirato-dímero hidrolasa (HDH, ID = 2679 y 5935) y un tercer gen *aaS* (ID = 3303), que codifica para la enzima acetoacetyl-CoA sintetasa. A su vez, se identificaron cuatro genes homólogos (ID = 4666, 5937, 3689, 7791) que codifican para posibles poli-(3-hidroxitirato) depolimerasa (conocidas como PHB depolimerasas) [250] involucrada en la degradación de PHB. Estos se encontraron en los plásmidos pCne512a y pCne512e, así como en los chr1 y chr2, respectivamente.

Tabla 11. Características y ubicación genómica de genes relacionados con el metabolismo de PHB en UYPR2.512.

Anotación Rast/Manual	Genes/proteínas relacionadas al metabolismo de PHB (ID)	Función asociada	Replicón
Síntesis	<i>bdhA</i> (1325, 1835)	Enoil-CoA hidratasa (EC 4.2.1.17)	Chr1
	<i>bdhA</i> (8588,8832, 9036, 9040)		Chr2
	<i>phaA</i> (1658, 1659)	3-cetoacil-CoA tiolasa (EC 2.3.1.16)	Chr1
	<i>phaB/phaB</i> (1652, 2604)	Acetoacetyl-CoA reductasa (EC 1.1.1.36)	Chr1
	<i>phaC/phaS</i> (1649)	Ácido polihidroxicanoico sintasa	Chr1
	<i>phaF</i> (2288)	Proteína asociada a gránulos de polihidroxicanoato	Chr1
	<i>phaR</i> (1653)	Represor de la síntesis de polihidroxicanoato	Chr1
Degradación	BHBD (1541) BHBD (5926)	D-β-hidroxicbutirato deshidrogenasa (EC 1.1.1.30)	Chr1 pCne512a
	Sin nombre asignado (4666) Sin nombre asignado (5937)	Precursor de poli(3-hidroxicbutirato) despolimerasa o PHB despolimerasa (EC 3.1.1.75)	pCne512a pCne512e (pSym) Chr1 Chr2
	Sin nombre asignado (3689) Sin nombre asignado (7791)		
	Aas (3303)	Acetoacetyl-CoA sintetasa (EC 6.2.1.16)	Chr1
	HDH (2679) HDH (5935)	D (-)-3-hidroxicbutirato oligómero hidrolasa (EC 3.1.1.22)	Chr1 pCne512a
Metabolismo	EH (1324, 1835, 2214, 2219) EH (8588, 8832, 8833,9043, 9158)	Enoil-CoA hidratasa (EC 4.2.1.17)	Chr1 Chr2
	HBP (3660, 3661) HBP (5925) HBP (3541)	D-β-hidroxicbutirato permeasa	Chr1 pCne512a Chr2

3.4.2 Genes relacionados con la interacción planta-microorganismo

Se buscaron genes en el genoma de UYPR2.512 relacionados con su capacidad de interactuar con plantas hospederas. La búsqueda incluyó genes relacionados con la quimiotaxis y la movilidad flagelar, y genes que codifiquen para sistemas de secreción. Además, se buscaron y anotaron genes relacionados a la promoción del crecimiento vegetal tales como la fijación biológica del nitrógeno, la síntesis de sideróforos y la modulación de hormonas vegetales.

3.4.2.1 Quimiotaxis y movilidad del flagelo

En el genoma de UYPR2.512 se encontraron todos los genes necesarios para la síntesis y el ensamblado del flagelo, codificados en el chr2 (Tabla 12). A su vez, se identificaron genes que codifican para posibles reguladores de la síntesis del flagelo, incluyendo *flhDC* (ID = 8497, 8498), *fliA* y *fleN* (ID = 8555, 8554), codificados mayoritariamente en el chr2 y en el plásmido pCne512b, así como en los otros plásmidos. Se encontraron además varios genes involucrados en la quimiotaxis tales como el operón *cheWRDBYZ* (ID = 8502 al 8507), genes que codifican para quimiorreceptores MCP (ID = 2718 y 8490), *aer* (ID = 8494), así como también dos copias del gen *cheA*, posiblemente implicados en la regulación de la movilidad flagelar.

Tabla 12. Características y ubicación genómica de genes/proteínas relacionadas con la quimiotaxis y en la movilidad flagelar en UYPR2.512.

Anotación RAST/Manual	Nombre gen/proteína (ID)	Función asociada	Replicón
Quimiotaxis	MCP (2718 y 8490)	Aceptora de metilo/sensora/transductora	Chr1 Chr2
	<i>aer</i> (8494)	Flavoproteína/sensora/receptora de aerotaxis	Chr2
	<i>cheAWRDBYZ</i> (8501 al 8507)	<i>cheA</i> : Histidina quinasa (EC 2.7.3.-) <i>cheW</i> : regulador positivo de la proteína CheA <i>cheR</i> : metil transferasa (EC 2.1.1.80) <i>cheD</i> : quimiotaxis <i>cheB</i> : Regulador en respuesta a quimiotaxis, glutamato metilesterasa (EC 3.1.1.61) <i>cheY</i> : Regulador de quimiotaxis, transmite señal del quimiorreceptor a los componentes del motor del flagelo <i>cheZ</i> : fosfatasa	Chr2
	<i>cheA</i> (1718) <i>cheV</i> (7942)	Histidina quinasa (EC 2.7.3.-) Involucrada en quimiotaxis (EC 2.7.3.-)	Chr1 Chr2
Movilidad flagelar	<i>flhBA</i> (8551 8552) <i>flhB</i> (8066)	Biosíntesis del flagelo	Chr2
	<i>flhDC</i> (8497 al 8498) <i>flhD</i> (7112) <i>flhC</i> (9292)	Activador transcripcional flagelar	Chr2
	<i>flhF</i> (8553)	Biosíntesis del flagelo	Chr2
	<i>fliCDSTFGHIJ</i> (8060 al 8072)	<i>fliC</i> : Flagelina <i>fliD</i> : Caperuza del filamento <i>fliS</i> : Biosíntesis del flagelo <i>fliT</i> : Biosíntesis del flagelo <i>fliE</i> : Complejo proteico cuerpo basal y gancho flagelar <i>fliF</i> : Componente proteico Anillo M <i>fliG</i> : Proteína de cambio de motor flagelar <i>fliH</i> : Ensamblaje del flagelo <i>fliI</i> : Asociada al cuerpo basal del flagelo (ATP sintasa) <i>fliJ</i> : Proteína flagelar	Chr2
	<i>fliR</i> (8768)	Biosíntesis del flagelo	Chr2
	<i>fliPONL</i> (8770 al 8774)	<i>fliP</i> : Biosíntesis flagelar <i>fliO</i> : Biosíntesis flagelar <i>fliN</i> : Proteína de cambio del motor flagelar y componente energético, de la familia <i>fliN/mopA/spaO</i> <i>fliL</i> : Asociada al cuerpo basal del flagelo	Chr2
	<i>motAB</i> (8499 al 8500)	<i>motA</i> : Rotación del motor del flagelo y generador de torque <i>motB</i> : Rotación del motor del flagelo	Chr2
	<i>fliN</i> (8554)	Regulador de la síntesis del flagelo	Chr2
	<i>fliA</i> (8555)	Factor sigma	Chr2
	<i>fliA</i> (6673)	Factor sigma	pCne512b

Anotación RAST/Manual	Nombre gen/proteína (ID)	Función asociada	Replicón
	<i>flgD</i> (8562)	Proteína de modificación del cuerpo basal	Chr2
	<i>flgH</i> (8566)	Componente proteico Anillo M	Chr2
	<i>flaG</i> (8061)	Proteína flagelar	Chr2

3.4.2.2 Genes involucrados en la interacción simbiótica y fijación biológica del nitrógeno

En UYPR2.512 los genes *nod*, *nif* y *fix*, se mostraron agrupados en una región de aproximadamente 36 Kb en el pCne512e (pSym) (Fig. 19 y tabla 13).

Esta isla simbiótica incluye genes *nodABCDEFGHIJSUQ* (ID= 4624 al 4632) implicados en la síntesis y exportación de los factores Nod, relacionados con los primeros pasos de la interacción simbiótica a su vez, se identificó el regulador transcripcional *nodD* (ID = 4633). La isla contiene también a los genes *nifHDK* (ID = 4660 al 4664) que codifican para la enzima Nitrogenasa y los *fixABCX* (ID = 4647 al 4650) relacionados con la fijación biológica del nitrógeno.

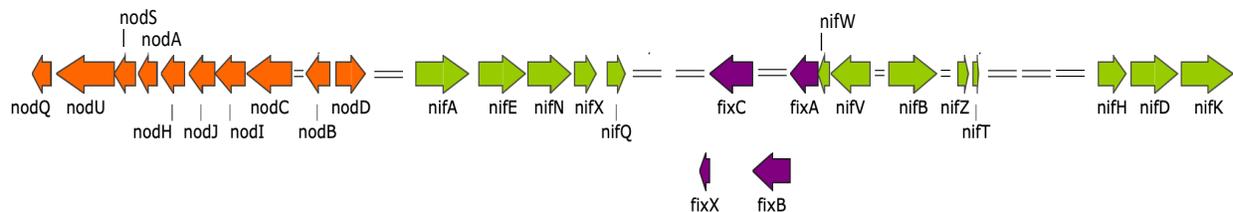


Figura 19. Ilustración de la isla simbiótica de UYPR2.512.

El operón *nod* se indica en color naranja, el operón *nif* en color verde, y parte del operón *fix* en color violeta, los cuales están incluidos en aprox. 36Kb. Realizado con el programa SNAPgene.

Fuera de la isla simbiótica se encontraron los genes *ccoPQONSI* (ID = 2745 al 2751, EC 1.9.3.1), ubicados en el chr1, los cuales codifican para cada una de las subunidades de la citocromo c oxidasa tipo Cbb3 (Tabla 13).

Tabla 13. Características y ubicación genómica de genes/proteínas relacionadas con la simbiosis y con la fijación biológica del nitrógeno en UYPR2.512.

Anotación Rast/Manual	Nombre del gen (ID)	Función asociada	Replicón
Simbiosis	<i>nodD</i> (4633)	Proteína de nodulación D (regulador transcripcional, familia LysR)	pCne512e (pSym)
	<i>nodQAHJICB</i> (4624 al 4632)	<i>nodQ</i> : Adenilil sulfato quinasa (EC 2.7.1.25) <i>nodU</i> : Carbamoiltransferasa <i>nodS</i> : N-metil transferasa <i>nodA</i> : Proteína A de nodulación <i>nodH</i> : Proteína hipotética <i>nodJ</i> : proteína permeasa (Transportador ABC de eflujo) <i>nodI</i> : proteína de unión a ATP (Transportador ABC de eflujo) <i>nodC</i> : Proteína C de nodulación <i>nodB</i> : Poliacárido desacetilasa	pCne512e (pSym)
Fijación biológica del nitrógeno	<i>nifAENXQ</i> (4636 al 4640, 4642 y 4643) <i>frdN</i> (4641)	<i>nifA</i> : Regulador transcripcional <i>nifE</i> : Síntesis y ensamblaje del cofactor FeMo, de la enzima nitrogenasa <i>nifN</i> : Síntesis y ensamblaje del cofactor FeMo, de la enzima nitrogenasa <i>nifX</i> : Síntesis y ensamblaje del cofactor FeMo de la enzima nitrogenasa <i>frdN</i> : Transferencia de electrones (4Fe-4S ferredoxina,) asociada a la enzima nitrogenasa <i>nifQ</i> : Síntesis del cofactor FeMo de la enzima nitrogenasa	pCne512e (pSym)
	<i>nifWVB</i> (4651 al 4653)	<i>nifW</i> : Proteína estabilizadora/protectora de la enzima nitrogenasa <i>nifV</i> : Homocitrato sintasa (EC 2.3.3.14) <i>nifB</i> : Síntesis y ensamblaje del cofactor FeMo, de la enzima nitrogenasa	pCne512e (pSym)
	<i>nifHDKZ</i> (4660 al 4665)	<i>nifH</i> : Síntesis de la dinitrogenasa reductasa Fe <i>nifD</i> : Síntesis de la cadena α de la dinitrogenasa FeMo (EC 1.18.6.1) <i>nifK</i> : Síntesis de la cadena β de la dinitrogenasa FeMo (EC 1.18.6.1) <i>nifZ</i> : Proteína estabilizadora/protectora de la enzima nitrogenasa	pCne512e (pSym)
	<i>nifT</i> (4655)	Función desconocida	pCne512e (pSym)
	<i>fixXCBA</i> (4647 al 4650)	<i>fixX</i> : Proteína tipo ferredoxina <i>fixC</i> : lavoproteína-quinona oxidoreductasa <i>fixB</i> : Subunidad α de la Flavoproteín <i>fixA</i> : Subunidad β de la Flavoproteína	pCne512e (pSym)
	<i>ccoPQONSI</i> (2645 al 2651)	<i>ccoPQONSI</i> : Citocromo c oxidasa tipo Cbb3	Chr1

Por otro lado, al realizar un estudio comparativo de secuencia del operón *nod* con un pariente cercano como ser la cepa LMG 19424 de *C. taiwanensis*, se observa una conservación en los genes *nodQUSAJICB*, precediendo al regulador *nodD* (Fig. 20). Sin

En el genoma de UYPR2.512 se identificaron tres genes; *tolC* (ID = 3321), *hlyD* (ID = 3121) y *hlyB* (ID = 3118), localizados en chr1, los cuales codificarían para las proteínas necesarias para formar un SSTI completo (Fig. 21 y Tabla 14).

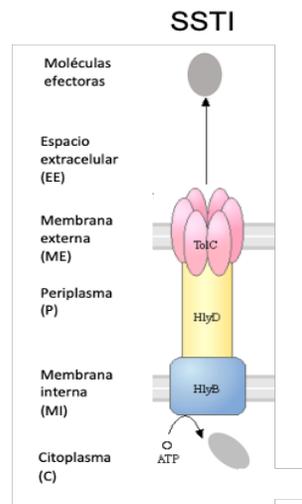


Figura 21. Sistema de secreción del tipo I (SSTI) en UYPR2.512.

Se representan las proteínas identificadas en UYPR2.512, HlyB, HlyD y TolC, ubicados en la membrana interna, en el periplasma y en la membrana externa, respectivamente. Tomado y adaptado de [224].

En el genoma de UYPR2.512 se encontraron genes que codifican para posibles proteínas que forman el núcleo del SSTII, incluidos en el operón *gsp* (ID = 4078, ID = 4081 al 4088, ID = 4089 al 4091). Se identificaron múltiples copias (22) del gen *gspA*, el cual codifica para una proteína esencial para el ensamblaje de la secretina (Tabla 14 y Fig. 22). Se observó que la secuencia del gen *gspA* se encuentra en un contexto genómico de integrasas y regiones repetidas conservado en otras especies del género *Cupriavidus*, sugiriendo su asociación con elementos móviles.

Involucrados en el sistema de translocación Sec, se identificaron los genes *secDF* (ID = 3582, 3583), *secE* (ID = 4056), *secG* (ID = 1273), *secY* (ID = 4019), y los genes *yajC* (ID = 3580) e *yidC* (ID = 148 al 149), los cuales codifican para proteínas ubicadas a través de la membrana. También se encontró el gen *secA* (ID = 3759), que codifica para una proteína con actividad ATPasa citoplásmica asociada a membrana, y el gen *secB* (ID = 586), que codifica para una chaperona que se une cooperativamente a la secuencia

señal de las proteínas secretoras para su translocación [251] (Tabla 14 y Fig. 22). A su vez, se identificaron los genes *tatABC* (ID = 3953, 3955, 3956) formando un operón codificado en el Chr1, los cuales conforman el sistema de translocación Tat. No se encontraron los genes *tatE* y *gspO*.

En relación con una de las subfamilias del SSTII, la cual utiliza el *pilus* de adherencia fuerte (sistema Tad), se identificaron varios genes involucrados en el ensamblaje del *pilus*, que forma parte del sistema, incluyendo *rcpA* (ID = 8441, 982), *tadA* (ID = 1198, 1199, 6748, 8443, 979), *tadB* (ID = 1200, 8444, 978), *tadC* (ID = 1201, 977), *tadV* (ID = 986) y *tadZ* (ID = 980), los cuales se ubicaron mayoritariamente en el chr1 (Tabla 12). Una de las copias del gen *tadA* se ubicó en el replicón (pCne512c), lo que sugiere su posible rol en la maquinaria conjugativa. Proteínas como TadA y TadZ, mostraron alta homología con proteínas involucradas en la maquinaria utilizada por el SSTIV, como las proteínas VirB11 (ubicada tanto en cromosomas como en el pCne512c) y CpaF, que tienen actividad ATPasa y estarían involucradas en el ensamblaje del *pilus* tipo IV. Las proteínas relacionadas al *pilus* no se indican en la Figura 22.

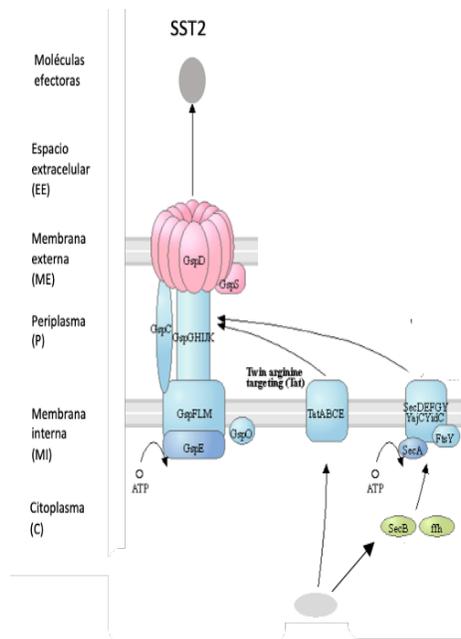


Figura 22. Sistema de secreción del tipo II (SSTII) en UYPR2.512.

Se representan las proteínas que forman parte del operón que codifica para el sistema general de secreción de proteínas o GSP (GspC, GspD, GspE, GspFLM, GspGHIJK), que forman parte del sistema Tat (TatABC) y del sistema Sec (SecABC) de translocación, en UYPR2.512, así como otras proteínas

accesorias a estos sistemas como ser YajC, YidC, FstY y Ffh. Las proteínas que conforman el sistema Sec como del sistema TaT, son las encargadas de transportar las proteínas desde el citoplasma hacia el periplasma, luego mediante la acción del complejo secretina, serán finalmente transportadas hacia el espacio extracelular. Tomado y adaptado de [224].

En el genoma de UYPR2.512 se identificaron varios genes que formarían un operón posiblemente involucrado en la síntesis del SSTIII, incluyendo varias proteínas accesorias ubicadas en la membrana interna y codificadas por los genes *yscSR* (ID = 5489, ID = 5490), *yscT* (ID = 5502), *yscU* (ID = 5494), *yscJ* (ID = 5497) y *yscN* (ID = 5500). También se identificaron genes que codifican la proteína secretina, codificada por *yscC* (ID = 5503) y el inyector o *pilus*, codificado por *yscP* (ID = 5492), que se encuentra en el espacio extracelular (Fig. 23 y Tabla 14). Las proteínas YscF, YscO, YscK, YscW, YscV, YscQ y YscL, no fueron identificadas en el genoma de UYPR2.512.

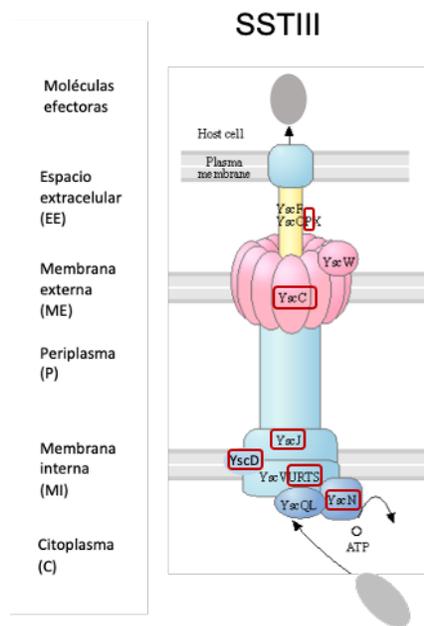


Figura 23. Sistema de secreción del tipo III (SSTIII) en UYPR2.512.

Los círculos rojos indican las proteínas identificadas en el genoma de UYPR2.512, que forman parte del complejo secretina YscC, del *pilus* YscP y de la plataforma de membrana interna YscJ, YscD, YscURTS, y YscN. Tomado y adaptado de [224].

Se identificaron varias copias del gen *virB11* (ID = 979, 1198, 1199, 6748, 8443), ubicados en ambos cromosomas y en el replicón pCne512b, los cuales codifican para

Se identificaron tres genes que codifican para sistemas de transportes de dos componentes (TPS), que incluyen *tpsA2* (ID = 1889, 1890) y *tpsB2* (ID = 1881), que forman parte de los SSTV (Tabla 14). Los genes *tpsA2* y *tpsB2* codifican para una exoproteína HecA, implicada en la adhesión de hemo y para la proteína de activación/secreción de hemolisina HecB, respectivamente conformando el sistema HecA/HecB. Este sistema presenta homología con los sistemas FhaB/FhaC y ShIA/ShIB, los cuales desempeñan un papel importante en la adhesión celular y formación de biopelículas [220].

Se encontraron varios genes que codifican para la maquinaria proteica que conforman el SSTVI, incluyendo dos copias de la proteína VgrG (ID = 1902, 8175), PAAR (ID = 8173), HcP (TssD) (ID = 1895), dos proteínas que forman parte del tubo exterior de inyección (codificadas por los genes *lip* (ID = 1915) y *lcmF* (ID = 8171), así como el gen *clpV* (TssH) (ID = 1894). Además, se identificaron genes formando parte del operón *imp*, incluyendo *impJCB* (ID = 1899 al 1901), *impAG* (ID = 1911 al 1913) e *impH* (ID = 1914), ubicados en el chr1, cuya función está asociada al SSTVI, en base a homología de secuencia, observándose una alta conservación proteica con otras especies del género *Cupriavidus*, y en su mayoría se asocian a proteínas hipotéticas (Fig. 25 y Tabla 14).

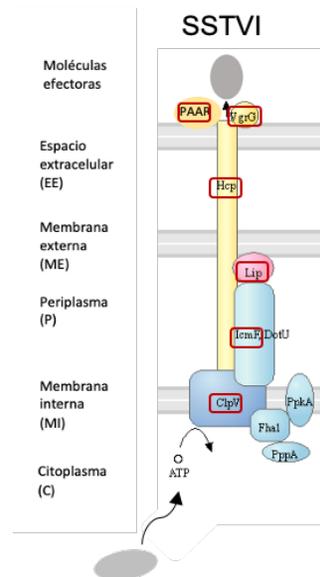


Figura 25. Sistema de secreción del tipo VI (SSTVI) en UYPR2.512.

El círculo rojo indica las proteínas identificadas en el genoma de UYPR2.512, que forman parte del dispositivo de punción de moléculas efectoras VgrG y PAAR, del doble tubo contráctil HcP (TssD), Lip, lcmF y de la plataforma de membrana ClpV (TssH). Tomado y adaptado de [224].

Tabla 14. Características y ubicación genómica de genes relacionados con los sistemas de secreción de moléculas en UYPR2.512.

Tipo de sistema de secreción (RAST)	Genes/proteínas (ID)	Características/Función asociada	Replicón
SSTI	<i>tolC</i> (3321), <i>hlyD</i> (3121), <i>hlyB</i> (3118)	<i>tolC</i> : Proteína de secreción de membrana externa <i>hlyD</i> : Proteína de secreción periplásmica <i>hlyB</i> : Transportador ABC/proteína de unión a ATP	Chr1
	<i>gspA</i> (205, 209, 5153, 5388) <i>gspA</i> (4473, 4479, 4800, 4804, 4899, 4946, 4952) <i>gspA</i> (5706, 5710, 5781, 5785) <i>gspA</i> (6211, 6215, 6379, 6383) <i>gspA</i> (5153, 5388, 5395, 5549)	Proteína de unión al ATP	Chr1 pCne512e (pSym) pCne512a pCne512d pCne512c
SSTII/GSP	<i>gspCDEFGHIJKLMN</i> (4078, 4089 al 4091, 4081 al 4088)	GspC; Proteína involucrada en la especificidad de sustrato, GspD; Función de extensión y retracción del pseudopilus GspEFLMN: complejo-plataforma de membrana GspGHIJK: pistón de las proteínas a través de la secretina	Chr1
SSTII	<i>flp</i> (8438, 988)	Pilina: ensamblaje del <i>pilus</i>	Chr2, Chr1
	<i>rcpC</i> (983, 984)	Ensamblaje del <i>pilus</i> , RcpC/CpaB	Chr1
SSTII/SSTIV	<i>rcpA</i> (8441, 982)	Secretina: asociada con el ensamblaje de la proteína del <i>pilus</i> Flp	Chr2, Chr1
	<i>tadA</i> (1198, 1199, 6748, 8443, 979)	TadA: ATPasa hidrolasa (TadA/VirB11/CpaF)	Chr1, Chr1, pCne512b
	<i>tadB</i> (1200, 8444, 978)	TadB: Ensamblaje del <i>pilus</i> Flp	Chr2
	<i>tadC</i> (1201, 977), <i>tadV</i> (986), <i>tadZ</i> (980)	TadC: Asociada con el ensamblaje de la proteína del <i>pilus</i> de adherencia Flp TadV: Prepilina: peptidasa, TadV/CpaA TadZ: ATPasa TadZ/CpaE, asociada con el ensamblaje de la proteína del <i>pilus</i> Flp	Chr1, Chr1, Chr1
SSTII/Sec	<i>secA</i> (3759), <i>secB</i> (586)	SecA: Proteína de translocación (translocasa) SecB: Proteína de exportación (mantiene el estado desplegado previo a la exportación)	Chr1, Chr1
	<i>secDFEGY</i> (3582, 3583, 4056, 1273, 4019)	Proteínas de translocación	Chr1
	<i>yajC</i> (3580)	Proteínas de translocación	Chr1

Tipo de sistema de secreción (RAST)	Genes/proteínas (ID)	Características/Función asociada	Replicón
	<i>yidC</i> (148 al 149)	Proteínas de translocación/chaperona	Chr1
	<i>ftsY</i> (621)	Receptor de reconocimiento de partículas señal	Chr1
	<i>ffH</i> (3700)	Proteína de reconocimiento de partícula señal (SRP) (EC 3.3.5.4)	Chr1
SSTII/Tat	<i>tatCBA</i> (3953 al 3956)	<i>tatCB</i> : Proteína de unión al péptido señal de la proteína a secretar <i>TatA</i> : Junto con <i>TatBC</i> , forman un canal para el transporte activo impulsado por un gradiente de protones	Chr1
SSTIII	<i>yscD</i> (5440) <i>yscSR</i> (5489 al 5490) <i>yscPU</i> (5492 y 5494) <i>yscJNTC</i> (5497, 5500, 5502 y 5503)	<i>yscD</i> : Proteína de membrana interna, involucrada en la exportación del flagelo <i>yscSR</i> : Proteína de secreción de membrana interna <i>yscP</i> : Proteína de secreción en el espacio extracelular <i>yscU</i> : Proteína de secreción de membrana interna <i>yscJ</i> : Puente de secreción entre la membrana interna y la lipoproteína de la membrana externa <i>yscN</i> : ATP sintasa citoplasmática (EC3.6.3.1.4) <i>yscT</i> : Proteína de secreción de membrana interna <i>yscC</i> : Proteína de secreción de membrana externa formadora de poro	pCne512c
SSTIV	<i>virB11</i> (979, 1198, 1199, 6748, 8443) <i>pilBCD</i> (3748 al 3750) <i>pilT</i> (3986) <i>pilPONM</i> (3989 al 3992) <i>pilWXY1E</i> (3285 al 3287), <i>pilWX</i> (675), <i>pilY1</i> (677), <i>pilE</i> (672)	ATP hidrolasa/ensamblaje del <i>pilus</i> . <i>pilB</i> : ATPasa, asociada a la formación de fimbrias <i>pilC</i> : Proteína asociada al ensamblaje de fimbrias <i>pilD</i> : Peptidasa pre- <i>pilus</i> (EC 3.4.23.43) Proteína de retracción <i>pilPONM</i> : Proteína de biogénesis del <i>pilus</i>	chr1 pCne512c chr2 Chr1 Chr1 Chr1
SSTV	<i>tpsB2</i> (1881), <i>tpsA2</i> (1889-1890)	<i>TpsB2</i> : Proteína de activación/secreción de hemolisina de la familia ShIB/FhaC/HecB <i>TpsA2</i> : Exoproteína implicada en la utilización o adhesión de hemo de la familia ShIA/HecA/FhaA	Chr1 Chr1
SSTVI	<i>vgrG</i> (1902 y 8175) PAAR (8173)	No caracterizada Facilita la secreción de efectores	Chr1, Chr2 Chr2

Tipo de sistema de secreción (RAST)	Genes/proteínas (ID)	Características/Función asociada	Replicón
	<i>clpV</i> y <i>hcP</i> (1894 al 1895)	ClpV: ATPasa desarma la estructura del sistema Hcp: ubicada en la parte interior del doble tubo	Chr2
	<i>impJ/vasE</i> (1899)	No caracterizada	Chr2
	<i>impC/evpB</i> y <i>impB/vipA</i> (1900 al 1901)	No caracterizada	Chr2
	<i>impAG/HDUF 879</i> (1911 al 1914)	No caracterizada	Chr2
	<i>liP</i> (1915)	Putativa lipoproteína	Chr2
	<i>lcmF</i> (8171)	Probable proteína de membrana YPO1482	Chr2

3.4.2.4 Modulación de fitohormonas

Mediante el análisis de los subsistemas presentes en el servidor web RAST, no se encontró ningún gen involucrado en la biosíntesis de AIA.

Se buscaron los genes relacionados con la síntesis de la hormona etileno y se observó la presencia del gen *acdS* (ID = 4617, 9710), cuyas dos copias se ubicaron en el pSym y en el chr2, respectivamente. El gen *acdS* codifica para enzima ACC desaminasa, la cual convierte el 1-aminociclopropano-1-carboxilato (ACC), precursor inmediato del etileno, en amoníaco y α -cetobutirato, los cuales pueden ser utilizados como fuente de nutrientes por la planta [253].

4 Discusión

4.1 Asignación de especie a la cepa UYPR2.512 de *Cupriavidus* sp.

Los primeros estudios de caracterización de la cepa UYPR2.512, proveniente de una colección de 42 aislamientos obtenidos de nódulos de angico (*Parapiptadenia rigida*), habían mostrado que pertenecía al género *Cupriavidus* [42]. Posteriormente en el año 2015, se secuenció su genoma (versión_1). Mediante estudios filogenéticos se comprueba que las especies más cercanas eran de cepas tipo, pertenecientes a las especies *C. necator*, *C. oxalaticus* y *C. taiwanensis* [254] y que formaba parte de un clado

monofilético que incluía otras cepas de rizobios aisladas a partir de *Mimosas* spp. nativas [165] y otros rizobios pertenecientes a *C. necator* [43].

La reconstrucción filogenómica (Fig. 12) y los valores de ANI (Tabla 4) obtenidos a partir de este trabajo sugieren fuertemente que UYPR2.512 pertenece a la especie *C. necator*. A su vez UYPR2.512, sería la cepa más divergente dentro del clado que integra a las cepas de *C. necator*, sugiriendo diferencias en variabilidad genética y evolutiva con respecto a las otras cepas. Esto podría tener implicaciones significativas, como por ejemplo confiriéndole la capacidad para adaptarse a distintas condiciones ambientales.

4.2 Análisis genómico

A partir de los resultados que confirmaron que UYPR2.512 es un representante de la especie *necator* presente en Uruguay, y en base a que solo se disponía de la versión borrador de su genoma, se realizó una nueva secuenciación con la tecnología de PacBio de forma de conocer con mas profundidad, contenido y composición de su genoma.

Al igual que como se había observado para las cepas N-1^T de *C. necator* y STM815 de *P. phymatum*, UYPR2.512 presentó un genoma multipartito formado en este caso por siete replicones, de los cuales 4 se lograron cerrar (Tabla 5).

El hecho de que 3 *contigs* no se lograran cerrar planteó la interrogante de que si eran moléculas circulares o lineales. Los plásmidos lineales se caracterizan por tener secuencias palindrómicas en sus extremos y/o proteínas unidas al extremo 5' terminal, las cuales facilitan su identificación [196]. Esta característica no se observó en los tres *contigs* que no se lograron cerrar (pCne512a, pCne512c y pCne512d). Sin embargo, sí se observó la presencia de genes relacionados con la partición/replicación como ser *parAB* y *repB*, sugiriendo que estas moléculas de ADN son plásmidos circulares (Tabla 6). Los restantes *contigs* (tig000000104, tig00000106 y tig00003540), estaban formados por un número muy bajo de lecturas y no presentaron genes relacionados con la replicación ni el mantenimiento de este, por lo que no fueron considerados en el resto del trabajo.

El tamaño final del nuevo genoma de UYPR2.512 es de 9,5 Mb, siendo similar al de la cepa N-1^T de *C. necator* (8,5 Mb) [255] y STM815^T de *P. phymatum* (8,7 Mb) [196] y

bastante mayor al de *C. taiwanensis* LMG1940^T (6,5 Mb) [164]. Los genomas grandes y multipartitos se encuentran en aproximadamente el 10% de los genomas bacterianos secuenciados [179] y especialmente entre los miembros de *Rhizobiales* y *Burkholderiales* [192] donde se ha demostrado que los replicones secundarios están enriquecidos en genes importantes para la adaptación bacteriana a nichos específicos [179].

Como se comentó previamente UYPR2.512, mostró en ambos cromosomas un %GC dentro del rango promedio observado en la familia de *Burkholderiaceae*, mientras que los plásmidos presentaron un % GC considerablemente menor. Esto le brindaría a los cromosomas de UYPR2.512 una mayor estabilidad, ya que es menos probable la rotura del enlace GC, en comparación con AT. A su vez, minimiza el costo metabólico del mantenimiento de genes no esenciales presentes en los plásmidos permitiéndole mayor flexibilidad adaptativa a diferentes entornos.

4.3 Características genómicas

4.3.1 Tolerancia a metales pesados

Se identificaron genes relacionados con la utilización, exportación y secuestro de metales, así como proteínas sensoras y reguladores transcripcionales de respuesta a las mismas. Se encontraron más de 50 proteínas involucradas en la exportación de cobre y a la tolerancia a diversos cationes divalentes (Tabla 8). En un trabajo previo realizado en nuestro Departamento, se demostró que la cepa UYPR2.512 era capaz de crecer en presencia de Zn⁺² 5mM, Cd⁺² 2mM, Co⁺² 2mM, Ni⁺² 2mM y Cu⁺² 2mM. Estas concentraciones son similares a las que tolera la cepa CH34 de *C. metallidurans*, conocida por su alta tolerancia a metales en el suelo [65]. La capacidad de UYPR2.512 de crecer en presencia de altas concentraciones de metales, también es comparable a la de la cepa *C. neocaledonicus* STM 6070, simbionte de *M. pudica*, aislada de suelos ricos en metales [185].

La peculiaridad de UYPR2.512 de poseer genes involucrados en el sentido y transporte del cobre, así como de otros metales, la vuelve una excelente candidata para asistir en el proceso de biorremediación o en combinación con plantas acumuladoras de metales pesados [65].

Se observó la presencia de genes posiblemente involucrados en la capacidad para adquirir y sintetizar sideróforos, en particular para el transporte y el reconocimiento del sideróforos tipo vibriobactina (Tabla 9). Esta capacidad, también observada en otros representantes del género *Cupriavidus* [78], le podría conferir a UYPR2.512 la habilidad de secuestrar sideróforos producidos por otras bacterias, favoreciendo el crecimiento de esta bacteria frente a otros competidores [256].

A su vez, se observaron dos copias de FUR, regulador transcripcional del metabolismo central del hierro. Una copia se encontró cercana a un gen que codifica para un transportador del tipo ABC, siendo este tipo de transportadores uno de los mecanismos más utilizados para la internalización del complejo Fe^{+3} -sideróforo al citoplasma [257]. La otra copia se encontró en un contexto genómico que no estaría relacionado directamente con la homeostásis del metal, sino que estaría involucrada en la respuesta al estrés oxidativo [248]. Estos resultados sugieren que la cepa UYPR2.512 tiene la capacidad de adquirir y utilizar el hierro en diversas condiciones, lo que puede ser importante para su crecimiento y supervivencia en diferentes ambientes.

4.3.2 Degradación de compuestos aromáticos

La capacidad de degradar compuestos aromáticos presentes en exudados radiculares les confiere a las bacterias una ventaja para la colonización de las raíces y su desarrollo en la rizósfera. La cepa UYPR2.512 presentó genes que codificarían enzimas involucradas en la vías periféricas y centrales de degradación de compuestos aromáticos (Tabla 10). Estas vías convergen en la formación de β -cetoadipato, un intermediario que puede incorporarse al metabolismo central, en forma de acetil-CoA o succinil-CoA, al igual que se observó en *C. pintaubonensis* y en otras cepas del tipo *necator* [89].

A modo de ejemplo, mutantes en genes involucrados en ambas ramas de la vía del β -cetoadipato, mostraron una reducción en la capacidad de degradar compuestos aromáticos presentes en los exudados radiculares [258].

Teniendo en cuenta que en la rizósfera gran parte de la materia orgánica está compuesta por lignina (polímero fenólico) y otras moléculas orgánicas, la presencia de genes involucrados en la degradación de diversos compuestos aromáticos, la darían una

ventaja a UYPR2.512 frente a otros competidores en dicho ambiente, como se ha observado en otras cepas del tipo *necator*, como ser C39, NH9 y JMP134 [89], [259].

4.3.3 Metabolismo de PHB

Distintas especies del género *Cupriavidus* son capaces de producir, almacenar, y biodegradar moléculas de PHB, en diferentes ambientes, lo cual les ha permitido sobrevivir en condiciones de estrés o en entornos competitivos. En particular, diversas cepas de *C. necator* son usadas para la obtención de estas moléculas como materia prima para la producción de plásticos biodegradables [260].

En este trabajo se identificaron los genes involucrados en la vía clásica de síntesis de PHB, en la acumulación de gránulos, en la degradación y en la regulación de su síntesis (Tabla 11). Esta organización genómica conservada se ha observado en diferentes cepas de *C. necator* como ser en JMP134 y en la cepa tipo N-1 [261], [262], así como también en otras especies del género *Cupriavidus*, como por ejemplo en *C. taiwanensis* y en *C. metallidurans* [263].

Por otro lado, estudios de microscopía realizados en nuestro Departamento, mostraron la presencia de gránulos de almacenamiento de PHB en UYPR2.512, lo cual confirma la presencia de dichos genes en el genoma (Fig. 26).

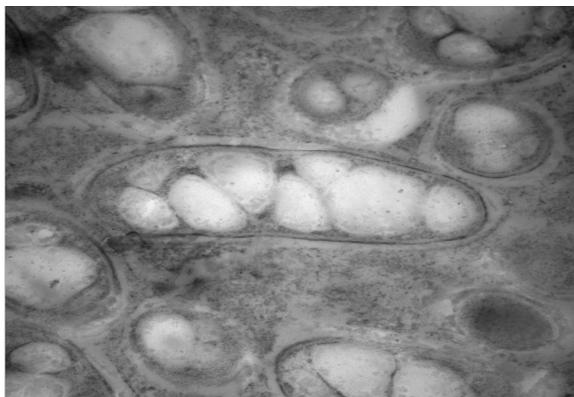


Figura 26. Imagen tomada mediante microscopía electrónica de transmisión (MET) de un bacteroide de la cepa UYPR2.512 en el interior de un nódulo de *M. pudica*.

Escala 2900x.

Estos resultados sugieren que UYPR2.512, sería capaz de producir y almacenar PHB, lo cual podría ser ventajoso por ejemplo en condiciones limitantes de crecimiento. Este

es el caso por ejemplo de la cepa ORS571 de *Azorhizobium caulinodans*, la cual en condiciones limitantes de oxígeno, es capaz de equilibrar la producción de poder reductor acumulado, redireccionándolo hacia la formación de PHB [264].

4.4 Características genómicas involucradas en la interacción de UYPR2.512 con sus hospederos

4.4.1 Quimiotaxis y movilidad bacteriana

En el genoma de UYP2.512 se identificaron varios genes codificantes para las proteínas quimiorreceptoras transmembrana del tipo MCP y Che (Tabla 12). En dicho contexto, se observó que mutantes en los receptores de quimiotaxis *mcpB* y *mcpC* de la cepa *R. leguminosarum biovar viciae* VF39SM, fueron menos competitivos en la nodulación en plantas de guisantes. Mientras en *S. meliloti* RU11, el quimiorreceptor McpU reconoce la prolina presente en los exudados de semillas de alfalfa, activando la quimiotaxis de la bacteria hacia la planta [265], [266].

La presencia de este quimiorreceptor en UYPR2.512 sería clave para mejorar la colonización de la planta hospedera en comparación con otras bacterias. También desempeñaría un papel fundamental en el establecimiento de una simbiosis efectiva, como se ha observado en otras interacciones simbióticas, por ejemplo, en *S. meliloti* y plantas de porotos [267].

La movilidad bacteriana impulsada por flagelos es un rasgo importante en la mayoría de los ambientes, donde los microorganismos deben explorar el hábitat en busca de recursos beneficiosos o evadir aquellos perjudiciales. En UYPR.512 están presentes los genes que codifican para las proteínas involucradas en la síntesis, ensamblado y el movimiento flagelar, e incluso reguladores transcripcionales, así como los genes asociados a la quimiotaxis. La presencia de genes involucrados en los procesos de la quimiotaxis y la movilidad impulsada por flagelos en UYPR2.512, constituyen funciones cruciales para la colonización bacteriana de las raíces, lo que sugiere una ventaja frente a otros competidores [268], [269].

4.4.2 Simbiosis y fijación biológica del nitrógeno

El análisis del genoma de UYPR2.512 mostró que los operones *nod*, *fix* y *nif* están fuertemente compactados (en aprox. 36 Kb) y conservados respecto a lo observado en otras cepas de rizobios pertenecientes al género *Cupriavidus* (Fig. 15). Se vió también que los genes *nod* se encuentran organizados en un solo operón que está precedido por una caja *nod*, de manera casi idéntica a lo observado en la cepa LMG 19424 de *C. taiwanensis* (Fig. 19) [194]. En esta organización se observan dos sitios de inicio de transcripción divergentes, uno para el gen que codifica el regulador transcripcional NodD y el otro para el operón *nod* que comprende los genes *nodBCIJHASUQ* [270]. Al alinear las secuencias de ambos modelos bacterianos, la región intergénica entre *nodD* y *nodB*, aunque es ligeramente diferente, contiene una caja *nod*, postulada como necesaria para la regulación del operón *nod* por NodD [194]. Esta ligera diferencia en las regiones intergénicas podría sin embargo, llevar a diferentes mecanismos de regulación.

A su vez, una segunda caja *nod* conservada se encontró en el genoma de UYPR2.512, al igual que en la cepa LMG19424 de *C. taiwanensis* [194]. Esta caja *nod* se encuentra a 195 pb aguas arriba del codón de inicio ATG de la copia del gen *noeM* (ID = 4812), la cual en la cepa LMG19424 de *C. taiwanensis* codifica para una proteína esteroles desaturasa, involucrada en la síntesis de moléculas atípicas de NFs [271], [272]. Sin embargo, UYPR2.512 contiene otra copia del gen *noeM* (ID = 4545), la cual codifica para una proteína hipotética. En LMG19424, se demostró que su expresión se induce por la presencia de luteolina, apigenina o exudados de raíces de *M. pudica* [272], [273]. Por lo cual se postula que la copia del gen *noeM* (ID = 4545), correspondería a la copia funcional, y debido a que no contiene una caja *nod* previa, su expresión no estaría directamente regulada por *nodD*.

En la cepa LMG 19424 de *C. taiwanensis* se ha demostrado que los NFs son quito-oligómeros pentaméricos sulfatados (sustituyente adicionado por el producto del gen *nodH*) en el extremo reductor y N-acilados por ácido vaccénico (C18:1) o ácido palmítico (C16:0) en su extremo no reductor. Además, la mayoría de las moléculas están sustituidas por un grupo N-metilo (sustituyente adicionado por el producto del gen *nodS*)

y carbamoilo (sustituyente adicionado por el producto del gen *nodU*) en el azúcar no reductor terminal [164]. La similitud de secuencias y de conservación de los genes *nodBCIJHASUQ* entre UYPR2.512 y LMG 19424 de *C. taiwanensis* observada, sugiere que un factor de nodulación de estructura similar les permitiría a ambas cepas, asociarse simbióticamente con el mismo hospedero. La caracterización de NFs producidos por UYPR2.512 nos permitiría evaluar dicha hipótesis.

Con el fin de asegurar la actividad de la nitrogenasa, los rizobios han desarrollado diferentes mecanismos de regulación, los cuales incluyen la presencia de citocromos tipo b y c, brindando un ambiente reducido y microaerófilico en el interior del nódulo[274], como es el caso de la oxidasa terminal del citocromo tipo Cbb3 (codificada por *fixNOQP*) [275].

Se ha reportado que los genes *fixNOQP* descritos en los rizobios son homólogos a los genes *ccoNOQP* de otras bacterias no fijadoras de nitrógeno, lo cual indica que ambos comparten un ancestro en común y tienen una similitud estructural o funcional [195]. Sin embargo, las diferencias en sus contextos genómicos y las funciones específicas que desempeñan en sus respectivos nichos ecológicos reflejan una capacidad adaptativa a diferentes ambientes.

En UYPR2.512 se observó que los genes *ccoPQON*, requeridos para la producción y ensamblaje de la oxidasa terminal, se ubicaron en el chr1, al igual que sus ortólogos *fixNOQP* de *C. taiwanensis* (Tabla 13). Por otro lado, los genes *fixNOQP* están ausentes en todas las cepas de *Paraburkholderia* analizadas hasta el momento [195]. Estos resultados sugieren algunos β -rizobios tendrían otro citocromo involucrado en la fijación del nitrógeno.

4.4.3 Sistemas de secreción

El análisis del genoma de la cepa UYPR2.512 mostró la presencia de secuencias codificantes para proteínas que conforman los SSTI, SSTII, SSTIII, SSTIV, SSTV y SSTVI (Tabla 14 y Fig. 22).

La cepa UYPR2.512 presentó todos los genes necesarios para completar el SSTI, lo cual le brinda el beneficio de secretar y liberar de forma eficiente moléculas involucradas en la obtención de nutrientes (lipasas, proteasas) [276], infectar el hospedero (toxinas), así como establecer interacciones simbióticas efectivas [277]. En dicho contexto, se ha visto que este tipo de sistema es capaz de secretar proteínas con actividad glicanasas (PlyA y PlyB), las cuales hidrolizan el exopolisacárido, componente fundamental en la formación de biopelículas, como se observó en *R. leguminosarum*, resultando clave para una simbiosis eficiente [278]. La proteína NodO (proteína de unión al Ca^{+2}), la cual estaría involucrada en el reclutamiento de NFs, así como también en la formación de los hilos de infección necesarios para el ingreso de la bacteria sería transportado por este sistema [279].

Se logró mapear casi la totalidad de los genes involucrados en el SSTII según lo reportado en la literatura, incluyendo los que codifican para el sistema general de secreción GSP, los que codifican para los sistemas de translocación SeC y TaT, lo cual además ha sido observado en otras bacterias del género *Cupriavidus* [78]. Asimismo, se identificó el clúster *tad*, el cual codifica para proteínas involucradas en el ensamblaje del *pilus* de adhesión. Este cluster se encuentra formando parte de una isla de colonización generalizada (WCI, de sus siglas en inglés *Widespread Colonization Island*), involucrada en formación de biopelículas y en la colonización, como se ha observado en bacterias del género *Burkholderia*, *Mesorhizobium*, y *Sinorhizobium* [209]. La diversidad de compuestos que secretan estos sistemas, incluyendo toxinas y enzimas hidrolíticas, tales como celulasas, elastasas, amilasas, proteasas, fosfatasas, nucleasas y lipasas, hace pensar que UYPR2.512 tendría la capacidad de secretar una amplia variedad de compuestos permitiendo su interacción con diversos organismos [280].

Formando parte del SSIII, se identificó el clúster *ysc*, el cual media el transporte de proteínas efectoras hacia el espacio extracelular. Varias de estas proteínas presentaron similitud estructural con las proteínas implicadas en la formación del flagelo bacteriano, lo que sugiere una relación evolutiva entre ambos sistemas y un mecanismo común de reconocimiento y exportación de moléculas [281]. Mediante esta interacción se logra reprimir el sistema inmunológico del hospedero, favoreciendo el establecimiento de la

infección y la nodulación. Contrariamente, las proteínas secretadas pueden ser reconocidas por proteínas receptoras de resistencia (R), y activar una inmunidad desencadenada por efectores (ETI, de sus siglas en inglés *Effector-Triggered Immunity*), bloqueando la infección, y resultando en una interacción incompatible [282]. Por otro lado, si bien la virulencia de las bacterias patógenas de las plantas depende en gran medida del funcionamiento del SSTIII, su importancia en la interacción simbiótica presenta muchas interrogantes y necesita ser más estudiada.

La organización genética del SSTIII en *Cupriavidus* no es universal. Su rol en la simbiosis no está claro, por un lado se sabe que no es esencial para establecer simbiosis con *M. pudica* y su eliminación en LMG 19424 de *C. taiwanensis* hace que también nodule a *Leucaena leucocephala* [283]. Asimismo, es uno de los sistemas que aparecen como más conservados al analizar los genomas de rizobios [284].

Se identificaron los genes que forman parte del operón *pil*, los cuales codifican para proteínas involucradas en la formación de las fimbrias y del T-*pilus*, que forman parte del SSTIV, y se ubicaron tanto en el cromosoma como en plásmidos. Estas estructuras forman parte de la maquinaria conjugativa de plásmidos encontrada en algunas bacterias y participan en la transferencia conjugativa de los plásmidos, así como también en la adhesión y colonización a las células del hospedero [240]. En dicho contexto, en un estudio en el que se expuso un cultivo de la cepa LMG 19424 de *C. taiwanensis* a exudados radiculares de *M. pudica*, se observó la sobreexpresión de los genes *pilVWXYE* y *pilQPONM*, los cuales codifican para proteínas involucradas en la adhesión de la raíz, indicando que pueden desempeñar un papel en la colonización de la planta hospedera [273]. Sin embargo, en cuanto al rol de los SSTIV en las interacciones simbióticas, actualmente se dispone de escasa información. En el caso de los rizobios, estos sistemas presentan una alta homología con las subunidades VirB/VirD4 identificadas en *Agrobacterium* [285]. La transcripción de los SSTIV está regulada por el sistema de dos componentes VirA/VirG, el cual es inducido por flavonoides, incluyendo la naringenina [286]. En relación con esto, se ha identificado la presencia de los genes *virD4* y del operón *virB1-B11* en el pSym de la cepa STM815 de *B. phymatum*, así como también en el pRALTA de *C. taiwanensis*, sugiriendo su posible participación en la

transferencia conjugativa [273]. A pesar de ello, aún se requiere de más investigaciones para determinar de manera precisa la función de los SSTIV en las simbiosis.

Hay muy poca información con relación a la participación del SSTV en las interacciones simbióticas. En el genoma de UYPR2.512 se identificaron proteínas involucradas en los sistemas TPS (HecA/HecB), los cuales presentaron homología con componentes pertenecientes a dichos sistemas en otras bacterias (Tabla 14 y Fig. 22) [287]. Se ha reportado que la exportación y adhesión del efector HecA de *Erwinia sp*, provoca la podredumbre en la planta *Nicotiana clevelandii* [288]. Se ha observado un efecto similar en *Xanthomonas sp*, al exportar la adhesinas Fha, jugando un papel crucial en la colonización de las superficies vegetales y la posterior infección del hospedero [289]. Por otro lado, en interacciones benéficas entre bacterias del género *Azospirillum* y sus plantas hospederas, se ha reportado el papel de HecA en la formación de biopelículas en la adhesión a las raíces de las plantas y en la producción factores de crecimiento que promueven el crecimiento vegetal [290]. La presencia en UYPR2.512 de los genes HecA/HecB sugiere su posible rol en la adhesión y colonización de otros organismos.

La mayoría de los estudios que brindan información sobre los componentes de los SSTVI, son en base a homología de secuencia y la mayoría de las proteínas no han sido caracterizadas. Formando parte del SSTVI de UYPR2.512, se identificaron genes que codifican para proteínas estructurales y para la maquinaria de inyección, los cuales también se han observado en otros rizobios [290] (Tabla 14 y Fig. 22). Se ha discutido la participación de dicho sistema de secreción en la modificación y manipulación de diversos procesos celulares que permiten que las bacterias colonicen y sobrevivan, influyendo en la adhesión, la estimulación de la internalización de la bacteria en la planta, reordenamiento del citoesqueleto y la evasión de la respuesta inmune innatas del huésped [291]. En dicho contexto, se observó que mutaciones en el SSTVI en *P. phymatum*, simbionte de *Mimosa spp*, afectan la formación de biopelículas, disminuyendo su agregación, perdiendo competitividad frente a otras bacterias colonizadoras [292]. Por lo cual, la presencia del SSTVI en la cepa UYPR2.512, le podría conferir la habilidad de mejorar la competencia frente a otras bacterias del suelo, capaces de nodular la misma planta hospedera. Por otro lado, se reportó un trabajo en el cual se

relacionó el SSTVI con la simbiosis, en el cual plantas de *P. vulgaris* fueron inoculadas con cepas mutantes de *Rhizobium etli*, defectuosas en genes que codifican para componentes estructurales del SSTVI (*tssD*, *tssA*, *tssM*), observándose la formación de nódulos de tamaño más pequeño en comparación con la cepa salvaje [293]. Sin embargo, la actividad de la nitrogenasa no se vio afectada, lo cual sugiere que dichas mutaciones podrían afectar eventos más tempranos.

4.4.4 Síntesis de fitohormonas y su regulación

La identificación de *acds*, gen *que* codifica para la enzima 1-aminociclopropano-1-carboxilato (ACC) deaminasa (EC 3.5.99.7) le podría conferir la capacidad a UYPR2.512 de promover el crecimiento vegetal de su planta hospedera, regulando los niveles de etileno, al igual que se ha observado en otros rizobios [294].

En resumen en este capítulo la re-secuenciación de UYPR2.512, mejoró su genoma, identificando cuatro replicones cerrados, incluyendo un plásmido con genes simbióticos *nod*, *nif* y *fix* en una isla conservada. Se confirmó que pertenece a la especie *necator*. El análisis genético reveló características de adaptación, como tolerancia a metales y degradación de compuestos aromáticos. Además, UYPR2.512 exhibe componentes proteicos relacionados con la simbiosis y la fijación biológica de nitrógeno, lo que evidencia su versatilidad y capacidad de adaptación a diferentes entornos.

CAPITULO II

Análisis transcriptómico de la respuesta de la cepa UYPR2.512 de *Cupriavidus necator* a la presencia de luteolina

Objetivo general

Identificar los cambios a nivel de la expresión génica provocados por la luteolina en UYPR2.512.

Objetivos específicos

1. Confirmar el efecto inductor de la luteolina en la expresión de los genes *nod* de UYPR2.512.
2. Describir los genes y las rutas metabólicas activadas o reprimidas en presencia de luteolina.

Resumen

Este capítulo se enfocó en el estudio de una de las primeras etapas de la interacción planta-bacteria, en donde el reconocimiento de moléculas, como los flavonoides presentes en los exudados radiculares, son claves. Partiendo de la información de que la luteolina es inductora de los genes *nod* en *C. taiwanensis*, se decidió evaluar el efecto de dicho flavonoide en UYPR2.512. Para ello, dicha cepa se transformó con el plásmido pCBM01 que contiene la fusión transcripcional *pnodB*₁₉₄₂₄-*lacZ* y con pcZ388 (plásmido control que contiene el gen *lacZ* sin promotor). Se evaluó la capacidad inductora del flavonoide mediante la medición de la actividad de la enzima β -galactosidasa.

Posteriormente, se analizó la respuesta a nivel transcripcional de UYPR2.512 a la exposición a luteolina mediante el RNA-Seq. Se utilizaron dos plataformas de secuenciación, la primera incluyó la utilización del Ion Torrent™ (*Ion PGM™ Sequencing Kit de Life Technologies™*), la cual se realizó en colaboración con el Departamento de Genómica del IIBCE y la segunda en la plataforma de Illumina en el servicio privado de

Macrogen Inc. Luego, con el fin de conocer las funciones y procesos metabólicos a los cuales se encuentran relacionados los *DEG* (genes diferencialmente expresados), se realizó un análisis de enriquecimiento funcional. Los resultados obtenidos permitieron generar un modelo integrativo en el cual se presentan las principales rutas metabólicas, funciones biológicas, reacciones químicas, enzimas y genes regulados en UYPR2.512 en respuesta a la exposición a luteolina.

1 Introducción

1.1 Transcriptómica. RNA-Seq como aproximación funcional

El transcriptoma es el conjunto completo de transcritos de ARN en una célula o tejido, e incluye el ARN ribosomal (ARNr), ARN mensajero (ARNm), ARN de transferencia (ARNt), micro ARN (miARN) y otros ARN no codificantes (ncARN). Su análisis tiene como objetivo cuantificar el nivel de expresión de cada gen del genoma en respuesta a cambios definidos. En las últimas décadas, se han desarrollado varias herramientas que permiten desentrañar los cambios transcripcionales globales, basados en la hibridación de ADNc (microarreglos de ADN) o en la secuenciación de ADNc (secuenciación de ARN o RNA-Seq) [295]. Actualmente, este último es el más utilizado, debido a que permite la secuenciación directa de ARN (o ADNc obtenido a partir del ARN). Esta tecnología ofrece una mayor resolución y especificidad, un mayor rango dinámico de niveles de expresión (> 4 órdenes de magnitud) y un menor requerimiento de material de entrada (ng), en comparación con los microarreglos [296].

El RNA-Seq incluye los siguientes pasos: extracción del ARN_{tot}, la depleción de los ARN ribosomales (ARNr), la obtención del ADNc a partir del ARNm, la adición de adaptadores (con secuencia única en uno, o ambos extremos) para la construcción de librerías de ADNc y su secuenciación. Finalmente, las secuencias obtenidas son analizadas mediante métodos bioinformáticos.

Debido a que el ARNr es el ARN predominante en el ARN_{tot}, el mismo debe ser eliminado o depletado, sino gran parte de las secuencias obtenidas serán de estas moléculas que, en la mayoría de las aplicaciones, son menos informativas que el ARNm.

El ARNr se puede eliminar inmediatamente después del aislamiento del ARN total o durante la preparación de bibliotecas de ADNc. Los primeros dos procedimientos se basan en la captura y remoción de las subunidades 16S y 23S del ARN ribosómico bacteriano presentes en el ARN_{tot}, mediante el uso de oligonucleótidos de captura. Dichos oligos bipartitos se unirán por complementariedad del extremo 5' a una región del ARNr y por su extremo 3' a un oligonucleótido que se encuentra unido covalentemente a esferas magnéticas. A diferencia de estas metodologías, las cuales depletan transcritos no deseados (ARNr) previo a la síntesis de las bibliotecas de ADNc, el fabricante NuGen, utiliza cebadores específicos contra transcritos no deseados en las bibliotecas de ADNc, evitando eventos de hibridación cruzada (consultado en www.agilent.com [297]).

1.2 Análisis de datos de secuencias obtenidas mediante RNA-Seq

Independientemente de la tecnología de secuenciación utilizada, como resultado se obtiene un conjunto de archivos, con un formato estándar *FASTQ* [298], que permite almacenar las secuencias. Los valores de calidad de la secuencia están codificados en la escala *Phred* [299], la cual da información de la probabilidad de error en la asignación de nucleótidos. Existen dos bases ASCII (la base 33 y la 64) principalmente utilizadas para codificar estos valores de calidad en los archivos *FASTQ*, que suelen ir de un valor de Q desde 0 hasta 40, que representan la calidad de cada nucleótido.

En la mayoría de los estudios de RNA-seq, los análisis de datos constan de los siguientes pasos: (1) control de calidad de las secuencias obtenidas, (2) alineamiento de las lecturas contra el genoma de referencia o ensamblado *de novo*, (3) cuantificación de las lecturas alineadas a los genes, (4) normalización de los datos e identificación de los genes con expresión diferencial y (5) análisis de enriquecimiento funcional e interpretación biológica. A continuación, se describirán brevemente, las etapas realizadas en este trabajo.

1. Control de calidad. El primer paso es comprobar la calidad de las lecturas obtenidas. Herramientas computacionales como el programa *FASTQC* [300], permiten visualizar, mediante valores estadísticos, la probabilidad de que la base asignada sea la correcta. A su vez, se deben eliminar las secuencias adaptadoras, así como también las

secuencias de mala calidad, siendo el programa *Trimmomatic* [301] uno de los más utilizados que permiten identificar, procesar y eliminar aquellas lecturas no aptas para poder continuar con el procesamiento de los datos.

2. Alineamiento de las lecturas contra un genoma de referencia. Una de las dificultades que presenta es la de lidiar con lecturas de zonas repetitivas, o regiones que no existen en la referencia, pero sí en la muestra. Para ello, uno de los programas utilizados es el *Bowtie2* [302], el cual evalúa la calidad del alineamiento e indica las diferencias encontradas contra el genoma de referencia.

3. Cuantificación de las lecturas obtenidas para cada gen. La cuantificación de las lecturas asignadas a cada gen se realiza con el fin de identificar los genes expresados diferencialmente, siendo el *feature Counts* uno de los programas más utilizados [303].

4. Normalización de los datos e identificación de los genes con expresión diferencial. Una vez que se estiman los niveles de expresión de los genes, con el fin de que las muestras sean comparables, los mismos deben ser normalizados. Esto puede deberse a diferencias en el número de lecturas generadas en cada secuenciación, lo cual produce cambios en el número de lecturas mapeadas en diferentes muestras. Los métodos de normalización incluyen el cálculo de la cantidad de secuencias por kb de transcrito por millón de secuencias mapeadas (RPKM, de sus siglas en inglés) [304] o por el cálculo de la cantidad de fragmentos transcritos por millones de secuencias mapeadas y (FPKM, de sus siglas en inglés) [305], los cuales normalizan el conteo de transcritos teniendo en cuenta la longitud de estos y la abundancia total de lecturas de la muestra. Luego, mediante programas que calculan las diferencias entre las muestras como ser *EdgeR* [306], se evalúa cómo los niveles de expresión cambian en diferentes condiciones. Como resultado se obtienen las lecturas normalizadas para cada gen, con su respectivo valor de tasa de cambio (FC, de sus siglas en inglés *Fold Change*), valores de probabilidad (valor de p y valor de q) y una tasa de asignaciones incorrectas (FDR, de sus siglas en inglés), lo cual permite ver cómo ha variado la expresión de los genes entre diferentes condiciones y el soporte estadístico que apoya estas diferencias.

5. Análisis de enriquecimiento funcional. Este análisis permite identificar, mediante un test-estadístico, las vías metabólicas afectadas, basándose en los datos de FC y los valores de probabilidad (valores de p) de todos los genes. El test de *Fisher* es uno de los más utilizados, en donde mediante una distribución binomial o hipergeométrica, se determina que vías metabólicas están enriquecidas significativamente. En este contexto se suelen utilizar anotaciones de diferentes bases de datos, tales como GO (*Gene Ontology*) [307], KEGG (*Kyoto Encyclopedia of Genes and Genome*) [308] y COG (*Cluster of Orthologous Group*) [309].

1.3 Transcritómica aplicada al estudio de la interacción rizobio-leguminosa

La disponibilidad de secuenciación y de genomas completos de rizobios y leguminosas, ha permitido desarrollar y aplicar enfoques de genómica funcional para comprender de manera integral cómo los rizobios se reorganizan y adaptan a nuevos entornos. En las últimas décadas, se han llevado a cabo varios estudios de genómica funcional sobre la interacción rizobio-leguminosa, utilizando enfoques de transcritómica [273], [310]. Estos estudios se han basado en analizar el perfil transcritómico de los rizobios en respuesta a la percepción de flavonoides y su relación con la especificidad de su hospedero [311]. En un estudio utilizando RNA-seq, se expuso a la cepa Ch24-10 de *R. phaseoli* a exudados radiculares de maíz (*Zea mays*) y frijol (*Phaseolus vulgaris*) durante 2 h. Se observó una sobre-expresión de los genes *nod* y de genes relacionados con el metabolismo de azúcares, aminoácidos, fitoalexinas y genes que codifican para un SSTIII. Estos resultados sugieren la capacidad adaptativa de los rizobios para colonizar tanto leguminosas como cereales [312].

Por otro lado, se han llevado a cabo pocos estudios de genómica funcional de la interacción entre los beta-rizobios y las leguminosas. Klonowska y col. en el 2018 realizaron un análisis comparativo del perfil transcritómico de las cepas *P. phymatum* STM815, *C. taiwanensis* LMG19424, y el alfa rizobio *R. mesoamericanum* STM3625, en presencia de exudados radiculares de *M. pudica*. En cepas de estas tres especies, se observó la sobre-expresión de los genes *nod* y de una hidroxilasa de ácidos grasos

recientemente identificada, lo cual sugiere que desempeñaría un papel durante la infección de la planta. Sin embargo, los resultados obtenidos confirmaron parcialmente las diferencias previamente observadas en cuanto a la capacidad competitiva de las tres cepas para infectar a *M. pudica* [273].

Al comparar los perfiles, se observó que en *P. phymatum*, en comparación con las otras dos cepas, se sobre-expresa *acdS* (que codifica la ACC desaminasa), un gen involucrado en la biosíntesis del ácido indol-3-acético (AIA), y otros genes relacionados con los SSTIV y SSTVI, lo que sugiere que *P. phymatum* podría ser la cepa más competente en esta interacción.

2 Materiales y métodos

2.1 Cepas bacterianas, medios y condiciones de cultivo

Las cepas de *E. coli* utilizadas fueron cultivadas en medio LB [313] con los requerimientos detallados en la tabla 15, e incubadas a 37 °C. Las cepas de *Cupriavidus*, LMG19424 y UYPR2.512, fueron cultivadas en medio LB (Anexos) y en medio mínimo M9 (Anexos) incubadas a 30 °C. En el caso de la cepa UYPR2.512 de *C. necator*, se agregó citrato 14 mM, como fuente de carbono [314].

Tabla 15. Cepas y plásmidos utilizados.

Cepas/plásmidos	Característica relevante	Referencia
<i>Escherichia coli</i> DH5α	Cloning host; F ⁻ λ ⁻ endA1 glnX44(AS) thiE1 recA1 relA1 spoT1 gyrA96(Nal ^R) rfbC1 deoR nupG Φ80(lacZΔM15) Δ(argF-lac)U169 hsdR17(rκ ⁻ mκ ⁺)	[48]
<i>Cupriavidus tawanensis</i> LMG19424	Cepa de referencia del género <i>Cupriavidus</i>	[315]
<i>Cupriavidus necator</i> UYPR2.512	Cepa utilizada en este trabajo	[42]
pRK600	Plásmido ayudante utilizado para conjugación, <i>oriV</i> (ColE1), RK2(<i>mob</i> + <i>tra</i> ⁺); Cm ^R	[316]
pCBM01	Contiene una región de 401 bp del promotor de <i>nodB</i> , Gm ^R , Tet ^R	[317], [318]
pCZ388	Plásmido control derivado del pLAFR6, que contiene el gen <i>lacZ</i> sin promotor, Gm ^R , Tet ^R	[317], [318]

2.2 Obtención de las cepas de UYPR2.512 de *C. necator* portando pCBM01 o pCZ388

Los plásmidos fueron introducidos en UYPR2.512 mediante conjugación triparental utilizando *E. coli* DH5 α (pCBM01) o *E. coli* DH5 α (pCZ388) como cepas donadoras y *E. coli* DH5 α (pRK600) como cepa auxiliar (Tabla 1). Para ello, las cepas bacterianas se cultivaron a 30 °C en un agitador rotatorio a 200 rpm en medio LB suplementado con el antibiótico adecuado. Los cultivos en fase exponencial se diluyeron 1/100 (v/v) en medio LB sin antibióticos y se incubaron hasta alcanzar una $DO_{600nm} = 0,2 - 0,3$. Para la conjugación, se mezclaron 200 μ L de cada cultivo, las células se centrifugaron durante 5 min a 2.300 g y se resuspendieron en 20 μ L de medio LB. Se colocaron alícuotas de 10 μ L de cada mezcla en medio sólido LB. Se incluyeron alícuotas que contenían cepas individuales como controles. Las placas se incubaron durante la noche a 30 °C. Las células se recogieron con un aza estéril, se resuspendieron en 1 mL de medio LB y 100 μ L de cada suspensión se sembraron en LB que contenía Ampicilina 100 μ g/mL (Ap), Tetraciclina 10 μ g/mL (Tc) y Nitrofurantoína 50 μ g/mL (Nf). La adición de Tc permitió la selección de cepas con el plásmido pCZ388 y pCBM01, mientras que la adición de Ap y Nf, la selección de la cepa UYPR2.512, debido a la inhibición del crecimiento de las cepas de *E. coli*. Los transconjugantes Nf^R, Ap^R, Tc^R se seleccionaron para determinar la presencia de plásmidos, utilizando los cebadores TetAF y TetAR [319] dirigidos al gen de resistencia a la tetraciclina. Se utilizaron cepas resistentes a antibióticos y que hayan amplificado mediante PCR los plásmidos de interés, para los ensayos de inducción de *pnodB19424*.

2.3 Evaluación de la expresión del *pnodB19424-lacZ* en UYPR2.512

Para evaluar la expresión del promotor del gen *nodB*, se utilizó el plásmido pCBM01 (Tc^R) que contiene la fusión transcripcional *pnodB19424-lacZ* [317], [318] o el plásmido pCZ388 (plásmido control que contiene el gen *lacZ* sin promotor). Cultivos de UYPR2.512 portando pCBM01 o pCZ388, se cultivaron ON a 30 °C utilizando un agitador rotatorio a 200 rpm, en 5 mL de medio mínimo M9, conteniendo citrato de sodio 14 mM como fuente de carbono. Finalizado este tiempo, se realizó una dilución 1/100 (v/v) en medio líquido

con luteolina 0,1, 0,5, 3,0 o 15 μM ($\geq 98\%$ (TLC), L9283, Sigma, Aldrich) y los cultivos se incubaron durante 18 h a 30 °C. Se incluyeron cultivos de control no inducidos para cada ensayo. Se realizaron triplicados biológicos independientes. La actividad β -galactosidasa se calculó según Miller [320], empleando la siguiente fórmula:

$$\text{Actividad relativa (Unidades Miller)} = \text{FD} (\text{Abs}_{420\text{nm}} / \text{DO}_{600\text{nm}} * \text{vol} * t)$$

Donde FD es factor de dilución de la muestra utilizada, $\text{DO}_{600\text{nm}}$: la Densidad óptica a 600nm, $\text{Abs}_{420\text{nm}}$, Absorbancia a 420nm, vol: volumen del extracto en mL, t: tiempo de incubación de la reacción en min.

Además, se realizaron experimentos durante 2 h de inducción. Los cultivos bacterianos crecidos ON se diluyeron 1/100 (v/v) en un medio líquido fresco y se cultivaron hasta que los cultivos alcanzaron una $\text{DO}_{600\text{nm}} = 0,5 - 0,6$. Seguidamente, se adicionó luteolina a los cultivos con una concentración final de 15 μM y se incubaron durante 2 h en las mismas condiciones.

2.4 Extracción y purificación de ARN total (ARNtot) de UYPR2.512

La cepa UYPR2.512 se cultivó en 5 mL de medio mínimo M9 [314] con citrato de sodio 14 mM como fuente de carbono [314], ON a 30 °C, con agitador rotatorio a 200 rpm. Transcurrido este tiempo, se realizó una dilución 1/100 (v/v) en medio líquido M9 y el cultivo se dejó crecer hasta que alcanzó la fase exponencial temprana ($\text{DO}_{600\text{nm}} = 0,5 - 0,6$). Posteriormente, se adicionaron luteolina 15 μM a los cultivos y se incubaron durante 2 h a 30 °C hasta alcanzar la fase exponencial tardía ($\text{DO}_{600\text{nm}} = 0,7 - 0,8$), se incluyeron controles no inducidos. El ARNtot se aisló mediante la utilización del kit comercial *PureLink RNA Mini* (Thermo Scientific), según las instrucciones del fabricante. La pureza y calidad del ARNtot extraído, se verificó cualitativamente en un gel de agarosa desnaturalizante 1% en buffer TAE (121 g/L Tris base, 28,55 mL/L ácido acético glacial 50 mL EDTA 0,5 M pH 8). Además, la verificación de la cantidad, integridad y calidad del ARN, se realizó utilizando un espectrofotómetro *Nanodrop 1000* (Thermo Scientific) y un bioanalizador *Agilent Technologies 2100* (*Agilent Technologies*). Los experimentos se

realizaron en tres réplicas biológicas independientes, tanto para el ARN_{tot} de la cepa de UYPR2.512 en presencia e luteolina como en ausencia de este.

2.5 Análisis transcriptómico a la presencia luteolina mediante RNA-Seq

Con el fin de conocer la respuesta transcripcional a la presencia del inductor luteolina de UYPR2.512, se realizaron análisis de RNA-Seq utilizando como primera aproximación la plataforma del *Ion Torrent*[™] (*Ion PGM*[™] *Sequencing Kit de Life Technologies*[™]) en colaboración con el Departamento de Genómica del IIBCE, y como segunda aproximación la plataforma de Illumina en Macrogen (Seúl, Korea). Los análisis de los transcriptomas obtenidos fueron realizados en colaboración con el Dr. Guillermo Eastman, del Departamento de Genómica del IIBCE.

2.5.1 Secuenciación mediante Ion Torrent

2.5.1.1 Enriquecimiento de ARN mensajero (ARN_m) o depleción de ARN ribosómico (ARN_r)

El ARN_m para secuenciar se obtuvo mediante la utilización de un kit comercial *MICROBExpress* (AMBION, Thermo Scientific), siguiendo las instrucciones del fabricante. Los experimentos se realizaron en tres réplicas biológicas independientes. Se utilizó un procedimiento para el enriquecimiento de ARN_m, en el cual se esperaba depletar aproximadamente un 80% del ARN_r según como lo sugiere el fabricante del kit. El procedimiento se basa en la captura y remoción de las subunidades 16S y 23S del ARN_r presentes en el ARN mediante el uso de oligonucleótidos de captura. Dichos oligos bipartitos se unirán por complementariedad del extremo 5' a una región del ARN_r y por su extremo 3' a un oligonucleótido que se encuentra unido covalentemente a esferas magnéticas. Debido a que el porcentaje de obtención de ARN_m enriquecido a partir de 10 µg de ARN es aproximadamente de entre 1-2,5 µg se realizó un primer paso de precipitación y de concentración. La verificación de la depleción del ARN_r fue realizada mediante la utilización el kit comercial *Agilent RNA 6000 Nano Kit Quick*, el cual tiene un rango cualitativo entre 25-6.000 nt y un rango cuantitativo entre 25-500 ng/µL. El ARN

obtenido, mayormente ARNm fue destinado para la construcción de librerías de ADN copia (ADNc).

2.5.1.2 Construcción de librerías de ADN copia (ADNc) a partir de los ARNm obtenidos

Las librerías de RNA-Seq fueron construidas mediante la utilización del kit comercial *Ion Total RNA-Seq Kit v2* (Thermo Scientific), siguiendo las instrucciones del fabricante. El protocolo consta de dos bloques que incluyen: un primer paso de fragmentación de los ARN (depletados en ARNr, con un rango esperado de fragmentación entre 100-200 nt) y purificación de los fragmentos, y un segundo bloque donde los fragmentos son ligados a adaptadores, retrotranscritos a ADNc, purificados, y se les adicionan códigos de barra o *barcodes*, para ser finalmente re-purificados y secuenciados.

2.5.1.3 Fragmentación del ARNm

Las muestras normalizadas (50ng/μL de cada ARNm depletado) fueron incubadas con la enzima de digestión RNAsa III (1U/μL) de AMBION (Thermo Scientific) durante 3 min a 65 °C. La correcta obtención de una población de transcritos fragmentados con el tamaño esperado fueron verificados mediante la utilización del kit comercial *Agilent RNA 6.000 Nano Kit Quick*, el cual tiene un rango cualitativo entre 25-6000 nt y un rango cuantitativo entre 25-500 ng/μL.

2.5.1.4 Amplificación de ADNc

Los ARNm fragmentados fueron sometidos a una retrotranscripción, una primera purificación, adición de oligos para indexar las secuencias o *barcodes* (lo que permite poder secuenciar al mismo tiempo múltiples muestras en una única corrida), amplificación del ADNc y purificación final, mediante la utilización del kit comercial *Ion Total RNA-Seq Kit v2* (Thermo Scientific). Cabe destacar que la realización de las purificaciones tiene el fin de seleccionar por tamaño los productos de los ADNc deseados. La obtención de una librería de ADNc de tamaño deseado, en un rango entre 100-230 nt fue evaluada mediante la utilización del *High Sensitivity DNA Kit*, con un rango cuantitativo de entre 5-500 pg/μL.

2.5.1.5 Secuenciación de las librerías

Las librerías fueron secuenciadas en un chip 318, en equipo *Ion Torrent*TM, de la Plataforma de Secuenciación del IIBCE.

2.5.2 Secuenciación mediante Illumina

En el caso de la secuenciación de las librerías de RNA-Seq mediante Illumina en Macrogen (Macrogen Inc. Korea), se utilizó un protocolo de preparación de bibliotecas de depleción de ARN ribosómico para generar librerías de transcriptomas completos siguiendo las recomendaciones del fabricante (Kit de eliminación de *ARNr Ribo-Zero* para bacterias y Kit de preparación de muestras de ARNtot *TruSeq*), y se utilizó la tecnología *Illumina NovaSeq (100 bp paired-end)*.

2.5.3 Análisis de las secuencias obtenidas (Ion Torrent / Illumina)

En ambas secuenciaciones la calidad de las lecturas de RNA-Seq crudas fueron analizadas usando *FastQC* [300].

En el caso de la secuenciación mediante Illumina, además de utilizar el *FastQC* se evaluó la calidad de las lecturas mediante la herramienta *phred score*, el cual brinda un valor de probabilidad, que indica si la base asignada a una posición dada en una lectura sea incorrecta. Se seleccionaron aquellas lecturas con un puntaje de calidad *phred score* > 30 (1/1000 la probabilidad de error en la identificación de la base), permitiendo filtrar las bases de baja calidad y mejorar la precisión del ensamblaje de la secuencia.

Con respecto al mapeo, en el caso de la secuenciación de las librerías de RNA-Seq mediante el Ion Torrent, las secuencias obtenidas fueron mapeadas contra la versión 1 del genoma de UYPR2.512 [254], utilizando el mediante el programa *CLC 3.7* [302]. En el caso de las librerías de RNA-Seq secuenciadas por Illumina fueron mapeadas contra la versión 2 del genoma de UYPR2.512, mediante la utilización del programa *Bowtie2* [302].

Con el fin de evaluar cómo se agrupan las muestras y su variación, en ambos casos se realizó un análisis de componentes principales (PCA, de sus siglas en inglés *Principal Component Analysis*) en R.

2.5.3.1 Análisis de la expresión diferencial de genes

Las secuencias obtenidas por Ion Torrent fueron analizadas mediante el programa CLC 3.7 [321].

En el caso de las secuencias obtenidas por Illumina, se utilizaron un conjunto de programas que corren en línea de comando, los cuales permiten un mejor control de la totalidad del proceso operativo. Luego de realizarse el mapeo contra la versión 2 , se realizó el conteo de las lecturas obtenidas sobre las secuencias codificantes (CDS), utilizando el programa *FeatureCounts* [303].

El análisis de expresión génica diferencial, se realizó utilizando el programa *edgeR* de R [306], el cual calcula el FC, valores de p, definiendo un valor de p ajustado o FDR [322]. Los genes diferencialmente expresados (DEG) fueron seleccionados según el criterio de $|FC| > 2$ y FDR o valor de p ajustado $< 0,05$. El valor de p ajustado menor a 0.05 indica que la probabilidad de que el resultado observado se deba al azar es inferior al 5%, o sea que el 5% de los genes diferencialmente expresados fueran falsos positivos

2.5.3.2 Análisis de categorías funcionales

La anotación funcional se realizó mediante la herramienta web *eggNOG-mapper* [323]. Con el objetivo de anotar las categorías funcionales, se utilizaron las bases de datos KEGG (*Kyoto Encyclopedia of Genes and Genome*) y COG (*Cluster of Orthologous Group*). El análisis de enriquecimiento funcional para los genes diferencialmente expresados (*DEG*) se realizó mediante la prueba exacta de *Fisher*, considerando un enriquecimiento significativo, si el valor de p $< 0,05$ en el programa *Python* [324].

2.5.4 Validación de *RNA-Seq* mediante q-PCR

2.5.4.1 Extracción de ARN total y síntesis de ADN copia

Para realizar los análisis de PCR cuantitativos, se repitió el experimento de inducción en presencia de luteolina y en ausencia de esta explicado previamente, utilizando esta vez cinco réplicas biológicas independientes. El ARN_{tot} se aisló de cultivos de la cepa UYPR2.512 inducidos y no inducidos, siguiendo el procedimiento previamente descrito. Posteriormente, se partió de 1ug de ARN el cual fue convertido a ADN copia, utilizando el kit comercial *High-Capacity cDNA Reverse Transcription* (Applied Biosystem) según las instrucciones del fabricante. Se incluyeron controles sin retrotranscribir, para cada réplica biológica.

Los cebadores para utilizar se diseñaron mediante el programa *Primer3* [325] y fueron sintetizados por Macrogen Inc. (Corea). La especificidad de los cebadores y la formación de dímeros de cebadores, se comprobaron mediante curvas de disociación [326]. La amplificación mediante PCR de un único amplicón para cada par de cebador sintetizados, se verificó mediante la realización de una electroforesis en un gel de agarosa 1%.

Las reacciones de q-PCR se realizaron en un termociclador táctil C1000 (*Bio-Rad*) utilizando *iQTM SYBR Green Supermix* (Bio-Rad). Las condiciones de las q-PCR fueron las siguientes: 5 min a 95 °C, 40 ciclos de 15 s a 95 °C, 30 s a 60 °C y 30 s a 72 °C.

Los análisis estadísticos se realizaron en el programa estadístico *InfoStat* [327]. La distribución normal de los datos se confirmó mediante la prueba de *Shapiro-Wilk* y se utilizó ANOVA unidireccional para las comparaciones de muestras. Las diferencias se consideraron estadísticamente significativas si el valor de $p < 0,05$.

2.5.4.2 Especificidad de los cebadores para la q-PCR

La eficiencia de amplificación de un par de cebadores se refiere a la tasa de duplicación del material genético en cada ciclo de amplificación. Una eficiencia en el rango entre 95-100% indica una duplicación completa en cada ciclo, lo cual aseguraría la unión más eficiente a las secuencias del ADN y la amplificación del fragmento de interés.

La eficiencia de los pares de cebadores se evaluó mediante la realización de reacciones de q-PCR independientes. Como molde se utilizaron diluciones seriadas 1/100 v/v de ADN_g de UYPR2.512 de concentración conocida (a partir de 50 ng/uL), y cebadores de concentración final de 500nM. A partir de estos datos se obtuvo una gráfica para cada par de cebadores, cuya pendiente correspondió a la eficiencia de la amplificación de la reacción. Se seleccionaron aquellos juegos de cebadores con un % de eficiencia mayor a 95%. La eficiencia se calcula mediante la siguiente fórmula:

$$\text{Eficiencia (\%)} = (10^{-1 / \text{pendiente}} - 1) * 100$$

Donde la pendiente se refiere a la pendiente negativa de la línea de regresión obtenida a partir de los valores de ciclo umbral (Ct) en función del logaritmo de la concentración de ADN inicial.

Durante la amplificación de la q-PCR, la fluorescencia se detecta después de cada ciclo de amplificación. A medida que se amplifica el fragmento de ADN de interés, la cantidad de producto acumulado aumenta y, por lo tanto, también aumenta la intensidad de fluorescencia registrada en cada ciclo. Los valores de RFU (eje de las y) representan esta intensidad de fluorescencia relativa en cada ciclo de amplificación (eje de las x).

La expresión relativa entre los genes normalizadores y los genes diana, se calcularon mediante la fórmula $\Delta\Delta\text{Ct}$:

$$\Delta\Delta\text{Ct} = 2^{-\Delta\text{Ct}}$$

Donde ΔCt es la resta del valor de Ct del gen de referencia al valor de Ct del gen de interés en cada muestra ($\Delta\text{Ct} = \text{Ct gen de interés} - \text{Ct gen de referencia}$).

En el caso de observarse disparidad en las eficiencias de los cebadores, los resultados se interpretaron mediante el método descrito por *Pfaff* [328].

Mediante este método se determina la expresión relativa de un gen de interés, en base a la eficiencia de la amplificación y la diferencia del Ct de la muestra de interés frente a una muestra control, en relación a la eficiencia y la diferencia de Ct de un gen de referencia.

La ecuación utilizada en el método de *Pfaff* es la siguiente:

$$\text{relación} = E_{\text{int}}^{\Delta\text{Ct control-muestra}} / E_{\text{nor}}^{\Delta\text{Ct control-muestra}}$$

Donde E_{int} es la eficiencia del gen de interés, E_{nor} es la eficiencia del gen normalizador, ΔCt (control-muestra) es la diferencia de Ct entre la muestra control y de la muestra problema para el gen de interés o el de referencia.

2.5.4.3 Selección de los genes a evaluar la expresión diferencial

Se seleccionaron varios genes que presentaron diferencia en la expresión con un alto número de lecturas, tanto en la condición de inducción, como en la condición control. A su vez, se seleccionaron seis genes para ser utilizados como genes normalizadores (*housekeeping*). La función de estos es normalizar los niveles de expresión génica de los genes en ambas condiciones, debido a que no presentan cambios en la expresión y a su vez corregir los sesgos técnicos y biológicos en los datos de RNA-Seq (Tabla 16).

Tabla 16. Secuencias nucleotídicas de cebadores de genes normalizadores, sobre-expresados y sub-expresados, evaluados para la validación del RNA-Seq, mediante q-PCR.

Expresión	Nombre del gen (ID)	Secuencia nucleotídica del juego de cebadores
Normalizador	<i>eloG</i>	(F) 5'-CTGTTCTACGTGTCAG-3' (R) 5'-TGATGTTGATGCGGTGTTTC-3'
Normalizador	<i>rpoH</i>	(F) 5'-CCGACCAAACCCTGAATAAC-3' (R) 5'-GCCTGGATATAGCTGTTCGAG-3'
Normalizador	<i>rpoN</i>	(F) 5'-TCAAGAACATCCAGCAAAGG-3' (R) 5'-CGTGTAACCCAGTGTATCG-3'
Normalizador	<i>ntrY</i>	(F) 5'-GCACACCGAGACTGTAGAATA-3' (R) 5'-CCGTGTTGTCATGATCTTC-3'
Normalizador	<i>acoA</i>	(F) 5'-GATGTGTTCAAGTCCTCGG-3' (R) 5'-GCAACTCTGTGATCAACGC-3'
Normalizador	<i>hisB</i>	(F) 5'-GCGACAAGAAAGGCATCAC-3' (F) 5'-CCGCATGGTTGACGAAGC-3'
Sobre-expresados	<i>nodU</i>	(F) 5'-GCTACCGAGGGAATATTAGGG-3' (R) 5'-GGTACTTGGATTGAATGGCG-3'
Sobre-expresados	<i>nodS</i>	(F) 5'-ACTGGTATTCGGATCGTCTC-3' (R) 5'-ACTTGTGTTGATGTGAACCGG-3'

Expresión	Nombre del gen (ID)	Secuencia nucleotídica del juego de cebadores
Sobre-expresados	<i>nodC</i>	(F) 5'-GACGATTCTCATGCTGACTGC-3' (R) 5'-GGTGTCCAGCGTAAGAAAGC-3'
Sobre-expresados	<i>nodJ</i>	(F) 5'-CTCGGTGAAATGACGTGGGC-3' (F) 5'-GTA ACTCGGCGCAATCGCG-3'
Sobre-expresados	<i>hcdP</i>	(F) 5'-CGGAATTGGATGTTATAACGCC-3' (R) 5'-GCTCCTGAAGACTTATCCAC-3'
Sub-expesados	<i>pvsA</i>	(F) 5'-GGTCTACTCCAACAGTGATC-3' (R) 5'-CATTTCCGCCTTGTTCTTG-3'
Sub-expesados	<i>exbD</i>	(F) 5'-TCCCCGTGATCAACCAC-3' (R) 5'-GCGCGATATTGCTTTCCAG-3'
Sub-expesados	<i>hemP</i>	(F) 5'-CATCGCATCCGTGAAGTC-3' (R) 5'-CGCAACGTGTAGACATCG-3'

3 Resultados

3.1 Evaluación de la expresión del gen *pnodB19424-lacZ* en UYPR2.512

La respuesta de los genes *nod* bacterianos a la presencia de luteolina se evaluó en cultivos de la cepa UYPR2.512 portando pCBM01 o pCZ388, mediante la determinación de la actividad de la β -galactosidasa. Como se muestra en la figura 27, luego de 18 h de incubación en presencia de luteolina, se observó un aumento significativo en la actividad de la β -galactosidasa en UYPR2.512 portando pCBM01, en comparación con UYPR2.512 portando pCZ388, en todas las concentraciones de luteolina ensayadas, con respecto a cultivos crecidos sin presencia de luteolina. A su vez se observó una correlación positiva entre la concentración de luteolina y mayor expresión.

Esto indica que la luteolina es un potente inductor de la expresión génica de los genes *nod* en UYPR2.512.

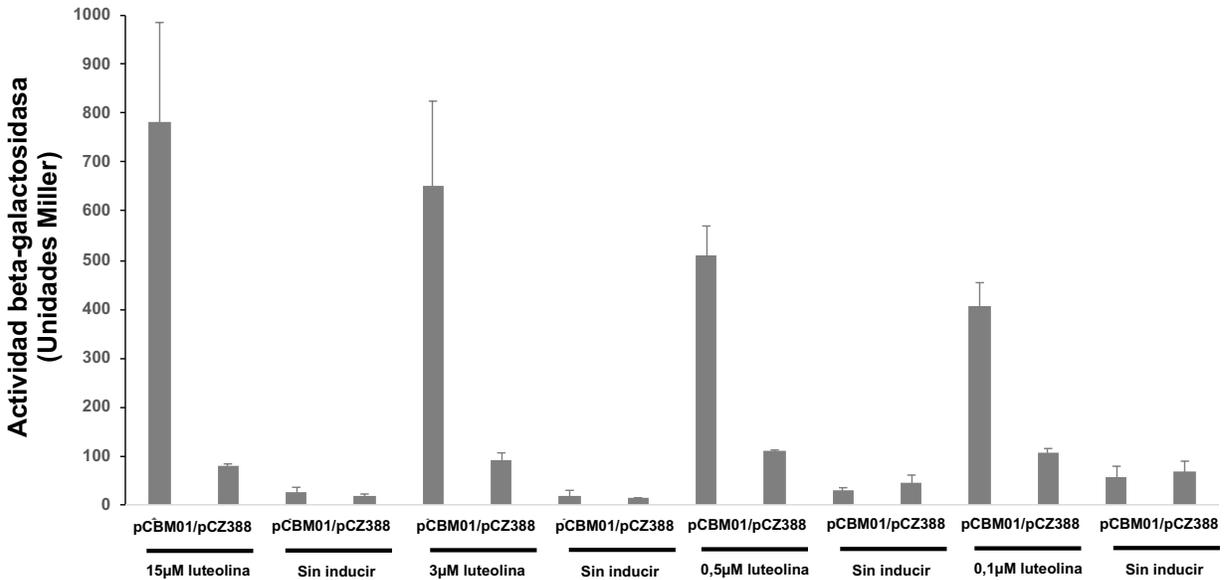


Figura 27. Respuesta de la región promotora del gen *nodB* a concentraciones decrecientes de luteolina en cultivos de UYPR2.512.

La determinación de la actividad β -galactosidasa se realizó en cultivos de la cepa UYPR2.512 portando el plásmido pCBM01, que contiene la fusión transcripcional *pnodB*_{19424-lacZ} o el plásmido pCZ388 (plásmido control que contiene un promotor *lacZ* sin promotor) [317]. En todos los casos, la expresión de *lacZ* se evaluó como actividad β -galactosidasa, registrada en unidades Miller, después de 18 h de incubación a 30 °C, en presencia de cantidades decrecientes de luteolina (15µM, 3µM, 0,5µM y 0,1µM). Los datos se expresan como la media de tres réplicas biológicas con sus errores estándar.

Se hizo un ensayo similar, con un menor tiempo de exposición (2 h) a luteolina (Fig. 28). Bajo estas condiciones, se observó la inducción en la expresión de la fusión *pnodB-lacZ* en los cultivos luego de 2 h de inducción, utilizando una concentración de luteolina de 15 µM. Estos resultados evidencian que la luteolina es un rápido inductor de la expresión génica en UYPR2.512.

Por lo tanto, se decidió utilizar una concentración de luteolina de 15 µM y un tiempo de inducción de 2 h, para estudiar los cambios globales a nivel transcripcional en UYPR2.512.

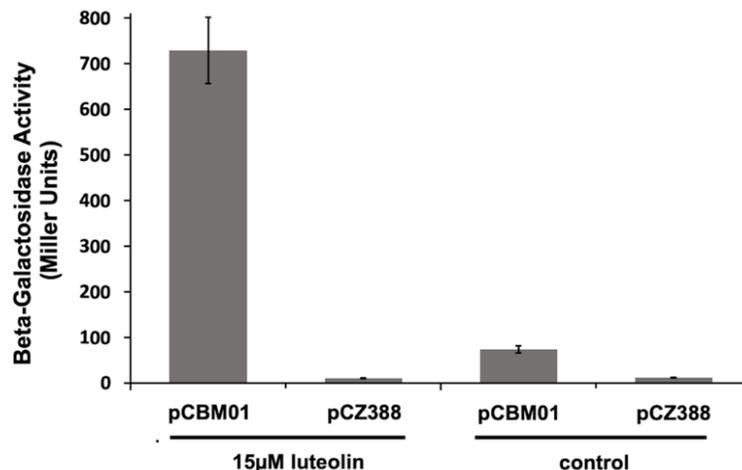


Figura 28. Respuesta de la región promotora del gen *nodB* en presencia de luteolina en cultivos de UYPR2.512.

La determinación de la actividad β -galactosidasa se realizó en cultivos de UYPR2.512 portando pCBM01 o pCZ388. La expresión de *lacZ* se evaluó como actividad β -galactosidasa registrada como unidades Miller después de incubación durante 2 h, a 30 °C en presencia de 15 μ M de luteolina, con respecto a un control sin el agregado de luteolina. Los datos se expresan como la media de tres réplicas biológicas con sus errores estándar.

3.2 Obtención del ARNtot bacteriano

Se lograron obtener entre 15 μ g y 20 μ g de ARNtot con valores de integridad (RIN) superiores a 7, lo que indica una alta calidad del ARNt obtenido. Además, el análisis del ARNtot mediante electroforesis en gel de agarosa demostró la ausencia de contaminación de ADNg en todas las muestras (Fig. 29). Por lo tanto, el ARNtot obtenido cumplió con los requisitos necesarios para proceder con la construcción de las librerías de RNA-Seq.

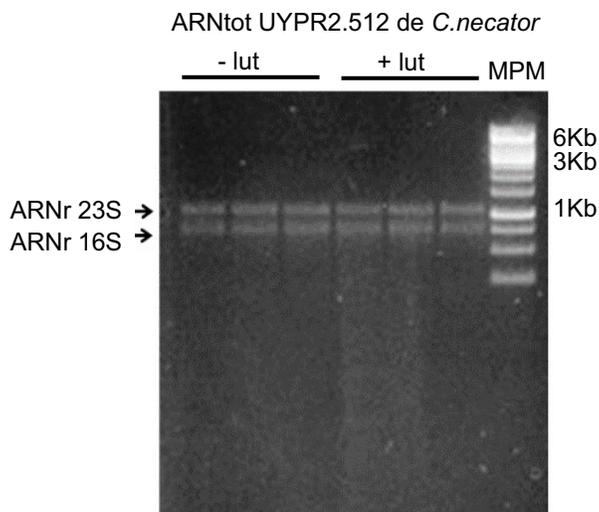


Figura 29. Extracción y purificación del ARNtot de UYPR2.512.

Gel agarosa 1%. ARNtot de UYPR2.512 suplementado con $15\mu\text{M}$ de luteolina. Las flechas indican las bandas correspondientes a las subunidades 16S y 23S del ARNr.

3.3 Construcción de librerías de *RNA-Seq* para *Ion Torrent*

3.3.1 Enriquecimiento del ARNm o depleción de los ARNr

Una vez obtenido el ARNtot, el siguiente paso en la construcción de las librerías consistió en enriquecer las muestras con ARNm, asegurando que las lecturas generadas posteriormente se correspondieran con ARN codificante en lugar del ARNr, que es el más abundante. Para ello, se aplicó un método de purificación basado en la eliminación de las subunidades 16S y 23S del ARNr, lo que resultó en una disminución en la cantidad de ARNr, como se aprecia en los cromatogramas mostrados en la figura 30. En dicha figura se presentan los perfiles de los fragmentos de ARN presentes en cada muestra luego de la depleción (muestras de 1 a 6), así como antes de la depleción (muestras 1R a 6R). La ausencia de los picos correspondientes a las subunidades del 16S y 23S del ARNr en las muestras tratadas indica que el proceso de depleción resultó exitoso. A continuación, se llevó a cabo la fragmentación de los ARNm obtenidos.

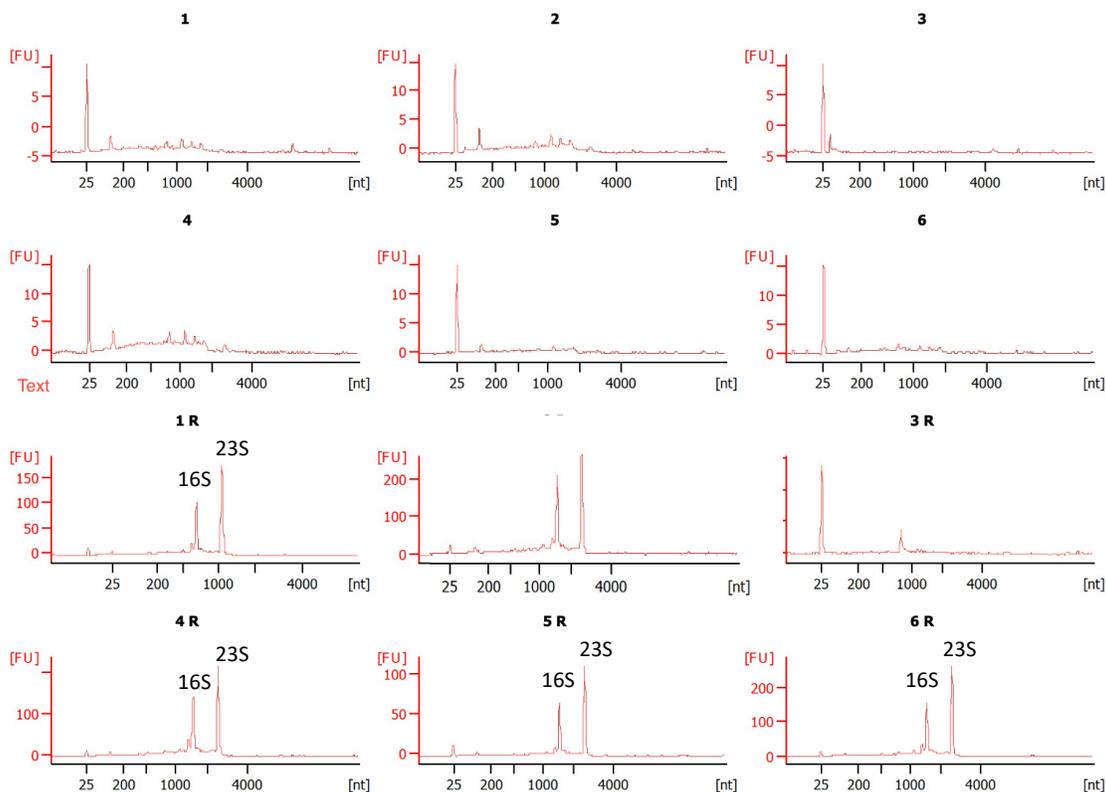


Figura 30. Análisis cualitativo mediante Bioanizador del ARNtot obtenido a partir del cultivo de UYPR2.512 en presencia y en ausencia de luteolina.

La numeración de 1 a 6, indican las muestras de ARNtot, luego de la depleción de ARNr, y la numeración de 1R hasta 6R, indican las muestras de ARNtot previo a la depleción de ARNr. En las muestras 1R hasta 6R, se indican los picos correspondientes a las subunidades del 16S y 23S del ARNr.

3.3.2 Fragmentación del ARNm

Luego del tratamiento con RNAsa, se observó un incremento en la población de fragmentos de ARN en el rango de 150 a 200 nt, lo cual indica la efectividad del tratamiento. Este efecto se puede apreciar en el desplazamiento de la curva azul (ARNm depletado y fragmentado) en comparación con la curva roja (ARNm depletado pero no fragmentado) en la figura 31. Es importante destacar que la obtención de fragmentos de ARNm en el rango de 200 nt es crucial para lograr una correcta ligación con los adaptadores utilizados durante el proceso de secuenciación.

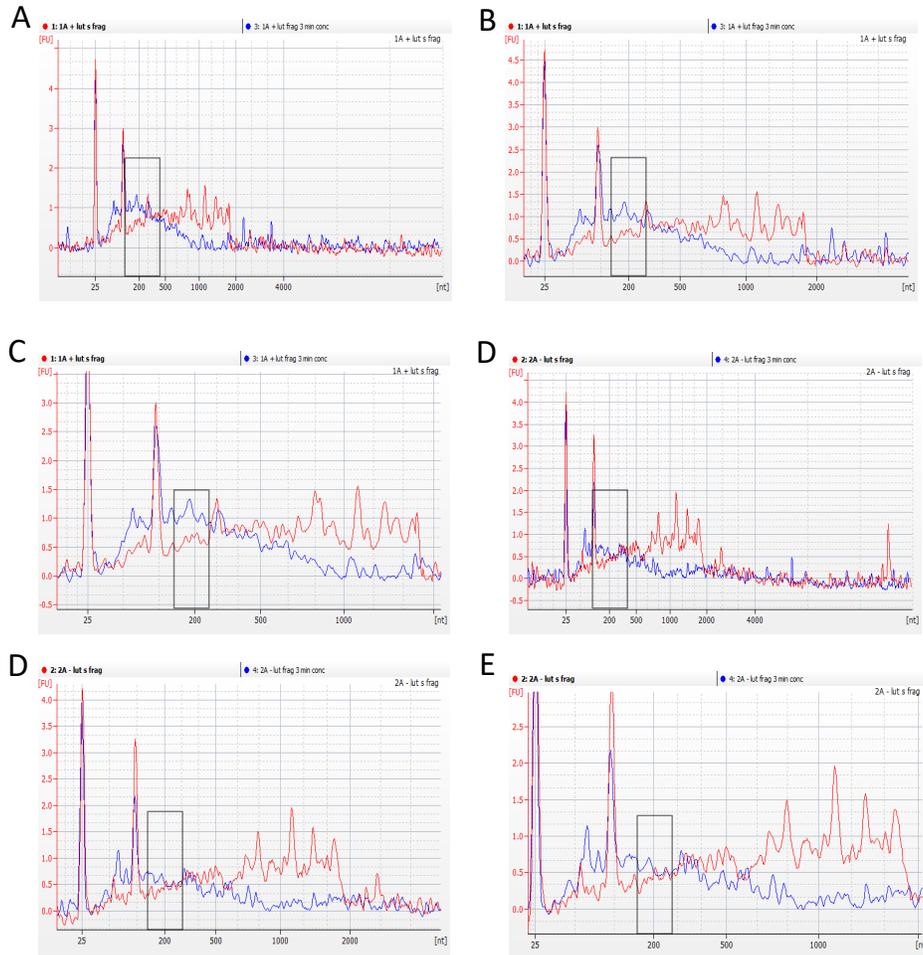


Figura 31. Cromatogramas de ARNm depletado y purificado de UYPR2.512.

Paneles A, B y C en ausencia de luteolina y paneles D, E y F, en presencia de luteolina. Sin fragmentar (curva roja) y luego de ser fragmentado (curva azul). El corrimiento en la curva azul hacia la derecha demuestra una correcta fragmentación del ARNm depletado del tamaño deseado. Los rectángulos indican los fragmentos de 200nt esperados, para cada una de las muestras, luego de la fragmentación enzimática.

3.3.3 Obtención del ADNc

Una vez obtenido el ARN enriquecido en ARNm y fragmentos del tamaño adecuado, el siguiente paso consistió en la obtención del ADNc. Para ello, se realizó una optimización de la cantidad de ARN requerida para obtener una alta proporción de fragmentos de ADNc con el tamaño deseado (Fig. 32). Se obtuvo un mayor porcentaje de fragmentos del tamaño esperado al utilizar una dilución 1:10 del ARN (Fig. 32-C), por lo tanto, se decidió continuar trabajando con esta dilución. Al aplicar estos parámetros, se logró

obtener un alto porcentaje de fragmentos de ADNc con el tamaño deseado (relación 60/160 nt igual a 7%).

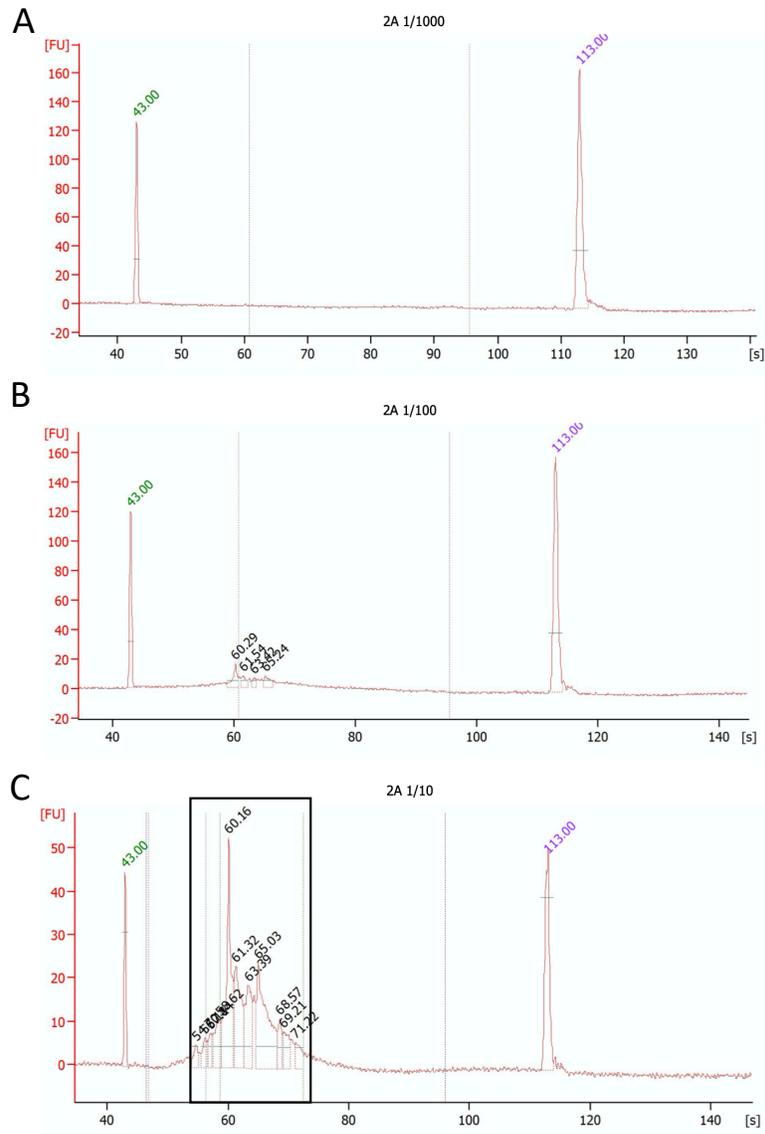


Figura 32. Librería de ADNc amplificada de UYPR2.512.

Se indican en: (A) ADNc dilución 1:1000, (B) dilución 1:100 y (C) dilución 1:10. El recuadro de color negro en C, indica la alta población de fragmentos de ADNc con un tamaño esperado entre 100-150 nt. Se muestra una sola muestra representativa. Se realizaron triplicados biológicos.

3.3.4 Análisis transcriptómico de UYPR2.512 en presencia de luteolina

3.3.4.1 Primera aproximación: Secuenciación en Ion Torrent

Mediante la secuenciación con la plataforma Ion Torrent se obtuvo un total aproximado de 250.000 de lecturas por muestra. Este resultó un número de lecturas muy inferior al esperado ya que se esperaba obtener aproximadamente 5 millones de lecturas/muestra. A su vez, el mapeo de las secuencias obtenidas contra el genoma de referencia indicó que, alrededor del 80-85% de las lecturas, correspondían a ARNr.

Se identificaron un total de 52 genes expresados diferencialmente (DEG), de los cuales 20 fueron sobre-expresados ($\log_{2}FC > 2$ y valor de p ajustado (FDR) $< 0,05$), y 32 sub-expresados ($\log_{2}FC < -2$ y valor de p ajustado $< 0,05$) (Tabla 17 y 18), respectivamente.

Tabla 17. Genes sobre-expresados en presencia de luteolina.

Se indican la descripción del gen, log FC, FC I/C (inducido versus control) y valor de p (FDR).

Descripción del gen	Nombre del gen	logFC	FC I/C	valor de p (FDR)
ABC-type multidrug transport system permease component	<i>lseI_0057</i>	4,54	23,20	0,05
Mobile element protein	NA	4,51	22,77	0,04
Putative protein-S-isoprenylcysteine methyltransferase	NA	3,67	12,71	3,78E-03
UDP-glucose 4-epimerase (EC 5.1.3.2)	NA	3,3	9,84	0,01
Hypothetical protein	NA	2,86	7,27	0,05
Probable transmembrane protein	NA	2,8	6,98	0,3
Regulatory protein LysR	NA	2,8	6,98	0,03
Hypothetical protein	NA	2,41	5,32	8,25E-03
Curli production assembly/transport component CsgG	<i>csqG</i>	2,39	5,25	0,02
Nodulation protein U (EC 2.1.3)	<i>nodU</i>	2,21	4,64	1,39E-05
FIG028593: membrane protein	NA	2,04	4,12	0,02
Nodulation ABC transporter NodI	<i>nodI</i>	1,98	3,96	0,02
2-hydroxychromene-2-carboxylate isomerase	<i>doxJ</i>	1,68	3,20	0,04
Vibrio ferritin amide bond forming protein PvsD/ Siderophore synthetase superfamily	NA	1,54	2,91	1,52E-03
Thioredoxin reductase (EC 1.8.1.9)	<i>trxB</i>	1,5	2,83	7,42E-03
Glucose-6-phosphate isomerase (EC 5.3.1.9)	NA	1,5	2,83	0,03
Vibrio ferritin decarboxylase protein PvsE	<i>pvsE</i>	1,21	2,32	0,04
FIG000325: clustered with transcription termination protein NusA	<i>nusA</i>	1,17	2,25	0,04
Histone-like protein	NA	1,13	2,19	0,03
Predicted ATPase related to phosphate starvation-inducible protein PhoH	<i>phoH</i>	1,04	2,05	0,02

Tabla 18. Genes sub-expresados en presencia de luteolina.

Se indican la descripción del gen, log FC, FC I/C (inducido versus control) y valor de p (FDR).

Descripción del gen	Nombre del gen	logFC	FC I/C	Valor de p (FDR)
Threonyl-tRNA synthetase (EC 6.1.1.3)	NA	-1,06	-2,09	0,03
FIG000859: hypothetical protein YebC	<i>yebC</i>	-1,22	-2,33	0,05
Acyl-CoA dehydrogenase (EC 1.3.8.7)	NA	-1,26	-2,39	0,03
Hypothetical protein	NA	-1,28	-2,43	0,04
Soluble lytic murein transglycosylase precursor (EC 3.2.1)	<i>slt</i>	-1,37	-2,58	0,04
Cytoplasmic axial filament protein CafA and Ribonuclease G (EC 3.1.4)	NA	-1,41	-2,66	0,05
IMP cyclohydrolase (EC 3.5.4.10) /Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3)	<i>purH/NA</i>	-1,47	-2,77	0,01
Glucans biosynthesis protein G precursor	<i>mdoG</i>	-1,48	-2,80	0,04
Anaerobic dehydrogenases typically selenocysteine-containing	<i>cgl0529</i>	-1,48	-2,80	0,04
FIG01285306: hypothetical protein	NA	-1,62	-3,08	8,18E-03
Outer membrane lipoprotein carrier protein LolA	<i>lolA</i>	-1,64	-3,11	0,04
Cell division protein	NA	-1,64	-3,11	0,04
Predicted transcriptional regulator for fatty acid degradation FadP TetR family	<i>fadP</i>	-1,68	-3,20	0,02
Phosphate:acyl-ACP acyltransferase PlsX	<i>plsX</i>	-1,73	-3,31	3,11E-03
Ribosomal large subunit pseudouridine synthase B (EC 4.2.1.70)	<i>rluB</i>	-1,79	-3,47	1,77E-03
Chaperonin 33kDa (Heat shock protein 33) (HSP33)	<i>hsp33</i>	-1,94	-3,83	0,03
Ribonucleotide reductase of class II (coenzyme B12-dependent) (EC 1.17.4.1)	<i>nrdJ</i>	-1,98	-3,95	9,35E-03
Iron-sulfur cluster regulator IscR	<i>iscR</i>	-2,08	-4,23	0,05
3-ketoacyl-CoA thiolase (EC 2.3.1.16)	<i>acaa2</i>	-2,18	-4,54	0,04
Cobalt/zinc/cadmium efflux RND transporter membrane fusion protein CzcB family	<i>czcB</i>	-2,18	-4,55	0,04
NADPH:quinone oxidoreductase	<i>nqr</i>	-2,24	-4,71	9,19E-03
Peptide chain release factor	<i>prfB</i>	-2,24	-4,71	0,01
Dioxygenases related to 2-nitropropane dioxygenase	<i>sy1_21710</i>	-2,28	-4,86	0,03
Nitrate/nitrite transporter	<i>narK</i>	-2,28	-4,86	0,03
Dipeptide-binding ABC transporter periplasmic substrate-binding component (TC 3.A.1.5.2)	<i>hbzs_101840</i>	-2,36	-5,13	0,03

Descripción del gen	Nombre del gen	logFC	FC I/C	Valor de p (FDR)
Putative helicase	NA	-2,37	-5,17	0,03
Hypothetical protein	NA	-2,59	-6,03	2,89E-03
Hypothetical protein	NA	-2,79	-6,91	0,04
Hypothetical protein	NA	-2,79	-6,91	0,04
Respiratory nitrate reductase beta chain (EC 1.7.99.4)	<i>rr42_m1771</i>	-3,25	-9,49	2,23E-04
Mobile element protein	NA	-5,19	-36,43	0,05

Dado el bajo nivel de cobertura de secuencia obtenido utilizando la plataforma Ion Torrent, se tomó la decisión de repetir el experimento utilizando la tecnología de Illumina para la secuenciación. En esta ocasión, los ARN extraídos se enviaron a MacroGen Inc, quienes generaron las librerías y llevaron a cabo la secuenciación de los ADNc. Se obtuvo un promedio de 27 millones de lecturas, cada una con una longitud de 100 pb por muestra. Se logró mapear más del 95% de las lecturas para cada muestra en el genoma de referencia, y más del 85% se alinearon con secuencias codificantes anotadas (Tabla 19).

Tabla 19. Características generales de las secuencias obtenidas.

Se muestra el total de lecturas obtenidas para cada muestra y el número de secuencias mapeadas contra el genoma de referencia en cada caso.

Muestra	Número total de lecturas secuenciadas en ambos extremos	Número de lecturas de ambos extremos mapeadas en el mismo transcrito	Número de lecturas de ambos extremos mapeadas en otro transcrito	Número de lecturas mapeadas en un solo extremo	% mapeo
Control 1	23.402.250	19.797.557	1.913.459	1.343.854	95,28%
Control 2	27.610.809	25.148.200	1.071.910	1.133.015	97,01%
Control 3	26.942.222	23.890.190	2.037.566	1.002.542	98,10%
Luteolina 1	24.894.361	22.507.441	1.396.324	921.207	97,87%
Luteolina 2	30.626.823	26.826.430	2.365.994	1.315.292	97,46%
Luteolina 3	28.420.362	25.549.257	1.656.514	1.002.682	97,49%

El PCA permitió discriminar las muestras según su condición, revelando una clara proximidad entre las muestras tratadas con luteolina (Fig. 33). La primera dimensión del PCA explicó el 88,6% de la variabilidad y separó de manera precisa las muestras de control 1 y 2 de las muestras tratadas con luteolina. La segunda dimensión completó la separación al distinguir las muestras de control 3. Estos resultados indican que existe una mayor variabilidad en la expresión génica entre las muestras de control en comparación con las muestras tratadas con luteolina.

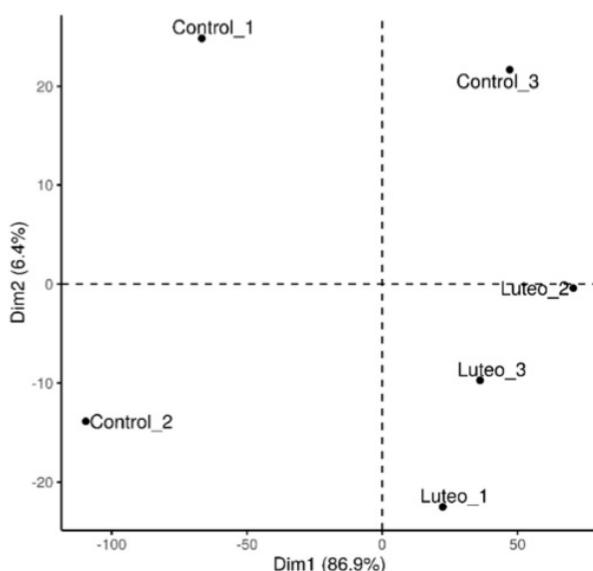


Figura 33. Análisis de componentes principales (PCA).

La dimensión 1 explica el 86.9% de la variabilidad y separa correctamente las muestras de control 1 y 2, de las tratadas con luteolina, mientras que la dimensión 2 separa las muestras de las muestras control.

3.3.5 Análisis de la expresión diferencial de genes

Se observaron un total de 145 genes diferencialmente expresados (DEG), en cultivos de UYPR2.512 en presencia de luteolina, de los cuales 73 genes corresponden a genes sub-expresados y 72 a genes sobre-expresados (Tabla S2 y S3 complementarias de los Anexos). El gráfico del volcán permitió identificar los genes expresados diferencialmente (Fig. 34). Por ejemplo, al graficar el $-\log_2FC$ en el eje horizontal (logaritmo del cambio en la expresión) y en el eje vertical su significancia estadística ($-\log_{10}$ valor de p ajustado), permitió identificar los genes que muestran cambios significativos en su expresión en presencia de luteolina en comparación con la condición control. Los puntos en la parte

superior del volcán (en color rojo y azul) indican genes con alta significancia estadística, lo que implica una diferencia en su expresión en comparación con el control.

A su vez, los puntos en los extremos izquierdo y derecho del volcán representaron los genes que experimentan los mayores cambios en su expresión. Estos genes pueden tener una mayor relevancia biológica debido a la magnitud de su cambio en comparación con los genes ubicados en el centro del volcán.

Por otro lado, los puntos ubicados en la región superior izquierda del volcán correspondieron a genes están significativamente sobre-expresados, mientras que los correspondientes a genes en la región superior derecha están significativamente sub-expresados, lo que sugiere que dichos genes pueden estar regulados de manera positiva o negativa, respectivamente, en respuesta a las condiciones experimentales.

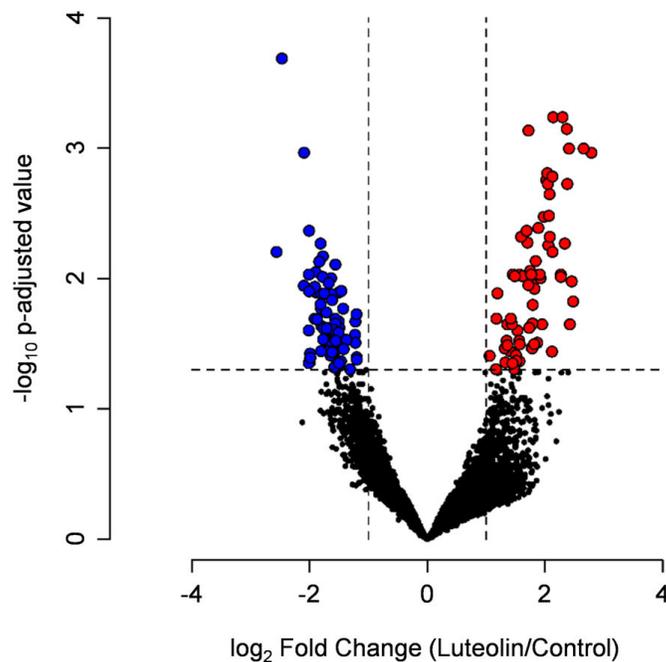


Figura 34. Diagrama de volcán mostrando los genes diferencialmente expresados (DEG) en presencia de luteolina.

En el diagrama se grafica la tasa de cambio (\log_2FC luteolina/control) versus la significancia estadística para cada gen ($-\log_{10}$ del valor p ajustado). Los puntos de color azul y de color rojo, corresponden a los genes que satisfacen los dos criterios utilizados de FC y el estadístico. En color negro, se indican genes que no satisfacen el criterio de FC, y se detectan con el mismo nivel tanto en presencia como en ausencia de luteolina.

Con el fin de visualizar patrones, tendencias y variaciones en la expresión de los genes en todas las muestras, se generó un *Heat map*. Se utilizaron diferentes colores para distinguir la expresión de cada gen, y la intensidad del color se relaciona con la tendencia de expresión de dicho gen.

En la Figura 35A se muestra el *Heat map* correspondiente a la expresión de 145 DEG, tanto sub-expresados como sobre-expresados en presencia de luteolina. Este análisis reveló la presencia de dos clusters. Un primer clúster que incluye los grupos LUT1, LUT2 y LUT3, y un segundo cluster que incluye los grupos CON1, CON2 y CON3. Estos clusteres agruparon los genes según su nivel de expresión. Además, el color verde y marrón representa los niveles de expresión de los genes sobre-expresados y sub-expresados, respectivamente. Al analizar la Figura 35A, se observó una diferencia en los niveles de expresión entre ambos clústeres, e incluso entre los triplicados biológicos de cada clúster, lo que indicó una regulación dinámica diferente entre genes sobre-expresados y sub-expresados.

Además, al graficar la distribución de los valores de expresión normalizada (FPKM) para cada *DEG*, se observó que los genes sub-expresados correspondían a genes altamente expresados en la condición de control (mediana de 867 FPKM), cuya expresión disminuyó en presencia de luteolina (mediana de 223 FPKM) (Fig. 35B). Por otro lado, los *DEG* sobre-expresados correspondían a genes con una expresión muy baja en la condición de control (mediana de 5,9 FPKM), cuya expresión se indujo en presencia de luteolina (mediana de 21 FPKM), aunque siempre por debajo de los valores de expresión normalizada de los genes reprimidos. Estos hallazgos indican que la luteolina ejerce un doble efecto regulador sobre la expresión génica en UYPR2.512, no solo induciendo la expresión de un número significativo de genes que estaban prácticamente inactivos, sino también reduciendo los niveles de ARNm de un número importante de genes que tenían una expresión alta en ausencia del flavonoide.

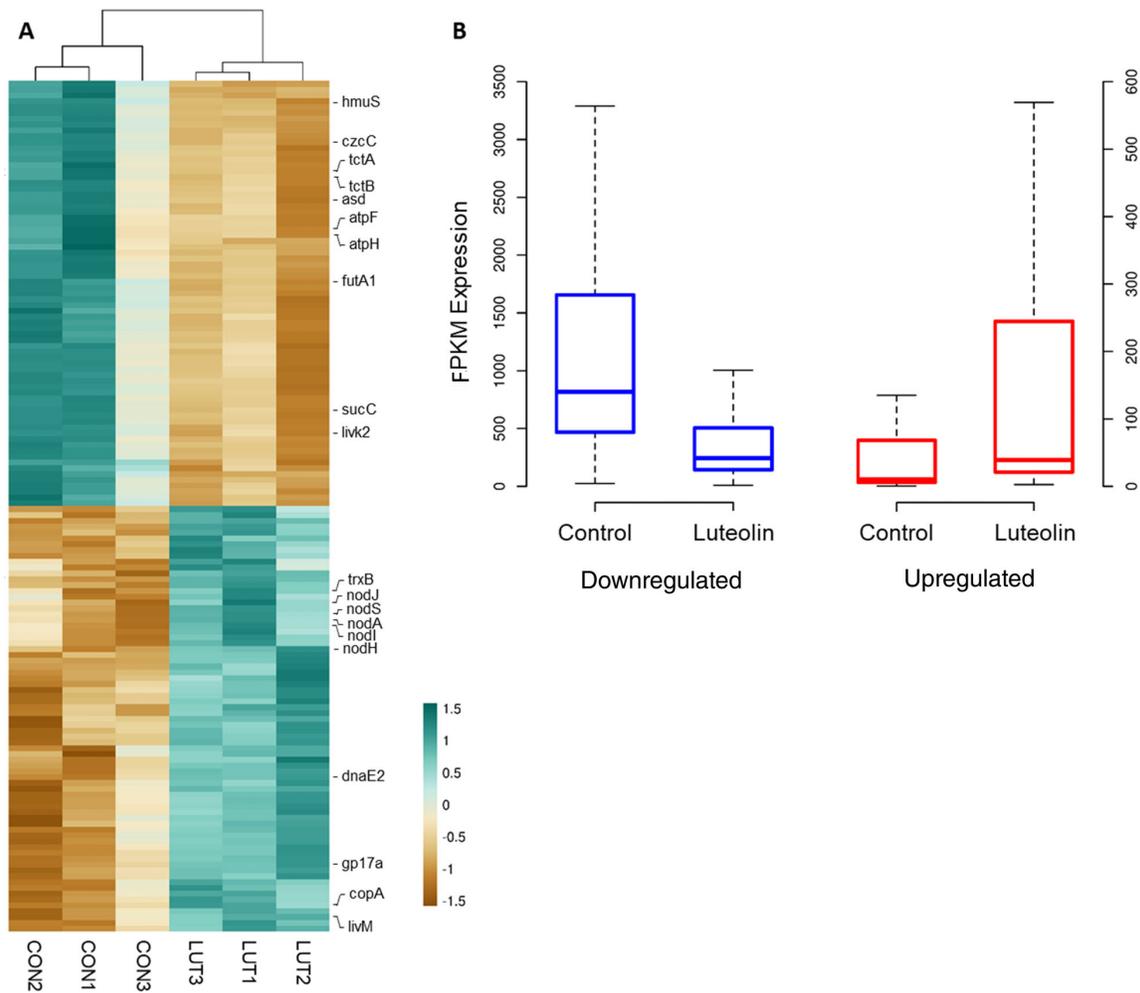


Figura 35. Comparativa de los niveles de expresión de los genes expresados diferencialmente.

Se indican en: (A) *Heat map* de la expresión de los 145 genes *DEG* en presencia de luteolina. En color verde y marrón se indican los niveles de expresión de los genes sobre-expresados y sub-expresados respectivamente. Se indica el nombre de los 10 genes con mayor tasa de cambio en cada condición. (B) Diagrama de caja de la expresión normalizada (FPKM) de genes sub-expresados en color azul y sobre-expresados en color rojo, en ausencia y presencia de luteolina. Se indican las medianas de los FPKM.

Además, se observó una distribución particular de los genes sobre-expresados y sub-expresados entre los replicones. El 86,3% de los genes sub-expresados se encontraron en el chr1, mientras que los genes sobre-expresados se distribuyen de manera más equitativa entre el chr1, el chr2 y el plásmido simbiótico (pSym) y en el resto de los replicones (25%, 33%, 22% y 20%, respectivamente) (Tabla 20). Esta distribución no es

aleatoria si consideramos el número de genes anotados en cada replicón. Los genes sub-expresados se enriquecieron a lo largo del chr1 (prueba exacta de *Fisher*, valor de $p < 1,0E-05$), mientras que los genes sobre-expresados tienen una fracción sobre representada de genes ubicados en el pSym (valor de $p = 4,0E-04$), una proporción esperada de genes del chr2 y una fracción subrepresentada de genes del chr1 (valor de $p = 1,7E-03$). La amplia distribución de genes inducidos *versus* reprimidos entre los replicones de UYPR2.512 sugiere una respuesta diferencial a la luteolina en diferentes procesos celulares.

Tabla 20. Distribución de genes sobre-expresados y sub-expresados

Se indican las cantidades de genes sobre-expresados y sub-expresados en cada uno de los replicones, y su porcentaje con respecto al total.

Replicón	Genes sobre-expresados		Genes sub-expresados	
chr1	18	25%	63	86,3%
chr2	24	33,3%	8	11,0%
pCne512e (pSym)	16	22,2%	1	1,4%
pCne512d	2	2,8%	1	1,4%
pCne512c	1	1,4%	0	0%
pCne512b	7	9,7%	0	0%
pCne512a	4	5,6%	0	0%
Total	72		73	

3.3.6 Análisis de enriquecimiento funcional

Con el objetivo de estudiar la respuesta de UYPR2.512 a la presencia de luteolina en diferentes procesos celulares, se clasificaron los DEG en categorías COG (Figura 36) y se asignaron a las vías KEGG (Tabla 21 y 22). Se logró categorizar el 83,5% de los DEG sub-expresados y el 65% de los DEG sobre-expresados. Los DEG anotados se distribuyeron en 20 categorías COG, lo cual indicó un amplio efecto de la luteolina en las funciones celulares de la cepa UYPR2.512. Al realizar el análisis de enriquecimiento basado en las categorías COG utilizando todos los DEG, se observó un enriquecimiento de las categorías C (Producción y conversión de energía) y P (Transporte y metabolismo

de iones inorgánicos) (Fig. 36). Por otro lado, al utilizar solo los DEG sub-expresados, las categorías C (Producción y conversión de energía) y E (Metabolismo y transporte de aminoácidos) fueron las enriquecidas, mientras que al utilizar los DEG sobre-expresados, la categoría Z (Citoesqueleto) fue la única enriquecida (Fig. 36). Sin embargo, se encontró que 26 genes sobre-expresados no se clasificaron en ninguna categoría COG, lo que sugiere que este análisis probablemente subestimó el número de funciones afectadas.

Con el objetivo de profundizar en el efecto de la luteolina sobre el metabolismo de UYPR2.512, se realizó un análisis de las vías KEGG en las que participan los DEG (Tabla 21 y 22). Los análisis de enriquecimiento revelaron que la mayoría de las vías enriquecidas entre los *DEG* sub-expresados pertenecían a la clase de Metabolismo, incluyendo el metabolismo de carbohidratos, energía, nucleótidos, aminoácidos, cofactores y vitaminas (Tabla 21). Estos resultados, junto con los análisis anteriores basados en categorías COG, sugieren que en UYPR2.512, la clase Metabolismo es uno de los principales procesos celulares afectados por la exposición a la luteolina. Por otro lado, la clase de Procesamiento de información ambiental fue la clase más representada entre las vías KO, de los DEG sobre-expresados (Tabla 22). Estos resultados respaldan la hipótesis de que la luteolina actúa como una fuerte señal química para UYPR2.512.

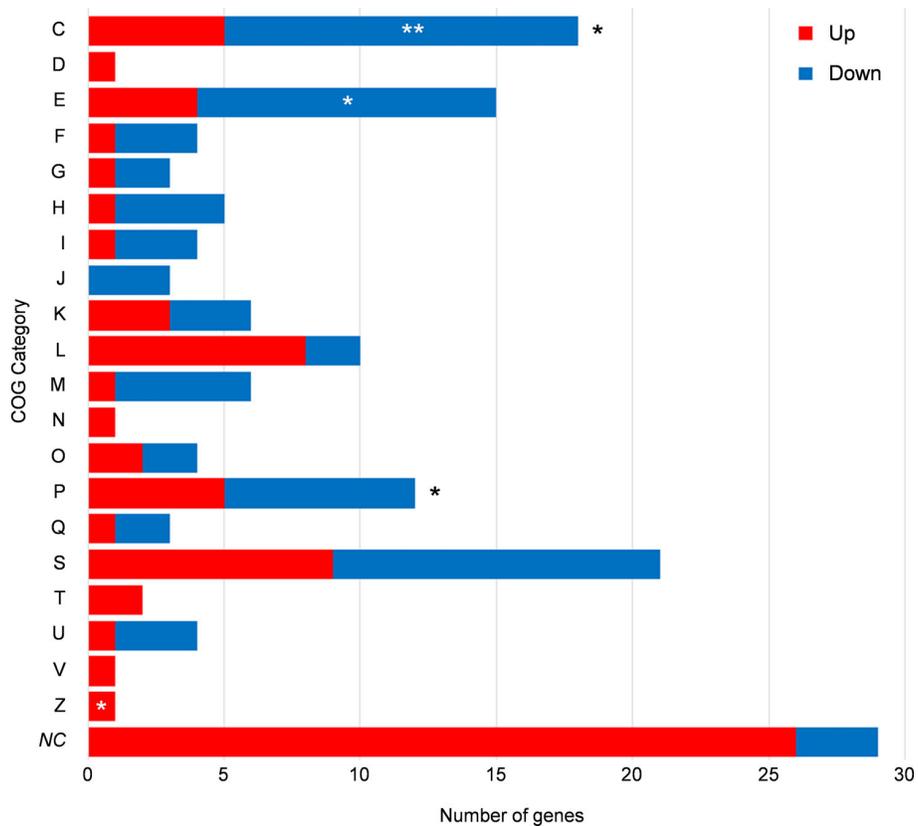


Figura 36. Análisis de enriquecimiento funcional de las categorías COG en los genes sobre-expresados y en sub-expresados.

Se muestran las categorías COG asignadas para genes sobre-expresados (en color rojo) y sub-expresados (en color azul). Las categorías de COG estadísticamente enriquecidas que consideran todos los *DEG*, se indican con un asterisco negro, mientras que las categorías específicas de COG estadísticamente enriquecidas para conjuntos de datos de genes sobre-expresados o sub-expresados se muestran con un asterisco blanco (prueba exacta de *Fisher*; *: <0,05; **: <0,01). Descripción de las categorías COG: C – Producción y conversión de energía; D – Control del ciclo celular y mitosis; E – Metabolismo de aminoácidos y transporte; F – Metabolismo de nucleótidos y transporte; G – Metabolismo de carbohidratos y transporte; H – Metabolismo de coenzimas; I – Metabolismo de lípidos; J – Traducción; K – Transcripción; L – Replicación y reparación; M – Biogénesis de pared celular/membrana/envoltura; N – Motilidad celular; O: modificación postransduccional, recambio de proteínas, funciones de chaperona; P – Transporte de iones inorgánicos y metabolismo; Q – Estructura Secundaria; S – Función desconocida; T – Transducción de señal; U – Tráfico intracelular y secreción; Z – Citoesqueleto; NC – No clasificado.

Tabla 21. Análisis de enriquecimiento funcional de las vías KO, en genes sub-expresados

Se muestran la clase, subclase y vía KO enriquecida. Se indica el número de genes asignados a cada vía en la lista de DEG y en el número de todos los genes del genoma completo de UYPR2.512 asignados a cada vía. El valor de p indica el soporte estadístico que apoya el enriquecimiento de cada vía, según prueba exacta de Fisher $p < 0,05$.

Clase	Subclase	Vía KO	ID KO	N.º DEG sub-expresados	N.º total de genes	valor de p
Metabolismo		Vías metabólicas	Ko01100	27	1397	1,64E-05
		Metabolismo del ácido 2-oxocarboxílico	Ko01210	4	69	2,74E-03
		Metabolismo del carbono	Ko01200	6	298	4,06E-02
	Mapas globales	Biosíntesis de aminoácidos	Ko01230	7	228	3,12E-03
		Biosíntesis de antibióticos	Ko01130	13	488	2,05E-04
		Metabolismo microbiano en diversos ambientes	Ko01120	13	738	9,99E-03
		Biosíntesis de metabolitos secundarios	Ko01110	17	575	4,47E-06
	Metabolismo de carbohidratos	Ciclo del citrato (TCA)	Ko00020	5	60	1,47E-04
		Metabolismo del piruvato	Ko00620	4	138	2,96E-02
		Fosforilación oxidativa	Ko00190	7	100	2,04E-05
		Fotosíntesis	Ko00195	4	21	2,58E-05
	Metabolismo de energía	Vías de fijación de carbono en procariontes	Ko00720	4	89	6,81E-03
		Metabolismo del azufre	Ko00920	4	94	8,24E-03
		Vías de fijación de carbono en organismos fotosintéticos	Ko00710	2	31	2,83E-02
	Metabolismo de nucleótidos	Metabolismo de purinas	Ko00230	4	141	3,17E-02

Clase	Subclase	Vía KO	ID KO	N.º DEG sub-expresados	N.º total de genes	valor de p
		Metabolismo de alanina, aspartato y glutamato	Ko00250	4	48	7,04E-04
	Metabolismo de aminoácidos	Metabolismo de cisteína y metionina	Ko00270	4	66	2,33E-03
		Metabolismo de arginina	Ko00220	3	36	3,44E-03
		Metabolismo de tirosina	Ko00350	3	60	1,43E-02
		Biosíntesis de valina, leucina e isoleucina	Ko00290	2	38	4,11E-02
		Metabolismo de otros aminoácidos	Metabolismo de aminoácidos ciano	Ko00460	2	20
	Metabolismo de cofactores y vitaminas	Biosíntesis de CoA y pantotenato	Ko00770	2	39	4,31E-02
Procesamiento de información Genética	Plegamiento, clasificación y degradación	Degradación de ARN	Ko03018	3	29	1,84E-03
Procesamiento de información del ambiente	Transporte de membrana	Transportadores ABC	Ko02010	7	245	4,63E-03

Tabla 22. Análisis de enriquecimiento funcional de las vías KO en genes sobre-expresados

Se muestran la clase, subclase y vía KO enriquecida. Se indica el número de genes asignados a cada vía en la lista de DEG y en el número de todos los genes del genoma completo de UYPR2.512 asignados a cada vía. El valor de p indica el soporte estadístico que apoya el enriquecimiento de cada vía, según prueba exacta de Fisher $p < 0,05$.

Clase	Subclase	Vía KO	ID KO	N.º DEG sobre-expresados	N.º total de genes	Valor de p
Metabolismo	Metabolismo energético	Fijación del carbono	Ko01710	2	31	1,72E-02
	Metabolismo de otros aminoácidos	Metabolismo de los compuestos de selenio	Ko01450	2	33	1,94E-02
	Metabolismo y degradación de compuestos xenobióticos	Degradación de bisfenol	Ko01363	1	1	6,52E-03
	Transporte de membrana	Transportadore de membrana	Ko01210	6	245	4,99E-03
Procesamiento de información del ambiente	Transducción de señales	Vías de señalización con la planta MAPK	Ko014016	1	6	3,85E-02
Procesamiento celular	Comunidad celular: procariotas	<i>Quorum sensing</i>	Ko012024	4	222	5,62E-02

A continuación, se describirá en profundidad la composición génica de los tres replicones pCne512e (pSym), chr1 y chr2, los cuales albergaron la mayoría de los genes diferencialmente expresados. Se describirá en particular aquellos genes relacionados a la interacción planta-bacteria y los que son característicos del género *Cupriavidus*.

3.3.7 Genes sobre-expresados en presencia de luteolina

Las características principales de los genes que se detallarán a continuación, se incluirán en la tabla 23.

3.3.7.1 Genes pertenecientes al pCne512e (pSym)

En presencia de luteolina, se observó la sobre-expresión de ocho genes del operón *nod*, que incluyen a *nodA* (ID = 4627), *nodC* (ID = 4631), *nodH* (ID = 4628), *nodQ* (ID = 4624), *nodS* (ID = 4626), *nodU* (ID = 4625), *nodJ* (ID = 4629) y *nodI* (ID = 4630) involucrados

en la biosíntesis y exportación de los factores *nod* (NFs) [109]. Al comparar las condiciones de control con las condiciones en presencia de luteolina, se observó un aumento en el número de lecturas asignadas a lo largo del operón *nod*, reflejando una amplia cobertura de lecturas (Fig. 37). En este contexto, todos los genes del operón *nod* mostraron una regulación positiva significativa en sus niveles de expresión (valor de p ajustado $< 0,05$), excepto el gen *nodB* (ID = 4632), el cual codifica para una polisacárido desacetilasa. En este caso se observó un FC = 2,43, pero con un valor p valor ajustado de 0,096, el cual es mayor al criterio de significancia estadística seleccionado (FDR $< 0,05$). Por otro lado, la expresión del gen *nodD* no presentó cambios significativos (FC = 1,14 y un valor de p ajustado = 0,78) en presencia de luteolina, lo cual ha sido descrito previamente en otros rizobios [329].

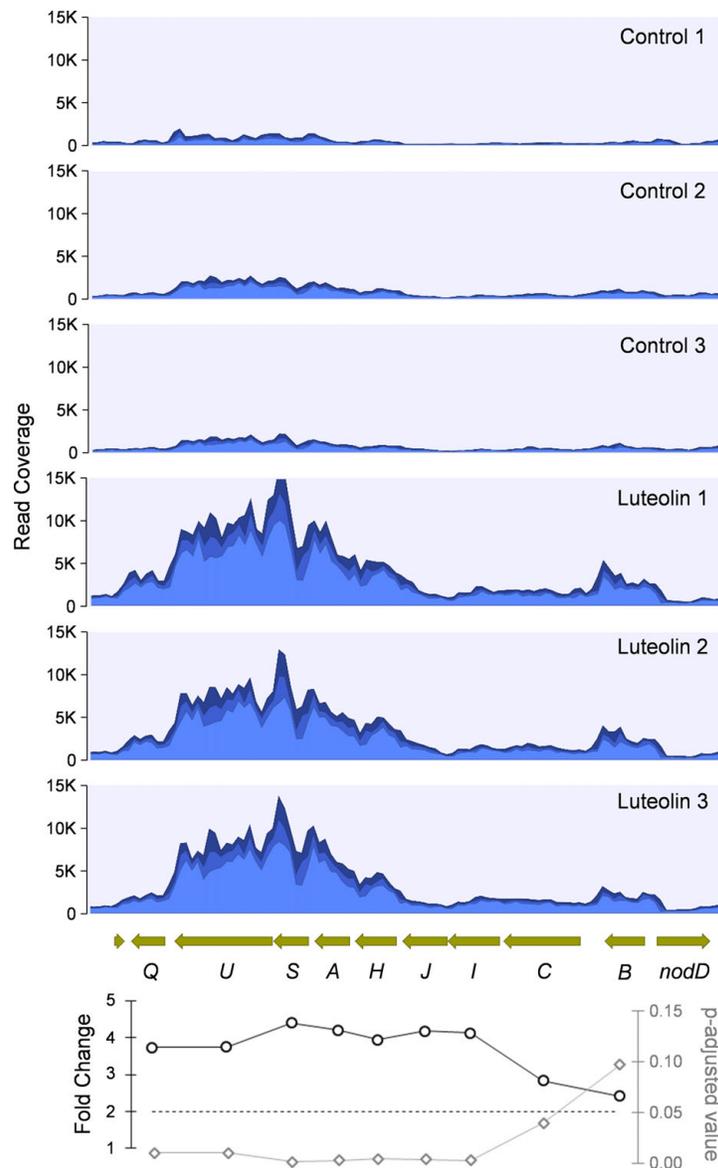


Figura 37. Comparación de la expresión de los genes *nod* en presencia y en ausencia de luteolina.

Se presenta la cobertura de lectura del operón *nod* en las seis muestras analizadas, utilizando la misma escala. Se puede observar una baja cobertura de lecturas en las tres muestras de control en comparación con las muestras en presencia de luteolina, donde las tres réplicas muestran un incremento en la cantidad de lecturas mapeadas en todos los genes del operón.

Aguas arriba del operón *nod*, se detectó sobre-expresión del gen *hdcP* (ID = 4622), el cual está anotado como una proteína con un dominio de hemeritina conservado, del cual se desconoce su función, pero comparte cierta similitud con reguladores transcripcionales de la familia Rrf2 [330]. Además, incluido en el pSym, se identificó un homólogo del gen *dnaE2* (ID = 4461), el cual codifica para la subunidad alfa de la ADN polimerasa III, una enzima involucrada en la reparación del ADN durante la respuesta SOS en bacterias [331].

En cuanto a los otros genes sobre-expresados, se observó que la mayoría estaban relacionados con secuencias de inserción, elementos móviles, transposasas e integrasas, características comunes en los plásmidos simbióticos de muchos rizobios [164], [332]. Sin embargo, la función biológica de estos genes está poco estudiada.

3.3.7.2 Genes pertenecientes al chr2

Entre los 24 genes ubicados en el chr2, se identificaron tres genes (ID = 6913, 6954 y 9794) que codifican para proteínas reguladoras posiblemente pertenecientes a la familia XpsR [333]. Estos tres genes están relacionados con la biosíntesis de un polisacárido rico en ramnosa, componente clave de los EPS, el cual desempeña un papel importante en los rizobios durante el estado de “vida libre” y en los primeros pasos de la interacción rizobio-leguminosa [334]. A su vez, se observó la sobre-expresión de un gen que codifica para una posible proteína extracelular con un dominio de laminina-G (ID = 9275), la cual desempeña un papel importante en la adhesión celular. Por lo tanto, esta proteína podría estar involucrada en la interacción con la planta y en los procesos de adhesión durante la simbiosis. También, se observó la inducción del gen *flaG* (ID = 8061), el cual forma parte de un operón conservado y se sugiere que está implicado en la biosíntesis, ensamblaje y movimiento de flagelos (ID = 8060 a 8073).

Además, se identificaron tres genes sobre-expresados implicados en el transporte de aminoácidos de cadenas ramificadas (BCAAs). Estos genes incluyen a *livM1* (ID = 8690), ubicado en el chr2, y *livM2* (ID = 5931) y *livH* (ID = 5932). Estos genes codifican para permeasas de los sistemas de transporte ABC de alta afinidad para aminoácidos de

cadena reamificadas (BCAAs, de sus siglas en inglés *Branched-Chain Amino Acids*). Este hallazgo sugiere que la luteolina podría modular el transporte de BCAAs en la cepa UYPR2.512, lo cual concuerda con la observación de que las categorías Metabolismo de aminoácidos y transportadores tipo ABC están enriquecidas entre los genes sobre-expresados.

Se detectaron dos genes adicionales sobre-expresados (ID = 8691 e ID = 8692) adyacentes a *livM1* (ID = 8690). El primero (ID = 8691) codifica para una posible proteína transmembrana anotada como S-isoprenil cisteína metiltransferasa, la cual estaría involucrada en modificaciones de lípidos y lipopolisacáridos [335] y en modificaciones de las membranas. Además, se identificó un gen *trxB* (ID = 8692), que codifica para una tiorredoxina reductasa conservada, implicada en la desintoxicación de especies reactivas de oxígeno, la modificación postraduccional de proteínas, el recambio de proteínas y las funciones de chaperonas [336].

Entre los genes sobre-expresados en presencia de luteolina, también se observaron genes implicados en el metabolismo de compuestos aromáticos. El gen *dapDO* (ID = 8648) el cual codifica para una 2,4'-dihidroxiacetofenona dioxigenasa bacteriana, la cual se coinduce con un gen que codifica para una posible oxidorreductasa (ID = 8647). Además, se encontraron genes que codifican para una supuesta proteína ligL con función de hidrolasa de ácido 2-pirona-4,6-dicarboxílico (ID = 8721) y para una probable vainillato desmetilasa, codificada por *vanA* (ID = 9750). Todos estos genes están implicados en la degradación de compuestos aromáticos [86], [89] y podrían desempeñar un papel importante en la colonización y supervivencia bacteriana en las rizosferas vegetales [258].

Además, un número significativo de genes sobre-expresados en el chr2 (32) codificaron proteínas hipotéticas según el servidor web *RAST*. Sin embargo, al menos seis de ellos (ID = 7264, 7265, 7292, 7323, 7331 y 7334) están relacionados con la biología de bacteriófagos cuando se analizan mediante *BLAST*.

3.3.7.3 Genes pertenecientes al chr1

Entre los otros genes sobre-expresados, se observó la inducción de *xdhC* (ID = 1238), que codifica para una subunidad hierro-azufre de la xantina deshidrogenasa, implicada en el metabolismo de las purinas y para una posible salicílico-CoA 5-hidroxisilasa/oxidorreductasa bifuncional codificada por el gen *abmA* (ID = 2893), implicada en la degradación del ácido salicílico [89]. Estas funciones estarían relacionadas con la degradación de compuestos aromáticos, complementando lo observado para varios genes ubicados en el chr2.

A su vez, se identificaron varios genes involucrados en el metabolismo de metales pesados. Dentro de ellos, los genes *chrA* (ID = 6559) y *chrB* (ID = 9271) han sido descritos como genes implicados en la resistencia al cromato, expulsándolo desde el citoplasma bacteriano [242]. Además, se observó la inducción de *actP* (ID = 49) que codifica un transportador de Cu^{+2} tipo P, el cual pertenece a la familia CPx [337], una proteína periplásmica Cu-ATPasas codificada por *cusF/copA* (COG5569; ID = 2613) y un homólogo de *copQ* (ID = 101), propuestos como involucrados en la resistencia al cobre en *C. metallidurans* [78]. Junto con *actP* (ID = 49), se indujo la expresión de *copZ* (ID = 50), que codifica para una chaperona de cobre, y *cueR* (ID = 48), que codifica para un regulador transcripcional sensible a Cu(I), lo que sugiere una respuesta coordinada de la homeostasis del cobre ante la presencia de luteolina. También se observó la inducción del *ttgR* (ID = 130), que codifica un regulador transcripcional miembro de la familia TetR, posiblemente implicado en la regulación de sistemas transportadores de extrusión de compuestos tóxicos [338].

Finalmente, se observaron tres genes (ID = 2335, ID = 2336 e ID = 2625) los cuales están involucrados en diferentes metabolismo bacterianos. El gen *ldh2* (ID = 2335) que codifica para la enzima L-lactato deshidrogenasa, responsable de la conversión reversible de piruvato en lactato) [339]. El gen *yhbL* (ID = 2336) codifica para un translocador de 2-oxoglutarato/malato, mientras que *bgIS* (ID = 2625) codifica para una supuesta glicosil hidrolasa extracelular de la Familia 3 [340]. Debido a la importancia de las enzimas hidrolíticas en la colonización de la rizosfera por microorganismos [341] sería

particularmente interesante investigar el papel del homólogo *bgIS* durante la colonización de las raíces de leguminosas por la cepa UYPR2.512.

Tabla 23. Genes sobre-expresados en presencia de luteolina.

Se indican replicón en el que se encuentran, ID del gen, nombre del gen y función asociada. ND corresponde a nombre del gen no determinado.

Replicón	ID del gen	Nombre del gen	Función asociada
pCne512e (pSym)	4624 al 4632	<i>nodA, nodC, nodH, nodQ, nodS, nodU, nodJ, nodI, nodB</i>	Biosíntesis y transporte de NFs
	4461	Homólogo a <i>hdcP</i>	Reparación del ADN
	6913, 6954, 9794	Sin nombre	Biosíntesis de polisacáridos (reconocimiento rizobio-leguminosa)
Chr2	9275	ND	Adhesión celular
	8061	<i>flaG</i>	Biosíntesis, ensamblaje y movimiento de flagelos
	8690	<i>livM1</i>	Sistemas de transporte ABC de alta afinidad para BCAAs
	8691	ND	Modificaciones de lípidos y lipopolisacáridos, y proteínas de membrana
	8692	<i>trxB</i>	Desintoxicación de especies reactivas de oxígeno, la modificación postraduccional de proteínas
pCne512a	5931, 5932	<i>livM2, livH</i>	Sistemas de transporte ABC de alta afinidad para BCAAs
Chr1	8648, 8721, 9750	<i>dapDO, ligL, vanA</i>	Metabolismo de compuestos aromáticos
	1238	<i>xdhC</i>	
	2893	<i>abmA</i>	
	6559, 9271	<i>chrA, chrB</i>	Metabolismo de metales pesados
	49	<i>actP</i>	
	2613, 101	<i>cusF/copA</i> y un homólogo de <i>copQ</i>	
50, 48, 130	<i>copZ, cueR, ttgR</i>		
	2335, 2336, 2625	<i>ldh2, yhbL, bgIS</i>	Metabolismo bacteriano

3.3.7.4 Genes sub-expresados en presencia de luteolina

Las características principales de los genes que se detallarán a continuación, se incluirán en la tabla 24. De los 73 genes sub-expresados, 63 estaban codificados en el chr1, ocho en el chr2 y uno en el plásmido pCne512e (Tabla 21). Este último codifica una posible proteína de unión al ADN bacteriano similar a las histonas (ID = 4557) (Tabla 24). Aunque la función exacta de estas proteínas no está clara, se ha informado que son capaces de

envolver el ADN y estabilizarlo frente a la desnaturalización causada por el estrés y las condiciones ambientales extremas [342].

Entre los 8 genes sub-expresados en el chr2, se encontró un homólogo del gen *metE* (ID = 7162), que codifica para una proteína involucrada en la biosíntesis de metionina, lo cual indica su participación en el metabolismo y transporte de aminoácidos. También se observó afectada la expresión de *tctA* y de *tctB* (ID = 7769 e ID = 7770), que junto con *tctC*, codifican para un sistema transportador de tricarboxilato tripartito (TTT) implicado en la adquisición de diferentes compuestos. Además, se detectó la sub-expresión de los homólogos de *czcB* (ID = 8788) y *czcC* (ID = 8787), que codifican para componentes del sistema de transporte de resistencia-nodulación-división, lo que sugiere que el transporte de diversos sustratos a través de las membranas se ve afectado en presencia de luteolina.

También se observó la sub-expresión de otros dos genes agrupados (ID = 9006 e ID = 9007). El primer gen codifica para una proteína hipotética involucrada en la formación de pseudouridina en el metabolismo bacteriano [343] y el otro gen codifica para una posible 2-amino-3-carboximuconato-6-semialdehído descarboxilasa (ID = 9007) implicada en el metabolismo de compuestos aromáticos [344].

Finalmente, se observó la sub-expresión del gen *fur* (ID = 8372), el cual codifica para una probable proteína reguladora de la captación de hierro. En dicho contexto, otros 10 genes relacionados con el metabolismo del hierro también se observaron sub-expresados, ubicados en el chr1. Estos incluyen un posible operón (ID = 2940 al 2943) que codifica para las proteínas TonB, ExbB y ExbD, las cuales están involucradas en el transporte activo de sustratos a través de la membrana externa [256], así como un probable receptor de membrana externa para la hemina.

En relación con la adquisición de hierro, se observó la sub-expresión del gen (ID = 1031) que codifica para un transportador de tipo ABC para el ión hierro férrico, un gen (ID = 1663) que codifica para una posible proteína de membrana regulada por hierro, y un homólogo de *fecA* (ID = 3656), el cual codifica para un receptor de sideróforo-hierro. También se observó la sub-expresión del gen *pvsA* (ID = 241) implicado en la biosíntesis

de sideróforos [345] del gen *bfd* (ID = 577) que codifica para una ferredoxina asociada a bacterioferritina y del gen *sod* (ID = 897) el cual codifica para una superóxido dismutasa de hierro.

Entre los genes sub-expresados, codificados en el chr1 se encontraron 9 genes relacionados con la producción de energía. Estos incluyen dos genes homólogos de *atpH* (ID = 18 e ID = 4231) y *atpF* (ID = 19 e ID = 4232), que codifican para la cadena delta de la ATP sintasa y para la subunidad F0 de la ATP sintasa, respectivamente. Además, se observó la sub-expresión de los genes *cox2* y *cyoA1* (ID = 597 e ID = 129), que codifican para el polipéptido II de la citocromo c oxidasa y para la subunidad II de la ubiquinol oxidasa del citocromo O, respectivamente. Estos resultados sugieren que el metabolismo energético se vio afectado en presencia de luteolina, que a su vez concuerda que las categorías COG que se observaron enriquecidas.

Particularmente, se observó que los genes *cysP* (ID = 2673) y *cysT* (ID = 2663), implicados en el metabolismo del azufre y la biosíntesis de cisteína (Aguilar-Barajas et al., 2011), así como el gen *ssuE* (ID = 2671), involucrado en la adquisición de azufre a partir de sulfonatos de alcanos [346], se encontraban sub-expresados, lo que sugiere que el metabolismo de azufre también se ve afectado por la presencia de luteolina.

Se observó la sub-expresión de varios genes que codifican para proteínas transportadoras luego de la exposición de UYPR2.512 a la luteolina. Entre ellos, se identificaron tres genes (ID = 119, ID = 399 e ID = 877) que codifican para posibles receptores periplásmicos con dominios Bug (llamados así debido a los genes de captación de *Bordetella*) (Antoine et al., 2000). Al menos uno de estos homólogos, junto con *tctB* (ID = 118), forma parte de un sistema TTT (*tctABC*) probablemente involucrado en el transporte de citrato [347].

Otra de las categorías COG que se vio enriquecida fue la del metabolismo de aminoácidos. En este sentido, se identificaron tres genes sub-expresados que codifican proteínas relacionadas con la adquisición de aminoácidos. Dos de ellos son homólogos

de *livK* (ID = 41 e ID = 3054), los cuales codifican componentes periplásmicos de los sistemas de transporte de BCAAs. Además, se identificó el gen *metQ* (ID = 2578), que codifica una proteína periplásmica de unión específica a metionina y el gen *takp* (ID = 71), que codifica para un sistema de transporte dependiente de unión a soluto.

A su vez, se observó la sub-expresión del gen *tyrB* (ID = 1380), que codifica una aminotransferasa de aminoácidos aromáticos, *fhaA* (ID = 619) y *maiA* (ID = 620). Estos últimos genes codifican para una fumarilacetoacetasa y para una maleilacetoacetato isomerasa, respectivamente, y están implicados en el metabolismo de la tirosina y la fenilalanina a través de la vía del homogentisato

Además, se encontraron sub-expresados un gen que codifican para una L-asparaginasa (ID = 2243), involucrada en la biosíntesis de glutamina, glutamato, aspartato y asparagina, y el gen *ilvC* (ID = 1260), que codifica para una proteína responsable de la biosíntesis de BCAA, así como del gen *leuA* (ID = 1265), el cual juega un papel en la biosíntesis de leucina, isoleucina y valina. Otros genes que se observaron sub-expresados incluyen argininosuccinato liasa (ID = 3365) y peptidil-prolil cis-trans isomerasa (ID = 1819), los cuales están implicados en la biosíntesis de arginina y prolina respectivamente, y finalmente *glyA* (ID = 3282) que está implicado en la biosíntesis de serina.

Se observó una disminución de la expresión de tres genes involucrados en el ciclo TCA en presencia de luteolina, un primer gen *mdh* (ID = 3063) que codifica para una malato deshidrogenasa, el gen *sdhA* (ID = 3059) que codifica para una subunidad de la flavoproteína succinato deshidrogenasa y *sucC* (ID = 825) que codifica para la subunidad β de la succinil-CoA ligasa (formadora de ADP). Además, se encontró una disminución en la expresión del gen *accaD* (ID = 3644) que codifica para una acetil-CoA carboxilasa, que cataliza el paso limitante de la biosíntesis de *nov*o de ácidos grasos.

En cuanto al metabolismo de los carbohidratos, se encontraron reprimidos en presencia de luteolina, un homólogo del gen *pck2* (ID = 107), que codifica una fosfoenolpiruvato-carboxiquinasa que cataliza la conversión de oxalacetato a fosfoenolpiruvato. Y el gen *paaA* (ID = 3804) que codifica para una proteína identificada como 2,3-dehidroadipil-CoA hidratasa, implicada en la degradación del fenilacetato (EC 4.2.1.17).

Finalmente, se encontraron dos genes sub-expresados, que incluyó a un gen que codifica el factor ρ (factor Rho) (*rho*; ID = 2831) implicado en la terminación de la transcripción, y el gen *nusB* (ID = 3298) que codifica para una proteína de terminación de la transcripción que forman parte de un complejo proteico considerado como un regulador del plegamiento del ARN ribosomal [348].

Tabla 24. Genes sub-expresados en presencia de luteolina.

Se indican replicón en el que se encuentran, ID del gen, nombre del gen, y función asociada. ND corresponde a nombre del gen no determinado.

Replicón	ID del gen	Nombre del gen	Función asociada
pCne512e	4557	ND	Estabilización del ADN en respuesta a estrés
chr2	7162	<i>metE</i> ,	Metabolismo y transporte de aminoácidos
	7769, 7770, 8788, 8787	<i>tctA, tctB, czcB, czcC</i>	Sistemas de transporte de diversos sustratos
	9006, 9007	ND	Metabolismo bacteriano
	9007	ND	Metabolismo de compuestos aromáticos
Chr1	8372, 2940 al 2943, 1031, 1663, 3656, 241, 577, 897	<i>Fur, tonB, exbB y exbD, sn, sn</i> , homólogo de <i>fecA, pvsA, bfd, sod</i>	Metabolismo de metales
	18 y 4231, 19 y 4232, 597, 129	dos genes homólogos a <i>atpH, atpF, cox2, cyoA1</i>	Producción de energía
	119, 399, 877, 118	<i>sn, tctB</i>	Proteínas transportadoras
	1380, 619, 620, 2578, 71	<i>tyrB, fhaA, maiA, metQ, takp</i>	Metabolismo y transporte de aminoácidos
	41, 3054	Homólogos a <i>livK</i>	Transporte de aminoácidos
	2243, 1260, 1265, 3365, 1819, 3282	<i>sn, ilvC, leuA, sn, sn, glyA</i>	Biosíntesis de aminoácidos
	3063, 3059, 825, 3644	<i>Mdh, sdhA, sucC, accaD</i>	Ciclo de ácidos tricarbónicos (TCA)
	2831, 3298	<i>Rho, nusB</i>	Transcripción del ARN
	107, 3804	<i>pck2, paaA</i>	Metabolismo de carbohidratos

3.4 Validación de RNA-Seq mediante q-PCR

En base a los criterios de selección previamente descritos como gen normalizador se seleccionó el gen *eloG* con una eficiencia ($E = 90\%$), mientras que dos genes sobre-expresados *nodS* ($E = 96,2\%$) y *hcdP* ($E = 92,3\%$), se seleccionaron para la cuantificación relativa del ARNm. La relación de expresión relativa entre el gen normalizador y los genes diana se calculó mediante la fórmula de *Pfaff*. Con el objetivo de validar la expresión génica diferencial en presencia de luteolina, se estudió la expresión del gen *nodS* y del gen *hcdP* mediante análisis de q-PCR.

Por otro lado, entre los *DEG* obtenidos mediante la secuenciación por la tecnología de Ion Torrent, cuatro genes sobre-expresados (*trx*, *odl*, *nodU*, *nodC*) y un gene sub-expresado (*czcB*), mostraron el mismo perfil de expresión luego de ser secuenciados mediante Illumina (Fig. 38). Como se observa en la figura 38, los valores de logFC mediante ambas metodologías de secuenciación, para los grupos de genes estudiados, mostraron un alta correlación $R^2 = 0.9975$.

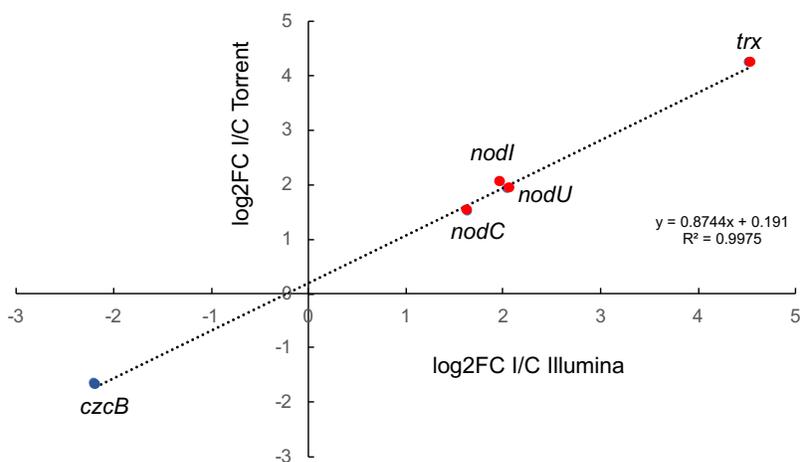


Figura 38. Análisis comparativo de DEG obtenidos mediante la secuenciación por Ion Torrent vs Illumina. En el gráfico se observa la relación entre los log2FC I/C mediante Ion Torrent (eje y) vs log2FC I/C mediante Illumina (eje x), de 5 genes sobre-expresados (color rojo) y un gen sub-expresado (en color azul). Los genes sobre-expresados corresponden a *trx*, *nodI*, *nodU*, *nodC* y un gen su-expresado *czcB*. La línea continua indica la tendencia, con una correlación $R^2 = 0.9975$.

3.5 Comparativa del análisis transcriptómico de la cepa UYPR2.512 de *C. necator* en presencia de luteolina mediante NGS

Al comparar los datos obtenidos mediante q-PCR con los resultados de RNA-Seq, se observó que la cantidad relativa de ARNm para ambos genes aumentó significativamente en presencia de luteolina (Fig. 39). En el caso del gen *nodS* mediante RNA-Seq el valor promedio de CPM (conteo de lecturas por millón) aumentó desde aprox. 200 a 900 y mediante q-PCR el Ct promedio aumentó desde 5 a 40 CPM, tanto en las muestras control, como en presencia de luteolina, respectivamente. Al estudiar la expresión del gen *hdcP*, mediante el análisis proveniente de RNA-Seq, se observó un aumento del valor promedio de CPM de 50 a 150 y mediante q-PCR el Ct promedio aumenta desde 1,8 a 9 CPM, en las muestras control y en presencia de luteolina, respectivamente.

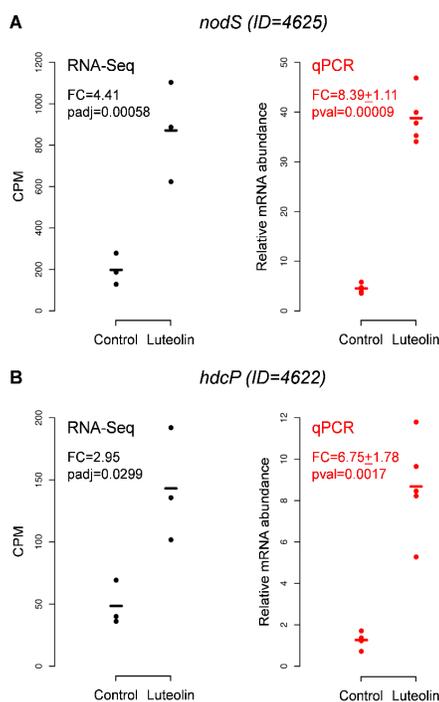


Figura 39. Comparación de la expresión de los genes *nodS* y *hdcP* de *C. necator* en presencia y en ausencia de luteolina, mediante RNA-Seq y q-PCR.

En los paneles A y B, se indican los datos obtenidos de RNA-Seq (CPM, recuentos de lecturas por millón) en color negro y en los paneles derechos los datos obtenidos de q-PCR (abundancia relativa de ARNm estimada a partir de los valores de Ct) en color rojo, en ambas condiciones experimentales (control y luteolina). Cada punto representa una réplica biológica (n=3 para RNA-Seq y n=5 para q-PCR) y los

segmentos de línea completos representan valores promedio. En A se muestran los valores de FC , valores de p ajustados para el gen *nodS*, y en B para el gen *hdcP*.

4 Discusión

Teniendo en cuenta que los flavonoides son moléculas clave involucradas en los primeros pasos de muchas interacciones entre rizobios y leguminosas, se decidió estudiar la respuesta de la cepa UYPR2.512 a la presencia de luteolina.

En un primer intento de explorar la respuesta a la exposición a dicha molécula, se analizó el efecto en la actividad del promotor del gen *nodB* (primer gen del operón). Para ello se utilizó una cepa derivada de UYPR2.512 que contenía el plásmido reportero pCBM01, observándose una fuerte inducción de la actividad β -galactosidasa luego de la exposición a luteolina (Fig. 28). Estos resultados evidencian que la proteína NodD de UYPR2.512 es capaz de reconocer la caja *nod* de la LMG19424 y activar la expresión de los genes *nod* en respuesta a la luteolina.

Luego, con el fin de comparar los niveles de expresión génica en cultivos en fase exponencial de UYPR2.512 en medio de cultivo M9, en ausencia y en presencia de luteolina, se realizó un análisis transcriptómico. Se utilizó como herramienta el RNA-Seq, aproximación utilizada para estudiar la respuesta a los flavonoides en alfa-rizobios [310] que aún no había sido utilizada en beta-rizobios.

Primeramente, se realizó RNA-Seq utilizando la tecnología de Ion Torrent. Se obtuvieron, un total de 250.000 lecturas por muestra mediante esta metodología, valor inferior al esperado. Una de las posibles causas, puede deberse a una insuficiente depleción de los ARN ribosomales, por lo cual una segunda ronda de depleción podría disminuir la presencia de dichas moléculas contaminantes, como se sugirió en un trabajo reportado por Klonowska [273]. También se podría implementar otra metodología de depleción del ARN ribosomal, la cual en vez de utilizar nucleótidos de captura y remoción contra las subunidades 16S y 23S del ARN ribosomal presentes en el ARN_{tot} (como se utilizó previamente en este trabajo), se utilizan cebadores contra transcritos específicos no deseados durante la preparación de las librerías de ADN_c, lo que aumenta la especificidad.

En base a lo observado se decidió repetir el experimento de RNA-Seq, pero esta vez mediante el uso de la plataforma Illumina. Entre ambas aproximaciones se logró obtener una versión mejorada del genoma de referencia (ver Capítulo II).

Mediante la plataforma de Illumina, al comparar los análisis de expresión diferencial de genes realizado se observaron 145 DEG (Fig. S2 y F2 suplementarias de Anexos) y entre los genes de alta expresión se identificaron los genes del operón *nod*.

A partir de los análisis de DEG, se observó una diferencia en los valores originales de expresión de los genes sobre-expresados y de los sub-expresados, lo cual indicó que la luteolina ejerce un fuerte efecto regulador de la expresión génica en UYPR2.512. En dicho contexto, la mayoría de los genes sub-expresados mostraron altos niveles de expresión en la condición de control y su expresión disminuyó luego de la exposición a la luteolina. A su vez, los genes sobre-expresados mostraron una expresión muy inferior en la condición de control y fueron inducidos por la presencia de luteolina (Fig. 35B). Lo que indicaría que hay un mayor efecto en la represión de genes que en la activación. Esto nos permite hipotetizar que en condiciones normales de cultivo los genes involucrados en funciones básicas, son reprimidos en presencia de una señal como ser la luteolina, estos genes disminuyen su expresión y a la vez aumenta la expresión de genes directamente involucrados a la respuesta a la luteolina.

Una segunda asimetría surge del análisis de la distribución de los genes sobre-expresados y sub-expresados entre los distintos replicones de la bacteria. En genomas compuestos, como es el caso de UYPR2.512, se ha demostrado que los cromosomas están enriquecidos en las funciones centrales, los cuales contienen la mayoría de los genes de mantenimiento celular básico, mientras que los cromosomas secundarios, como ser los crómidos y los plásmidos, están enriquecidos en el metabolismo secundario, el transporte y las funciones reguladoras [184]. Dada la distribución observada (Tabla 20), es esperable que los procesos celulares centrales estén enriquecidos entre los genes sub-expresados, mientras que las funciones codificadas por los replicones secundarios se encuentren enriquecidos entre los genes sobre-expresados. Los resultados obtenidos concuerdan con esta hipótesis, ya que los análisis de enriquecimiento funcional indicaron que la mayoría de los genes sub-expresados

pertenecen a la clase de metabolismo, mientras que los genes sobre-expresados se asocian con la clase de procesamiento de información ambiental (Fig. 36, Tabla 21 y 22).

Cabe destacar que la mayoría de los genes sub-expresados no pudieron ser clasificados, lo que sugiere que estos genes codifican para funciones no caracterizadas, lo que evidencia aún más la importancia de su caracterización, ya que podrían proporcionar nuevos conocimientos sobre la interacción simbiótica entre *Cupriavidus* y sus hospederos.

Al explorar los datos de RNA-Seq, primero se analizaron los niveles de expresión de los genes *nod*, debido a su importancia clave durante muchas interacciones entre rizobios y leguminosas. Teniendo en cuenta el nivel de conservación de la secuencia entre las cajas *nod* de UYPR2.512 y de LMG19424, era de esperar que la luteolina active la inducción de la expresión de los genes *nod* en UYPR2.512, al igual a lo observado en LMG19424 [194], a pesar de las pequeñas diferencias observadas entre ambas regiones intergénicas *nodD-nodB* (ver Capítulo II). En dicho escenario, los resultados de RNA-Seq confirmaron los resultados obtenidos con el uso de genes reporteros de que la luteolina puede actuar como un inductor *in vivo* de los genes *nod* en UYPR2.512. Esto sugiere que, en respuesta a la luteolina, UYPR2.512 sería capaz de producir moléculas de NFs sulfatadas con una estructura similar a las producidas por LMG19424 de *C. taiwanensis* [194] y refuerza la hipótesis de que los flavonoides también son moléculas clave en la comunicación entre β -rizobios y sus hospederos.

Al igual que como se describe en la cepa LMG19424 de *C. taiwanensis* [194], se encontró una segunda caja *nod* conservada en el genoma de UYPR2.512 (ver resultados en Capítulo II). Esta caja *nod* se ubica a 195 pb corriente arriba del codón de inicio ATG del gen *noeM* (ID = 4812), el cual se encuentra anotado como un gen que codifica para la proteína esteroles desaturasa, involucrado en la síntesis de moléculas NFs atípicas en LMG19424 de *C. taiwanensis* [194], [272]. Además, se demostró que su expresión es inducida por la presencia de luteolina o de exudados radiculares de *M. pudica* [272], [273]. Sin embargo, no se observó una expresión diferencial del gen *noeM* en

UYPR2.512, en este estudio. Este resultado podría sugerir diferentes mecanismos de regulación entre ambos simbioses e indica que el gen *noeM* de UYPR2.512 no tendría un papel esencial en sus hospederos de nodulación, lo cual también se observó en LMG19412 de *C. taiwanensis* [272].

Además de los NFs, los LPS y EPS en rizobios tienen un rol muy importante durante la interacción rizobio-leguminosa, así como también cuando los rizobios se encuentran en el estadio de “vida libre” [148], [334]. En dicho contexto, se observó que la presencia de luteolina indujo la expresión de tres genes (ID = 6913, 6954 y 9794) que codifican homólogos de XpsR en UYPR2.512. Se ha demostrado que XpsR está implicado en la regulación de la biosíntesis de EPS en *Ralstonia solanacearum* [333] donde se requiere para una completa virulencia [349]. Además, se propuso que *xpsR* es uno de los reguladores de virulencia de *R. solanacearum*, utilizado también por *C. taiwanensis* por estar involucrado en la simbiosis con *M. pudica* [166], pero no se ha investigado su papel en profundidad. En base a estos resultados, se sugiere que la luteolina modula la producción de EPS a través de la acción de XpsR en UYPR2.512.

Entre otras funciones, los flavonoides juegan un papel importante como molécula señal repelente o atrayente para bacterias patógenas o beneficiosas en la rizósfera [116]. A su vez, los flagelos de los rizobios son necesarios para la competencia de la rizosfera y la colonización de las raíces [350]. En dicho contexto, se ha observado la inducción de genes flagelares en respuesta a flavonoides específicos en *B. japonicum* cuando se exponen a genisteína [351] y también en cultivos de *C. taiwanensis*, *P. phymatum* y *R. mimosarum* cultivados en presencia de exudados radiculares de *M. pudica* [273].

En la cepa UYPR2.512 se observó la inducción del gen *flaG* (ID = 8061). Dado el papel de los flagelos, estos resultados sugieren que la luteolina podría estar activando la síntesis flagelar, actuando como atrayente para este flavonoide, afectando su competitividad, como se ha observado en otros rizobios [352].

En relación con esto, se observó la sub-expresión de un homólogo del gen *yciL* (ID = 9006), el cual estaría involucrado en la formación de biopelículas en *B. cepacia* [353], lo

que sugiere además que la luteolina induce en UYPR2.512 un cambio en su metabolismo, lo cual se refleja en su movilidad.

Se observó la inducción de la expresión de genes supuestamente involucrados en la unión del calcio, como el gen *rapA* (ID = 4076), el cual codifica para una hipotética proteína de unión a calcio extracelular. Si bien los homólogos de este gen se conservan entre *Cupriavidus* y organismos estrechamente relacionados, no hay evidencia experimental de su función. Sin embargo, el homólogo distante *casA*, identificado en *Rhizobium etli* codifica para una proteína similar a la calmodulina, que es secretada y es capaz de unirse a Ca^{2+} , desempeñando un papel clave en el desarrollo de los bacteroides durante la simbiosis con *P. vulgaris* [354]. Curiosamente, en el genoma de UYPR2.512, *rapA* está agrupado con genes que codifican proteínas de la vía Sec o SSTII (ID = 4078 a 4091) que podrían estar involucradas en la secreción de esta proteína. De forma complementaria, la co-inducción de un gen (ID = 9275) que codifica una proteína hipotética que contiene un dominio laminina-G, conocido por su capacidad para unir calcio [355], confirma la participación de Ca^{2+} durante la interacción *Cupriavidus-Mimosa*.

Una gran cantidad de genes relacionados con la homeostasis de metales, en particular con la homeostasis del cobre, fueron regulados en presencia de luteolina. En dicho contexto, el cobre es un cofactor esencial para las oxidasas respiratorias del hemo-cobre y para las enzimas involucradas en algunas reacciones redox. A su vez, podría ser altamente tóxico cuando se superan las concentraciones intracelulares óptimas, por lo cual resulta esencial una regulación controlada de la homeostasis del cobre [356]. La inducción de la expresión de genes involucrados con la resistencia a Cu^{2+} como el gen *copQ* (ID = 101 y 9522), genes de unión que incluyó a *copZ* (ID = 50) y *cusF* (ID = 2613) y de eflujo *actP* (ID = 49) indicaron que UYPR2.512 activa el eflujo de iones Cu^{2+} en respuesta a la luteolina. En ese sentido, se ha demostrado que CopZ modula el eflujo de Cu^{2+} mediado por la ATPasa en *Listeria monocytogenes* y el gen *cueR* codifica para un activador transcripcional del eflujo de cobre en *E. coli* [357], lo que sugiere que la homeostasis del cobre en UYPR2.512 estaría regulado por estos factores transcripcionales. Estos resultados son interesantes debido a la conocida afinidad del

género *Cupriavidus* por el cobre y su capacidad para crecer en presencia de altas concentraciones de metales, especialmente para la cepa UYPR2.512 de *C. necator*, LMG19424 de *C. taiwanensis* y UYMMa02A de *Cupriavidus* sp. cómo se comentó en el Capítulo I.

Sin embargo, el nexo entre la homeostasis del cobre y la presencia del flavonoide luteolina no está claro actualmente. Una posible explicación a ello estaría relacionada con la capacidad de los flavonoides de quelar diversos iones metálicos, y volverse tóxico para la célula [358]. De hecho, se observó la sobre-expresión de genes de resistencia al cromato implicados en el eflujo, podría deberse a que la luteolina provoca el aumento de genes involucrados en la resistencia a este metal [359].

Sin embargo, la principal evidencia de una respuesta coordinada de homeostasis de metales proviene del análisis del regulón de hierro. Al igual que el cobre, el hierro es un nutriente esencial para las bacterias que en exceso podría volverse tóxico. Por lo tanto, las bacterias y en particular los rizobios han desarrollado sofisticados mecanismos de homeostasis del hierro [83]. Se observó que la luteolina reprime la expresión de *pvsA* de UYPR2.512, presuntamente implicado en la biosíntesis y el transporte de sideróforos. Al analizar el genoma, se observa que el gen *pvsA* forma parte de un operón junto con *pvsB*, *pvsC*, *pvsD* y *pvsE* (ID = 238 al 241) cuyos homólogos están implicados en la síntesis y exportación de los sideróforos de tipo α -hidroxicarboxilato xantoferrina y vibrioferrina en *Xanthomonas* y *Vibrio*, respectivamente [345], [360].

Estas observaciones sugieren que UYPR2.512 sería capaz de producir un sideróforo similar en condiciones limitantes de hierro. En el caso de *Xanthomonas* y *Vibrio*, el gen *pvuA* codifica para el receptor de la membrana externa responsable de la captación del sideróforo férrico. Sin embargo, no se observó un homólogo del gen *pvuA* en el genoma de UYPR2.512. En su lugar, se identificó un gen putativo receptor de citrato férrico (*fecA* ID = 3656) cuya expresión también estaba sub-expresada en presencia de luteolina. Estos resultados proponen la interrogante de que si este homólogo de *fecA* sería el responsable del transporte del sideróforo producido por UYPR2.512 o por el contrario le otorga a esta bacteria la capacidad de transportar sideróforos producidos por otros

microorganismos. En este caso restaría por identificar el transportador del sideróforo del tipo vibroferrina producido por UYPR2.512.

Curiosamente, también se encontró que un homólogo a *fur* (regulador transcripcional de la homeostasis de hierro) (ID = 8372) estaba sub-expresado en respuesta a la luteolina. En muchas bacterias, Fur es un regulador central capaz de controlar la expresión de muchos genes en respuesta a la disponibilidad de hierro [361]. Sin embargo, se ha demostrado que en α -rizobios los homólogos a Fur no participan en este tipo de regulación, sino que están involucrados en la homeostasis de manganeso [362], [363]. En base a estas observaciones, el gen *fur* de UYPR2.512 podría regular los sistemas de captación de hierro y/o de manganeso. Además, podría afectar otros procesos biológicos, evitando el estrés oxidativo y respuestas a condiciones ambientales no favorables, como se ha observado específicamente en la cepa JMP134 de *C. pinatubonensis* [364].

A pesar de no tener evidencia experimental, se postula que UYPR2.512 podría transportar hemina a través de un sistema de adquisición constituido por los genes *tonB*, *exbB*, *exbD* y *hemP* (IDs = 2940 al 2943), los cuales se encuentran bajo el mismo esquema de regulación, y permitiría que esta bacteria lo utilice como hierro nutricional como en otros rizobios [365].

Dos genes adicionales, *bfd* (ID = 577) y *sodB* (ID = 897), completarían el regulón de hierro afectado por luteolina. El gen *bfd*, codifica para una ferredoxina asociada a bacterioferritina cuya función en algunas bacterias es la liberación de hierro de la bacterioferritina en condiciones limitantes de hierro [366]. El gen *sodB* codifica para una proteína superóxido dismutasa con funciones críticas en la resistencia al estrés y durante las interacciones bacterias-huésped [367].

Las bacterias ambientales tienen la capacidad de hacer frente a condiciones adversas incluyendo la presencia de compuestos tóxicos como xenobióticos, metales pesados, antibióticos y especies de oxígeno altamente reactivas [368], [369]. Los principales mecanismos de resistencia están coordinados por reguladores de estrés globales que involucran la expresión de vías biodegradación y bombas de desintoxicación. La presencia de luteolina induce en UYPR2.512 varios mecanismos de resistencia al estrés,

incluida la expresión de *acrR* (ID = 130), implicado en la señalización global del estrés y la regulación de los sistemas de extrusión de múltiples fármacos [370].

En relación con esto, la presencia de luteolina indujo la expresión de genes que codifican para homólogos a las proteínas tiorredoxina y flavodoxina reductasa (ID = 8692 y 9750), pertenecientes a una familia de enzimas de importancia en los mecanismos de detoxificación celular y eliminación de compuestos xenobióticos [371].

Finalmente, se observó que la presencia de luteolina en las condiciones de cultivo utilizadas disminuye la expresión de varios genes altamente expresados pertenecientes al metabolismo central, en particular relacionados con el ciclo de los ácidos tricarboxílicos (TCA, de sus siglas en inglés *Tricarboxylic Acid Cycle*) y el metabolismo de aminoácidos. La bacteria modelo UYPR2.512 es capaz de crecer en medio mínimo M9 [314], utilizando citrato como fuente de carbono, energía y amoníaco como única fuente de nitrógeno. Estos nutrientes deben transportarse al interior de la célula para ser metabolizados y utilizados para producir precursores para la biosíntesis de componentes básicos celulares (por ejemplo, lípidos, aminoácidos, purinas, pirimidinas y vitaminas). En la mayoría de los organismos, si no en todos, el metabolismo del carbono y el nitrógeno están coordinados [34].

En presencia de luteolina, se observó la sub-expresión de genes implicados en la biosíntesis de oxalacetato y fumarato, dos de los principales precursores del metabolismo de aminoácidos, lo cual podría explicar en parte la disminución de la expresión de muchos genes relacionados con el metabolismo de los aminoácidos [372]. Por otro lado, la disminución en la síntesis de acetil-CoA a través de represión del gen que codifica para la enzima acetil-CoA carboxilasa podría estar afectando la biosíntesis de ácidos grasos. Además, es importante considerar que, entre los efectos de la luteolina, también se observó la síntesis de *novo* de NFs [373] la cual requeriría nuevas moléculas precursoras (oligosacáridos, lípidos, quitinas), y una disminución de la biosíntesis del sideróforo similar a la vibrioferrina que aliviaría la carga metabólica de moléculas precursoras (alanina, ácido cítrico, etanolamina y ácido 2-cetoglutarico) [345], [374].

A partir de todos los resultados observados, se propone un modelo integrativo, el cual resume las principales rutas metabólicas, funciones biológicas, reacciones químicas, enzimas y genes regulados a nivel transcriptómico en UYPR2.512 en respuesta a la exposición de luteolina (Fig. 40). Como se ha observado, la mayoría de los genes reprimidos se ubican en el Chr1 y están implicados principalmente en el metabolismo, transporte de carbono y aminoácidos. Esto sugiere que luego de la exposición a la luteolina, UYPR2.512 redirige su metabolismo, causando una disminución en la expresión de genes altamente expresados relacionados con el metabolismo, a la misma vez que aumenta la expresión de genes casi no expresados. Los genes inducidos codifican muchas funciones probablemente involucradas en la interacción con la planta hospedera. Más allá de la síntesis y exportación de NFs, también se evidenció la implicancia de la luteolina en la regulación de EPS, flagelos, colonización de las raíces, homeostasis de metales y respuesta frente al estrés. Esto permite concluir que los genes regulados descritos en este capítulo representan las principales adaptaciones que ocurren en UYPR2.512, durante los primeros pasos del proceso de nodulación entre β -rizobios y leguminosas.

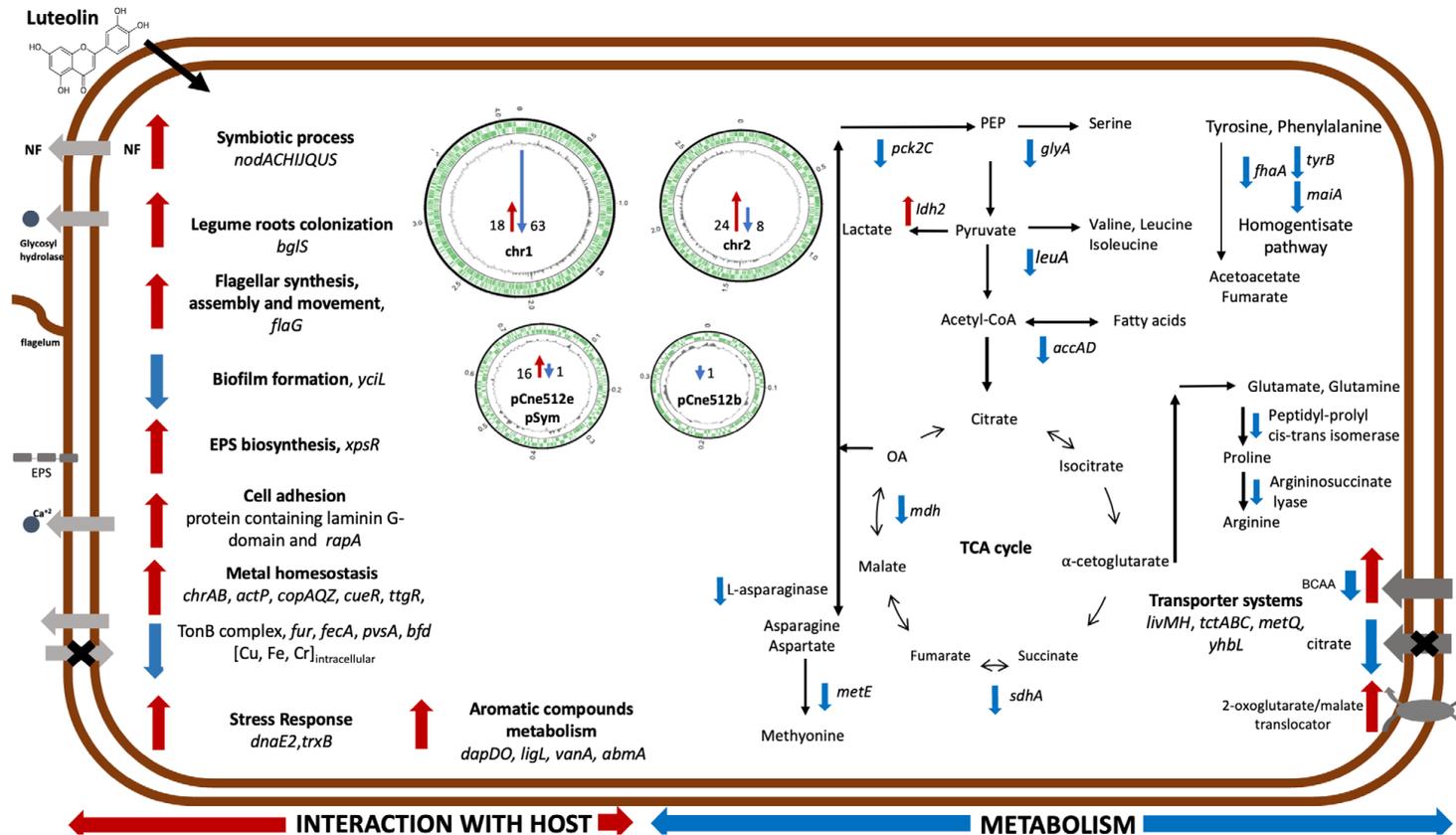


Figura 40. Modelo de la respuesta transcriptómica de la cepa UYPR2.512 frente al agregado de luteolina.

La ilustración muestra las rutas metabólicas, las funciones biológicas, las reacciones químicas, las enzimas involucradas y los genes regulados a nivel transcriptómico en UYPR2.512 de *C. necator* en respuesta a la exposición al flavonoide luteolina. Las flechas rojas y azules indican procesos regulados, genes sobre-expresados (dirección hacia arriba) o sub-expresados (dirección hacia abajo). TCA: ciclo de los ácidos tricarboxílicos; NFs: factores de nodulación; PEP: fosfoenol-piruvato; AO: oxaloacetato; BCAA: aminoácido de cadena ramificada; EPS: exopolisacáridos.

CAPÍTULO III

Identificación de las señales involucradas en el diálogo molecular entre la cepa UYPR2.512 *Cupriavidus necator* y *Mimosa pudica*.

Objetivo general

Identificación de las señales involucradas en el diálogo molecular entre UYPR2.512 y *M. pudica*.

Objetivos específicos

1. Identificación de flavonoides involucrados en la interacción de UYPR2.512 con *M. pudica*.
2. Caracterización de metabolitos presentes en los exudados radiculares de *M. pudica*.

Resumen

Los flavonoides desempeñan un papel esencial en el reconocimiento entre la mayoría de los rizobios y leguminosas hospederas, los cuales al unirse a la proteína reguladora NodD inducen la expresión de los genes *nod* en rizobios. En el capítulo II de esta tesis se demostró que la luteolina es inductor de la expresión de los genes *nod* en UYPR2.512, por lo cual en este capítulo nos enfocamos en identificar los flavonoides en los exudados radiculares de *M. pudica*. Se observó que los exudados radiculares de *M. pudica* provocaron un aumento significativo en la actividad de la β -galactosidasa en UYPR2.512 portando el plásmido pCBM01, sugiriendo la presencia de moléculas inductoras. Para identificar estas moléculas se emplearon técnicas de química analítica como HPLC-UV-ELSD, HPLC-ESI-MS/MS y GC-MS.

Se identificaron 68 metabolitos, de los cuales 40 se observaron tanto en ausencia como en presencia de la cepa UYPR2.512, 19 en ausencia de UYPR2.512 y 9 metabolitos fueron específicos en presencia de UYPR2.512. La mayoría de estos 68 metabolitos pertenecieron a las familias químicas de ácidos carboxílicos/dicarboxílicos, aminoácidos y polisacáridos. Entre los 9 metabolitos observados como únicos de la interacción, se identificaron el ácido benzoico, ácido xilónico, cisteína y el mio-inositol. Sin embargo, su rol en la interacción debe ser estudiado en profundidad.

A su vez, se evaluó la respuesta a otros flavonoides, y se identificaron la crisina, miricetina, apigenina, como inductores de los genes *nod*.

Mediante los experimentos realizados en este capítulo podemos concluir que los flavonoides crisina, apigenina, miricetina y moléculas no flavonoides como el ácido cumárico son inductores de los genes *nod* en UYPR2.512. Y que, a su vez, ácidos fenólicos como el ácido benzoico, el cual fue detectado en los exudados radiculares de *M. pudica*, es inductor de los genes *nod* en otras bacterias y podría estar involucrado en los primeros pasos de la interacción simbiótica. Estos hallazgos muestran que en los primeros pasos del reconocimiento entre la bacteria y la planta estarían involucradas una variedad de compuestos de diferente naturaleza química.

1 Introducción

1.1 Metabolómica

La metabolómica se define como un enfoque analítico para la identificación y cuantificación de pequeñas moléculas en un sistema biológico, conocidas como metabolitos (generalmente con un peso molecular inferior a 1.500 Da) [375].

A su vez, los metabolitos surgen como producto final de las reacciones o intermediarios de las rutas metabólicas que ocurren en el organismo. Al ser el último paso en el flujo de información que va desde el genotipo hasta el fenotipo (o función) de un individuo, los productos metabólicos pueden considerarse como la última respuesta a los cambios ambientales y a la composición genética de los individuos, brindando información acerca de su actividad bioquímica.

El auge de la metabolómica tuvo lugar a finales de los años 90 y desde entonces el uso de la metabolómica ha crecido exponencialmente, siendo una herramienta muy utilizada en diferentes áreas como ser en el estudio de enfermedades [376], el desarrollo de fármacos [377], el análisis medioambiental [375] y la toxicología [378], entre otros.

Por otro lado, la metabolómica se divide en dos enfoques, la metabolómica dirigida y la no dirigida. La metabolómica dirigida, se basa en la medición de un número definido de moléculas pequeñas caracterizadas químicamente, para obtener datos cuantitativos, como por ejemplo los relacionados con una vía metabólica específica [379], [380].

Por otro lado, la metabolómica no dirigida, tiene como objetivo analizar de manera integral y no selectiva, las moléculas contenidas en una muestra, sean conocidas o sean nuevos metabolitos [380]. Con el fin de identificar dichas moléculas se han utilizado diferentes técnicas de la metabolómica. A continuación, se realizará una descripción de alguna de las técnicas de separación más utilizadas en metabolómica.

1.2 Técnicas de separación cromatográficas aplicada a la metabolómica

1.2.1 Cromatografía líquida de alta presión (HPLC)

La técnica de HPLC es un tipo de cromatografía de líquidos, que se realiza a alta presión y consiste en inyectar la muestra en una fase móvil líquida y hacerla pasar a través de una fase estacionaria inmisible con la fase líquida y que se mantiene empaquetada en una columna. Generalmente, se utilizan columnas con fases estacionarias no polares o con baja polaridad (por ejemplo, hidrocarburos) y como fase móvil se utilizan solventes relativamente polares (por ejemplo, agua, metanol o acetonitrilo). Esta técnica permite separar un amplio rango de compuestos orgánicos, desde pequeñas moléculas hasta péptidos y proteínas. Los componentes de la solución migran de acuerdo con las interacciones no-covalentes de los compuestos con la columna y la fase móvil actúa de portador de la muestra. A medida que los compuestos eluyen de la columna, son detectados por diferentes detectores según la naturaleza de los compuestos a determinar.

1.2.2 Cromatografía de gases (GC)

La cromatografía de gases es otra técnica de separación, en la cual se utiliza como fase móvil un gas inerte (Helio y/o Hidrógeno) purificado y regulado, y como fase estacionaria, una columna en la que se produce la separación de los compuestos dependiendo de las interacciones físicas y/o químicas entre los compuestos de la muestra y la fase estacionaria. Además, cuenta con un inyector el cual volatiliza previamente las muestras líquidas. Las fases estacionarias se diferencian de acuerdo con su tamaño, polaridad y estabilidad térmica, lo cual determinará su capacidad para interactuar de forma selectiva con diferentes solutos. Algunas características deseables en las fases estacionarias incluyen tener un amplio rango de temperatura (entre -60 °C y 400 °C), baja presión de vapor, químicamente inerte, baja viscosidad en las condiciones de trabajo utilizadas y una alta adherencia, lo cual evita que sea arrastrada por la fase móvil. Las más utilizadas incluyen hidrocarburos, polisiloxanos, polifenilésteres, poliésteres y polietilenglicol, entre otros.

1.3 Métodos de detección para Cromatografía líquida de alta presión (HPLC) y Cromatografía de gases (GC)

Los detectores acoplados a cromatógrafos incluyen; el de índice de refracción, electroquímicos, de fluorescencia, de arreglo de diodos (DAD) y ultravioleta-visible (UV-Vis), los cuales se basan en las propiedades físicas o químicas de los analitos, y detectores de masas como los espectrómetros. La espectrometría de masas es una técnica de alta sensibilidad y exactitud, basada en la fragmentación de la muestra, en la separación y en el registro de los iones moleculares producidos según su relación masa/carga (m/z) en un sistema a vacío. Luego de que la muestra es inyectada, atraviesa la fuente de ionización, la cual carga a las moléculas y las fragmenta, generando los iones que serán luego separados en el analizador, en función de un patrón de ionización característico para cada metabolito, que se utiliza posteriormente para su identificación [381].

Esta técnica puede aplicarse tanto al análisis cualitativo como al análisis cuantitativo de una muestra, ya que proporciona las herramientas necesarias para la identificación de sustancias, tanto a partir de sus iones, fragmentos que se producen al romper la molécula

analizada (caracterización estructural), como empleando el valor de su masa determinado con una elevada exactitud (composición elemental) [382].

A su vez, hay detectores de masas que trabajan en configuración tándem o híbrida, conocida como MS/MS, en la que dos o más analizadores de masas se acoplan mediante un paso de reacción adicional para aumentar su capacidad de analizar muestras químicas. Un uso común de la MS en tándem es el análisis de biomoléculas, como proteínas y péptidos. Las moléculas de una muestra dada se ionizan y el primer espectrómetro (designado MS1) separa estos iones por su relación masa-carga (m/z). Los iones de una relación m/z particular provenientes de MS1 se seleccionan y luego se dividen en iones de fragmentos más pequeños. Esto se logra a través de la fragmentación de iones por disociación inducida por colisión (CID) en instrumentos en el espacio en tándem. Estos fragmentos luego se introducen en el segundo espectrómetro de masas (MS2), que a su vez separa los fragmentos por su relación m/z y los detecta [383].

Las variantes en la espectrometría de masas se deben fundamentalmente al uso de distintas fuentes de ionización, como ser Ionización electrónica (EI, de sus siglas en inglés *Electronic Ionization*), ionización por electrospray (ESI, de sus siglas en inglés *Electrospray Ionization*) [384], ionización química a presión atmosférica (APCI, de sus siglas en inglés *Chemical Ionization at Atmospheric Pressure*), entre otras, diferenciándose en el rango de peso molecular y en la polaridad de los analitos que detectan. Entre ellos, a menudo se prefiere ESI para perfilar metabolitos "desconocidos", ya que este enfoque de ionización forma iones moleculares intactos y ayuda a la identificación inicial. En dicho contexto, la introducción de técnicas APCI, permitió expandir el número de compuestos que pueden ser analizados por HPLC-MS. En APCI, las moléculas del analito primero son ionizadas (modo positivo y modo negativo) a presión atmosférica y los iones del analito son separados mecánicamente y electrostáticamente de las moléculas neutras.

A su vez, las variantes en la espectrometría de masa están dadas en base a los distintos analizadores de masas, los cuales consisten en una cámara de vacío donde se aplica algún tipo de campo eléctrico o magnético para poder ver las diferencias de

comportamiento de los iones al moverse a través de la cámara. Existen varios tipos de analizadores de masas, los más utilizados son los analizadores de Tiempo de Vuelo (Q-TOF, de sus siglas en inglés *Time Of Flight*) [384], Simple Cuadrupolos (QS, de sus siglas en inglés *Simple Quadrupoles*) y las trampas iónicas, como ser Quadrupole Ion Trap QIT, de sus siglas en inglés) [385] y Orbitrap [386]. El Q-TOF utiliza un campo eléctrico para acelerar los iones y medir el tiempo que toman para alcanzar el detector, dando como resultado la relación m/z . Los QS, constan de cuatro varillas de metal enfrentadas en pares, sobre los que se aplica una corriente continua y otra alterna, permitiendo esto crear un campo eléctrico controlado que desvía selectivamente los iones que lo atraviesan, pudiéndose filtrar de esta manera los iones con un margen muy pequeño de m/z , siendo éstos los únicos que llegarán al detector. Las trampas iónicas, funcionan de un modo similar a los cuadripolos, pero en vez de desviar los iones que no se encuentran en un determinado margen de m/z , éstos se confinan y almacenan en una cámara, y posteriormente son liberados de forma selectiva. En dicho contexto los analizadores Q-TOF son los más utilizados debido a que brindan información de masa molecular exacta (alta resolución) [384].

1.4 Métodos de identificación para Cromatografía líquida de alta presión (HPLC-MS) y Cromatografía de gases (GC-MS)

Luego que se conoce la relación masa m/z , y partiendo de la premisa de que moléculas con estructuras similares producen patrones de fragmentos similares, es posible comparar los patrones obtenidos en bases de datos tales como *Global Nature Product Social* (GNPS) [387] y SciFinder [388].

Utilizando herramientas capaces de agrupar los espectros según su similitud, como el programa Molecular Networking del CytoScape [389], se generan clústeres que posibilitan una visualización general del interactoma, es decir, el conjunto completo de interacciones moleculares. Esto permite, además, la identificación cualitativa de los compuestos contribuyentes y la cuantificación de las cantidades absolutas de metabolitos involucrados en los procesos estudiados.

Se obtienen clústeres, los cuales permiten una visualización general del conjunto completo de las interacciones moleculares o interactoma e incluso identificar los compuestos contribuyentes (cualitativos) y cuantificar las cantidades absolutas de metabolitos que participan en los procesos estudiados (cuantitativos).

A su vez, la identificación definitiva, así como la cuantificación de cada molécula debe hacerse mediante el empleo de estándares.

Algunos ejemplos de sus aplicaciones incluyen el análisis de pesticidas y herbicidas en la industria agrícola, identificación de compuestos aromáticos, aceites, azúcares en la industria alimenticia [390], presencia de alcoholes, ácidos orgánicos en diferentes procesos industriales [391], y de fármacos en la industria farmacéutica [392], entre otros.

1.5 La metabolómica aplicada al estudio de la simbiosis rizobio-leguminosa

Mediante la utilización de transcriptómica y proteómica se han logrado identificar y caracterizar genes y proteínas cruciales para la simbiosis [393]. Sin embargo, estos estudios no proporcionan evidencia directa de que los genes y proteínas detectados estén realmente activos metabólicamente bajo las diversas condiciones estudiadas. En dicho contexto, la metabolómica aplicada a la simbiosis rizobio-leguminosa permite identificar los metabolitos involucrados y las vías metabólicas clave que se activan o modifican en respuesta a la interacción, lo que conduce a una mejor comprensión de los mecanismos moleculares y las reacciones químicas implicadas en esta asociación. El mayor inconveniente que presenta esta aproximación es la dificultad para diferenciar los metabolitos que provienen de la bacteria o de la planta, especialmente en muestras que contienen una mezcla de metabolitos de ambos organismos, lo cual resulta desafiante.

1.5.1 Comunicación en la rizósfera

Durante el desarrollo de las plantas, estas exudan a través de sus raíces una gran variedad de compuestos hacia la rizosfera que las rodea. Las plantas liberan hasta un 40% de los productos de la fotosíntesis como exudados. Estos pueden clasificarse en compuestos de bajo peso molecular (ácidos orgánicos, azúcares y flavonoides), representando gran parte de la diversidad, y compuestos de alto peso molecular menos diversos, como mucílago (polisacáridos) y proteínas [394]. Se ha demostrado que la

cantidad y composición de los exudados depende de diversos factores como ser: la especie vegetal, el estado de desarrollo de la planta, y a las condiciones bióticas y abióticas [395].

Los metabolitos secretados por las raíces participan en varios procesos en la rizósfera. Al modular la composición de los exudados de las raíces, las plantas pueden modificar las propiedades del suelo para adaptarse y asegurar su supervivencia en condiciones adversas. Utilizan diversas estrategias como: (1) modificar el pH del suelo para solubilizar nutrientes, (2) quelar compuestos tóxicos, (3) atraer microbiota beneficiosa o (4) liberar sustancias tóxicas para patógenos, entre otras [396].

Muchos componentes de los exudados de las raíces sirven como fuentes de carbono y nitrógeno para los microorganismos de la rizósfera y también desempeñan un papel en los procesos de señalización que regulan las interacciones planta-bacteria [396]. Por esta razón, tanto las raíces como sus exudados son candidatos ideales para abordar el estudio de la simbiosis rizobio-leguminosa desde una perspectiva metabolómica. Tecnologías como la cromatografía líquida acoplada a espectrometría de masas en tándem (HPLC-MS/MS) y UHPLC-MS/MS [397], la cromatografía de gases acoplada a espectrometría de masas (GC/MS) [398], son las más utilizadas para el estudio de metabolitos de una muestra biológica.

La presencia de flavonoides y otros metabolitos implicados en la interacción entre rizobios y leguminosas se ha abordado desde la metabolómica. Se ha detectado la presencia de flavonoides en raíces y nódulos [399], así como en exudados radiculares de soja [400]. En este contexto, la presencia de cromóforos en la estructura de los flavonoides les otorga la capacidad de absorber la luz en la región del espectro UV-Vis. Estos presentan dos bandas de absorción: la banda I corresponde a la fracción cinamoilo, cuyo espectro oscila entre 300 y 380 nm, y la banda II corresponde al benzoilo, cuyo espectro oscila entre 240 y 280 nm. Estas características espectrales facilitan su identificación.

Con el fin de identificar las moléculas señal involucradas durante la interacción entre de *C. necator* UYPR2.512 y su hospedero *M. pudica*, en este capítulo nos enfocamos en el

estudio de la composición de los metabolitos presentes en los exudados radiculares, mediante varias metodologías.

2 Materiales y métodos

2.1 Obtención de exudados radiculares

Las semillas de *M. pudica* se obtuvieron comercialmente del proveedor Outsidepride Seeds, LLC (Oregón, EE. UU.). Primeramente, las semillas se deshidrataron con etanol 70%, se secaron en papel de filtro y se trataron con H₂SO₄ concentrado durante 10 min, seguido de 5 enjuagues de 1 min cada uno con agua estéril. Luego fueron sumergidas en hipoclorito de sodio 4% durante 5 min, seguido de 6 lavados de 1 min con agua estéril. Las semillas esterilizadas superficialmente se pregerminaron en placas de agua con agar:agua 0,8% (p/v), a 30 °C durante 48 h.

Las semillas pregerminadas se sembraron en frascos de vidrio de 250 mL, con 20 semillas por frasco, y se agregaron 20 mL de agua estéril junto con una rejilla de acero inoxidable para el soporte de las plántulas. Las plantas se incubaron durante 5 días bajo un fotoperíodo de 16 h de luz y 8 h de oscuridad, a 26 °C.

Luego, se recolectó la solución de agua que contenía los exudados de las raíces en tubos cónicos de plástico de 50 mL, los que se centrifugaron durante 5 min a 12.000 rpm para romper los desechos celulares y los sobrenadantes se filtraron mediante filtros de nitrocelulosa con poros de 0,45 µm (Millipore, Estados Unidos). Los exudados de raíces filtrados y esterilizados se utilizaron en lugar de agua para preparar medio M9 modificado (Anexos).

2.2 Evaluación de la expresión del gen *pnodB19424-lacZ*

Con el fin de evaluar el efecto inductor de los exudados radiculares se utilizó el mismo procedimiento utilizado en el punto 2.3 de Materiales y métodos del capítulo II, adaptando el protocolo a la utilización de placas de 96 pocillos. Los cultivos bacterianos de crecimiento ON se normalizaron a una concentración final de 10⁸ cel/mL en medio mínimo M9, y 20 µL de esta suspensión fueron traspasados a una placa de 96 pocillos conteniendo 180 µL de medio mínimo fresco M9 y M9 modificado (Anexos), conteniendo

éste en vez de agua el mismo volumen de exudados radiculares frescos colectados. Como control se utilizó medio de cultivo sin inocular. Se realizaron cuatuplicados técnicos. Como control se utilizó la cepa LMG19424 de *C. taiwanensis* conteniendo los mismos plásmidos [194].

2.3 Búsqueda de moléculas involucradas en la interacción entre UYPR2.512-*M. pudica*

2.3.1 Estrategia I. Búsqueda de flavonoides en exudados radiculares mediante técnicas de HPLC-DAD

2.3.1.1 Colecta y obtención de exudados radiculares

Semillas de *M. pudica* se esterilizaron y germinaron mediante la metodología detallada en 2.1. Las semillas se sembraron en frascos de vidrio de 250 mL (50 semillas por frasco) conteniendo 50 mL de medio de cultivo Howieson 1/10 (Anexos) [401]. Las semillas en los frascos que contenían este sistema fueron inoculadas con una suspensión bacteriana de UYPR2.512 a una concentración de 10^7 ufc/semilla y se incubaron durante 3 y 16 días. Después de este periodo, se recogieron los exudados. Como controles se utilizaron plántulas sin inocular y medio inoculado sin plantas. Los exudados liofilizados se mantuvieron en un desecador hasta su uso.

2.3.1.2 Búsqueda de compuestos flavonoides

Con el fin de evaluar la presencia de una mezcla de flavonoides en los exudados inoculados con UYPR2.512, los liofilizados fueron reconstituidos en dimetilsulfóxido (DMSO) y se realizó una curva de diluciones seriadas que fueron luego analizadas mediante lector de placas multimodo Varioskan Flash del IIBCE, realizando un barrido espectral desde 230 a 400 nm, que permitiera observar los 2 picos característicos de compuestos polifenólicos, correspondientes a las fracciones benzoilo y cinamoilo. Luego, las muestras fueron disueltas en buffer ácido fosfórico 0,5% pH 2 e inyectadas en el HPLC-DAD (Shimadzu, Kioto, Japón) equipado con una bomba de disolvente LC-6AD, un controlador de sistema SCL 10AVB, un horno de columna CTO-10ASVP, un inyector Rheodyne modelo 7725, un detector de matriz de diodos SPD-M10AVP, un software

Clase VP para adquisición de datos. Como fase estacionaria se utilizó una columna analítica C18 HPLC Column (5 μ m, 250 \times 4.6 mm, Phenomenex).

El análisis fue llevado a cabo utilizando el siguiente gradiente (fase móvil: A = metanol 90%; fase móvil B = ácido fórmico 0,1%, pH 2,65): 0-30 min 35-55 %A; 30-55 min 55 %A; 55-56 min 55-35 %A; 56-70 min 35%A.

2.3.1.3 Análisis de los cromatogramas obtenidos

Los cromatogramas se analizaron utilizando el software *Empower Pro*, a la longitud de onda óptima para cada metabolito buscado. Con el fin de identificar la presencia de flavonoides se realizaron análisis comparativos a partir de los cromatogramas de flavonoides utilizados como estándares. Como estándares se utilizaron los siguientes flavonoides con las correspondientes longitudes de onda óptimas: luteolina (λ =375 nm), genisteína (λ = 259 nm), daizeína (λ = 248 nm), naringenina (λ = 290 nm), catequina (λ = 278 nm) y quercetina (λ = 375 nm) [402].

2.3.2 Estrategia II-A y II-B. Búsqueda de flavonoides mediante HPLC-ESI-MS/MS

Para este ensayo, se utilizaron los exudados radiculares colectados después de 16 días de incubación.

2.3.2.1 Extracción de moléculas presentes en los exudados radiculares

Los liofilizados fueron macerados con mortero en un mínimo volumen de metanol calidad HPLC y con el fin de mejorar la extracción, aproximadamente 5 mL de cada muestra fueron colocadas en baño de ultrasonido durante 10 min, filtradas en algodón y el volumen colectado fue traspasado a tubos limpios.

Las muestras fueron concentradas en vacío en un rotavapor a 40 °C durante 2 h. Los metabolitos presentes en los exudados radiculares, se sometieron a una etapa de extracción en fase sólida (SPE) usando una columna C18 (500mg). Previamente las columnas fueron activadas de la siguiente forma: con 1 mL de metanol 100% y con 1 mL de agua MilliQ 3 veces consecutivas. Se adicionó el extracto disuelto en agua en los cartuchos y destinadas a fraccionamiento secuencial en agua 100%, metanol 50%, metanol 100% y acetona 100% para su elución [403].

Las fracciones fueron secadas mediante aeración, se determinaron sus pesos y se mantuvieron en frío hasta su análisis. Las muestras fueron disueltas en acetonitrilo a una concentración de 1mg/mL y fueron analizadas en primer lugar por HPLC-DAD-ELSD y aquellas muestras que presentaron perfiles cromatográficos con picos únicos y reproducibles fueron luego reanalizadas por HPLC-ESI-MS/MS. Además de las muestras de los exudados radiculares en presencia o ausencia de bacterias, se analizaron los extractos del pellet celular, del sobrenadante del cultivo y del medio de cultivo sin inocular. Se realizaron cuadruplicados biológicos. Las muestras se denominaron con las siguientes abreviaciones: ERMP MeOH 50% o ERMP MeOH 100% si eran exudados radiculares de *M. pudica* extraídos en metanol 50% o 100%, respectivamente y ERMP-UYPR2.512 MeOH% 50 o ERMP-UYPR2.512 MeOH 100% si eran exudados radiculares de *M. pudica* obtenidos en presencia de UYPR2.512 y extraídos en metanol 50% o 100%, respectivamente.

2.3.2.2 Obtención y fraccionamiento de extractos celulares

Se partió de un cultivo de UYPR2.512 crecido en 200 mL de medio líquido M9 1X suplementado con citrato 14 mM. Como control se utilizó el medio de cultivo sin inocular. A partir del cultivo primario se realizó una dilución 1/100 en medio fresco y se incubó a 30 °C, con agitación mecánica a 150 rpm durante 7 días. Finalizado el tiempo, el cultivo se centrifugó 4.000 rpm durante 10 min, se traspasó el sobrenadante a una bola de decantación, y se destinó para la partición con acetato de etilo. Para ello, se le adicionó 250 mL de acetato de etilo, se realizaron 3 lavados consecutivos, y se colectó la fracción orgánica en busca de los compuestos de interés. La fase orgánica se rota-evaporó a 164 mbar hasta evaporación total del solvente, se recogió en el mínimo volumen de una mezcla de acetato de etilo y agua (1:1) y se concentró hasta sequedad. El extracto fue colectado en acetato de etilo, metanol 100% y agua, se concentró nuevamente a vacío en *speedvac*, se resuspendió en 1mL de agua y se fraccionó en columna SPE (C18), al igual que en el punto anterior. Las fracciones de los distintos solventes fueron colectadas y concentradas.

El pellet celular se resuspendió en 3 ml de metanol 100%, se colocó en baño ultrasónico durante 10 min y se centrifugó durante 10 min a 5000 rpm. El solvente se evaporó hasta

sequedad, se resuspendió en 1 mL de agua y se fraccionó en columna SPE C18 al igual que en el punto anterior. Las muestras se eluyeron con 1 mL de metanol/agua 9:1 (v/v) y se concentraron hasta sequedad. El medio de cultivo sin inocular se analizó directamente en el HPLC-MS/MS, sin realizar las etapas de fraccionamiento. Las muestras fueron almacenadas a 4°C hasta ser analizadas.

2.3.2.3 Análisis cromatográficos cualitativo de metabolitos

Todas las muestras fueron disueltas en 50% metanol a una concentración de 1 mg/mL y se analizaron mediante HPLC-MS/MS en un sistema Shimadzu UFLC (dos bombas de disolvente LC20AD, horno de columna CTO20A ajustado a 40 °C, muestreador automático SIL20A y controlador CBM20A) acoplado a un espectrómetro de masas con un analizador de captura de electrones con trampa de iones (AmaZon SL), ubicados en la Facultad de Ciencias Farmacéuticas de Riberáo Preto, Brasil. Un volumen de 20 µL de cada muestra se pasó por una columna Phenomenex C18 (250 mm x 4.6 mm, partículas de 5 µm) y se eluyeron con un flujo de 1 mL/min de una mezcla de agua (disolvente A) y acetonitrilo (disolvente B), ambos con un 0,1% de ácido fórmico. El gradiente utilizado se realizó de la siguiente manera: 0-5 min (10% B), 5-60 min (10% a 100% B), 60-70 min (100% B), 70-75 min (100% a 10% B) y 75-85 min (10% B). El espectrómetro de masas se hizo funcionar en modo de ionización positiva con los siguientes parámetros: capilar de 3.5 kV, desplazamiento de placa final 500V, nebulizador de 60 psi, gas seco (N₂) con un caudal de 10 mL/min, temperatura de secado de 330 °C, auto MS/MS adquiriendo datos entre m/z 50 y 1300, promedio de 3 espectros, modo UltraScan para MS/MS, adquisición de tasa espectral de 3 espectros/s, exclusión de un ion dado después de 3 espectros y liberado después de 30s. El espectrómetro de masas fue controlado por el software *Hystar* (Bruker Daltonics Inc., EE. UU.).

2.3.2.4 Procesamiento de datos y análisis de los perfiles cromatográficos

Los datos de HPLC-MS/MS de los modos de ionización positiva y negativa, fueron convertidos al formato.mzXML, mediante la utilización del software *MSconvert* (Proteowizard Software Foundation, Palo Alto, CA, EE. UU.) y exportados como tabla .csv y como archivo .mgf, para ser analizados posteriormente.

Los perfiles cromatográficos de fragmentación fueron analizados en comparación con las moléculas anotadas en la librería del GNPS (*Global Natural Products Source*) [387]. La red molecular GNPS organiza los compuestos según sus características de espectrometría de masas en tándem (MS/MS) en grupos y compara los espectros con bibliotecas espectrales. A partir de las masas obtenidas del HPLC-MS/MS se generó una red molecular, la cual fue analizada mediante el programa *Cytoscape* versión 3.8 [389]. Los análisis se realizaron en base a parámetros previamente utilizados en GNPS por diferentes grupos de investigación [404].

2.3.3 Estrategia III. Búsqueda de metabolitos en exudados radiculares mediante GC/MS

2.3.3.1 Obtención de exudados radiculares

Se realizó el mismo procedimiento que en el punto 2.3 de Materiales y Métodos de este capítulo. En esta ocasión los frascos contenían 40 mL de agua estéril, se colocaron 40 semillas por frasco y fueron incubados durante 10 días. Finalizado este tiempo, los exudados fueron filtrados y liofilizados.

2.3.3.2 Preparación de extractos a partir de los exudados radiculares colectados

Los exudados liofilizados colectados en el punto 2.3.3.1 fueron resuspendidos en 5 mL de metanol anhidro 100%, la re-suspensión fue homogenizada mediante vortex e incubada a 4 °C ON. Finalizado este tiempo, la suspensión se centrifugó durante 2 min a 14.000 rpm. Se obtuvo un pellet insoluble en metanol anhidro 100%, y una fracción metanólica de 3,45 mL. Las fracciones metanólicas (de entre 200-500 µL) se concentraron 4°C a vacío y se resuspendieron en 80 µL de metoxiamina hidroxidloruro (20mg/mL) disuelto en piridina y se incubaron durante 90 min a 40 °C. Luego se añadieron 80 µL del derivatizante N-metil-N-(trimetilsilil) trifluoroacetamida (MSTFA, 15222 - Sigma-Aldrich). La reacción se llevó a cabo durante 30 minutos a 40 °C. Después de este tiempo, las muestras estuvieron listas para su inyección en el GC-MS.

El análisis cromatográfico se llevó a cabo en el equipo GCMS-QP 2010 ULTRA (Shimadzu), equipado con un auto-inyector (AOC-20i, Shimadzu), una pre-columna GC *Guard column* (0,25 μm ×5 m, RESTEK cat#1000), una columna HP-5MS (0,25 μm ×30 m, Agilent, 19091SS - 433) y un flujo de gas Helio de 28,9 mL por min, ubicado en la plataforma de química analítica del IIBCE. La temperatura inicial del horno fue de 80 °C, mantenida por 2 min, se incrementó luego a 290 °C (8 °C/min), se mantuvo por 3 min y finalmente se incrementó a 310 °C (20 °C/min) y se mantuvo por 10 min adicionales. La temperatura del inyector y detector fue de 300 °C. El volumen de inyección fue de 1 μL y el modo de inyección fue el *splitless*. Para la detección se empleó la ionización por impacto electrónico a 70 eV, y se escanearon los m/z entre 50-850.

2.3.3.3 Identificación cualitativa de metabolitos

Para la identificación de metabolitos, se analizó la presencia de los iones totales (TIC), el cual permite obtener un espectro de masa asociado para cada molécula. Los valores de las masas obtenidos se compararon con la base de datos de espectros de masa presente en la librería del equipo (*NIST and Wiley libraries*). Como criterio de comparación de espectros entre los picos obtenidos y las bases de datos utilizadas, se utilizó una similitud > 90 e índices de retención con una tolerancia de +/- 10. A su vez, los valores de las áreas de cada pico se normalizaron con respecto al ribitol (estándar interno), el cual fue adicionado a todas las muestras.

Los cromatogramas se analizaron con el software GC-MS Solutions de Shimadzu. Para la confirmación de algunos de los metabolitos identificados se utilizaron estándares de aminoácidos, para lo cual se empleó una mezcla de aminoácidos calidad estándar analítico (Sigma-Aldrich; AAS18). A su vez, se utilizaron estándares de azúcares (sacarosa, fructosa, D(+) glucosa, L(-) arabinosa), de polialcoholes (mio-inositol, xilitol, D(-), xilosa, L(-) arabitol), ácidos orgánicos (ácido glucónico), ácidos carboxílicos (ácido oxálico) de concentraciones conocidas y como estándar interno se utilizó ribitol 2mg/mL, el cual se adicionó a cada una de las muestras. Los estándares fueron disueltos en agua, concentrados por evaporación a baja presión en un concentrador *speedvac* y derivatizados con la misma metodología utilizada para los extractos de los exudados radiculares. La concentración se normalizó en función del estándar interno.

2.3.3.4 Estudio del efecto inductor de exudados radiculares de *M. pudica* en URPR2.512

Se empleó la misma metodología descrita en el punto 2.3 de Materiales y Métodos. En esta ocasión, los cultivos fueron inducidos en presencia de medio mínimo M9 suplementado con la fracción metanólica al 4% (v/v). Como control positivo, se utilizó luteolina 15 μ M.

2.4 Estudio del efecto inductor de flavonoides y otros compuestos fenólicos en URPR2.512

Se utilizó la misma metodología utilizada en el punto 2.3. Los cultivos fueron inducidos en presencia de flavonoides de concentración 15 μ M, que incluyeron daidzeína (isoflavononas), naringenina (flavanona), catequeina (flavanoles), miricetina, crisina, apigenina (flavonas) y quercetina (flavonoles), y compuestos fenólicos como el reverastrol (estibeno), el ácido curmárico (derivado de caído cinámico). En todos los casos los reactivos utilizados fueron de calidad estándar (98% de pureza, Sigma-Aldrich), disueltos en DMS o en agua, según correspondiese en cada caso. como control positivo se utilizó luteolina 15 μ M. Se realizaron triplicados técnicos.

3 Resultados

3.1 Los exudados radiculares de *M. pudica* inducen la expresión del *pnodB19424-lacZ* en UYPR2.512

Al utilizar exudados frescos de *M. pudica* en cultivos de UYPR2.512 se observó un aumento en la actividad de la β -galactosidasa de los cultivos de UYPR2.512 portando pCBM01 así como en la cepa control LMG19424 portando pCBM01. Estos resultados indican que los exudados obtenidos a partir de *M. pudica* contienen inductores de la expresión de los genes *nod* de *Cupriavidus*.

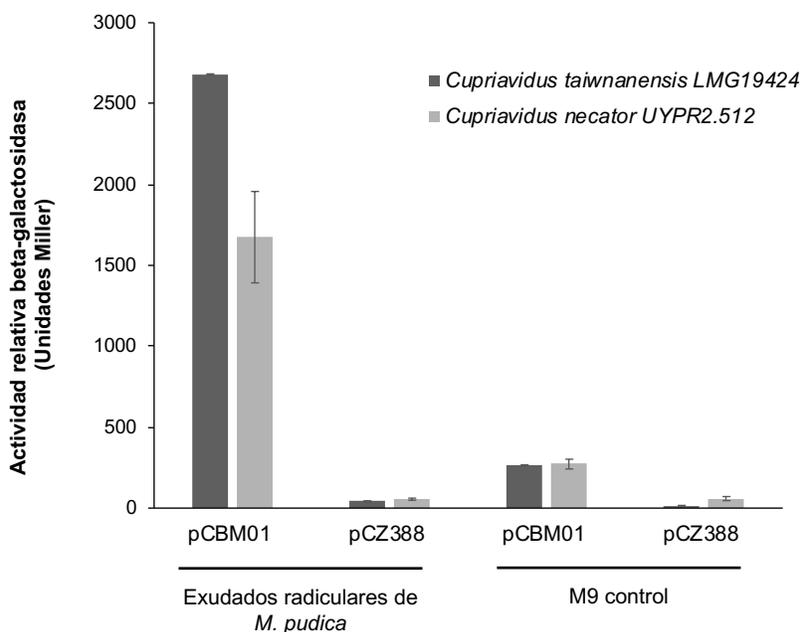


Figura 41. Respuesta de la región promotora del gen *nodB* a la presencia de exudados radiculares frescos en cultivos de UYPR2.512.

La determinación de la actividad β -galactosidasa se realizó en cultivos de UYPR2.512 portando los plásmidos pCBM01 o pCZ388 en medio mínimo M9 modificado, luego de un período de incubación durante 16 h a 30 °C. Como control se utilizó la cepa LMG19424 de *C. taiwanensis*. La expresión de *lacZ* se evaluó como actividad β -galactosidasa registrada como unidades Miller. Los exudados fueron colectados luego de 10 días de incubación. Los datos se expresan como la media de tres réplicas biológicas con sus errores estándar.

Con el afán de identificar las moléculas presentes en la interacción entre UYPR2.512 y su leguminosa hospedera *M. pudica*, se decidió utilizar un enfoque metabolómico basado en el estudio de las moléculas presentes en los exudados radiculares.

3.2 Búsqueda de moléculas involucradas en la interacción entre UYPR2.512-*M. pudica*

3.2.1 Estrategia I. Búsqueda de flavonoides mediante HPLC-DAD

La primera estrategia consistió en analizar los metabolitos presentes en los exudados radiculares de *M. pudica* en presencia del simbionte UYPR2.512, mediante una aproximación metabolómica dirigida. En particular, se buscó evidenciar la presencia de flavonoides. Al realizar un espectro de barrido de absorción de los exudados (de 230 a

400 nm), se observó solamente uno de los 2 picos máximos característicos de los flavonoides, en el entorno de los 250 nm, el cual correspondería a la fracción cinamoilo (254 nm). Luego mediante HPLC-DAD, se observó la presencia de compuestos con picos máximos de absorción en el entorno de 250 nm y 260 nm (Figura 42), los cuales podrían corresponderse al ácido ferúlico y al ácido gálico, ambos precursores de flavonoide [405], [406].

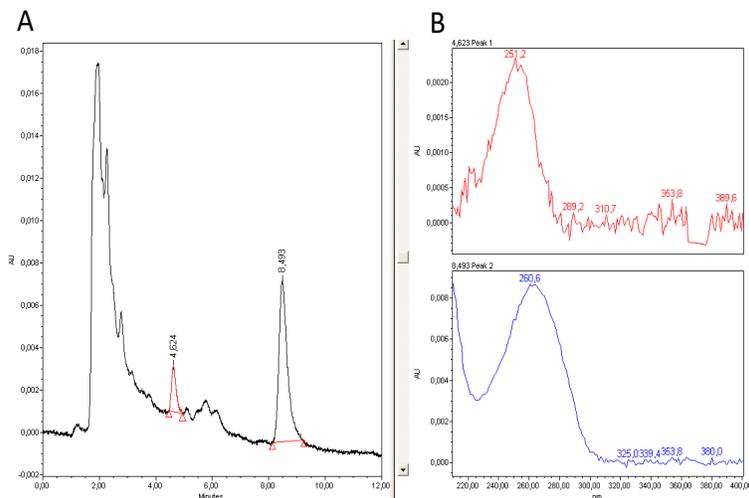


Figura 42. Perfil cromatográfico de exudados radiculares de plántulas de *M. pudica* inoculadas con UYPR2.512.

En A: Cromatograma HPLC-DAD (Unidades de absorbancia vs RT (tiempo de retención, min)), y B: Espectros ultravioletas de cada pico individual encontrado. Se observa un pico de absorbancia máximo en el entorno de 250 nm. Se indica el perfil de una muestra representativa.

En base a estos resultados se repitió el experimento con un mayor número de plantas (40 plantas), y aumentando el tiempo de interacción entre los simbiosntes a 20 días. El análisis por HPLC-DAD mostró un cromatograma con mayor cantidad de picos a la longitud de onda de 254 nm, región del espectro de absorción ultravioleta en la que absorben los compuestos aromáticos, como los presentes en los polifenoles esperados en las muestras procesadas (Figura 43 y 44). A su vez, se observó una mayor cantidad de picos en presencia de la bacteria, en comparación con el control (Figura 43 y 44).

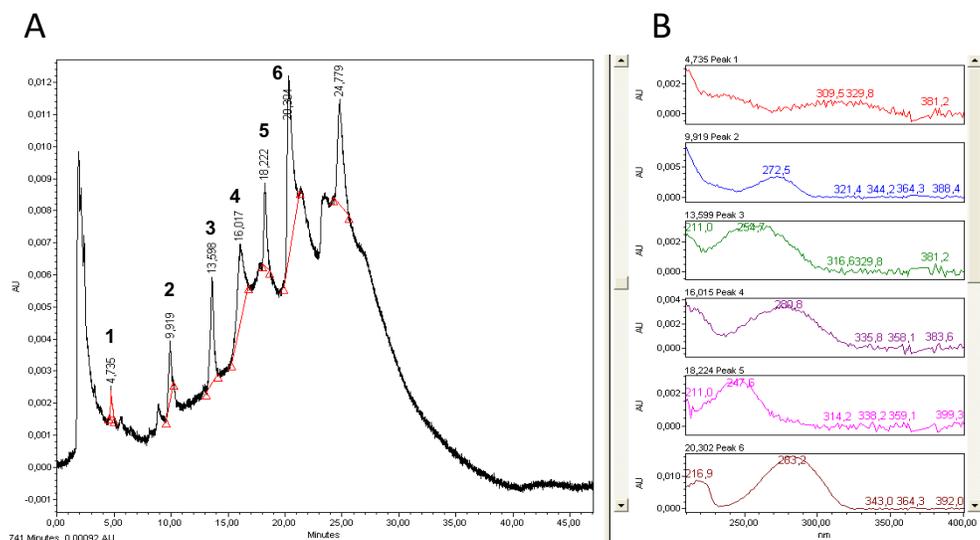


Figura 43. Perfil cromatográfico de exudados radiculares de plántulas de *M. pudica* inoculadas con UYPR2.512.

Se indica en A: Cromatograma HPLC-DAD C18 de los exudados radiculares, en modo gradiente utilizando como fase móvil ácido fosfórico 0,5% pH 2 y metanol 90%, registrado a 254 nm y en B: De arriba hacia abajo se indican los espectros ultravioletas de 6 picos individuales observados en el cromatograma en A,

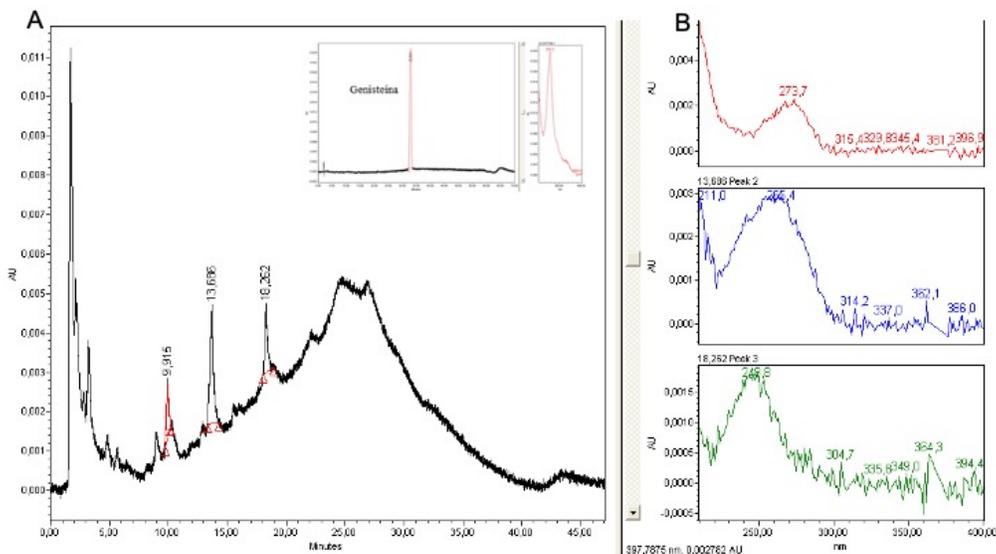


Figura 44. Perfil cromatográfico de exudados radiculares de plántulas de *M. pudica*.

Se indica en A: Cromatograma HPLC-DAD C18 de la muestra control, en modo gradiente utilizando como fase móvil ácido fosfórico 0,5% pH 2 y metanol 90%, registrado a 254 nm y en B: Espectros ultravioleta de los picos encontrados. En el panel A se incluye el perfil cromatográfico del flavonoide genisteína ($\lambda = 259$ nm). El estudio comparativo indicaría un pico muy similar a dicho flavonoide.

Como se muestra en ambas figuras, se observó la presencia de un pico con características muy similares al de la genisteína en el entorno de los 260 nm, un flavonoide inductor de los genes *nod* de rizobios [407]. Sin embargo, no coincidió su tiempo de retención, con el observado en la muestra analizada. No fue posible identificar el resto de los picos correspondientes, ya que los mismos tampoco coincidieron con los tiempos de retención, ni el espectro con los de los estándares de flavonoides comerciales disponibles (Fig. 45).

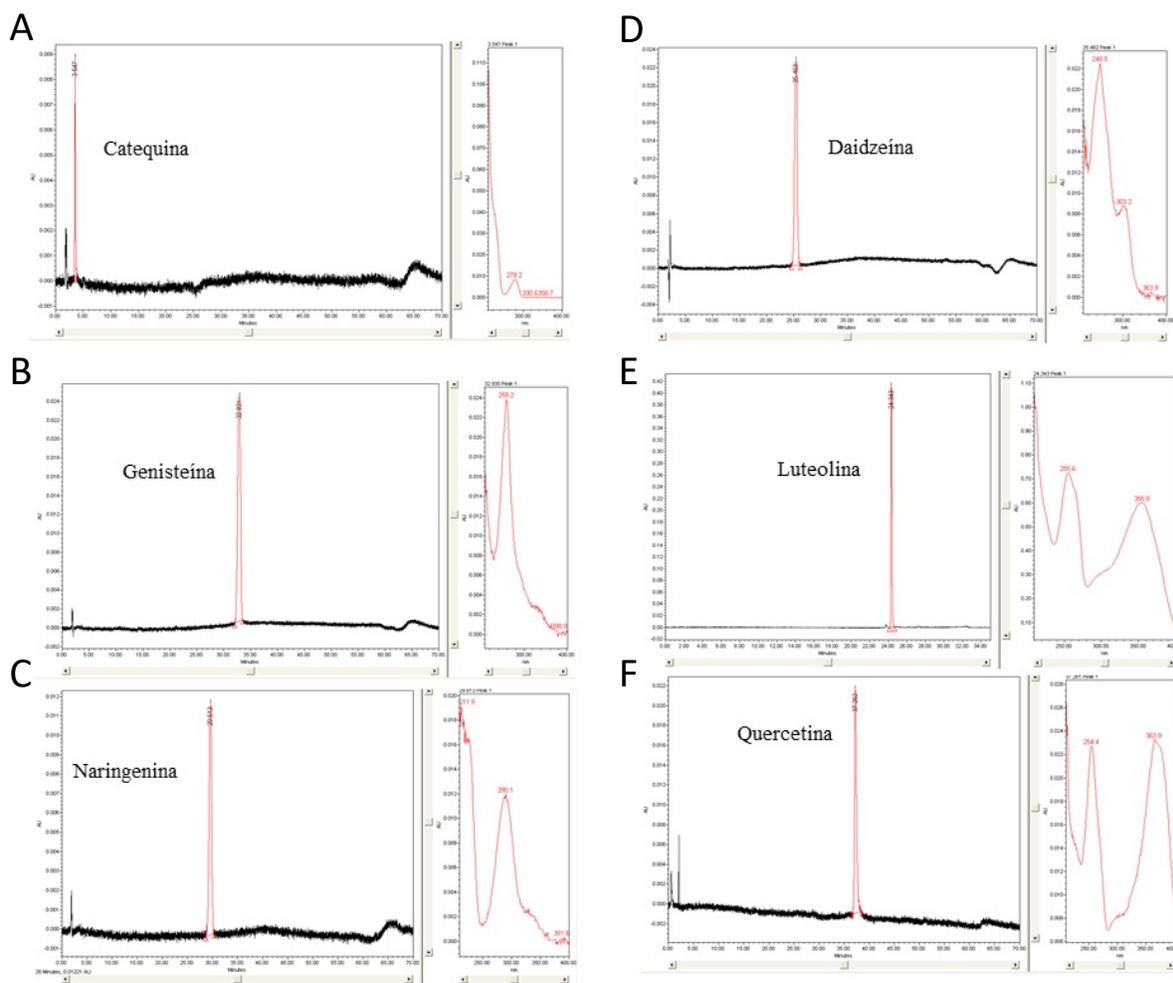


Figura 45. Perfiles cromatográficos de flavonoides utilizados como estándares.

A la izquierda se indican los cromatogramas HPLC-DAD C18 y a la derecha los espectros UV de su respectivo pico, para catequina ($\lambda = 278$ nm), daidzeína ($\lambda = 248$ nm), genisteína ($\lambda = 259$ nm), luteolina ($\lambda = 375$ nm), naringenina ($\lambda = 290$ nm), y quercetina ($\lambda = 375$ nm).

Con el objetivo de profundizar en la identificación de los compuestos presentes en los exudados radiculares decidimos realizar un nuevo análisis mediante HPLC-UV, HPLC-DAD-ELSD y HPLC-ESI-MS/MS, en el contexto de una pasantía realizada en la Facultad de Ciencia Farmacéutica de la Universidad de Sao Paulo, Ribeirao Preto, Brasil, bajo la dirección de la Dra. Mónica Puppo y financiada por la ANII.

3.2.2 Estrategia II-A. Espectros obtenidos mediante HPLC-UV y HPLC-DAD-ELSD

Con la finalidad de evidenciar la presencia de flavonoides en los extractos provenientes de exudados radiculares en presencia o ausencia de la bacteria, se realizó un primer análisis exploratorio mediante HPLC-UV. Tal como había ocurrido anteriormente (estrategia I), no se observó ninguna de las 2 bandas características de los flavonoides en el UV-Vis. Sin embargo, al comparar los espectros de los exudados radiculares en presencia de la bacteria (ERMP-5V12), con el correspondiente exudado sin la bacteria (ERMP), se observaron picos diferenciales (Figura 46, panel A y B).

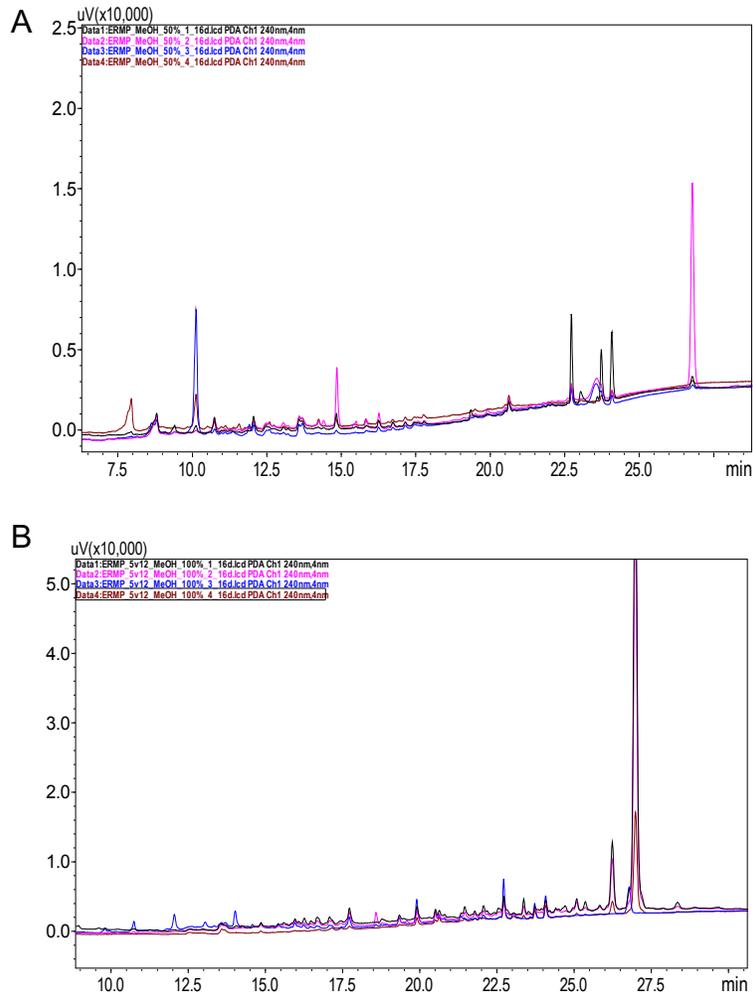


Figura 46. Perfiles cromatográficos obtenidos mediante HPLC-DAD-ELSD.

Se indica en A: Perfiles cromatográficos de la fracción metanólica de ERMP MEOH 50% y en B: Perfiles cromatográficos de la fracción metanólica de ERMP-UYPR2.512 MEOH 100%, incubadas durante 16 días. Perfiles cromatográficos registrados a 240 nm. En cada panel se muestran 4 cromatogramas en distintos colores, correspondientes a los cuatriplicados biológicos. Para el análisis se consideraron aquellas muestras que presentaron picos únicos y reproducibles en los cuatriplicados biológicos.

A partir de estos datos, se seleccionaron las muestras ERMP-UYPR2.512 MeOH 100% y ERMP MeOH 50% como control, por contener picos diferenciales, para su posterior análisis mediante HPLC-ESI-MS/MS. A su vez, se incluyeron extractos del pellet celular, del sobrenadante del cultivo bacteriano, del sobrenadante del medio de cultivo y del solvente utilizado.

3.2.3 Estrategia II-B. Espectros obtenidos mediante HPLC-ESI-MS/MS

A partir de los espectros de HPLC-ESI-MS/MS de los extractos de los exudados de *M. pudica* y de los exudados de *M. pudica* en presencia de UYPR2.512, se detectaron 27 picos, los cuales corresponderían a compuestos únicos en cada condición (Fig. 47C).

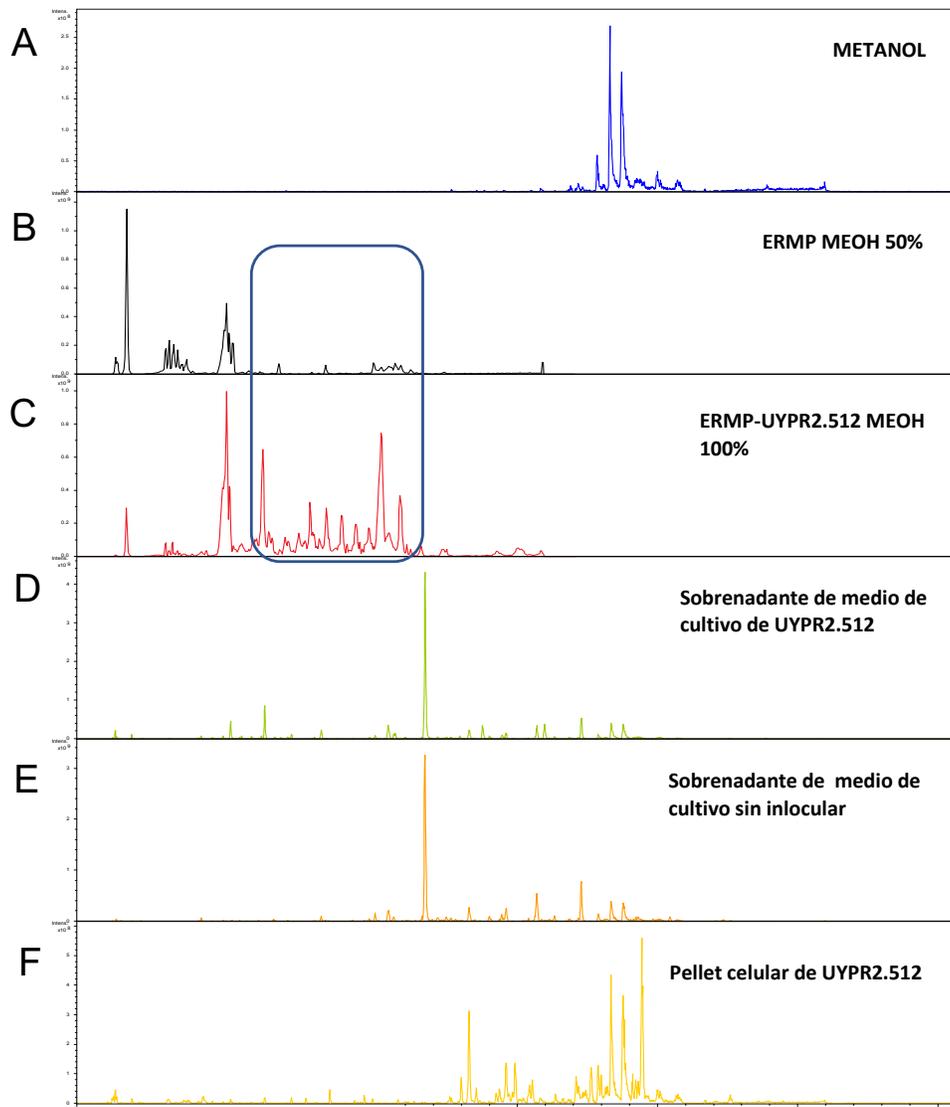


Figura 47. Cromatogramas de los extractos de exudados radiculares obtenidos mediante HPLC-ESI-MS/MS.

Se muestran los siguientes perfiles: A: Metanol (solvente), B: ERMP MEOH 50% (exudado de la planta en metanol 50%), C: ERMP-UYPR2.512 MEOH 100% (exudado de la planta en presencia de la bacteria en

metanol 100%), D: sobrenadante de medio de cultivo de UYPR2.512, E: sobrenadante del medio de cultivo sin inocular con UYPR2.512 y F: pellet celular de UYPR2.512. Se muestra una corrida representativa de cada extracto. El recuadro azul resalta la presencia de picos correspondientes a compuestos expresados en cada uno de los extractos. En eje de las ordenadas se indica intensidad de fluorescencia y en eje de las abscisas el tiempo de retención.

El análisis detallado de ambos extractos (Figura 48), mostró la presencia de 9 picos encontrados únicamente en presencia de la bacteria, lo que sugiere que estos corresponderían a compuestos que podrían estar implicados en la interacción simbiótica.

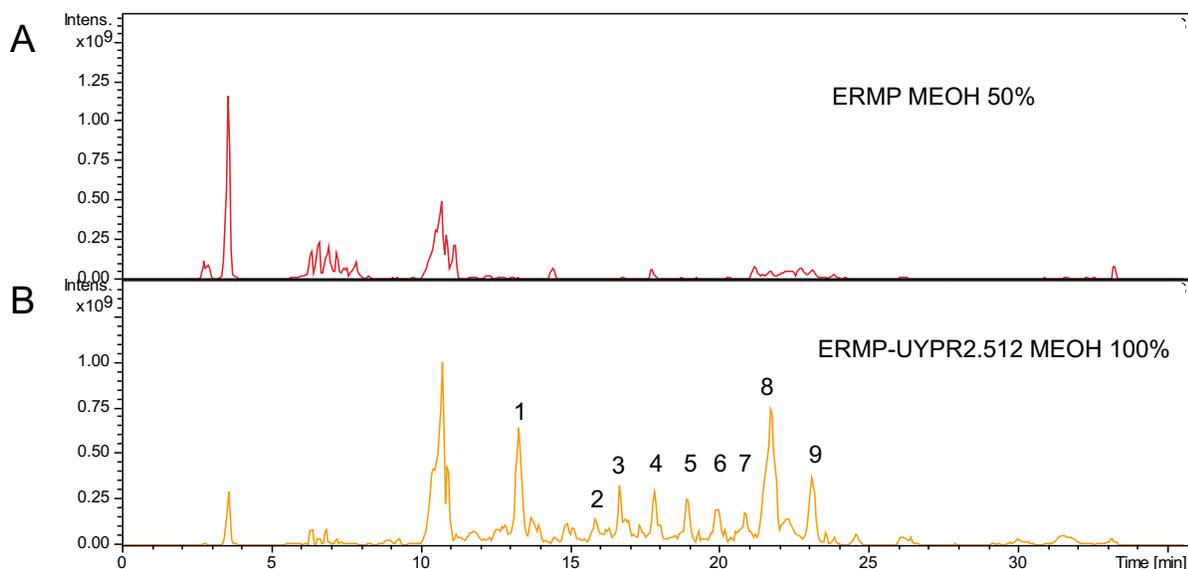


Figura 48. Perfiles obtenidos mediante HPLC-ESI-MS/MS de los extractos de exudados radiculares de *M. pudica*.

Se muestran los datos de una muestra representativa de cada extracto. Se indica en A: el exudado de ERMP MEOH 50% y en B: ERMP-UYPR2.512 MEOH 100%. Los números indican los picos correspondientes a compuestos únicos presentes en los exudados radiculares de la planta en presencia de la bacteria. En eje de las ordenadas se indica intensidad de fluorescencia y en eje de las abscisas el tiempo de retención.

3.3 Caracterización de los picos observados

Con la finalidad de caracterizar los picos diferenciales observados, los patrones de espectrometría de masas obtenidos fueron analizados de acuerdo con su similitud,

dando origen a grupos de espectros (*cluster*), partiendo de la premisa de que moléculas con estructuras similares producen patrones de fragmentos similares.

La búsqueda de similitudes en la base de datos de GNPS permitió identificar solamente un clúster (Fig. 49, recuadro azul) el cual tendría nodos que corresponderían al flavonoide miricetina-3-rutinoso.

Por otro lado, la gran cantidad de singletones observadas sería indicativo de la presencia de moléculas únicas, reflejando la compleja composición química de las muestras analizadas.

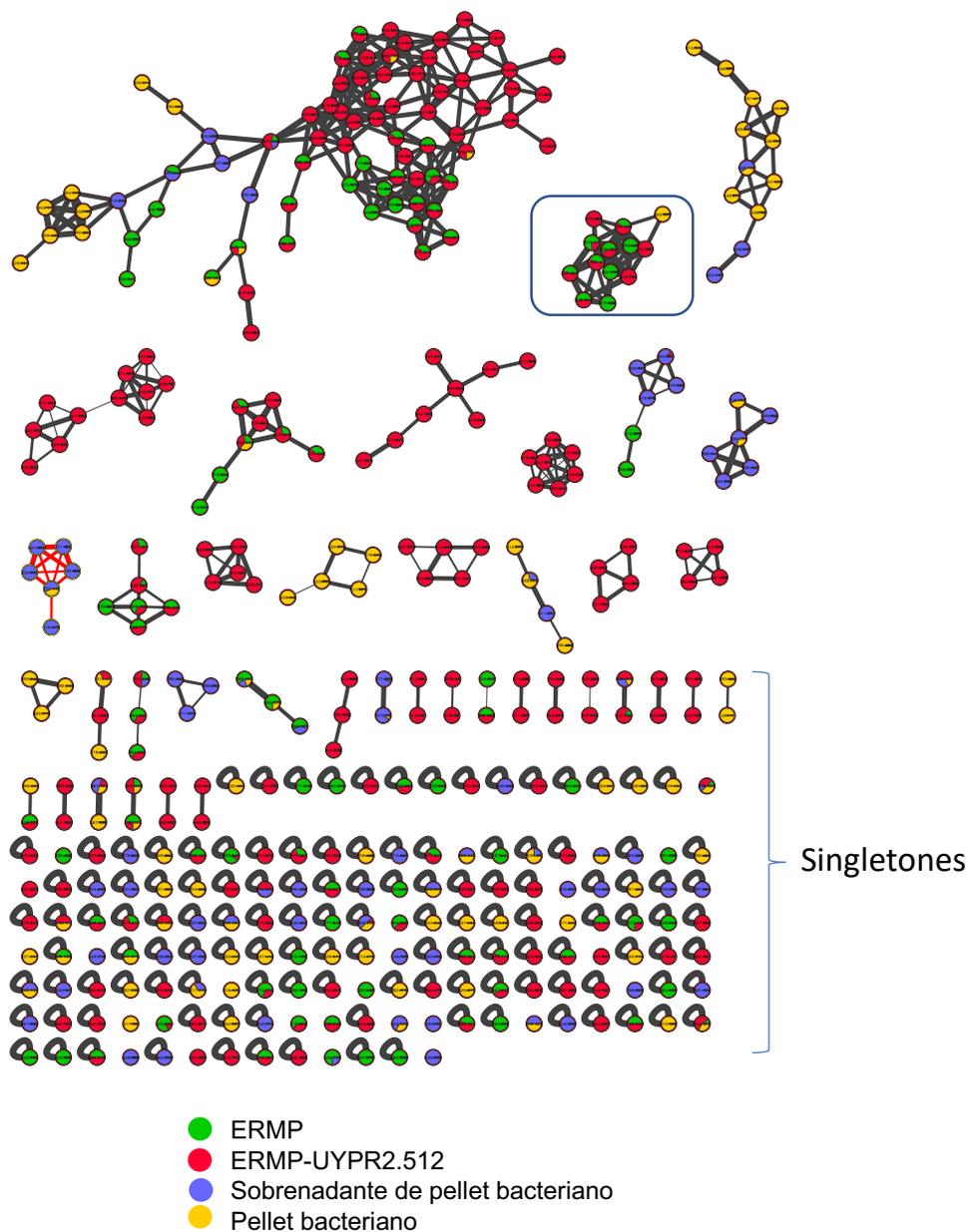


Figura 49. Red molecular generada a partir de los espectros obtenidos mediante HPLC-ESI-MS/MS.

Los nodos se conectan mediante líneas negras, en función a la similitud espectral. Cada nodo indica la masa (m/z) del compuesto, y los colores indican el origen de ese compuesto. En el recuadro azul se resalta un agrupamiento (*cluster*) de interés en el cual de acuerdo a la base comparativa del GNPS presentaría dos nodos anotados como correspondientes al flavonoide miricetina-3-rutinoso. El paréntesis azul indica los singletones. Imagen obtenida mediante la utilización del programa Cytoscape.

Al realizar un análisis detallado manual, en el cual se comparó los valores de m/z de las moléculas obtenidos mediante HPLC MS/MS ($m/z = 626.635$ y $m/z = 626.57$) y los

valores de m/z de dichas moléculas anotados como miricetina-3-rutinosa por el GNPS ($m/z = 627.15$), se observó que las diferencias de las masas resultaron ser mayores a 0,5 Da, por lo cual no fue posible confirmar que se tratara de este flavonoide.

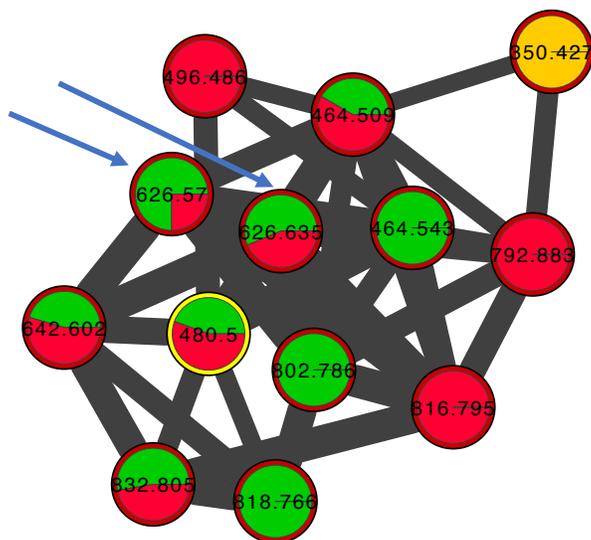


Figura 50. Clúster de interés con nodos anotados como correspondientes a flavonoides según la librería del GNPS.

Los porcentajes dentro de los nodos indican la masa molecular masa/carga (m/z) de los compuestos identificados y los colores indican el origen de ese compuesto. En color verde el compuesto que proviene de ERMP, en color rojo de ERMP-UYPR2.512 y en amarillo de pellet bacteriano. Las flechas azules indican los nodos anotados como miricetina 3-rutinosa en el GNPS.

Por otro lado, con el fin de confirmar que realmente los nodos no correspondían al flavonoide miricetina 3-rutinosa, un análisis mediante HPLC de alta resolución realizado en la Facultad de Ciencia Farmacéutica de la Universidad de Sao Paulo, confirmó que esas moléculas no corresponderían a flavonoides. Por lo cual la miricetina-3-rutinosa estaría mal anotada en la base de datos del GNPS.

A su vez, como se observa en la figura 49, existen varios compuestos que fueron producidos únicamente durante la interacción planta-bacteria (nodos de color rojo). Sin embargo, debido que las moléculas no fueron anotadas por el GNPS, se sugiere que se trataría de compuestos nuevos o no descritos anteriormente y sería necesario aislarlos para determinar su estructura química

3.4 Estrategia III. Identificación de metabolitos presentes en exudados radiculares mediante GC/MS

Debido a que no se logró evidenciar la presencia de flavonoides en los exudados radiculares de *M. pudica*, se analizó su composición usando una estrategia de metabolómica no dirigida. Mediante análisis por GC-MS se identificaron entre 255-369 picos en los ERMP y entre 288-364 en los ERMP-UYPR2.512, correspondientes a posibles metabolitos presentes en las muestras analizadas (Tabla 25 y Fig. 50).

Tabla 25. Cantidad de metabolitos identificados mediante GC-MS para cada una de las de las fracciones metanólicas analizadas

Las siglas ERMP corresponde a exudados radiculares de *M. pudica*, y ERMP-UYPR2.512 corresponde a exudados radiculares de *M. pudica* en presencia de UYPR2.512.

Muestra	Nº de picos detectados
ERMP (1)	255
ERMP (2)	369
ERMP (3)	340
ERMP (4)	323
ERMP-UYPR 2.512 (1)	288
ERMP-UYPR2.512 (2)	364
ERMP-UYPR 2.512 (3)	306
ERMP-UYPR2.512 (4)	328

Los perfiles cromatográficos correspondientes a los cuadruplicados biológicos de los exudados obtenidos en ausencia o en presencia de UYPR2.512 presentaron variaciones tanto en el número de picos detectados como en su definición (Fig. 51).

Con la finalidad de asignar un identificador a los picos obtenidos mediante GC/MS, se estableció como criterio del análisis la asignación de picos a metabolitos cuya similitud con la base de datos fuera superior al 90%, con un índice de retención (IR) $< 0,1$ y que estuvieran presentes en al menos 1 de las 4 muestras analizadas para cada condición.

Siguiendo estos criterios, se identificaron un total de 68 metabolitos presentes en los exudados radiculares de *M. pudica* independientemente de la ausencia o presencia de UYPR2.512 (Tabla 26). De estos, 40 metabolitos se mantuvieron constante en presencia de UYPR2.512, 19 solo se encontraron en ausencia de UYPR2.512 y 9 fueron específicos en presencia de UYPR2.512. Estos 9 metabolitos, que solamente se detectaron cuando ambos simbiontes están presentes, podrían estar involucrados en la interacción simbiótica. De acuerdo a su estructura química, estos metabolitos se clasificaron en 4 familias químicas; ácidos carboxílicos (5 metabolitos), ácidos aromáticos (2 metabolitos), aminoácidos (1 metabolito) y polialcoholes (1 metabolito) (Tabla 26).

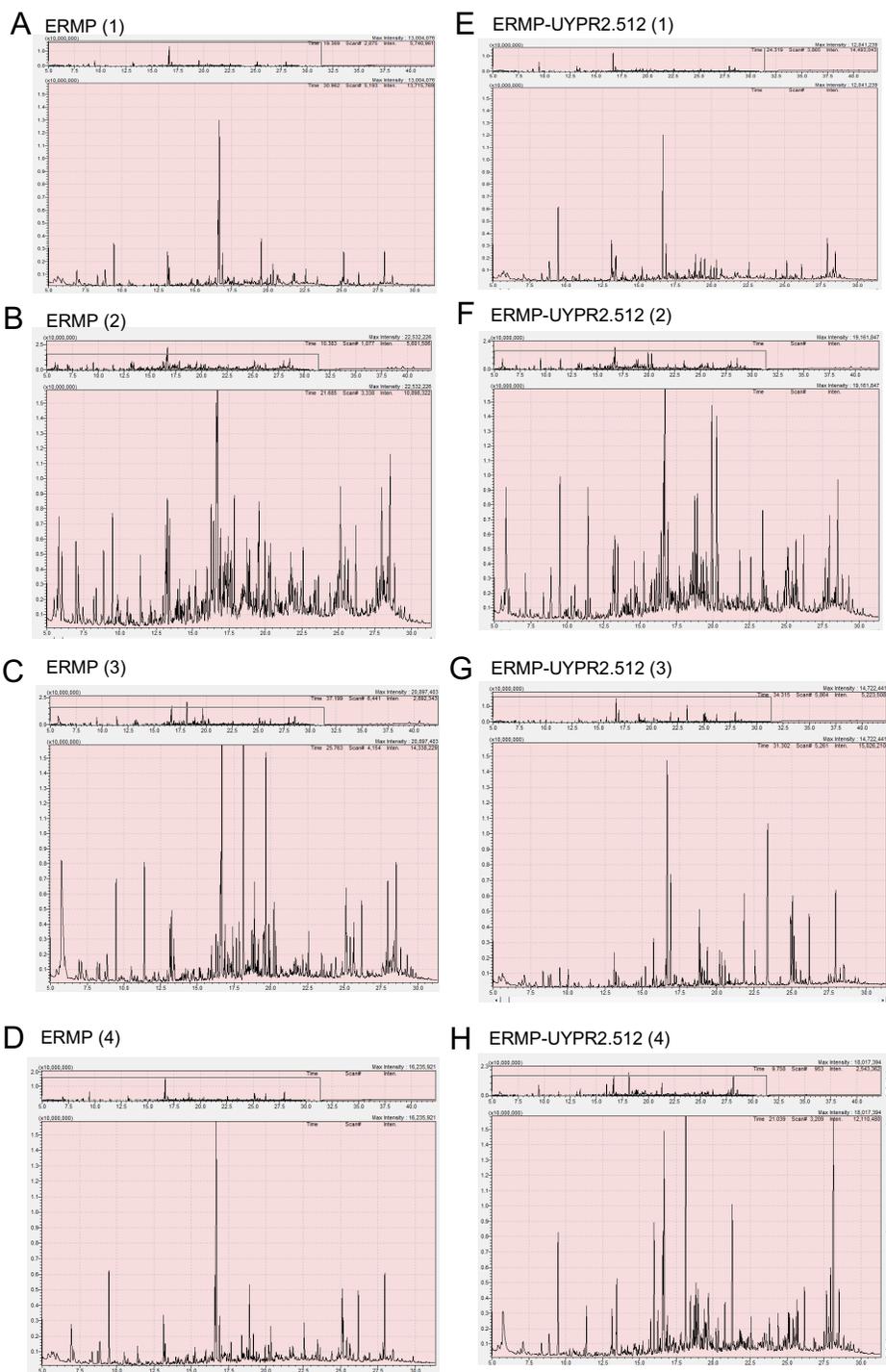


Figura 51. Cromatogramas de los exudados analizados mediante GC-MS.

A: exudados radiculares de *M. pudica* (ERMP) (1), B: ERMP (2), C: ERMP (3), D: ERMP (4), E: exudados radiculares de *M. pudica* en presencia de UYPR2.512 (ERMP-UYPR2.512) (1), F: ERMP-UYPR2.512 (2), G: ERMP-UYPR2.512 (3), H: ERMP-UYPR2.512 (4).

Tabla 26. Metabolitos detectados en exudados radiculares de *M. pudica* en ausencia y/o en presencia de UYPR2.512.

Los metabolitos resaltados en negrita fueron confirmados por estándares. Resaltados en color amarillo se indican los metabolitos presentes independientemente de la presencia de UYPR2.512, en color turquesa los metabolitos presentes únicamente en ausencia de UYPR2.512, y en color naranja los metabolitos detectados únicamente en presencia de UYPR2.512.

Familia Química	Metabolitos detectados en exudados radiculares de <i>M. pudica</i> en ausencia y en presencia de UYPR2.512 (Similitud >90 y IR<0,1)
Acidos alditólicos	D-Arabinonic acid, 2,3,5-tris-O-(trimethylsilyl)-,gamma-lactone (CAS) 2,3,5-TRI-O-TRIMETHYLSILYL-ARABINONO-1,4-LACTONE
Acidos aromáticos	Benzenepropanoic acid, alpha-[(trimethylsilyl)oxy]-, trimethylsilyl ester (CAS) BETA.-PHENYLLACTIC ACID-DITMS
Acidos aromáticos	Gluconic Acid/Galactonic acid
Acidos aromáticos	Galactaric Acid-HEXATMS
Acidos aromáticos	Benzoic acid, 4-[(trimethylsilyl)oxy]-, trimethylsilyl ester (CAS) PARA-HYDROXYBENZOIC-DITMS
Acidos aromáticos	Benzoic acid, 3-methoxy-4-[(trimethylsilyl)oxy]-, trimethylsilyl ester (CAS) BISTRIMETHYLSILYL VANILLIC ACID
Acidos carboxílicos	Propanoic acid, 2-[(trimethylsilyl)oxy]-, trimethylsilyl ester (CAS) LACTIC ACID-DITMS
Acidos carboxílicos	Acetic acid, [(trimethylsilyl)oxy]-, trimethylsilyl ester (CAS) GLYCOLIC ACID-DITMS
Acidos carboxílicos	Erythronic acid-.gamma.-lactone, bis-O-(trimethylsilyl)
Acidos carboxílicos	PARABANIC ACID, BIS-O-(TRIMETHYLSILYL)
Acidos carboxílicos	Heptanedioic acid, bis(trimethylsilyl) ester (CAS) BISTRIMETHYLSILYL PIMELIC ACID
Acidos carboxílicos	Xylonic acid, 2,3,5-tris-O-(trimethylsilyl)-, .gamma.-lactone, D- (CAS) 2,3,5-TRI-O-TRIMETHYLSILYL-XYLONO-1,4-LACTONE
Acidos carboxílicos	2-Piperidinecarboxylic acid, 1-(trimethylsilyl)-, trimethylsilyl ester (CAS) PIPECOLIC-DITMS
Acidos dicarboxílicos	Ethanedioic acid, bis(trimethylsilyl) ester (CAS) OXALIC ACID-DITMS
Acidos dicarboxílicos	Propanedioic acid, bis(trimethylsilyl) ester (CAS) BISTRIMETHYLSILYL MALONIC ACID
Acidos dicarboxílicos	Butanedioic acid, bis(trimethylsilyl) ester
Acidos dicarboxílicos	Fumaric acid
Acidos grasos	Bis-O-trimethylsilyl-palmitinic acid-glycerin-(1)-monoester
Acidos grasos	Hexadecanoic acid, trimethylsilyl ester (CAS) PALMITIC ACID-MONOTMS
Ácidos grasos	Octadecanoic acid, trimethylsilyl ester
Acidos orgánicos	Lactic acid dimer, bis bis(trimethylsilyl)-
Acidos orgánicos	Butanedioic acid, [(trimethylsilyl)oxy]-, bis(trimethylsilyl) ester (CAS) MALIC ACID 3TMS
Acidos orgánicos	Citric acid
Acidos orgánicos	Gluconic Acid, lactone
Acidos orgánicos	Tartaric acid
Alcoholes	1,2,3-butantriol, tris-O-(trimethylsilyl)"
Aldehídos	Butanal, 2,3,4-tris[(trimethylsilyl)oxy]-, (R*,R*)- (CAS) ERYTHROSE-TRITMS

Familia Química	Metabolitos detectados en exudados radiculares de <i>M. pudica</i> en ausencia y en presencia de UYPR2.512 (Similitud >90 y IR<0,1)
Alditoles	Pentitol, 1-desoxy-tetrakis-O-(trimethylsilyl)
Alditoles	D-erythro-Pentitol, 2-deoxy-1,3,4,5-tetrakis-O-(trimethylsilyl)- (CAS)
Amidas	Urea, N,N'-bis(trimethylsilyl)- (CAS) Bis(trimethylsilyl)urea
Amidas	9-Octadecenamide, (Z)- (CAS) OLEOAMIDE
Aminoácidos	Threonine
Aminoácidos	Valine
Aminoácidos	Glycine
Aminoácidos	Serine
Aminoácidos	L-Proline, 5-oxo-1-(trimethylsilyl)-, trimethylsilyl ester (CAS) PYROGLUTAMIC ACID-DITMS
Aminoácidos	L-aspartic
Aminoácidos	Glutamine
Aminoácidos	Trimethylsilyl ester of N,N-bis(trimethylsilyl)-cysteine
Azúcares/polisacáridos	Pyrimidine, 2,4-bis[(trimethylsilyl)oxy]- (CAS) URACIL, O2,O4-BIS(TRIMETHYLSILYL)
Azúcares/polisacáridos	D-Glucopyranose, 4-O-[2,3,4,6-tetrakis-O-(trimethylsilyl)-.beta.-D-galactopyranosyl]-1,2,3,6-tetrakis-O-(trimethylsilyl)-
Azúcares/polisacáridos	D-xylose
Azúcares/polisacáridos	L-arabinose
Azúcares/polisacáridos	D-Galactose, 6-deoxy-2,3,4,5-tetrakis-O-(trimethylsilyl)-, O-methyloxime (CAS) TETRAKISTRIMETHYLSILYL ETHER DESOXY-6-GALACTOSE
Azúcares/polisacáridos	Fructose
Azúcares/polisacáridos	Sacarosa-OCTATAMS
Azúcares/polisacáridos	Cellobiose-OCTATAMS
Carbazoles	9H-Carbazole, 9-phenyl- (CAS) N-Phenylcarbazole
Esteres	Dipropylene glycol dibenzoate
Eteres	3,6,9-Trioxa-2,10-disilaundecane, 2,2,10,10-tetramethyl- (CAS) DIETHYLENE GLYCOL BISTRIMETHYLSILYL ETHER
Eteres	3,6,9,12-Tetraoxa-2,13-disilatetradecane, 2,2,13,13-tetramethyl- (CAS) TRIETHYLENGLYCOL, BIS-O-(TRIMETHYLSILYL)
Eteres	3,8-Dioxa-2,9-disiladecane, 2,2,9,9-tetramethyl-5,6-bis[(trimethylsilyl)oxy]-, (R*,S*)- (CAS) Erythritol, 1,2,3,4-tetrakis-O-(t
Eteres	3,8-Dioxa-2,9-disiladecane, 2,2,9,9-tetramethyl-
Eteres	2-Monostearin trimethylsilyl ether
Fosfato de silanol	Silanol, trimethyl-, phosphate (3:1)
Hidroxiácidos	Benzeneacetic acid, alpha.-[(trimethylsilyl)oxy]-, trimethylsilyl ester (CAS) BISTRIMETHYLSILYL MANDELIC ACID
No se logró identificar	Pentonic acid, 2-deoxy-3,4,5-tris-O-(trimethylsilyl)-, trimethylsilyl ester (CAS) 1-OXO-1,3,4,5-TETRAKIS(TRIMETHYLSILOXY)PENTAN
Pipecolinas	Ethyl 1-methylpipecolate
Piridinas	4-Hydroxypyridine, trimethylsilyl ether
Polialcoholes	3,7-Dioxa-2,8-disilanonane, 2,2,8,8-tetramethyl- (CAS) 1-3-PROPANEDIOL-DI-TMS

Familia Química	Metabolitos detectados en exudados radiculares de <i>M. pudica</i> en ausencia y en presencia de UYPR2.512 (Similitud >90 y IR<0,1)
Polialcoholes	Bis-O-(trimethylsilyl)-1,4-butandiol-1-D1"
Polialcoholes	Xylitol
Polialcoholes	L(-)arabitol
Polialcoholes	Galactitol, 1,2,3,4,5,6-hexakis-O-(trimethylsilyl)- (CAS) GALACTITOL-HEXATMS
Polialcoholes	Myo-inositol
Polialcoholes	Dulcitol, hexakis(trimethylsilyl) ether
Silanos	Trimethyl-[2-[2-[2-(2-trimethylsilyloxyethoxy)ethoxy]ethoxy]ethoxy]silane
Siliconas	Pentasiloxane, dodecamethyl- (CAS) Dodecamethylpentasiloxane

3.4.1 La fracción metanólica presenta metabolitos inductores del *pnodB19424-lacZ* en UYPR2.512

Después de 16 horas de incubación de los cultivos en presencia de la fracción metanólica de los exudados radiculares de *M. pudica*, se observó un aumento significativo en la actividad de la β -galactosidasa en UYPR2.512 portando pCBM01, en comparación con UYPR2.512 portando pCZ388. Además, la inducción de la expresión fue aún mayor en comparación con la inducción provocada por la luteolina (Fig. 52). lo cual indica que la fracción metanólica de los exudados radiculares aún contiene inductores de la expresión génica de los genes *nod* en UYPR2.512.

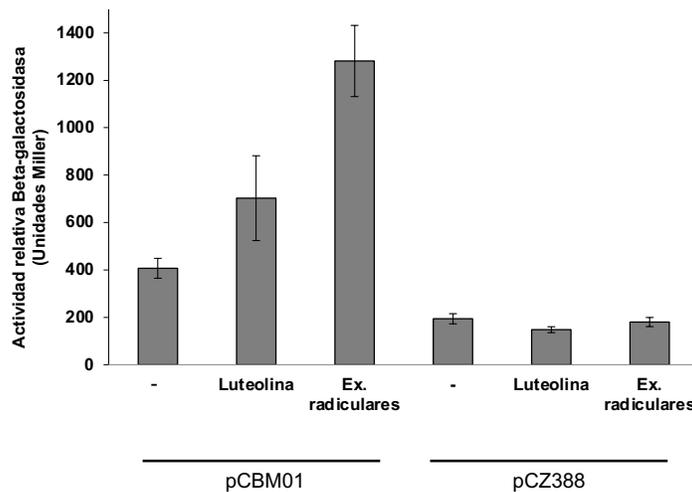


Figura 52. Respuesta de la región promotora del gen *nodB* a la presencia de exudados radiculares en cultivos de UYPR2.512.

La determinación de la actividad β -galactosidasa se realizó en cultivos de UYPR2.512 portando los plásmidos pCBM01 o pCZ388 en medio mínimo M9 modificado, luego de un período de incubación de 16h a 30 °C. Como control se utilizó luteolina 15 μ M. La expresión de *lacZ* se evaluó como actividad β -

galactosidasa registrada como unidades Miller. Los exudados fueron colectados luego de 10 días de incubación. La barra en el eje de las x corresponde los cultivos portando cada uno de los plásmidos sin inducir. Los datos se expresan como la media de tres réplicas biológicas con sus errores estándar.

3.4.2 Flavonoides y otras moléculas son inductores de los genes *nod* en cultivos de UYPR2.512

A partir de los resultados presentados en el punto 3.1, los cuales demostraron que los exudados presentaban moléculas responsables de la inducción de los genes *nod* y con el fin de conocer cuál es la naturaleza de esas moléculas, se estudió el efecto de flavonoides y otros polifenoles, reportados como inductores de genes *nod* en otros rizobios [116], [134], [135], [138], [146], [408]. Se estudió el efecto en la inducción de los cultivos de UYPR2.512, luego de 18 h de inducción (Fig. 53).

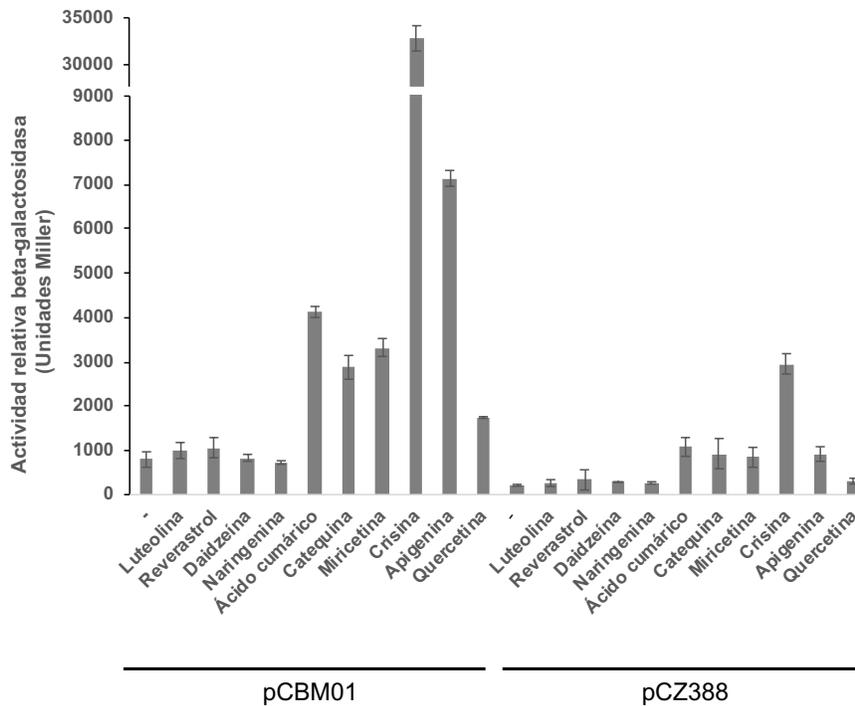


Figura 53. Respuesta de la región promotora del gen *nodB* en presencia de flavonoides y otros polifenoles en cultivos de UYPR2.512.

La determinación de la actividad β -galactosidasa se realizó en cultivos de UYPR2.512 portando los plásmidos pCBM01 o pCZ388. La expresión de *lacZ* se evaluó como actividad β -galactosidasa registrada como unidades Miller luego de un período de incubación de 18 h a 30 °C, en presencia de los flavonoides

La barra en el eje de las x corresponde los cultivos portando cada uno de los plásmidos sin inducir. Los datos se expresan como la media de duplicados biológicos con sus errores estándar.

Se observó un aumento de la actividad β -galactosidasa de cultivos de UYPR2.512 portando pCBM01 en presencia de todos los compuestos evaluados. Particularmente, las flavonas crisina (11X), miricetina (7X) y apigenina, (4X) produjeron un mayor aumento de la expresión en cultivos de UYPR2.512 portando pCBM01, con respecto al del resto de los flavonoides, e incluso al compararlo con luteolina (3,8X). Con respecto a otros polifenoles, la presencia de ácido cumárico (3,8X) (precursor de flavonoides), resultó ser un inductor más potente que el reverastrol (3X). No se observó un efecto en el crecimiento de UYPR2.512 para ninguno de los compuestos evaluados como inductores (datos no mostrados).

4 Discusión

4.1 Respuesta de los genes *nod* en UYPR2.512 a exudados radiculares, flavonoides y otros compuestos

Varios trabajos han demostrado que los exudados radiculares contienen una gran variedad de moléculas, algunas de las cuales son una señal clave para la interacción simbiótica, ya sea actuando como moléculas quimioatrayentes hacia los rizobios o favoreciendo el proceso de simbiosis [409], [410]. En dicho contexto, demostramos que los exudados radiculares de *M. pudica* son capaces de inducir la expresión de los genes *nod* en UYPR2.512 indicando que contienen una o más moléculas inductoras de los genes *nod* (Fig. 52).

Como se comentó previamente los flavonoides unidos a las proteínas NodD, inducen la expresión de los genes *nod* tanto en alfa como en beta-rizobios. Los resultados presentados en el capítulo II, muestran efectivamente el efecto inductor de la luteolina sobre los genes *nod* en UYPR2.512. Luego, nos surgió el interrogante de que posiblemente existieran otros flavonoides y/o otras moléculas, que tal vez también podrían estar involucradas en dicha interacción. En este capítulo se evaluó el efecto

inductor de otros flavonoides, y de otros compuestos fenólicos presentes en los exudados radiculares de *M. pudica*.

De los compuestos evaluados, las flavonas crisina y apigenina, resultaron ser inductores muy potentes de la expresión de los genes *nod* (Figura 53). Estos flavonoides son inductores en otros rizobios, como ser en los géneros *Rhizobium* y *Mesorhizobium*, y se ha visto su implicancia en la simbiosis de los mismos con sus plantas hospederas [408], [411]. A su vez, resultaron inductores, aunque en menor medida, la flavona miricetina y la catequina, compuestos que se clasifican dentro de la familia de los flavanoles. De las tres flavonas evaluadas, crisina y apigenina presentan mayor similitud a nivel estructural en comparación con la miricetina. Las diferencias en la capacidad inductora entre las flavonas e incluso con los flavanoles, permiten suponer que existiría una cierta especificidad o preferencia de reconocimiento, mediada por la estructura, de estos compuestos al regulador NodD. A pesar de que se ha demostrado que la habilidad de unión de NodD a la caja *nod*, y la posterior inducción de los genes *nod*, está influenciada por los flavonoides [116], [138], se tiene muy poca información sobre la localización de los sitios de unión a los flavonoides y sobre los mecanismos de activación de NodD.

En un estudio en el que realizaron modelados computacionales en base a homología de secuencias aminoacídicas de proteínas NodD, con otros reguladores transcripcionales de la familia LysR, se predijo la estructura secundaria de la proteína en presencia de luteolina [412]. Según las predicciones, los residuos de asparagina 135 y 153 del extremo C-terminal de NodD, estarían involucrados en la unión a la luteolina mediante la formación de enlaces de hidrógeno entre los grupos aceptores de hidroxilos de los residuos de asparagina y los grupos donadores de hidroxilos de la luteolina. Estos resultados nos permitirían suponer que la presencia de un mayor número de grupos donadores de hidrógeno en el flavonoide, estabilizarían la unión a NodD. Sin embargo, en nuestro caso, observamos una menor inducción del *pnodB* en presencia de miricetina, que tiene 3 grupos hidroxilos en el anillo B, con respecto a crisina que no contiene ninguno y con apigenina que tiene un grupo hidroxilo, lo cual sugiere que la estabilidad de la unión estaría afectada por otras uniones.

En otro estudio *in-silico*, realizado por Bag y col. [413], se determinaron los diferentes valores de afinidad de unión (kcal/mol) entre la proteína NodD de *R. leguminosarum* bv *trifolii* y 19 tipos distintos de flavonoides, donde una afinidad de unión más negativa sería indicativa de una interacción más fuerte. En dicho estudio, se determinó que la unión a los diferentes flavonoides estaría afectada en función del número de formación de enlaces de hidrógeno, como en el número de interacciones hidrofóbicas. A modo de ejemplo según este estudio, en la formación de enlaces de hidrógeno entre NodD y flavonoides como ser la luteolina, quercetina y apigenina, estarían involucrados los residuos de treonina 203, asparagina 7, fenilalanina 48 y en la formación de enlaces hidrofóbicos arginina 202 (D), leucina 204, prolina 206, serina 207, triptofano 211, donde en conjunto estabilizarían la unión.

En este sentido, estudios de modelado computacional utilizando la secuencia de la proteína NodD de UYPR2.512 podrían aportar más información y confirmar nuestros resultados.

Los niveles de respuesta de *pnodB19424-lacZ* en UYPR2.512 a los distintos tipos de flavonoides reflejan cierta especificidad del sistema, sin embargo, también se observó una respuesta significativa a otros compuestos fenólicos, sugiriendo que la proteína NodD de UYPR2.512, presenta flexibilidad en su especificidad. Los flavonoides han sido las primeras señales moleculares descritas y las más estudiadas en la interacción rizobio-leguminosa y, como se comentó previamente, los tipos de flavonoides varían entre las asociaciones simbióticas. Se ha reportado que isoflavonoides como la genisteína y la daídzeína están involucrados en la simbiosis entre *Bradyrhizobium* y *Glycine max* (soja) [138], y en el caso de *R. leguminosarum* bv. *viciae* es estimulado por el flavonoide hesperetina secretada por las raíces de *Pisum sativum* [414]. En dicho contexto, se ha reportado la presencia de ácidos fenólicos (cumárico, ferúlico y cafeico) en los exudados radiculares de *L. japonicum*, los cuales inducen la expresión de los genes *nod* en cepas de *Mesorhizobium* [411].

Esto nos llevó a hipotetizar que además de los EPS, SSTIII y las lectinas, otras moléculas como las aquí observadas podrían estar involucradas en las etapas de reconocimiento [100], [150], [159].

4.2 Compuestos presentes en los exudados de *M. pudica*

Con el fin de describir la composición de los exudados radiculares de *M. pudica*, se utilizó un sistema de hidroponia en el que se incubaron las plantas solas o en presencia de su par simbiótico, la cepa UYPR2.512 de *C. necator*. Los exudados fueron colectados y su composición fue analizada mediante técnicas analíticas (HPLC-DAD, HPLC-ESI-MS/MS y GC-MS) ampliamente utilizadas en la actualidad [397], [398], [399], [400]. Gran parte de nuestros esfuerzos estuvieron enfocados en la búsqueda de flavonoides, los cuales se encuentran en concentraciones de nanomolar en distintos exudados [118]. Aunque no se obtuvieron indicios claros de su presencia, si se observó la presencia de compuestos con un máximo de absorción, que coincidirían con precursores de flavonoides (Fig. 42) [405], [406].

El empleo de HPLC-ESI-MS/MS permitió la identificación cualitativa de una serie de picos que corresponderían a compuestos presentes únicamente en presencia de UYPR2.15 (Fig. 47 y 48), lo cual sugirió que podría tratarse de compuestos específicos producidos durante la interacción simbiótica. Entre estos, en una primera instancia se identificó la miricetina 3-rutinosa, aunque estudios posteriores descartaron que se tratara de dicho compuesto, sugiriendo que estaríamos en presencia de un compuesto aún no identificado. Sin embargo, se observó que la miricetina es inductora de la expresión de los genes *nod* en UYPR2.512, por lo que no podemos descartar que esté involucrada en la interacción de UYPR2.512 con su hospedero *M. pudica* (Fig. 53). La miricetina pertenece a la subfamilia de las flavonas, la cual presenta alta similitud estructural con la quercetina, debido a la presencia de un tercer grupo hidroxilo en el anillo B y ha sido detectada en plantas leguminosas de la familia de las *Fabaceae* [415]. Se ha reportado que la quercetina induce la expresión de los genes *nod* en rizobios [416], sin embargo, hasta la actualidad no hay trabajos que evidencien el rol de la miricetina durante la interacción rizobio-leguminosa. El hecho de no haber detectado flavonoides en los exudados, aún cuando las técnicas analíticas utilizadas fueron de alta sensibilidad, nos permite descartar el rol de estas moléculas en la interacción. Los flavonoides generalmente son detectados en tejidos como las raíces de las plantas, en los que están a mayor concentración, en comparación a los exudados radiculares. La realización de

experimentos que involucren un mayor número de plantas, o estudiar la presencia de dichas moléculas en las raíces nos permitiría evaluar posibles flavonoides implicados en esta interacción.

Una tercera aproximación empleada para identificar los metabolitos presentes en los exudados radiculares y en la interacción entre UYPR2.512 y *M. pudica*, fue su análisis mediante GC-MS. En este caso analizamos la fracción metanólica de los exudados radiculares, cuya capacidad de inducción de los genes *nod* fue confirmada.

En los exudados de *M. pudica* se identificaron cualitativamente 68 metabolitos presentes, tanto en ausencia como en presencia de UYPR2.512. Estos metabolitos pertenecen a una gran variedad de familias químicas que incluyeron mayoritariamente polisacáridos, aminoácidos, polialcoholes y ácidos carboxílicos/dicarboxílicos (carboxilatos), y en menor medida ácidos fenólicos y ácidos grasos, entre otros (Tabla 27), similar a lo reportado por Klonowska y col. [273].

Estas familias químicas de compuestos han sido identificadas en exudados de plantas, tanto a partir de semillas como de raíces y se ha reportado su participación en las interacciones simbióticas como moléculas señalizadoras con otros microorganismos [135], [266], [417], [418], [419]. En el trabajo realizado por Compton y col. en el 2021, se identificaron diversos compuestos con capacidad quimioatrayente sobre bacterias de la rizósfera, siendo ciertos aminoácidos y algunos carboxilatos los más abundantes [417]. Por ejemplo, la cepa 2011 de *S. meliloti* es atraída fuertemente por determinados aminoácidos, y débilmente atraída frente a aminoácidos ácidos como el glutamato y aspartato. En contraste a lo observado en la cepa USDA110 de *B. japonicum*, donde el aspartato y el glutamato actúan como los atrayentes potentes, lo cual refleja cierta especificidad en la interacción [420]. Los aminoácidos también pueden cumplir otras funciones relacionadas al metabolismo de la planta, mediante la modificación del pH del entorno al liberar protones (H^+) o hidroxilos (OH^-), lo que la beneficiaría en la disponibilidad de nutrientes para la planta [421].

Entre los compuestos detectados en los exudados radiculares de *M. pudica*, se identificaron el ácido palmítico y el ácido glucónico. Los ácidos orgánicos son moléculas esenciales que desempeñan roles cruciales en casi todos los organismos, y se los ha

relacionado fundamentalmente con las interacciones de las plantas y su entorno. Estudios han demostrado que el ácido palmítico inhibe el crecimiento de patógenos vegetales en la rizósfera, favoreciendo el crecimiento de las plantas [422].

Por otro lado, al ácido glucónico, se lo ha relacionado con su rol en la solubilización de fosfatos y en la regulación de la producción de compuestos antifúngicos en bacterias rizosféricas, lo que puede contribuir al crecimiento de las plantas y a su protección contra enfermedades de raíz, como se observó en la cepa CHA0 de *Pseudomonas fluorescens* [423].

Entre los metabolitos identificados durante la interacción UYPR2.512 y *M. pudica*, se encuentran el ácido benzoico y el ácido vanilínico (Tabla 27). Se ha demostrado que compuestos aromáticos como los ácidos fenólicos (ácido cumárico, isoferúlico, cafeico, y benzoico, entre otros), actúan como moléculas señal durante la simbiosis rizobio-leguminosa [411], [424]. El ácido benzoico y el ácido vanilínico pueden provenir de la degradación de la lignina, uno de los componentes estructurales más abundantes en las células vegetales [425]. Resulta interesante la capacidad metabólica de UYPR2.512 de utilizar al ácido benzoico como fuente de carbono [426] ya que contiene los genes *benABCD*. También se ha demostrado la capacidad quimioatrayente de compuestos fenólicos, que incluyen ácido vanilínico, gálico, ferúlico y cinámico entre otros, en interacciones simbióticas entre plantas y hongos [427]. La identificación de estos compuestos en los exudados en presencia de UYPR2.512 sugiere su posible participación en la interacción simbiótica tal como se ha observado en otros modelos [411].

Vale la pena resaltar el posible rol del mio-inositol durante esta interacción, el cual es una fuente de carbono comúnmente detectada en el suelo y se encuentra presente en una amplia variedad de alimentos, incluida las legumbres [428]. Este polialcohol tiene diversas funciones celulares entre las que se destacan la de ser un componente de la membrana celular, participar en el metabolismo de carbohidratos, en la señalización celular tanto en células eucariotas como en procariontes y particularmente el inositol fosforilado, es la principal molécula de almacenamiento de fósforo en plantas [429]. También, se ha reportado su utilización como fuente de carbono y energía, por rizobios

como *S. meliloti* y *R. leguminosarum*, lo que les confiere la capacidad de competir de manera más eficientes frente a otras bacterias en la nodulación de huéspedes como *M. sativa* [430]. Sin embargo, su rol como inductor de los genes *nod* no ha sido estudiado.

CONCLUSIONES

Capítulo I

La re-secuenciación de UYPR2.512 permitió mejorar la calidad del genoma disponible. De acuerdo al análisis de esta última secuenciación esta bacteria tendría un genoma compuesto por al menos 4 replicones, incluyendo un plásmido de 770 kb que contiene los genes simbióticos, agrupados en una isla simbiótica.

Análisis *in-silico* indican que la cepa pertenece a la especie *Cupriavidus necator*. La anotación del genoma nos permitió describir en profundidad sus principales características genéticas y funcionales relacionadas al género *Cupriavidus* y a las relacionadas a su interacción con el hospedero, reflejando su adaptación ambiental y su versatilidad metabólica.

Capítulo II

El análisis genómico funcional mediante RNA-Seq reveló la respuesta transcripcional de UYPR2.512 a la luteolina, un flavonoide inductor de los genes nod en UYPR2.512. Los genes de mantenimiento fisiológico, ubicados en el chr1, exhibieron una reducción en los niveles de expresión comparado con el control. Por otro lado, los genes que exhibieron un incremento en los niveles de expresión se encuentran dispersos en todos los replicones, aunque especialmente ubicados en el plásmido pSym, que alberga los genes simbióticos.

Se identificaron 145 DEG, 72 sobre-expresados y 73 sub-expresados. Entre las principales categorías funcionales COG, se encontraron aquellas relacionadas con la respuesta al estrés y detoxificación, el metabolismo central, la homeostasis de metales y los genes vinculados con la nodulación. Esta regulación genética podría representar los primeros pasos en el proceso de simbiosis en respuesta a señales químicas liberadas por las plantas leguminosas.

Capítulo III.

La metodología empleada no logró identificar flavonoides en los exudados de *M. pudica*. Sin embargo, se demostró que estos exudados radiculares contienen moléculas inductoras de los genes *nod* en UYPR2.512, las cuales no parecen ser flavonoides. Se observó la presencia cualitativa de picos correspondientes a compuestos únicos en presencia de UYPR2.512, aunque no pudieron ser identificados. Se confirmó que, además de la luteolina, otros flavonoides como la crisina, apigenina, ácido cumárico y miricetina, son inductores de los genes *nod* en UYPR2.512, lo que sugiere su función como moléculas señal en las etapas iniciales de la interacción con sus hospederos. La variación en la composición química de los exudados radiculares de *M. pudica* en respuesta a la presencia de UYPR2.512 y la identificación de metabolitos diferenciales reflejaron la alta complejidad de la interacción. Estos hallazgos sugieren una interacción simbiótica diversa y compleja, con la participación de diversas moléculas inductoras en la comunicación entre la cepa UYPR2.512 y *M. pudica*.

ANEXOS

1 Medios de cultivos

1.1 Medio de cultivo líquido mínimo M9

Componente	Cantidad por litro
Sales M9 5x *	200 mL
MgSO ₄ 1M	2 mL
Solución 20 % (p/v) de la fuente de carbono	20 mL
CaCl ₂ 1M	0,1 mL
H ₂ O c.s.p 1 litro	
Sales 5x *	
Na ₂ HPO ₄ .7H ₂ O	64 g
KH ₂ PO ₄	15 g
NaCl	2,5 g
NH ₄ Cl	5,0 g
H ₂ O c.s.p 1 litro	

*Las sales y el resto de los componentes se esterilizan por separado y luego se prepara el medio. Como fuente de carbono se utilizó citrato de sodio tribásico dihidratado 14mM (SIGMA #087K0055).

Nota: El medio mínimo M9 modificado corresponde a la misma preparación pero en lugar de utilizarse agua destilada, se utiliza el agua recolectada conteniendo los exudados radiculares de *M. pudica*, previamente esterilizada por filtración con filtro de 0,45 µm o la fracción metanólica en la cual los exudados fueron extraídos.

1.2 Medio de cultivo líquido Luria – Bertani (LB)

Componente	Cantidad por litro
Triptona	10 g
Extracto de levadura	5 g
NaCl ₂	10 g
H ₂ O c.s.p 1 litro	

1.3 Medio de cultivo líquido Howelson 10X

Componente	Cantidad por litro
K ₂ SO ₄	0,174 g
KH ₂ PO ₄	0,27 g
CaSO ₄	1,36 g
MgSO ₄ . 7H ₂ O	0,49 g
Sol. de micronutrientes*	1mL
FeSO ₄	0,055 g
H ₂ O c.s.p 1 litro	

*Solución de micronutrientes 100X

Componente	Cantidad por litro
H ₃ BO ₃	1,85 g
Na ₂ MoO ₄ .2H ₂ O	0,066 g
ZnSO ₄ ·7H ₂ O	2,87 g
MnSO ₄ ·H ₂ O	0,17 g
CaSO ₄ .5H ₂ O	0,5 g
CaCl ₂	0,240 g
H ₂ O c.s.p 1 litro	

2 Tablas suplementarias

2.1 Capítulo I.

Tabla S1. Cepas de 71 genomas representantes de la familia *Burkholderiaceae* utilizadas para el análisis filogenómico.

Cepa	Catálogo de genomas ensamblados (GCA)
<i>Chitinimonas taiwanensis</i> DSM 18899	GCA_900119825.1
<i>Chitinimonas koreensis</i> DSM 17726	GCA_000428465.1
<i>Limnobacter</i> sp. CACIAM 66H1	GCA_001601825.1
<i>Lautropia mirabilis</i> ATCC51599	GCA_000186425.1
<i>Lautropia</i> sp. SCN 66-9	GCA_001725315.1
<i>Pandoraea thiooxydans</i>	GCA_001931675.1
<i>Pandoraea norimbergensis</i>	GCA_001465545.3
<i>Pandoraea apista</i>	GCA_001027265.1
<i>Pandoraea oxalatirovans</i>	GCA_000972785.1
<i>Robbsia andropogonis</i> Ba3549	GCA_000566705.1
<i>Glometribacter</i> sp. 1016415	GCA_000424705.1
<i>Burkholderia vietnamiensis</i>	GCA_001718775.1
<i>Burkholderia ambifaria</i> AMMD	GCA_000203915.1
<i>Burkholderia stabilis</i>	GCA_001742165.1
<i>Burkholderia lata</i>	GCA_000012945.1
<i>Paraburkholderia acidipaludis</i> NBRC 101816	GCA_000684975.1
<i>Cabelleroniazhejiangensis</i>	GCA_000773055.1
<i>Cabelleronia calidae</i>	GCA_900044055.2
<i>Cabelleronia megalochromosomata</i>	GCA_001580565.1
<i>Cabelleronia jiangsuensis</i>	GCA_000648925.1
<i>Polynulcoebacter</i> sp.MWH-K35W1	GCA_002206635.1
<i>Ralstonia solanacearum</i>	GCA_000749995.1
<i>Ralstonia mannitollytica</i>	GCA_000953875.1
<i>Ralstonia pickettii</i> 12D	GCA_000023425.1
<i>Ralstonia insidiosa</i>	GCA_001663855.1
<i>Cupriavidus gilardii</i> CR3	GCA_001281465.1
<i>Cupriavidus</i> sp. HPC	GCA_000307735.1
<i>Cupriavidus gilradii</i>	GCA_001658125.1
<i>Cupriavidus</i> sp.WS	GCA_000395345.1
<i>Cupriavidus basilensis</i> KFG708	GCA_000876015.1
<i>Cupriavidus basilensis</i> B-8	GCA_000282815.1
<i>Cupriavidus</i> sp. SK-3	GCA_000611145.2
<i>Cupriavidus basilensis</i> 4G11	GCA_000832305.1
<i>Cupriavidus basilensis</i> OR16	GCA_000243095.2
<i>Cupriavidus</i> sp. YR651	GCA_900101625.1

Cepa	Catálogo de genomas ensamblados (GCA)
<i>Cupriavidus</i> sp. OV096	GCA_900115455.1
<i>Cupriavidus</i> sp. OV038	GCA_900112215.1
<i>Cupriavidus</i> sp. D384	GCA_001652915.1
<i>Cupriavidus</i> pauculus	GCA_000974605.1
<i>Cupriavidus</i> sp. BIS7	GCA_000292345.1
<i>Cupriavidus</i> sp. SHE	GCA_000812445.1
<i>Cupriavidus metallidurans</i> NE12	GCA_000709065.1
<i>Cupriavidus</i> sp. HMR-1	GCA_000319775.1
<i>Cupriavidus metallidurans</i> CH34	GCA_000196015.1
<i>Cupriavidus metallidurans</i> NBRC101272	GCA_001598775.1
<i>Cupriavidus metallidurans</i> NA4	GCA_000709045.1
<i>Cupriavidus metallidurans</i> NDB3NO24	GCA_001543455.1
<i>Cupriavidus metallidurans</i> NA1	GCA_000709025.1
<i>Cupriavidus metallidurans</i> H1130	GCA_000496715.1
<i>Cupriavidus</i> sp. USMAA2-4	GCA_001854305.1
<i>Cupriavidus</i> sp. USMAA1020	GCA_001854325.1
<i>Cupriavidus</i> sp. USMAAHM13	GCA_001854285.1
<i>Cupriavidus pinatubonensis</i> JMP134	GCA_000203875.1
<i>Cupriavidus</i> sp. UYMMa02A	GCA_001725945.1
<i>Cupriavidus</i> sp. IDO	GCA_000812465.1
<i>Cupriavidus necator</i> A3-1	GCA_000744095.1
<i>Cupriavidus taiwanensis</i> LMG19424	GCA_000069785.1
<i>Cupriavidus taiwanensis</i> STM 6018	GCA_000472465.1
<i>Cupriavidus taiwanensis</i> STM 6070	GCA_000372525.1
<i>Cupriavidus alkaliphilus</i>	GCA_900094595.1
<i>Cupriavidus nantongensis</i>	GCA_001598055.1
<i>Cupriavidus oxalaticus</i> NBRC 13593	GCA_001592245.1
<i>Cupriavidus</i> sp. amp6	GCA_000426345.1
<i>Cupriavidus</i> sp. NH9	GCA_002011925.1
<i>Cupriavidus</i> sp. UYPR2.512	GCA_000379565.1
<i>Cupriavidus necator</i> N-1	GCA_000219215.1
<i>Cupriavidus</i> sp. SK-4	GCA_000611125.1
<i>Cupriavidus necator</i> PHE3-6	GCA_001853325.1
<i>Cupriavidus necator</i> NBRC 102504	GCA_001598755.1
<i>Cupriavidus</i> sp. GA3-3	GCA_000389805.1
<i>Cupriavidus necator</i> H16	GCA_000209285.1

2.2 Capítulo II.

Tabla S2. Información resumida cuantitativa y cualitativa de los genes sub-expresados en UYPR2.512 en presencia de luteolina.

Se indican la ID de gen, número de lecturas promedio en la condición control (C), número de lecturas promedio en presencia de luteolina (L), log FC, valor p, FRD, descripción del gen, nombre del gen y categoría COG asignada.

ID gen	Promedio_C (CPM)	Promedio_L (CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre del gen	Categoría COG
18	885.41	221.49	-1.9990	6.20E-04	4.40E-02	ATP synthase delta chain	<i>atpH</i>	C
19	1012.12	254.35	-1.9925	4.89E-04	3.77E-02	ATP synthase F0 sector subunit b	NA	C
41	1512.09	532.12	-1.5067	2.62E-04	2.59E-02	ABC-type branched-chain amino acid transport systems, periplasmic component	NA	E
71	1422.63	462.79	-1.6201	1.02E-04	1.46E-02	TRAP transporter solute receptor, unknown substrate 6	<i>takP</i>	C
107	2081.99	730.79	-1.5104	3.83E-04	3.26E-02	Phosphoenolpyruvate carboxykinase [GTP]	<i>pckG</i>	H
118	7605.87	2050.42	-1.8912	8.08E-05	1.28E-02	Tripartite tricarboxylate transporter TctC family	<i>tctC</i>	S
119	726.28	215.78	-1.7510	3.34E-04	3.00E-02	Tripartite tricarboxylate transporter TctB family	<i>tctB</i>	S
241	1034.41	335.50	-1.6244	8.98E-05	1.32E-02	Vibrio ferritin ligase/carboxylase protein PvsA	<i>pvsA</i>	F
399	2842.58	843.80	-1.7522	8.65E-05	1.30E-02	BUG/TctC family periplasmic protein	NA	S
577	214.13	63.07	-1.7633	2.33E-04	2.42E-02	Bacterioferritin-associated ferredoxin	<i>bfr</i>	P
597	2179.08	740.93	-1.5563	3.32E-04	3.00E-02	Cytochrome c oxidase polypeptide II	<i>cox2</i>	C
619	331.55	108.10	-1.6167	4.73E-04	3.67E-02	Fumarylacetoacetase	<i>fhaA</i>	Q

ID gen	Promedio_C (CPM)	Promedio_L (CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre del gen	Categoría COG
620	327.13	86.88	-1.9127	6.62E-05	1.16E-02	Maleylacetoacetate isomerase @ Glutathione S-transferase, zeta	<i>maiA</i>	G
825	4305.37	1069.03	-2.0098	7.55E-05	1.25E-02	Succinyl-CoA ligase [ADP-forming] beta chain	<i>sucC</i>	F
877	3187.00	973.61	-1.7108	2.35E-04	2.42E-02	BUG/TctC family periplasmic protein	<i>tctC</i>	S
897	867.04	323.46	-1.4225	4.22E-04	3.47E-02	Superoxide dismutase [Fe]	<i>sodB</i>	C
1031	867.71	203.04	-2.0955	1.33E-06	1.08E-03	Ferric iron ABC transporter, iron-binding protein	<i>futA1</i>	P
1142	193.60	82.96	-1.2224	1.79E-04	2.14E-02	3-oxoacyl-[acyl-carrier protein] reductase	<i>fabG</i>	IQ
1260	2903.09	1002.45	-1.5341	6.02E-04	4.30E-02	Ketol-acid reductoisomerase (NADP(+))	<i>ilvC</i>	H
1265	969.82	420.99	-1.2039	1.44E-04	1.88E-02	2-isopropylmalate synthase	<i>leuA</i>	E
1270	3160.02	1038.40	-1.6056	2.80E-04	2.72E-02	Polyribonucleotide nucleotidyltransferase	<i>pnp</i>	J
1271	264.82	113.06	-1.2278	2.76E-04	2.70E-02	Quinone oxidoreductase	<i>nqo</i>	C
1294	283.35	94.86	-1.5787	4.64E-04	3.63E-02	Cytochrome O ubiquinol oxidase subunit II	<i>cyoA1</i>	C
1317	227.00	80.18	-1.5012	7.83E-05	1.28E-02	Glutathione S-transferase	<i>gstI</i>	O
1346	232.39	81.84	-1.5056	1.76E-04	2.13E-02	Putative periplasmic cytochrome type-C oxidoreductase signal peptide protein	<i>cytc2</i>	C

ID gen	Promedio_C (CPM)	Promedio_L (CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre del gen	Categoría COG
1347	203.87	58.28	-1.8064	1.26E-04	1.70E-02	Putative periplasmic cytochrome type-C oxidoreductase signal peptide protein	<i>cytC1</i>	C
1380	1519.00	443.84	-1.7750	2.21E-05	6.76E-03	Biosynthetic Aromatic amino acid aminotransferase alpha @ Aromatic-amino-acid aminotransferase	<i>tyrB</i>	E
1663	308.36	132.11	-1.2228	3.52E-04	3.11E-02	putative iron-regulated membrane protein	NA	S
1665	88.12	25.26	-1.8023	2.13E-04	2.31E-02	hypothetical protein	NA	NonClass
1666	96.09	28.36	-1.7604	9.02E-05	1.32E-02	hypothetical protein	NA	NonClass
1679	8876.55	2196.67	-2.0147	2.48E-04	2.50E-02	Uncharacterized MFS-type transporter	NA	EGP
1819	650.60	207.84	-1.6462	5.13E-04	3.89E-02	Peptidyl-prolyl cis-trans isomerase	NA	M
2243	2075.96	642.71	-1.6915	8.03E-05	1.28E-02	L-asparaginase	NA	EJ
2529	208.26	80.37	-1.3735	3.14E-04	2.95E-02	FIG027190: Putative transmembrane protein	NA	NonClass
2578	167.85	56.74	-1.5645	1.64E-04	2.03E-02	Methionine ABC transporter substrate-binding protein	<i>metQ</i>	M
2663	231.05	83.84	-1.4623	7.37E-05	1.24E-02	Sulfate transport system permease protein CysT	<i>cysT</i>	P
2671	233.81	82.87	-1.4962	6.83E-04	4.68E-02	FMN reductase (NADPH)	<i>ssuE</i>	S
2673	342.08	121.26	-1.4961	2.27E-04	2.39E-02	Sulfate and thiosulfate binding protein CysP	<i>cysP</i>	P

ID gen	Promedio_C (CPM)	Promedio_L (CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre del gen	Categoría COG
2831	1699.76	553.04	-1.6199	4.56E-04	3.63E-02	Transcription termination factor Rho	<i>rho</i>	K
2940	645.46	173.22	-1.8977	3.37E-05	8.92E-03	putative TonB-dependent receptor	<i>tonB</i>	M
2941	1257.61	425.57	-1.5632	1.90E-04	2.21E-02	MotA/TolQ/ExbB proton channel family protein	<i>exbB</i>	U
2942	650.13	209.15	-1.6362	5.08E-05	9.94E-03	Biopolymer transport protein ExbD/TolR	<i>exbD1</i>	U
2943	1442.65	260.13	-2.4714	8.38E-08	2.05E-04	Hemin uptake protein	<i>hmuS</i>	P
3008	1942.86	638.63	-1.6051	2.41E-04	2.46E-02	Ribonuclease E	<i>rne</i>	J
3046	1668.07	442.32	-1.9150	1.63E-04	2.03E-02	Aspartate-semialdehyde dehydrogenase	<i>asd</i>	E
3054	1324.35	403.97	-1.7129	2.27E-04	2.39E-02	Branched-chain amino acid ABC transporter, amino acid-binding protein (TC 3,A,1,4,1)	<i>braC</i>	E
3059	4255.82	1420.76	-1.5828	7.07E-04	4.81E-02	Succinate dehydrogenase flavoprotein subunit	<i>sdhA</i>	C
3063	4924.29	1574.52	-1.6450	3.31E-04	3.00E-02	Malate dehydrogenase	<i>mdh</i>	C
3067	266.97	81.24	-1.7163	1.37E-04	1.82E-02	hypothetical protein	NA	S
3189	495.29	178.96	-1.4686	5.99E-04	4.30E-02	hypothetical protein	<i>fmt2</i>	S
3268	349.56	100.14	-1.8035	8.63E-05	1.30E-02	Alkyl hydroperoxide reductase subunit C-like protein	<i>lsfA</i>	O
3282	2256.00	762.11	-1.5657	3.00E-04	2.88E-02	Serine hydroxymethyltransferase	<i>glyA</i>	E

ID gen	Promedio_C (CPM)	Promedio_L (CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre del gen	Categoría COG
3298	186.42	65.80	-1.5023	6.48E-04	4.47E-02	Transcription termination protein NusB	<i>nusB</i>	K
3365	583.82	253.91	-1.2012	5.37E-04	4.01E-02	Argininosuccinate lyase (ASL or ASAL, 4.3.2.1)	<i>argH</i>	E
3420	930.90	315.16	-1.5625	2.79E-05	7.80E-03	Outer membrane vitamin B12 receptor BtuB	<i>btuB</i>	H
3487	1189.78	294.70	-2.0134	3.67E-05	9.31E-03	Branched-chain amino acid ABC transporter, amino acid-binding protein (TC 3,A,1,4,1)	<i>livk2</i>	E
3544	258.29	112.62	-1.1974	5.69E-04	4.19E-02	3-methyl-2-oxobutanoate hydroxymethyltransferase (EC 2.1. 2.11)	<i>panB</i>	H
3591	865.57	321.20	-1.4302	1.27E-04	1.70E-02	Adenylosuccinate lyase @ SAICAR lyase	<i>adsl</i>	F
3644	216.38	66.63	-1.6992	4.50E-04	3.61E-02	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	<i>accAD</i>	I
3656	3268.80	915.50	-1.8361	2.58E-05	7.42E-03	Ferrichrome-iron receptor	<i>fecA</i>	P
3804	438.41	127.75	-1.7789	4.76E-05	9.71E-03	2,3-dehydroadipyl-CoA hydratase	<i>paaA</i>	I
4231	904.30	223.77	-2.0148	6.36E-04	4.45E-02	ATP synthase delta chain	<i>atpF</i>	C
4232	1012.72	255.06	-1.9893	5.46E-04	4.05E-02	ATP synthase F0 sector subunit b	NA	C
4557	256.26	75.12	-1.7703	3.08E-04	2.93E-02	Histone-like bacterial DNA-binding protein	<i>hup</i>	L
6192	12.11	4.87	-1.3124	7.46E-04	4.97E-02	Mobile element protein	NA	L

ID gen	Promedio_C (CPM)	Promedio_L (CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre del gen	Categoría COG
7162	10522.16	2992.74	-1.8139	1.13E-04	1.58E-02	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	<i>metE</i>	E
7769	3993.29	932.28	-2.0987	6.37E-05	1.13E-02	Tripartite tricarboxylate transporter TctA family	<i>tctA</i>	S
7770	1328.59	224.84	-2.5629	1.98E-05	6.24E-03	Tripartite tricarboxylate transporter TctB family	<i>tctB</i>	S
8372	223.07	63.41	-1.8146	1.53E-05	5.39E-03	Ferric uptake regulation protein FUR	<i>fur</i>	K
8787	900.08	223.51	-2.0097	1.00E-05	4.30E-03	Heavy metal RND efflux outer membrane protein, CzcC family	<i>czcC</i>	MU
8788	650.13	203.59	-1.6750	5.75E-05	1.08E-02	Cobalt/zinc/cadmium efflux RND transporter, membrane fusion protein, CzcB family	<i>czcB</i>	M
9006	13.21	3.60	-1.8729	1.68E-04	2.06E-02	hypothetical protein	NA	S
9007	57.90	16.55	-1.8066	4.45E-04	3.60E-02	2-amino-3-carboxymuconate-6-semialdehyde decarboxylase	NA	S

Tabla S3. Información resumida cuantitativa y cualitativa de los genes sobre-expresados en UYPR2.512 en presencia de luteolina.

Información resumida cuantitativa y cualitativa de los genes sobre-expresados en UYPR2.512 en presencia de luteolina. Se indican la ID de gen, número de lecturas promedio en la condición control (C), número de lecturas promedio en presencia de luteolina (L), log *FC*, valor p, FRD, descripción del gen, nombre del gen y categoría *COG* asignada.

ID gen	Promedio _C(CPM)	Promedio _L(CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre de gen	Categoría COG
48	71.25	160.87	1.1748	1.58E-04	2.03E-02	Cu(I)-responsive transcriptional regulator	<i>cueR</i>	K
49	179.21	736.96	2.0398	2.06E-06	1.55E-03	Lead, cadmium, zinc and mercury transporting ATPase; Copper-translocating P-type ATPase	<i>copA</i>	P
50	62.47	188.75	1.5949	1.21E-05	4.78E-03	Copper chaperone	<i>copZ</i>	P
101	131.31	498.50	1.9247	5.05E-05	9.94E-03	Hypothetical protein	NA	<i>NonClass</i>
130	95.88	315.86	1.7199	5.98E-07	7.32E-04	Transcriptional regulator, AcrR family	<i>acrR</i>	K
463	13.25	48.49	1.8693	3.52E-04	3.11E-02	ABC transporter, substrate-binding protein (cluster 4, leucine/isoleucine/valine/benzoate)	NA	E
508	17.15	47.83	1.4792	7.40E-04	4.96E-02	Hypothetical protein	NA	<i>NonClass</i>
1238	135.58	282.69	1.0599	5.20E-04	3.91E-02	Xanthine dehydrogenase iron-sulfur subunit / Xanthine dehydrogenase, FAD binding subunit	<i>xdhA</i>	C
2104	1.88	5.50	1.5316	2.51E-04	2.50E-02	Hypothetical protein	NA	<i>NonClass</i>
2159	3.42	8.71	1.3491	3.27E-04	3.00E-02	Hypothetical protein	NA	<i>NonClass</i>
2335	7.93	30.73	1.9520	2.00E-04	2.25E-02	L-lactate dehydrogenase	<i>ldh</i>	C
2336	2.59	6.62	1.3595	2.01E-04	2.25E-02	2-oxoglutarate/malate translocator	<i>ybhl</i>	P
2427	1.14	3.36	1.5680	4.19E-05	9.31E-03	Hypothetical protein	NA	<i>NonClass</i>
2613	1.37	3.66	1.4044	4.45E-04	3.60E-02	Hypothetical periplasmic protein	<i>cusF</i>	S
2625	9.75	33.35	1.7725	4.02E-04	3.39E-02	Beta-glucosidase	<i>bgIS</i>	G

ID gen	Promedio _C(CPM)	Promedio _L(CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre de gen	Categoría COG
2851	3.45	10.66	1.6208	4.51E-05	9.60E-03	Kef-type K ⁺ transport systems, predicted NAD-binding component	NA	P
2893	153.66	351.11	1.1920	8.35E-05	1.30E-02	Probable bifunctional hydroxylase/oxidoreductase	<i>abmA</i>	CH
4076	42.41	116.17	1.4540	4.15E-05	9.31E-03	Putative calcium binding signal peptide protein	NA	DTZ
4399	6.08	18.06	1.5663	5.96E-04	4.30E-02	Mobile element protein	NA	L
4400	3.86	20.09	2.3797	3.19E-06	1.88E-03	Retron-type RNA-directed DNA polymerase	NA	L
4461	3.88	15.86	2.0230	2.69E-06	1.76E-03	Error-prone repair homolog of DNA polymerase III alpha subunit	<i>dnaE2</i>	L
4622	48.46	143.13	1.5625	3.22E-04	3.00E-02	Hypothetical protein	<i>VP0501</i>	S
4624	80.65	301.45	1.9023	3.89E-05	9.31E-03	Adenylylsulfate kinase	<i>nodQ</i>	F
4625	624.71	2350.95	1.9120	3.91E-05	9.31E-03	Carbamoyltransferase	<i>nodU</i>	O
4626	197.60	871.21	2.1405	3.54E-07	5.78E-04	N-methyl transferase nodS	<i>nodS</i>	Q
4627	163.98	692.47	2.0783	4.14E-06	2.25E-03	Nodulation protein A	<i>nodA</i>	S
4628	125.56	495.89	1.9816	6.86E-06	3.36E-03	Hypothetical protein	<i>nodH</i>	S
4629	38.94	163.34	2.0685	6.40E-06	3.30E-03	Efflux ABC transporter, permease protein	<i>nodJ</i>	U
4630	54.04	223.53	2.0485	3.26E-06	1.88E-03	Efflux ABC transporter, ATP-binding protein	<i>nodI</i>	V
4631	117.59	334.27	1.5072	5.07E-04	3.88E-02	Nodulation protein C	<i>nodC</i>	M
4694	2.25	7.87	1.7926	1.15E-04	1.59E-02	Hypothetical protein	NA	<i>NonClass</i>
4865	0.50	1.78	1.7846	4.12E-04	3.43E-02	Mobile element protein	<i>ISPlu13C</i>	L
4874	1.43	4.67	1.7027	1.41E-05	5.30E-03	Hypothetical protein	NA	<i>NonClass</i>

ID gen	Promedio _C(CPM)	Promedio _L(CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre de gen	Categoría COG
4904	0.41	2.84	2.7903	1.30E-06	1.08E-03	Hypothetical protein	NA	<i>NonClass</i>
5529	4.54	11.34	1.3179	4.14E-04	3.43E-02	Mobile element protein	NA	L
5805	1.51	5.11	1.7504	3.21E-05	8.73E-03	Hypothetical protein	NA	<i>NonClass</i>
5931	3.72	13.79	1.8883	8.77E-06	4.09E-03	High-affinity branched-chain amino acid transport system permease protein LivH (TC 3,A,1,4,1)	<i>livH</i>	E
5932	6.41	21.95	1.7747	2.05E-04	2.26E-02	Branched-chain amino acid transport system permease protein LivM (TC 3,A,1,4,1)	<i>livM</i>	E
5949	1.21	3.29	1.4411	2.02E-04	2.25E-02	Hypothetical protein	NA	<i>NonClass</i>
6134	1.19	6.61	2.4551	5.48E-05	1.05E-02	Hypothetical protein	NA	<i>NonClass</i>
6459	0.96	2.45	1.3273	6.26E-04	4.41E-02	Hypothetical protein	NA	<i>NonClass</i>
6559	2.52	6.76	1.4185	1.60E-04	2.03E-02	Chromate transport protein ChrA	<i>chrA</i>	P
6655	7.26	31.63	2.1197	4.63E-04	3.63E-02	Hypothetical protein	NA	<i>NonClass</i>
6700	6.05	29.75	2.2987	3.34E-07	5.78E-04	Putative DNA helicase	NA	L
6791	5.69	20.11	1.8181	3.64E-04	3.18E-02	Hypothetical protein	NA	<i>NonClass</i>
6833	0.31	1.82	2.4788	1.06E-04	1.50E-02	Hypothetical protein	NA	<i>NonClass</i>
6836	7.75	41.27	2.4098	1.03E-06	1.01E-03	Hypothetical protein	NA	L
6837	1.24	3.48	1.4876	4.40E-05	9.58E-03	Hypothetical protein	NA	<i>NonClass</i>
6913	0.18	0.99	2.3711	5.09E-07	7.11E-04	Structural elements; cell exterior; surface polysaccharides/antigens	NA	<i>NonClass</i>
6915	0.78	2.26	1.5161	5.86E-04	4.28E-02	Hypothetical protein	NA	<i>NonClass</i>

ID gen	Promedio _C(CPM)	Promedio _L(CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre de gen	Categoría COG
6954	0.89	3.94	2.1263	2.36E-06	1.65E-03	Structural elements; cell exterior; surface polysaccharides/antigens	NA	NonClass
7043	1.61	5.82	1.8434	2.47E-05	7.34E-03	Hypothetical protein	NA	NonClass
7264	5.13	32.50	2.6585	9.98E-07	1.01E-03	Hypothetical protein	<i>gp17a</i>	S
7265	1.48	8.05	2.4242	1.95E-04	2.24E-02	Hypothetical protein	NA	L
7292	5.80	20.08	1.7910	6.19E-05	1.12E-02	Hypothetical protein	NA	K
7323	1.98	6.90	1.7883	1.89E-04	2.21E-02	Hypothetical protein	NA	NonClass
7331	2.78	11.65	2.0582	1.64E-05	5.55E-03	Hypothetical protein	NA	S
7334	0.20	1.11	2.3376	1.54E-05	5.39E-03	Hypothetical protein	NA	NonClass
7946	91.33	249.33	1.4488	6.45E-04	4.47E-02	Hypothetical protein	NA	NonClass
8061	2.91	10.35	1.8220	6.97E-05	1.20E-02	Flagellar protein FlaG	<i>flagG</i>	N
8647	12.73	37.79	1.5705	3.69E-04	3.19E-02	Oxidoreductase, short-chain dehydrogenase/reductase family	NA	I
8648	13.87	46.02	1.7311	2.23E-04	2.39E-02	2,4'-dihydroxyacetophenone dioxygenase	<i>dapDO</i>	S
8690	29.84	130.54	2.1284	1.92E-05	6.24E-03	Branched-chain amino acid transport system permease protein LivM (TC 3,A,1,4,1)	<i>livM</i>	E
8691	20.93	262.95	3.6518	2.49E-16	1.22E-12	Putative protein-S-isoprenylcysteine methyltransferase	NA	O
8692	95.43	1788.87	4.2285	8.32E-25	8.14E-21	Thioredoxin reductase	<i>trxB</i>	C
8721	100.81	226.51	1.1676	7.36E-04	4.96E-02	Putative 2-pyrone-4,6-dicarboxylic acid hydrolase	<i>ligL</i>	S
9271	4.61	14.87	1.6882	1.01E-05	4.30E-03	Chromate resistance protein ChrB	<i>chrB</i>	S

ID gen	Promedio _C(CPM)	Promedio _L(CPM)	logFC	Valor p	FDR	Descripción del gen	Nombre de gen	Categoría COG
9275	7.97	20.51	1.3628	3.78E-04	3.25E-02	Hypothetical protein	NA	T
9522	27.01	88.95	1.7189	6.15E-05	1.12E-02	Hypothetical protein. 60% identity with copQ Copper resistance protein	NA	<i>NonClass</i>
9525	27.94	135.00	2.2712	4.12E-05	9.31E-03	Hypothetical protein	NA	<i>NonClass</i>
9526	2.69	13.07	2.2699	4.72E-05	9.71E-03	Hypothetical protein	NA	<i>NonClass</i>
9750	4.96	16.87	1.7659	4.16E-05	9.31E-03	Flavodoxin reductases (ferredoxin-NADPH reductases) family 1; Vanillate O-demethylase oxidoreductase	<i>vanA</i>	C
9794	5.09	21.60	2.0821	1.22E-05	4.78E-03	Structural elements; cell exterior; surface polysaccharides/antigens	NA	<i>NonClass</i>

BIBLIOGRAFÍA

- [1] S. Burén, G. López-Torrejón, and L. M. Rubio, “Extreme bioengineering to meet the nitrogen challenge,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 36, pp. 8849–8851, Sep. 2018, doi: 10.1073/pnas.1812247115.
- [2] X. Mu and Y. Chen, “The physiological response of photosynthesis to nitrogen deficiency,” *Plant Physiology and Biochemistry*, vol. 158, pp. 76–82, Jan. 2021, doi: 10.1016/j.plaphy.2020.11.019.
- [3] L. Philippot and J. C. Germon, “Contribution of Bacteria to Initial Input and Cycling of Nitrogen in Soils,” in *Microorganisms in Soils: Roles in Genesis and Functions*, Berlin/Heidelberg: Springer-Verlag, pp. 159–176. doi: 10.1007/3-540-26609-7_8.
- [4] F. J. de Bruijn, “The Quest for Biological Nitrogen Fixation in Cereals: A Perspective and Prospective,” in *Biological Nitrogen Fixation*, Wiley, 2015, pp. 1087–1101. doi: 10.1002/9781119053095.ch108.
- [5] S. Burén, G. López-Torrejón, and L. M. Rubio, “Extreme bioengineering to meet the nitrogen challenge,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 36, pp. 8849–8851, Sep. 2018, doi: 10.1073/pnas.1812247115.
- [6] “byjus.com/biology.”
- [7] R. Sammauria, S. Kumawat, P. Kumawat, J. Singh, and T. K. Jatwa, “Microbial inoculants: potential tool for sustainability of agricultural production systems,” *Arch Microbiol*, vol. 202, no. 4, pp. 677–693, May 2020, doi: 10.1007/s00203-019-01795-w.
- [8] P. Irisarri *et al.*, “Editorial: Maximizing Nitrogen Fixation in Legumes as a Tool for Sustainable Agriculture Intensification,” *Frontiers in Agronomy*, vol. 3, Nov. 2021, doi: 10.3389/fagro.2021.796717.
- [9] T. Rütting, H. Aronsson, and S. Delin, “Efficient use of nitrogen in agriculture,” *Nutr Cycl Agroecosyst*, vol. 110, no. 1, pp. 1–5, Jan. 2018, doi: 10.1007/s10705-017-9900-8.
- [10] C. Franche, K. Lindström, and C. Elmerich, “Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants,” *Plant Soil*, vol. 321, no. 1–2, pp. 35–59, Aug. 2009, doi: 10.1007/s11104-008-9833-8.
- [11] L. Lloret and E. Martínez-Romero, “[Evolution and phylogeny of rhizobia].,” *Rev Latinoam Microbiol (1958)*, vol. 47, no. 1–2, pp. 43–60, 2005.
- [12] S. Fields, “Global Nitrogen: Cycling out of Control,” *Environ Health Perspect*, vol. 112, no. 10, Jul. 2004, doi: 10.1289/ehp.112-a556.
- [13] K. Guo, J. Yang, N. Yu, L. Luo, and E. Wang, “Biological nitrogen fixation in cereal crops: Progress, strategies, and perspectives,” *Plant Commun*, vol. 4, no. 2, p. 100499, Mar. 2023, doi: 10.1016/j.xplc.2022.100499.
- [14] J. Natzke, J. Noar, and J. M. Bruno-Bárcena, “Azotobacter vinelandii Nitrogenase Activity, Hydrogen Production, and Response to Oxygen Exposure,” *Appl Environ Microbiol*, vol. 84, no. 16, Aug. 2018, doi: 10.1128/AEM.01208-18.
- [15] A. Nonaka *et al.*, “Accessory Proteins of the Nitrogenase Assembly, NifW, NifX/NafY, and NifZ, Are Essential for Diazotrophic Growth in the Nonheterocystous Cyanobacterium *Leptolyngbya boryana*,” *Front Microbiol*, vol. 10, Mar. 2019, doi: 10.3389/fmicb.2019.00495.

- [16] R. Dixon and D. Kahn, “Genetic regulation of biological nitrogen fixation,” *Nat Rev Microbiol*, vol. 2, no. 8, pp. 621–631, Aug. 2004, doi: 10.1038/nrmicro954.
- [17] D. B. Layzell, L. Diaz del Castillo, S. Hunt, M. Kuzma, O. Van Cauwenberghe, and I. Oresnik, “The Regulation of Oxygen and Its Role in Regulating Nodule Metabolism,” 1993, pp. 393–398. doi: 10.1007/978-94-017-2416-6_39.
- [18] W. A. Rice and E. A. Paul, “The organisms and biological processes involved in asymbiotic nitrogen fixation in waterlogged soil amended with straw,” *Can J Microbiol*, vol. 18, no. 6, pp. 715–723, Jun. 1972, doi: 10.1139/m72-114.
- [19] C. Poza-Carrión, C. Echavarrri-Erasun, and L. M. Rubio, “Regulation of nif Gene Expression in *Azotobacter vinelandii*,” in *Biological Nitrogen Fixation*, Wiley, 2015, pp. 99–108. doi: 10.1002/9781119053095.ch9.
- [20] M. David *et al.*, “Cascade regulation of nif gene expression in *Rhizobium meliloti*,” *Cell*, vol. 54, no. 5, pp. 671–683, Aug. 1988, doi: 10.1016/S0092-8674(88)80012-6.
- [21] W. Zhang, Y. Chen, K. Huang, F. Wang, and Z. Mei, “Molecular Mechanism and Agricultural Application of the NifA–NifL System for Nitrogen Fixation,” *Int J Mol Sci*, vol. 24, no. 2, p. 907, Jan. 2023, doi: 10.3390/ijms24020907.
- [22] R. H. M. Op den Camp *et al.*, “Nonlegume *Parasponia andersonii* Deploys a Broad *Rhizobium* Host Range Strategy Resulting in Largely Variable Symbiotic Effectiveness,” *Molecular Plant-Microbe Interactions®*, vol. 25, no. 7, pp. 954–963, Jul. 2012, doi: 10.1094/MPMI-11-11-0304.
- [23] S. G. Pueppke and W. J. Broughton, “*Rhizobium* sp. Strain NGR234 and *R. fredii* USDA257 Share Exceptionally Broad, Nested Host Ranges,” *Molecular Plant-Microbe Interactions®*, vol. 12, no. 4, pp. 293–318, Apr. 1999, doi: 10.1094/MPMI.1999.12.4.293.
- [24] M. Andrews and M. E. Andrews, “Specificity in Legume-Rhizobia Symbioses,” *Int J Mol Sci*, vol. 18, no. 4, p. 705, Mar. 2017, doi: 10.3390/ijms18040705.
- [25] J.-H. Keet, A. G. Ellis, C. Hui, and J. J. Le Roux, “Legume–rhizobium symbiotic promiscuity and effectiveness do not affect plant invasiveness,” *Ann Bot*, vol. 119, no. 8, pp. 1319–1331, Jun. 2017, doi: 10.1093/aob/mcx028.
- [26] S. L. Dwivedi *et al.*, “Advances in Host Plant and *Rhizobium* Genomics to Enhance Symbiotic Nitrogen Fixation in Grain Legumes,” 2015, pp. 1–116. doi: 10.1016/bs.agron.2014.09.001.
- [27] M. M. Klock, L. G. Barrett, P. H. Thrall, and K. E. Harms, “Host promiscuity in symbiont associations can influence exotic legume establishment and colonization of novel ranges,” *Divers Distrib*, vol. 21, no. 10, pp. 1193–1203, Oct. 2015, doi: 10.1111/ddi.12363.
- [28] Q. Wang, J. Liu, and H. Zhu, “Genetic and Molecular Mechanisms Underlying Symbiotic Specificity in Legume-Rhizobium Interactions,” *Front Plant Sci*, vol. 9, Mar. 2018, doi: 10.3389/fpls.2018.00313.
- [29] L. Celador-Lera, E. Menéndez, A. Peix, J. M. Igual, E. Velázquez, and R. Rivas, “*Rhizobium zeae* sp. nov., isolated from maize (*Zea mays* L.) roots,” *Int J Syst Evol Microbiol*, vol. 67, no. 7, pp. 2306–2311, Jul. 2017, doi: 10.1099/ijsem.0.001944.
- [30] A. Soenens and J. Imperial, “Novel, non-symbiotic isolates of *Neorhizobium* from a dryland agricultural soil,” *PeerJ*, vol. 6, p. e4776, May 2018, doi: 10.7717/peerj.4776.
- [31] C. Masson-Boivin, E. Giraud, X. Perret, and J. Batut, “Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes?,” *Trends Microbiol*, vol. 17, no. 10, pp. 458–466, Oct. 2009, doi: 10.1016/j.tim.2009.07.004.

- [32] W. F. Chen, E. T. Wang, Z. J. Ji, and J. J. Zhang, “Recent development and new insight of diversification and symbiosis specificity of legume rhizobia: mechanism and application,” *J Appl Microbiol*, vol. 131, no. 2, pp. 553–563, Aug. 2021, doi: 10.1111/jam.14960.
- [33] A. Hördt *et al.*, “Analysis of 1,000+ Type-Strain Genomes Substantially Improves Taxonomic Classification of Alphaproteobacteria,” *Front Microbiol*, vol. 11, Apr. 2020, doi: 10.3389/fmicb.2020.00468.
- [34] K. Lindström and S. A. Mousavi, “Effectiveness of nitrogen fixation in rhizobia,” *Microb Biotechnol*, vol. 13, no. 5, pp. 1314–1335, Sep. 2020, doi: 10.1111/1751-7915.13517.
- [35] P. M. de Lajudie *et al.*, “Minimal standards for the description of new genera and species of rhizobia and agrobacteria,” *Int J Syst Evol Microbiol*, vol. 69, no. 7, pp. 1852–1863, Jul. 2019, doi: 10.1099/ijsem.0.003426.
- [36] P. Estrada-de los Santos *et al.*, “Whole genome analyses suggests that Burkholderia sensu lato contains two additional novel genera (Mycetohabitans gen. nov., and Trinickia gen. nov.): Implications for the evolution of diazotrophy and nodulation in the Burkholderiaceae,” *Genes (Basel)*, vol. 9, no. 8, 2018, doi: 10.3390/genes9080389.
- [37] X. Liu *et al.*, “Diversity and Geographic Distribution of Microsymbionts Associated With Invasive Mimosa Species in Southern China,” *Front Microbiol*, vol. 11, Oct. 2020, doi: 10.3389/fmicb.2020.563389.
- [38] R. Platero *et al.*, “Novel Cupriavidus Strains Isolated from Root Nodules of Native Uruguayan Mimosa Species,” *Appl Environ Microbiol*, vol. 82, no. 11, pp. 3150–3164, Jun. 2016, doi: 10.1128/AEM.04142-15.
- [39] R. P. N. Mishra *et al.*, “Genetic diversity of Mimosa pudica rhizobial symbionts in soils of French Guiana: investigating the origin and diversity of Burkholderia phymatum and other beta-rhizobia,” *FEMS Microbiol Ecol*, vol. 79, no. 2, pp. 487–503, Feb. 2012, doi: 10.1111/j.1574-6941.2011.01235.x.
- [40] R. Melkonian *et al.*, “The geographical patterns of symbiont diversity in the invasive legume *Mimosa pudica* can be explained by the competitiveness of its symbionts and by the host genotype,” *Environ Microbiol*, vol. 16, no. 7, pp. 2099–2111, Jul. 2014, doi: 10.1111/1462-2920.12286.
- [41] P. Gyaneshwar *et al.*, “Legume-Nodulating Betaproteobacteria: Diversity, Host Range, and Future Prospects,” *Molecular Plant-Microbe Interactions®*, vol. 24, no. 11, pp. 1276–1288, Nov. 2011, doi: 10.1094/MPMI-06-11-0172.
- [42] C. Taulé *et al.*, “New Betaproteobacterial Rhizobium Strains Able To Efficiently Nodulate Parapiptadenia rigida (Benth.) Brenan,” *Appl Environ Microbiol*, vol. 78, no. 6, pp. 1692–1700, Mar. 2012, doi: 10.1128/AEM.06215-11.
- [43] K. da Silva, L. A. Florentino, K. B. da Silva, E. de Brandt, P. Vandamme, and F. M. de Souza Moreira, “Cupriavidus necator isolates are able to fix nitrogen in symbiosis with different legume species,” *Syst Appl Microbiol*, vol. 35, no. 3, pp. 175–182, May 2012, doi: 10.1016/j.syapm.2011.10.005.
- [44] G. Garau, R. J. Yates, P. Deiana, and J. G. Howieson, “Novel strains of nodulating Burkholderia have a role in nitrogen fixation with papilionoid herbaceous legumes adapted to acid, infertile soils,” *Soil Biol Biochem*, vol. 41, no. 1, pp. 125–134, Jan. 2009, doi: 10.1016/j.soilbio.2008.10.011.
- [45] L. Moulin, A. Munive, B. Dreyfus, and C. Boivin-Masson, “Nodulation of legumes by members of the β -subclass of Proteobacteria,” *Nature*, vol. 411, no. 6840, pp. 948–950, Jun. 2001, doi: 10.1038/35082070.

- [46] A. Sawana, M. Adeolu, and R. S. Gupta, “Molecular signatures and phylogenomic analysis of the genus Burkholderia: proposal for division of this genus into the emended genus Burkholderia containing pathogenic organisms and a new genus Paraburkholderia gen. nov. harboring environmental species,” *Front Genet*, vol. 5, Dec. 2014, doi: 10.3389/fgene.2014.00429.
- [47] P. Vandamme, J. Goris, W.-M. Chen, P. de Vos, and A. Willems, “Burkholderia tuberum sp. nov. and Burkholderia phymatum sp. nov., Nodulate the Roots of Tropical Legumes,” *Syst Appl Microbiol*, vol. 25, no. 4, pp. 507–512, Jan. 2002, doi: 10.1078/07232020260517634.
- [48] W. M. Chen *et al.*, “Ralstonia taiwanensis sp. nov., isolated from root nodules of Mimosa species and sputum of a cystic fibrosis patient.,” *Int J Syst Evol Microbiol*, vol. 51, no. 5, pp. 1729–1735, Sep. 2001, doi: 10.1099/00207713-51-5-1729.
- [49] P. Vandamme, “Taxonomy of the genus Cupriavidus: a tale of lost and found,” *Int J Syst Evol Microbiol*, vol. 54, no. 6, pp. 2285–2289, Nov. 2004, doi: 10.1099/ijms.0.63247-0.
- [50] T. Feng, K. H. Kim, J. Oh, and C. O. Jeon, “Cupriavidus lacunae sp. nov., isolated from pond-side soil,” *Antonie Van Leeuwenhoek*, vol. 112, no. 4, pp. 543–551, Apr. 2019, doi: 10.1007/s10482-018-1187-5.
- [51] Md. A. Hashem, Md. A. Momen, M. Hasan, Md. S. Nur-A-Tomal, and Md. H. R. Sheikh, “Chromium removal from tannery wastewater using Syzygium cumini bark adsorbent,” *International Journal of Environmental Science and Technology*, vol. 16, no. 3, pp. 1395–1404, Mar. 2019, doi: 10.1007/s13762-018-1714-y.
- [52] I. Arroyo-Herrera *et al.*, “Cupriavidus agavae sp. nov., a species isolated from Agave L. rhizosphere in northeast Mexico,” *Int J Syst Evol Microbiol*, vol. 70, no. 7, pp. 4165–4170, Jul. 2020, doi: 10.1099/ijsem.0.004263.
- [53] A. Ravi *et al.*, “Plant Beneficial Features and Application of Paraburkholderia sp. NhPBG1 Isolated from Pitcher of Nepenthes hamblack,” *Probiotics Antimicrob Proteins*, vol. 13, no. 1, pp. 32–39, Feb. 2021, doi: 10.1007/s12602-020-09665-4.
- [54] J. Butler, S. D. Kelly, K. J. Muddiman, A. Besinis, and M. Upton, “Hospital sink traps as a potential source of the emerging multidrug-resistant pathogen Cupriavidus pauculus: characterization and draft genome sequence of strain MF1,” *J Med Microbiol*, vol. 71, no. 2, Feb. 2022, doi: 10.1099/jmm.0.001501.
- [55] G. KENICER, “Legumes of the World. Edited by G. Lewis, B. Schrire, B. MacKinder & M. Lock. Royal Botanic Gardens, Kew. 2005. xiv + 577pp., colour photographs & line drawings. ISBN 1 900 34780 6. £55.00 (hardback).,” *Edinb J Bot*, vol. 62, no. 3, pp. 195–196, May 2006, doi: 10.1017/S0960428606190198.
- [56] N. Azani *et al.*, “A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny: The Legume Phylogeny Working Group (LPWG),” *Taxon*, vol. 66, no. 1, pp. 44–77, Feb. 2017, doi: 10.12705/661.3.
- [57] P. Cardoso *et al.*, “Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components,” *J Biogeogr*, vol. 41, no. 4, pp. 749–761, Apr. 2014, doi: 10.1111/jbi.12239.
- [58] E. Velazquez, L. R. Silva, and A. Peix, “Legumes: A Healthy and Ecological Source of Flavonoids,” *Curr Nutr Food Sci*, vol. 6, no. 2, pp. 109–144, May 2010, doi: 10.2174/157340110791233247.
- [59] D. Rubiales and A. Mikic, “Introduction: Legumes in Sustainable Agriculture,” *CRC Crit Rev Plant Sci*, vol. 34, no. 1–3, pp. 2–3, Jun. 2015, doi: 10.1080/07352689.2014.897896.

- [60] J. I. Sprent, J. K. Ardley, and E. K. James, "From North to South: A latitudinal look at legume nodulation processes," *South African Journal of Botany*, vol. 89, pp. 31–41, Nov. 2013, doi: 10.1016/j.sajb.2013.06.011.
- [61] F. B. dos Reis Jr *et al.*, "Nodulation and nitrogen fixation by *Mimosa* spp. in the Cerrado and Caatinga biomes of Brazil," *New Phytologist*, vol. 186, no. 4, pp. 934–946, Jun. 2010, doi: 10.1111/j.1469-8137.2010.03267.x.
- [62] F. B. dos Reis Jr *et al.*, "Nodulation and nitrogen fixation by *Mimosa* spp. in the Cerrado and Caatinga biomes of Brazil," *New Phytologist*, vol. 186, no. 4, pp. 934–946, Jun. 2010, doi: 10.1111/j.1469-8137.2010.03267.x.
- [63] G. N. Elliott *et al.*, "*Burkholderia* spp. are the most competitive symbionts of *Mimosa*, particularly under N-limited conditions," *Environ Microbiol*, vol. 11, no. 4, pp. 762–778, Apr. 2009, doi: 10.1111/j.1462-2920.2008.01799.x.
- [64] F. B. dos Reis Jr *et al.*, "Nodulation and nitrogen fixation by *Mimosa* spp. in the Cerrado and Caatinga biomes of Brazil," *New Phytologist*, vol. 186, no. 4, pp. 934–946, Jun. 2010, doi: 10.1111/j.1469-8137.2010.03267.x.
- [65] R. Platero *et al.*, "Novel Cupriavidus Strains Isolated from Root Nodules of Native Uruguayan Mimosa Species," *Appl Environ Microbiol*, vol. 82, no. 11, pp. 3150–3164, Jun. 2016, doi: 10.1128/AEM.04142-15.
- [66] C. H. Calvey *et al.*, "Improving growth of Cupriavidus necator H16 on formate using adaptive laboratory evolution-informed engineering," *Metab Eng*, vol. 75, pp. 78–90, Jan. 2023, doi: 10.1016/j.ymben.2022.10.016.
- [67] O. J. Kweon, Y. K. Lim, H. R. Kim, T.-H. Kim, S. Ha, and M.-K. Lee, "Isolation of a novel species in the genus Cupriavidus from a patient with sepsis using whole genome sequencing," *PLoS One*, vol. 15, no. 5, p. e0232850, May 2020, doi: 10.1371/journal.pone.0232850.
- [68] D. C. Gillan *et al.*, "Paleomicrobiology to investigate copper resistance in bacteria: isolation and description of Cupriavidus necator B9 in the soil of a medieval foundry," *Environ Microbiol*, vol. 19, no. 2, pp. 770–787, Feb. 2017, doi: 10.1111/1462-2920.13645.
- [69] L.-N. Sun *et al.*, "Cupriavidus nantongensis sp. nov., a novel chlorpyrifos-degrading bacterium isolated from sludge," *Int J Syst Evol Microbiol*, vol. 66, no. 6, pp. 2335–2341, Jun. 2016, doi: 10.1099/ijsem.0.001034.
- [70] A. Kugler *et al.*, "Bioremediation of copper in sediments from a constructed wetland ex situ with the novel bacterium Cupriavidus basilensis SRS," *Sci Rep*, vol. 12, no. 1, p. 17615, Oct. 2022, doi: 10.1038/s41598-022-20930-0.
- [71] Z. Zhang, W. Deng, S. Wang, L. Xu, L. Yan, and P. Liao, "First case report of infection caused by Cupriavidus gilardii in a non-immunocompromised Chinese patient," *IDCases*, vol. 10, pp. 127–129, 2017, doi: 10.1016/j.idcr.2017.10.009.
- [72] A. Klonowska *et al.*, "Biodiversity of Mimosa pudica rhizobial symbionts (Cupriavidus taiwanensis, Rhizobium mesoamericanum) in New Caledonia and their adaptation to heavy metal-rich soils," *FEMS Microbiol Ecol*, vol. 81, no. 3, pp. 618–635, Sep. 2012, doi: 10.1111/j.1574-6941.2012.01393.x.
- [73] P. Chandrangsu, C. Rensing, and J. D. Helmann, "Metal homeostasis and resistance in bacteria," *Nat Rev Microbiol*, vol. 15, no. 6, pp. 338–350, Jun. 2017, doi: 10.1038/nrmicro.2017.15.

- [74] N. S. Makkar and L. E. Casida, “Cupriavidus necator gen. nov., sp. nov.; a Nonobligate Bacterial Predator of Bacteria in Soil,” *Int J Syst Bacteriol*, vol. 37, no. 4, pp. 323–326, Oct. 1987, doi: 10.1099/00207713-37-4-323.
- [75] J. M. Argüello, D. Raimunda, and T. Padilla-Benavides, “Mechanisms of copper homeostasis in bacteria,” *Front Cell Infect Microbiol*, vol. 3, 2013, doi: 10.3389/fcimb.2013.00073.
- [76] L. Macomber and J. A. Imlay, “The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 20, pp. 8344–8349, May 2009, doi: 10.1073/pnas.0812808106.
- [77] L. Maertens, N. Leys, J.-Y. Matroule, and R. Van Houdt, “The Transcriptomic Landscape of *Cupriavidus metallidurans* CH34 Acutely Exposed to Copper,” *Genes (Basel)*, vol. 11, no. 9, p. 1049, Sep. 2020, doi: 10.3390/genes11091049.
- [78] P. J. Janssen *et al.*, “The Complete Genome Sequence of *Cupriavidus metallidurans* Strain CH34, a Master Survivalist in Harsh and Anthropogenic Environments,” *PLoS One*, vol. 5, no. 5, p. e10433, May 2010, doi: 10.1371/journal.pone.0010433.
- [79] S. Maes *et al.*, “Biological Recovery of Platinum Complexes from Diluted Aqueous Streams by Axenic Cultures,” *PLoS One*, vol. 12, no. 1, p. e0169093, Jan. 2017, doi: 10.1371/journal.pone.0169093.
- [80] F. Montero-Silva, N. Durán, and M. Seeger, “Synthesis of extracellular gold nanoparticles using *Cupriavidus metallidurans* CH34 cells,” *IET Nanobiotechnol*, vol. 12, no. 1, pp. 40–46, Feb. 2018, doi: 10.1049/iet-nbt.2017.0185.
- [81] E. Fabiano, G. Gualtieri, C. Pritsch, G. Polla, and A. Arias, “Extent of high-affinity iron transport systems in field isolates of rhizobia,” *Plant Soil*, vol. 164, no. 2, pp. 177–185, Jul. 1994, doi: 10.1007/BF00010069.
- [82] F. Noya, A. Arias, and E. Fabiano, “Heme compounds as iron sources for nonpathogenic *Rhizobium* bacteria,” *J Bacteriol*, vol. 179, no. 9, pp. 3076–3078, May 1997, doi: 10.1128/jb.179.9.3076-3078.1997.
- [83] E. Fabiano and M. R. O’Brian, “Mechanisms and Regulation of Iron Homeostasis in the Rhizobia,” 2012, pp. 41–86. doi: 10.1007/978-94-007-5267-2_3.
- [84] I. J. Schalk, M. Hannauer, and A. Braud, “New roles for bacterial siderophores in metal transport and tolerance,” *Environ Microbiol*, vol. 13, no. 11, pp. 2844–2854, Nov. 2011, doi: 10.1111/j.1462-2920.2011.02556.x.
- [85] A. M. Timofeeva, M. R. Galyamova, and S. E. Sedykh, “Bacterial Siderophores: Classification, Biosynthesis, Perspectives of Use in Agriculture,” *Plants*, vol. 11, no. 22, p. 3065, Nov. 2022, doi: 10.3390/plants11223065.
- [86] D. Pérez-Pantoja *et al.*, “Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales,” *Environ Microbiol*, vol. 14, no. 5, pp. 1091–1117, May 2012, doi: 10.1111/j.1462-2920.2011.02613.x.
- [87] C. Ramos, I. Fernández, M. E. Suárez-Ojeda, and J. Carrera, “Inhibition of the anammox activity by aromatic compounds,” *Chemical Engineering Journal*, vol. 279, pp. 681–688, Nov. 2015, doi: 10.1016/j.cej.2015.05.071.
- [88] J. D. Haddock, “Aerobic Degradation of Aromatic Hydrocarbons: Enzyme Structures and Catalytic Mechanisms,” in *Handbook of Hydrocarbon and Lipid Microbiology*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 1057–1069. doi: 10.1007/978-3-540-77587-4_74.

- [89] D. Pérez-Pantoja, R. De La Iglesia, D. H. Pieper, and B. González, *Metabolic reconstruction of aromatic compounds degradation from the genome of the amazing pollutant-degrading bacterium Cupriavidus necator JMP134*, vol. 32, no. 5. 2008. doi: 10.1111/j.1574-6976.2008.00122.x.
- [90] S. Castro-Sowinski, S. Burdman, O. Matan, and Y. Okon, “Natural Functions of Bacterial Polyhydroxyalkanoates,” 2010, pp. 39–61. doi: 10.1007/978-3-642-03287-5_3.
- [91] F. Ertan, B. Keskinler, and A. Tanriseven, “Exploration of Cupriavidus necator ATCC 25207 for the Production of Poly(3-hydroxybutyrate) Using Acid Treated Beet Molasses,” *J Polym Environ*, vol. 29, no. 7, pp. 2111–2125, Jul. 2021, doi: 10.1007/s10924-020-02020-2.
- [92] M. E. Grigore, R. M. Grigorescu, L. Iancu, R.-M. Ion, C. Zaharia, and E. R. Andrei, “Methods of synthesis, properties and biomedical applications of polyhydroxyalkanoates: a review,” *J Biomater Sci Polym Ed*, vol. 30, no. 9, pp. 695–712, Jun. 2019, doi: 10.1080/09205063.2019.1605866.
- [93] B. McAdam, M. Brennan Fournet, P. McDonald, and M. Mojicevic, “Production of Polyhydroxybutyrate (PHB) and Factors Impacting Its Chemical and Mechanical Characteristics,” *Polymers (Basel)*, vol. 12, no. 12, p. 2908, Dec. 2020, doi: 10.3390/polym12122908.
- [94] A. Soumare *et al.*, “Exploiting Biological Nitrogen Fixation: A Route Towards a Sustainable Agriculture,” *Plants*, vol. 9, no. 8, p. 1011, Aug. 2020, doi: 10.3390/plants9081011.
- [95] S. M. Beyan, E. Wolde-meskel, and F. D. Dakora, “An assessment of plant growth and N₂ fixation in soybean genotypes grown in uninoculated soils collected from different locations in Ethiopia,” *Symbiosis*, vol. 75, no. 3, pp. 189–203, Jul. 2018, doi: 10.1007/s13199-018-0540-9.
- [96] C. Masson-Boivin and J. L. Sachs, “Symbiotic nitrogen fixation by rhizobia — the roots of a success story,” *Curr Opin Plant Biol*, vol. 44, pp. 7–15, Aug. 2018, doi: 10.1016/j.pbi.2017.12.001.
- [97] M. Bright and S. Bulgheresi, “A complex journey: transmission of microbial symbionts,” *Nat Rev Microbiol*, vol. 8, no. 3, pp. 218–230, Mar. 2010, doi: 10.1038/nrmicro2262.
- [98] M. Robledo *et al.*, “Role of Rhizobium endoglucanase CelC2 in cellulose biosynthesis and biofilm formation on plant roots and abiotic surfaces,” *Microb Cell Fact*, vol. 11, no. 1, p. 125, Dec. 2012, doi: 10.1186/1475-2859-11-125.
- [99] L. G. Castellani *et al.*, “Exopolysaccharide Characterization of Rhizobium favelukesii LPU83 and Its Role in the Symbiosis With Alfalfa,” *Front Plant Sci*, vol. 12, Feb. 2021, doi: 10.3389/fpls.2021.642576.
- [100] P. L. De Hoff, L. M. Brill, and A. M. Hirsch, “Plant lectins: the ties that bind in root symbiosis and plant defense,” *Molecular Genetics and Genomics*, vol. 282, no. 1, pp. 1–15, Jul. 2009, doi: 10.1007/s00438-009-0460-8.
- [101] L. A. Weston and U. Mathesius, “Flavonoids: Their Structure, Biosynthesis and Role in the Rhizosphere, Including Allelopathy,” *J Chem Ecol*, vol. 39, no. 2, pp. 283–297, Feb. 2013, doi: 10.1007/s10886-013-0248-5.
- [102] A. Kereszt, P. Mergaert, and E. Kondorosi, “Bacteroid Development in Legume Nodules: Evolution of Mutual Benefit or of Sacrificial Victims?,” *Molecular Plant-Microbe Interactions®*, vol. 24, no. 11, pp. 1300–1309, Nov. 2011, doi: 10.1094/MPMI-06-11-0152.

- [103] G. E. D. Oldroyd, J. D. Murray, P. S. Poole, and J. A. Downie, “The Rules of Engagement in the Legume-Rhizobial Symbiosis,” *Annu Rev Genet*, vol. 45, no. 1, pp. 119–144, Dec. 2011, doi: 10.1146/annurev-genet-110410-132549.
- [104] A. F. Haag *et al.*, “Molecular insights into bacteroid development during *Rhizobium*-legume symbiosis,” *FEMS Microbiol Rev*, p. n/a-n/a, Oct. 2012, doi: 10.1111/1574-6976.2012.12003.
- [105] E. Larrainzar *et al.*, “Hemoglobins in the legume–*Rhizobium* symbiosis,” *New Phytologist*, vol. 228, no. 2, pp. 472–484, Oct. 2020, doi: 10.1111/nph.16673.
- [106] X. Perret, C. Staehelin, and W. J. Broughton, “Molecular Basis of Symbiotic Promiscuity,” *Microbiology and Molecular Biology Reviews*, vol. 64, no. 1, pp. 180–201, Mar. 2000, doi: 10.1128/MMBR.64.1.180-201.2000.
- [107] P. del Cerro *et al.*, “Regulatory nodD1 and nodD2 genes of *Rhizobium tropici* strain CIAT 899 and their roles in the early stages of molecular signaling and host-legume nodulation,” *BMC Genomics*, vol. 16, no. 1, p. 251, Dec. 2015, doi: 10.1186/s12864-015-1458-8.
- [108] S. Kelly, J. T. Sullivan, Y. Kawaharada, S. Radutoiu, C. W. Ronson, and J. Stougaard, “Regulation of Nod factor biosynthesis by alternative NodD proteins at distinct stages of symbiosis provides additional compatibility scrutiny,” *Environ Microbiol*, vol. 20, no. 1, pp. 97–110, Jan. 2018, doi: 10.1111/1462-2920.14006.
- [109] V. Poinso, M. B. Crook, S. Erdn, F. Maillet, A. Bascaules, and J.-M. Ané, “New insights into Nod factor biosynthesis: Analyses of chitooligomers and lipo-chitooligomers of *Rhizobium* sp. IRBG74 mutants,” *Carbohydr Res*, vol. 434, pp. 83–93, Nov. 2016, doi: 10.1016/j.carres.2016.08.001.
- [110] S. R. Long, “*Rhizobium* symbiosis: nod factors in perspective.,” *Plant Cell*, vol. 8, no. 10, pp. 1885–1898, Oct. 1996, doi: 10.1105/tpc.8.10.1885.
- [111] T. Laeremans, “Isolation and Characterization of the *Rhizobium tropici* Nod Factor Sulfation Genes,” *Molecular Plant-Microbe Interactions*, vol. 9, no. 6, p. 492, 1996, doi: 10.1094/MPMI-9-0492.
- [112] P. Mergaert *et al.*, “**Fucosylation and arabinosylation of Nod factors in *Azorhizobium caulinodans* : involvement of *nolK nodZ* as well as *noeC* and/or downstream genes ,**” *Mol Microbiol*, vol. 21, no. 2, pp. 409–419, Jul. 1996, doi: 10.1046/j.1365-2958.1996.6451366.x.
- [113] V. Poinso, M. B. Crook, S. Erdn, F. Maillet, A. Bascaules, and J.-M. Ané, “New insights into Nod factor biosynthesis: Analyses of chitooligomers and lipo-chitooligomers of *Rhizobium* sp. IRBG74 mutants,” *Carbohydr Res*, vol. 434, pp. 83–93, Nov. 2016, doi: 10.1016/j.carres.2016.08.001.
- [114] E. Giraud *et al.*, “Legumes Symbioses: Absence of *Nod* Genes in Photosynthetic Bradyrhizobia,” *Science (1979)*, vol. 316, no. 5829, pp. 1307–1312, Jun. 2007, doi: 10.1126/science.1139548.
- [115] B. Daubech *et al.*, “*noeM*, a New Nodulation Gene Involved in the Biosynthesis of Nod Factors with an Open-Chain Oxidized Terminal Residue and in the Symbiosis with *Mimosa pudica*,” *Molecular Plant-Microbe Interactions®*, vol. 32, no. 12, pp. 1635–1648, Dec. 2019, doi: 10.1094/MPMI-06-19-0168-R.
- [116] M. C. Peck, R. F. Fisher, and S. R. Long, “Diverse Flavonoids Stimulate NodD1 Binding to *nod* Gene Promoters in *Sinorhizobium meliloti*,” *J Bacteriol*, vol. 188, no. 15, pp. 5417–5427, Aug. 2006, doi: 10.1128/JB.00376-06.

- [117] M. L. F. Ferreyra, P. Serra, and P. Casati, “Recent advances on the roles of flavonoids as plant protective molecules after <scp>UV</scp> and high light exposure,” *Physiol Plant*, vol. 173, no. 3, pp. 736–749, Nov. 2021, doi: 10.1111/ppl.13543.
- [118] A. Mouradov and G. Spangenberg, “Flavonoids: a metabolic network mediating plants adaptation to their real estate,” *Front Plant Sci*, vol. 5, Nov. 2014, doi: 10.3389/fpls.2014.00620.
- [119] G. Agati, E. Azzarello, S. Pollastri, and M. Tattini, “Flavonoids as antioxidants in plants: Location and functional significance,” *Plant Science*, vol. 196, pp. 67–76, Nov. 2012, doi: 10.1016/j.plantsci.2012.07.014.
- [120] J.-M. Lü, P. H. Lin, Q. Yao, and C. Chen, “Chemical and molecular mechanisms of antioxidants: experimental approaches and model systems,” *J Cell Mol Med*, vol. 14, no. 4, pp. 840–860, Apr. 2010, doi: 10.1111/j.1582-4934.2009.00897.x.
- [121] L. J. Sweetlove and A. R. Fernie, “The Spatial Organization of Metabolism Within the Plant Cell,” *Annu Rev Plant Biol*, vol. 64, no. 1, pp. 723–746, Apr. 2013, doi: 10.1146/annurev-arplant-050312-120233.
- [122] A. R. Fernie and T. Tohge, “The Genetics of Plant Metabolism,” *Annu Rev Genet*, vol. 51, no. 1, pp. 287–310, Nov. 2017, doi: 10.1146/annurev-genet-120116-024640.
- [123] A. Domaszewska-Szostek, M. Puzianowska-Kuźnicka, and A. Kuryłowicz, “Flavonoids in Skin Senescence Prevention and Treatment,” *Int J Mol Sci*, vol. 22, no. 13, p. 6814, Jun. 2021, doi: 10.3390/ijms22136814.
- [124] K. Yonekura-Sakakibara, Y. Higashi, and R. Nakabayashi, “The Origin and Evolution of Plant Flavonoid Metabolism,” *Front Plant Sci*, vol. 10, Aug. 2019, doi: 10.3389/fpls.2019.00943.
- [125] S. Alseekh, L. Perez de Souza, M. Benina, and A. R. Fernie, “The style and substance of plant flavonoid decoration; towards defining both structure and function,” *Phytochemistry*, vol. 174, p. 112347, Jun. 2020, doi: 10.1016/j.phytochem.2020.112347.
- [126] L. Perez de Souza, K. Garbowicz, Y. Brotman, T. Tohge, and A. R. Fernie, “The Acetate Pathway Supports Flavonoid and Lipid Biosynthesis in Arabidopsis,” *Plant Physiol*, vol. 182, no. 2, pp. 857–869, Feb. 2020, doi: 10.1104/pp.19.00683.
- [127] K. M. Davies *et al.*, “The Evolution of Flavonoid Biosynthesis: A Bryophyte Perspective,” *Front Plant Sci*, vol. 11, Feb. 2020, doi: 10.3389/fpls.2020.00007.
- [128] J. Zhao and R. A. Dixon, “The ‘ins’ and ‘outs’ of flavonoid transport,” *Trends Plant Sci*, vol. 15, no. 2, pp. 72–80, Feb. 2010, doi: 10.1016/j.tplants.2009.11.006.
- [129] E. Butelli *et al.*, “Noemi Controls Production of Flavonoid Pigments and Fruit Acidity and Illustrates the Domestication Routes of Modern Citrus Varieties,” *Current Biology*, vol. 29, no. 1, pp. 158-164.e2, Jan. 2019, doi: 10.1016/j.cub.2018.11.040.
- [130] M. Wang, J. Firrman, L. Liu, and K. Yam, “A Review on Flavonoid Apigenin: Dietary Intake, ADME, Antimicrobial Effects, and Interactions with Human Gut Microbiota,” *Biomed Res Int*, vol. 2019, pp. 1–18, Oct. 2019, doi: 10.1155/2019/7010467.
- [131] T. Tohge, L. Perez de Souza, and A. R. Fernie, “On the natural diversity of phenylacylated-flavonoid and their in planta function under conditions of stress,” *Phytochemistry Reviews*, vol. 17, no. 2, pp. 279–290, Apr. 2018, doi: 10.1007/s11101-017-9531-3.
- [132] T. Louveau and A. Osbourn, “The Sweet Side of Plant-Specialized Metabolism,” *Cold Spring Harb Perspect Biol*, vol. 11, no. 12, p. a034744, Dec. 2019, doi: 10.1101/cshperspect.a034744.

- [133] D. M. W. Ochieno *et al.*, “Rhizobium-Linked Nutritional and Phytochemical Changes Under Multitrophic Functional Contexts in Sustainable Food Systems,” *Front Sustain Food Syst*, vol. 4, Jan. 2021, doi: 10.3389/fsufs.2020.604396.
- [134] U. A. Hartwig, C. A. Maxwell, C. M. Joseph, and D. A. Phillips, “Chrysoeriol and Luteolin Released from Alfalfa Seeds Induce nod Genes in Rhizobium meliloti,” *Plant Physiol*, vol. 92, no. 1, pp. 116–122, Jan. 1990, doi: 10.1104/pp.92.1.116.
- [135] K. K. Compton, S. B. Hildreth, R. F. Helm, and B. E. Scharf, “An Updated Perspective on Sinorhizobium meliloti Chemotaxis to Alfalfa Flavonoids,” *Front Microbiol*, vol. 11, Oct. 2020, doi: 10.3389/fmicb.2020.581482.
- [136] S. Subramanian, G. Stacey, and O. Yu, “Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*,” *The Plant Journal*, vol. 48, no. 2, pp. 261–273, Oct. 2006, doi: 10.1111/j.1365-313X.2006.02874.x.
- [137] N. Nouwen, D. Gargani, and E. Giraud, “The Modification of the Flavonoid Naringenin by *Bradyrhizobium* sp. Strain ORS285 Changes the nod Genes Inducer Function to a Growth Stimulator,” *Molecular Plant-Microbe Interactions*®, vol. 32, no. 11, pp. 1517–1525, Nov. 2019, doi: 10.1094/MPMI-05-19-0133-R.
- [138] C.-W. Liu and J. Murray, “The Role of Flavonoids in Nodulation Host-Range Specificity: An Update,” *Plants*, vol. 5, no. 3, p. 33, Aug. 2016, doi: 10.3390/plants5030033.
- [139] C. Amadou *et al.*, “Genome sequence of the β -rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia,” *Genome Res*, vol. 18, no. 9, pp. 1472–1483, Sep. 2008, doi: 10.1101/gr.076448.108.
- [140] L. G. Castellani *et al.*, “Exopolysaccharide Characterization of Rhizobium favelukesii LPU83 and Its Role in the Symbiosis With Alfalfa,” *Front Plant Sci*, vol. 12, Feb. 2021, doi: 10.3389/fpls.2021.642576.
- [141] S. Acosta-Jurado, F. Fuentes-Romero, J.-E. Ruiz-Sainz, M. Janczarek, and J.-M. Vinardell, “Rhizobial Exopolysaccharides: Genetic Regulation of Their Synthesis and Relevance in Symbiosis with Legumes,” *Int J Mol Sci*, vol. 22, no. 12, p. 6233, Jun. 2021, doi: 10.3390/ijms22126233.
- [142] E. Limpens and T. Bisseling, “Signaling in symbiosis,” *Curr Opin Plant Biol*, vol. 6, no. 4, pp. 343–350, Aug. 2003, doi: 10.1016/S1369-5266(03)00068-2.
- [143] E. B. Madsen *et al.*, “A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals,” *Nature*, vol. 425, no. 6958, pp. 637–640, Oct. 2003, doi: 10.1038/nature02045.
- [144] S. Radutoiu *et al.*, “Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases,” *Nature*, vol. 425, no. 6958, pp. 585–592, Oct. 2003, doi: 10.1038/nature02039.
- [145] A. Broghammer *et al.*, “Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 34, pp. 13859–13864, Aug. 2012, doi: 10.1073/pnas.1205171109.
- [146] J. Mierziak, K. Kostyn, and A. Kulma, “Flavonoids as Important Molecules of Plant Interactions with the Environment,” *Molecules*, vol. 19, no. 10, pp. 16240–16265, Oct. 2014, doi: 10.3390/molecules191016240.
- [147] S. Radutoiu *et al.*, “LysM domains mediate lipochitin–oligosaccharide recognition and Nfr genes extend the symbiotic host range,” *EMBO J*, vol. 26, no. 17, pp. 3923–3935, Sep. 2007, doi: 10.1038/sj.emboj.7601826.

- [148] A. Skorupska, M. Janczarek, M. Marczak, A. Mazur, and J. Król, “Rhizobial exopolysaccharides: genetic control and symbiotic functions,” *Microb Cell Fact*, vol. 5, no. 1, p. 7, Dec. 2006, doi: 10.1186/1475-2859-5-7.
- [149] J. E. M. M. Wong *et al.*, “Structural signatures in EPR3 define a unique class of plant carbohydrate receptors,” *Nat Commun*, vol. 11, no. 1, p. 3797, Jul. 2020, doi: 10.1038/s41467-020-17568-9.
- [150] Y. Kawaharada *et al.*, “Receptor-mediated exopolysaccharide perception controls bacterial infection,” *Nature*, vol. 523, no. 7560, pp. 308–312, Jul. 2015, doi: 10.1038/nature14611.
- [151] Y. Kawaharada *et al.*, “Differential regulation of the Epr3 receptor coordinates membrane-restricted rhizobial colonization of root nodule primordia,” *Nat Commun*, vol. 8, no. 1, p. 14534, Feb. 2017, doi: 10.1038/ncomms14534.
- [152] H.-P. Cheng and G. C. Walker, “Succinoglycan Is Required for Initiation and Elongation of Infection Threads during Nodulation of Alfalfa by *Rhizobium meliloti*,” *J Bacteriol*, vol. 180, no. 19, pp. 5183–5191, Oct. 1998, doi: 10.1128/JB.180.19.5183-5191.1998.
- [153] K. M. Jones, “Increased Production of the Exopolysaccharide Succinoglycan Enhances Sinorhizobium meliloti 1021 Symbiosis with the Host Plant Medicago truncatula,” *J Bacteriol*, vol. 194, no. 16, pp. 4322–4331, Aug. 2012, doi: 10.1128/JB.00751-12.
- [154] S. J. Kelly *et al.*, “Conditional Requirement for Exopolysaccharide in the *Mesorhizobium–Lotus* Symbiosis,” *Molecular Plant-Microbe Interactions®*, vol. 26, no. 3, pp. 319–329, Mar. 2013, doi: 10.1094/MPMI-09-12-0227-R.
- [155] C. Zipfel and G. E. D. Oldroyd, “Plant signalling in symbiosis and immunity,” *Nature*, vol. 543, no. 7645, pp. 328–336, Mar. 2017, doi: 10.1038/nature22009.
- [156] D. Chinchilla *et al.*, “A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence,” *Nature*, vol. 448, no. 7152, pp. 497–500, Jul. 2007, doi: 10.1038/nature05999.
- [157] X. Zhang, P. N. Dodds, and M. Bernoux, “What Do We Know About NOD-Like Receptors in Plant Immunity?,” *Annu Rev Phytopathol*, vol. 55, no. 1, pp. 205–229, Aug. 2017, doi: 10.1146/annurev-phyto-080516-035250.
- [158] R. B. Abramovitch, J. C. Anderson, and G. B. Martin, “Bacterial elicitation and evasion of plant innate immunity,” *Nat Rev Mol Cell Biol*, vol. 7, no. 8, pp. 601–611, Aug. 2006, doi: 10.1038/nrm1984.
- [159] A. Teulet *et al.*, “The rhizobial type III effector ErnA confers the ability to form nodules in legumes,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 43, pp. 21758–21768, Oct. 2019, doi: 10.1073/pnas.1904456116.
- [160] J. Liu, J. H. Rice, N. Chen, T. J. Baum, and T. Hwezi, “Synchronization of Developmental Processes and Defense Signaling by Growth Regulating Transcription Factors,” *PLoS One*, vol. 9, no. 5, p. e98477, May 2014, doi: 10.1371/journal.pone.0098477.
- [161] K. da Silva, L. A. Florentino, K. B. da Silva, E. de Brandt, P. Vandamme, and F. M. de Souza Moreira, “Cupriavidus necator isolates are able to fix nitrogen in symbiosis with different legume species,” *Syst Appl Microbiol*, vol. 35, no. 3, pp. 175–182, May 2012, doi: 10.1016/j.syapm.2011.10.005.
- [162] R. Platero *et al.*, “Novel Cupriavidus Strains Isolated from Root Nodules of Native Uruguayan Mimosa Species,” *Appl Environ Microbiol*, vol. 82, no. 11, pp. 3150–3164, Jun. 2016, doi: 10.1128/AEM.04142-15.

- [163] S. E. De Meyer *et al.*, “High-quality permanent draft genome sequence of the Parapiptadenia rigida-nodulating Burkholderia sp. strain UYPR1.413,” *Stand Genomic Sci*, vol. 10, no. 1, p. 31, Dec. 2015, doi: 10.1186/s40793-015-0018-9.
- [164] C. Amadou *et al.*, “Genome sequence of the β -rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia,” *Genome Res*, vol. 18, no. 9, pp. 1472–1483, Sep. 2008, doi: 10.1101/gr.076448.108.
- [165] A. Iriarte, R. Platero, V. Romero, E. Fabiano, and J. R. Sotelo-Silveira, “Draft genome sequence of *Cupriavidus* UYMMa02A, a novel beta-rhizobium species,” *Genome Announc*, vol. 4, no. 6, 2016, doi: 10.1128/genomeA.01258-16.
- [166] C. Clerissi *et al.*, “Parallels between experimental and natural evolution of legume symbionts,” *Nat Commun*, vol. 9, no. 1, pp. 1–13, 2018, doi: 10.1038/s41467-018-04778-5.
- [167] F. Sanger, S. Nicklen, and A. R. Coulson, “DNA sequencing with chain-terminating inhibitors,” *Proceedings of the National Academy of Sciences*, vol. 74, no. 12, pp. 5463–5467, Dec. 1977, doi: 10.1073/pnas.74.12.5463.
- [168] A. M. Maxam and W. Gilbert, “A new method for sequencing DNA.,” *Proceedings of the National Academy of Sciences*, vol. 74, no. 2, pp. 560–564, Feb. 1977, doi: 10.1073/pnas.74.2.560.
- [169] Q. Long, C. Wu, X. Wang, L. Jiang, and J. Li, “A Multiobjective Genetic Algorithm Based on a Discrete Selection Procedure,” *Math Probl Eng*, vol. 2015, pp. 1–17, 2015, doi: 10.1155/2015/349781.
- [170] E. R. Mardis, “A decade’s perspective on DNA sequencing technology,” *Nature*, vol. 470, no. 7333, pp. 198–203, Feb. 2011, doi: 10.1038/nature09796.
- [171] J. Shendure and H. Ji, “Next-generation DNA sequencing,” *Nat Biotechnol*, vol. 26, no. 10, pp. 1135–1145, Oct. 2008, doi: 10.1038/nbt1486.
- [172] C. S. Pareek, R. Smoczynski, and A. Tretyn, “Sequencing technologies and genome sequencing,” *J Appl Genet*, vol. 52, no. 4, pp. 413–435, Nov. 2011, doi: 10.1007/s13353-011-0057-x.
- [173] D. C. Koboldt, K. M. Steinberg, D. E. Larson, R. K. Wilson, and E. R. Mardis, “The Next-Generation Sequencing Revolution and Its Impact on Genomics,” *Cell*, vol. 155, no. 1, pp. 27–38, Sep. 2013, doi: 10.1016/j.cell.2013.09.006.
- [174] M. Kchouk, J. F. Gibrat, and M. Elloumi, “Generations of Sequencing Technologies: From First to Next Generation,” *Biology and Medicine*, vol. 09, no. 03, 2017, doi: 10.4172/0974-8369.1000395.
- [175] S. Goodwin, J. D. McPherson, and W. R. McCombie, “Coming of age: ten years of next-generation sequencing technologies,” *Nat Rev Genet*, vol. 17, no. 6, pp. 333–351, Jun. 2016, doi: 10.1038/nrg.2016.49.
- [176] “www.illumina.com.”
- [177] E. E. Schadt, S. Turner, and A. Kasarskis, “A window into third generation sequencing,” *Hum Mol Genet*, vol. 20, no. 4, pp. 853–853, Feb. 2011, doi: 10.1093/hmg/ddq481.
- [178] A. Rhoads and K. F. Au, “PacBio Sequencing and Its Applications,” *Genomics Proteomics Bioinformatics*, vol. 13, no. 5, pp. 278–289, Oct. 2015, doi: 10.1016/j.gpb.2015.08.002.
- [179] G. C. diCenzo and T. M. Finan, “The Divided Bacterial Genome: Structure, Function, and Evolution,” *Microbiology and Molecular Biology Reviews*, vol. 81, no. 3, Sep. 2017, doi: 10.1128/MMBR.00019-17.

- [180] P. W. Harrison, R. P. J. Lower, N. K. D. Kim, and J. P. W. Young, “Introducing the bacterial ‘chromid’: not a chromosome, not a plasmid,” *Trends Microbiol*, vol. 18, no. 4, pp. 141–148, Apr. 2010, doi: 10.1016/j.tim.2009.12.010.
- [181] D. R. Boer *et al.*, “Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains,” *EMBO J*, vol. 28, no. 11, pp. 1666–1678, Jun. 2009, doi: 10.1038/emboj.2009.125.
- [182] S. Hiraga, “DYNAMIC LOCALIZATION OF BACTERIAL AND PLASMID CHROMOSOMES,” *Annu Rev Genet*, vol. 34, no. 1, pp. 21–59, Dec. 2000, doi: 10.1146/annurev.genet.34.1.21.
- [183] P. Koper, K. Żebracki, M. Marczak, A. Skorupska, and A. Mazur, “RepB proteins of the multipartite *Rhizobium leguminosarum* bv. *trifolii* genome discriminate between centromere-like *parS* sequences for plasmid segregational stability,” *Mol Microbiol*, vol. 102, no. 3, pp. 446–466, Nov. 2016, doi: 10.1111/mmi.13472.
- [184] G. C. diCenzo, A. Mengoni, and E. Perrin, “Chromids Aid Genome Expansion and Functional Diversification in the Family *Burkholderiaceae*,” *Mol Biol Evol*, vol. 36, no. 3, pp. 562–574, Mar. 2019, doi: 10.1093/molbev/msy248.
- [185] A. Klonowska *et al.*, “Novel heavy metal resistance gene clusters are present in the genome of *Cupriavidus neocaledonicus* STM 6070, a new species of *Mimosa pudica* microsymbiont isolated from heavy-metal-rich mining site soil,” *BMC Genomics*, vol. 21, no. 1, p. 214, Dec. 2020, doi: 10.1186/s12864-020-6623-z.
- [186] P. Alviz-Gazitua, R. E. Durán, F. A. Millacura, F. Cárdenas, L. A. Rojas, and M. Seeger, “*Cupriavidus metallidurans* CH34 Possesses Aromatic Catabolic Versatility and Degrades Benzene in the Presence of Mercury and Cadmium,” *Microorganisms*, vol. 10, no. 2, p. 484, Feb. 2022, doi: 10.3390/microorganisms10020484.
- [187] H. H. Zahran, “Plasmids impact on rhizobia-legumes symbiosis in diverse environments,” *Symbiosis*, vol. 73, no. 2, pp. 75–91, Oct. 2017, doi: 10.1007/s13199-017-0476-5.
- [188] C. Sánchez-Cañizares *et al.*, “Genomic Diversity in the Endosymbiotic Bacterium *Rhizobium leguminosarum*,” *Genes (Basel)*, vol. 9, no. 2, p. 60, Jan. 2018, doi: 10.3390/genes9020060.
- [189] J. Yang, L. Lan, Y. Jin, N. Yu, D. Wang, and E. Wang, “Mechanisms underlying legume–rhizobium symbioses,” *J Integr Plant Biol*, vol. 64, no. 2, pp. 244–267, Feb. 2022, doi: 10.1111/jipb.13207.
- [190] J. P. Ramsay, J. T. Sullivan, G. S. Stuart, I. L. Lamont, and C. W. Ronson, “Excision and transfer of the *Mesorhizobium loti* R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS,” *Mol Microbiol*, vol. 62, no. 3, pp. 723–734, Nov. 2006, doi: 10.1111/j.1365-2958.2006.05396.x.
- [191] C. Masson-Boivin, E. Giraud, X. Perret, and J. Batut, “Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes?,” *Trends Microbiol*, vol. 17, no. 10, pp. 458–66, Oct. 2009, doi: 10.1016/j.tim.2009.07.004.
- [192] B. A. Geddes, J. Kearsley, R. Morton, G. C. diCenzo, and T. M. Finan, *The genomes of rhizobia*, vol. 94. Elsevier Ltd, 2020. doi: 10.1016/bs.abr.2019.09.014.
- [193] C. Rodríguez-Espérón *et al.*, “Nodulation in the absence of nod genes induction: alternative mechanisms involved in the symbiotic interaction between *Cupriavidus* sp. UYMMa02A and *Mimosa pudica*,” *Environmental Sustainability*, vol. 6, no. 3, pp. 383–401, Jul. 2023, doi: 10.1007/s42398-023-00286-5.

- [194] C. Amadou, G. Pascal, and S. Mangenot, “Genome sequence of the β -rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia,” *Genome Res*, vol. 18, no. 9, pp. 1472–1483, Sep. 2008, doi: 10.1101/gr.076448.108.
- [195] S. E. De Meyer *et al.*, “Symbiotic Burkholderia Species Show Diverse Arrangements of *nif/fix* and *nod* Genes and Lack Typical High-Affinity Cytochrome *cbb3* Oxidase Genes.,” *Mol Plant Microbe Interact*, vol. 29, no. 8, pp. 609–19, 2016, doi: 10.1094/MPMI-05-16-0091-R.
- [196] L. Moulin *et al.*, “Complete Genome sequence of *Burkholderia phymatum* STM815T, a broad host range and efficient nitrogen-fixing symbiont of *Mimosa* species,” *Stand Genomic Sci*, vol. 9, no. 3, pp. 763–774, Mar. 2014, doi: 10.4056/sigs.4861021.
- [197] M. Lardi, Y. Liu, G. Purtschert, S. B. de Campos, and G. Pessi, “Transcriptome analysis of *Paraburkholderia phymatum* under Nitrogen starvation and during symbiosis with *Phaseolus Vulgaris*,” *Genes (Basel)*, vol. 8, no. 12, 2017, doi: 10.3390/genes8120389.
- [198] M. A. Matilla, F. Velando, D. Martín-Mora, E. Monteagudo-Cascales, and T. Krell, “A catalogue of signal molecules that interact with sensor kinases, chemoreceptors and transcriptional regulators,” *FEMS Microbiol Rev*, vol. 46, no. 1, Jan. 2022, doi: 10.1093/femsre/fuab043.
- [199] M. A. Matilla and T. Krell, “Targeting motility and chemotaxis as a strategy to combat bacterial pathogens,” *Microb Biotechnol*, Jun. 2023, doi: 10.1111/1751-7915.14306.
- [200] K. Wuichet and I. B. Zhulin, “Origins and Diversification of a Complex Signal Transduction System in Prokaryotes,” *Sci Signal*, vol. 3, no. 128, Jun. 2010, doi: 10.1126/scisignal.2000724.
- [201] B. L. Carroll and J. Liu, “Structural Conservation and Adaptation of the Bacterial Flagella Motor,” *Biomolecules*, vol. 10, no. 11, p. 1492, Oct. 2020, doi: 10.3390/biom10111492.
- [202] D. L. Guzmán-Herrador, A. Fernández-Gómez, and M. Llosa, “Recruitment of heterologous substrates by bacterial secretion systems for transkingdom translocation,” *Front Cell Infect Microbiol*, vol. 13, Mar. 2023, doi: 10.3389/fcimb.2023.1146000.
- [203] A. Filloux, “Bacterial protein secretion systems: Game of types,” *Microbiology (N Y)*, vol. 168, no. 5, May 2022, doi: 10.1099/mic.0.001193.
- [204] K. Kanonenberg, O. Spitz, I. N. Erenburg, T. Beer, and L. Schmitt, “Type I secretion system—it takes three and a substrate,” *FEMS Microbiol Lett*, vol. 365, no. 11, Jun. 2018, doi: 10.1093/femsle/fny094.
- [205] M. Nivaskumar and O. Francetic, “Type II secretion system: A magic beanstalk or a protein escalator,” *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1843, no. 8, pp. 1568–1577, Aug. 2014, doi: 10.1016/j.bbamcr.2013.12.020.
- [206] S. P. Howard *et al.*, “Structure and assembly of pilotin-dependent and -independent secretins of the type II secretion system,” *PLoS Pathog*, vol. 15, no. 5, p. e1007731, May 2019, doi: 10.1371/journal.ppat.1007731.
- [207] T. Palmer and B. C. Berks, “Moving folded proteins across the bacterial cell membrane,” *Microbiology (N Y)*, vol. 149, no. 3, pp. 547–556, Mar. 2003, doi: 10.1099/mic.0.25900-0.
- [208] L. Craig, K. T. Forest, and B. Maier, “Type IV pili: dynamics, biophysics and functional consequences,” *Nat Rev Microbiol*, vol. 17, no. 7, pp. 429–440, Jul. 2019, doi: 10.1038/s41579-019-0195-4.
- [209] M. Tomich, P. J. Planet, and D. H. Figurski, “The *tad* locus: postcards from the widespread colonization island,” *Nat Rev Microbiol*, vol. 5, no. 5, pp. 363–375, May 2007, doi: 10.1038/nrmicro1636.

- [210] D. A. Chiavelli, J. W. Marsh, and R. K. Taylor, “The Mannose-Sensitive Hemagglutinin of *Vibrio cholerae* Promotes Adherence to Zooplankton,” *Appl Environ Microbiol*, vol. 67, no. 7, pp. 3220–3225, Jul. 2001, doi: 10.1128/AEM.67.7.3220-3225.2001.
- [211] K. H. Piepenbrink, “DNA Uptake by Type IV Filaments,” *Front Mol Biosci*, vol. 6, Feb. 2019, doi: 10.3389/fmolb.2019.00001.
- [212] S. Wagner, I. Grin, S. Malmshemer, N. Singh, C. E. Torres-Vargas, and S. Westerhausen, “Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into eukaryotic host cells,” *FEMS Microbiol Lett*, vol. 365, no. 19, Oct. 2018, doi: 10.1093/femsle/fny201.
- [213] S. S. Abby and E. P. C. Rocha, “The Non-Flagellar Type III Secretion System Evolved from the Bacterial Flagellum and Diversified into Host-Cell Adapted Systems,” *PLoS Genet*, vol. 8, no. 9, p. e1002983, Sep. 2012, doi: 10.1371/journal.pgen.1002983.
- [214] A. J. Bogdanove *et al.*, “ **Unified nomenclature for broadly conserved *hrp* genes of phytopathogenic bacteria** ,” *Mol Microbiol*, vol. 20, no. 3, pp. 681–683, May 1996, doi: 10.1046/j.1365-2958.1996.5731077.x.
- [215] C. Marie, W. J. Broughton, and W. J. Deakin, “Rhizobium type III secretion systems: legume charmers or alarmers?,” *Curr Opin Plant Biol*, vol. 4, no. 4, pp. 336–342, Aug. 2001, doi: 10.1016/S1369-5266(00)00182-5.
- [216] T. R. D. Costa, L. Harb, P. Khara, L. Zeng, B. Hu, and P. J. Christie, “Type IV secretion systems: Advances in structure, function, and activation,” *Mol Microbiol*, vol. 115, no. 3, pp. 436–452, Mar. 2021, doi: 10.1111/mmi.14670.
- [217] Y. G. Li, B. Hu, and P. J. Christie, “Biological and Structural Diversity of Type IV Secretion Systems,” *Microbiol Spectr*, vol. 7, no. 2, Apr. 2019, doi: 10.1128/microbiolspec.PSIB-0012-2018.
- [218] E. Fan, N. Chauhan, D. B. R. K. G. Udatha, J. C. Leo, and D. Linke, “Type V Secretion Systems in Bacteria,” *Microbiol Spectr*, vol. 4, no. 1, Jan. 2016, doi: 10.1128/microbiolspec.VMBF-0009-2015.
- [219] F. Jacob-Dubuisson, J. Guérin, S. Baelen, and B. Clantin, “Two-partner secretion: as simple as it sounds?,” *Res Microbiol*, vol. 164, no. 6, pp. 583–595, Jul. 2013, doi: 10.1016/j.resmic.2013.03.009.
- [220] I. Meuskens, A. Saragliadis, J. C. Leo, and D. Linke, “Type V Secretion Systems: An Overview of Passenger Domain Functions,” *Front Microbiol*, vol. 10, May 2019, doi: 10.3389/fmicb.2019.01163.
- [221] F. Boyer, G. Fichant, J. Berthod, Y. Vandenbrouck, and I. Attree, “Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources?,” *BMC Genomics*, vol. 10, no. 1, p. 104, Dec. 2009, doi: 10.1186/1471-2164-10-104.
- [222] J. Wang, M. Brodmann, and M. Basler, “Assembly and Subcellular Localization of Bacterial Type VI Secretion Systems,” *Annu Rev Microbiol*, vol. 73, no. 1, pp. 621–638, Sep. 2019, doi: 10.1146/annurev-micro-020518-115420.
- [223] Y. Cherrak, N. Flaugnatti, E. Durand, L. Journet, and E. Cascales, “Structure and Activity of the Type VI Secretion System,” *Microbiol Spectr*, vol. 7, no. 4, Jul. 2019, doi: 10.1128/microbiolspec.PSIB-0031-2019.
- [224] M. Kanehisa, “KEGG: Kyoto Encyclopedia of Genes and Genomes,” *Nucleic Acids Res*, vol. 28, no. 1, pp. 27–30, Jan. 2000, doi: 10.1093/nar/28.1.27.

- [225] P. Rice, I. Longden, and A. Bleasby, “EMBOSS: the European Molecular Biology Open Software Suite,” *Trends Genet*, vol. 16, no. 6, pp. 276–7, Jun. 2000, doi: 10.1016/s0168-9525(00)02024-2.
- [226] B. Contreras-Moreira and P. Vinuesa, “GET_HOMOLOGUES, a Versatile Software Package for Scalable and Robust Microbial Pangenome Analysis,” *Appl Environ Microbiol*, vol. 79, no. 24, pp. 7696–7701, Dec. 2013, doi: 10.1128/AEM.02411-13.
- [227] K. Katoh and D. M. Standley, “MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability,” *Mol Biol Evol*, vol. 30, no. 4, pp. 772–780, Apr. 2013, doi: 10.1093/molbev/mst010.
- [228] A. E. Teschendorff, Y. Wang, N. L. Barbosa-Morais, J. D. Brenton, and C. Caldas, “A variational Bayesian mixture modelling framework for cluster analysis of gene-expression data,” *Bioinformatics*, vol. 21, no. 13, pp. 3025–3033, Jul. 2005, doi: 10.1093/bioinformatics/bti466.
- [229] M. C. LaFave, G. K. Varshney, and S. M. Burgess, “GeIST: a pipeline for mapping integrated DNA elements,” *Bioinformatics*, vol. 31, no. 19, pp. 3219–3221, Oct. 2015, doi: 10.1093/bioinformatics/btv350.
- [230] S. Mirarab, R. Reaz, Md. S. Bayzid, T. Zimmermann, M. S. Swenson, and T. Warnow, “ASTRAL: genome-scale coalescent-based species tree estimation,” *Bioinformatics*, vol. 30, no. 17, pp. i541–i548, Sep. 2014, doi: 10.1093/bioinformatics/btu462.
- [231] S. Koren, B. P. Walenz, K. Berlin, J. R. Miller, N. H. Bergman, and A. M. Phillippy, “Canu: Scalable and accurate long-read assembly via adaptive κ -mer weighting and repeat separation,” *Genome Res*, vol. 27, no. 5, pp. 722–736, 2017, doi: 10.1101/gr.215087.116.
- [232] R. K. Aziz *et al.*, “The RAST Server: Rapid Annotations using Subsystems Technology,” *BMC Genomics*, vol. 9, no. 1, p. 75, Dec. 2008, doi: 10.1186/1471-2164-9-75.
- [233] Y. Moriya, M. Itoh, S. Okuda, A. C. Yoshizawa, and M. Kanehisa, “KAAS: an automatic genome annotation and pathway reconstruction server,” *Nucleic Acids Res*, vol. 35, no. Web Server, pp. W182–W185, May 2007, doi: 10.1093/nar/gkm321.
- [234] M. Krzywinski *et al.*, “Circos: An information aesthetic for comparative genomics,” *Genome Res*, vol. 19, no. 9, pp. 1639–1645, Sep. 2009, doi: 10.1101/gr.092759.109.
- [235] J. Goris, K. T. Konstantinidis, J. A. Klappenbach, T. Coenye, P. Vandamme, and J. M. Tiedje, “DNA–DNA hybridization values and their relationship to whole-genome sequence similarities,” *Int J Syst Evol Microbiol*, vol. 57, no. 1, pp. 81–91, Jan. 2007, doi: 10.1099/ijs.0.64483-0.
- [236] F. CASSE, C. BOUCHER, J. S. JULLIOT, M. MICHEL, and J. DENARIE, “Identification and Characterization of Large Plasmids in *Rhizobium meliloti* using Agarose Gel Electrophoresis,” *J Gen Microbiol*, vol. 113, no. 2, pp. 229–242, Aug. 1979, doi: 10.1099/00221287-113-2-229.
- [237] W. F. Fricke, B. Kusian, and B. Bowien, “The Genome Organization of *Ralstonia eutropha* Strain H16 and Related Species of the Burkholderiaceae,” *Microb Physiol*, vol. 16, no. 1–2, pp. 124–135, 2009, doi: 10.1159/000142899.
- [238] N. Dubarry, F. Pasta, and D. Lane, “ParABS Systems of the Four Replicons of *Burkholderia cenocepacia*: New Chromosome Centromeres Confer Partition Specificity,” *J Bacteriol*, vol. 188, no. 4, pp. 1489–1496, Feb. 2006, doi: 10.1128/JB.188.4.1489-1496.2006.

- [239] M. Kamruzzaman, A. Y. Wu, and J. R. Iredell, “Biological Functions of Type II Toxin-Antitoxin Systems in Bacteria,” *Microorganisms*, vol. 9, no. 6, p. 1276, Jun. 2021, doi: 10.3390/microorganisms9061276.
- [240] C. Smillie, M. P. Garcillán-Barcia, M. V. Francia, E. P. C. Rocha, and F. de la Cruz, “Mobility of Plasmids,” *Microbiology and Molecular Biology Reviews*, vol. 74, no. 3, pp. 434–452, Sep. 2010, doi: 10.1128/MMBR.00020-10.
- [241] T. von Rozycki and D. H. Nies, “Cupriavidus metallidurans: evolution of a metal-resistant bacterium,” *Antonie Van Leeuwenhoek*, vol. 96, no. 2, pp. 115–139, Aug. 2009, doi: 10.1007/s10482-008-9284-5.
- [242] C. Díaz-Pérez, C. Cervantes, J. Campos-García, A. Julián-Sánchez, and H. Riveros-Rosas, “Phylogenetic analysis of the chromate ion transporter (CHR) superfamily,” *FEBS Journal*, vol. 274, no. 23, pp. 6215–6227, Dec. 2007, doi: 10.1111/j.1742-4658.2007.06141.x.
- [243] Y. Yang *et al.*, “Structural visualization of transcription activated by a multidrug-sensing MerR family regulator,” *Nat Commun*, vol. 12, no. 1, p. 2702, May 2021, doi: 10.1038/s41467-021-22990-8.
- [244] N. L. Brown, J. V. Stoyanov, S. P. Kidd, and J. L. Hobman, “The MerR family of transcriptional regulators,” *FEMS Microbiol Rev*, vol. 27, no. 2–3, pp. 145–163, Jun. 2003, doi: 10.1016/S0168-6445(03)00051-2.
- [245] M. R. O’Brian, “Perception and Homeostatic Control of Iron in the Rhizobia and Related Bacteria,” *Annu Rev Microbiol*, vol. 69, no. 1, pp. 229–245, Oct. 2015, doi: 10.1146/annurev-micro-091014-104432.
- [246] F. Battistoni, R. Platero, F. Noya, A. Arias, and E. Fabiano, “Intracellular Fe content influences nodulation competitiveness of Sinorhizobium meliloti strains as inocula of alfalfa,” *Soil Biol Biochem*, vol. 34, no. 5, pp. 593–597, May 2002, doi: 10.1016/S0038-0717(01)00215-2.
- [247] E. M. Brear, D. A. Day, and P. M. C. Smith, “Iron: an essential micronutrient for the legume-rhizobium symbiosis,” *Front Plant Sci*, vol. 4, 2013, doi: 10.3389/fpls.2013.00359.
- [248] G. Rudolph, H. Hennecke, and H. M. Fischer, “Beyond the fur paradigm: Iron-controlled gene expression in rhizobia,” *FEMS Microbiol Rev*, vol. 30, no. 4, pp. 631–648, 2006, doi: 10.1111/j.1574-6976.2006.00030.x.
- [249] M. Ellermann and J. C. Arthur, “Siderophore-mediated iron acquisition and modulation of host-bacterial interactions,” *Free Radic Biol Med*, vol. 105, pp. 68–78, Apr. 2017, doi: 10.1016/j.freeradbiomed.2016.10.489.
- [250] R. Mitra, T. Xu, G. Chen, H. Xiang, and J. Han, “An updated overview on the regulatory circuits of polyhydroxyalkanoates synthesis,” *Microb Biotechnol*, vol. 15, no. 5, pp. 1446–1470, May 2022, doi: 10.1111/1751-7915.13915.
- [251] L. Fröderberg, E. N. G. Houben, L. Baars, J. Luirink, and J.-W. de Gier, “Targeting and Translocation of Two Lipoproteins in Escherichia coli via the SRP/Sec/YidC Pathway,” *Journal of Biological Chemistry*, vol. 279, no. 30, pp. 31026–31032, Jul. 2004, doi: 10.1074/jbc.M403229200.
- [252] G. G. Sgro *et al.*, “Bacteria-Killing Type IV Secretion Systems,” *Front Microbiol*, vol. 10, May 2019, doi: 10.3389/fmicb.2019.01078.
- [253] R. P. Singh, G. M. Shelke, A. Kumar, and P. N. Jha, “Biochemistry and genetics of ACC deaminase: a weapon to ‘stress ethylene’ produced in plants,” *Front Microbiol*, vol. 6, Sep. 2015, doi: 10.3389/fmicb.2015.00937.

- [254] S. E. De Meyer *et al.*, “High-quality permanent draft genome sequence of the parapiptadenia rigida-nodulating cupriavidus sp. Strain UYPR2.512,” *Stand Genomic Sci*, vol. 10, no. APRIL2015, pp. 1–8, 2015, doi: 10.1186/1944-3277-10-13.
- [255] A. Poehlein, B. Kusian, B. Friedrich, R. Daniel, and B. Bowien, “Complete Genome Sequence of the Type Strain Cupriavidus necator N-1,” *J Bacteriol*, vol. 193, no. 18, pp. 5017–5017, Sep. 2011, doi: 10.1128/JB.05660-11.
- [256] N. Noinaj, M. Guillier, , Travis J. Barnard, and S. K. Buchanan, “TonB-Dependent Transporters: Regulation, Structure, and Function,” *Annu Rev Microbiol*, vol. 64, no. 1, pp. 43–60, Oct. 2010, doi: 10.1146/annurev.micro.112408.134247.
- [257] I. J. Schalk and L. Guillon, “Fate of ferrisiderophores after import across bacterial outer membranes: different iron release strategies are observed in the cytoplasm or periplasm depending on the siderophore pathways,” *Amino Acids*, vol. 44, no. 5, pp. 1267–1277, May 2013, doi: 10.1007/s00726-013-1468-2.
- [258] T. Ledger *et al.*, “Aromatic compounds degradation plays a role in colonization of Arabidopsis thaliana and Acacia caven by Cupriavidus pinatubonensis JMP134,” *Antonie Van Leeuwenhoek*, vol. 101, no. 4, pp. 713–723, May 2012, doi: 10.1007/s10482-011-9685-8.
- [259] Z. Xie *et al.*, “Whole Genome Sequence Analysis of Cupriavidus necator C39, a Multiple Heavy Metal(loid) and Antibiotic Resistant Bacterium Isolated from a Gold/Copper Mine,” *Microorganisms*, vol. 11, no. 6, p. 1518, Jun. 2023, doi: 10.3390/microorganisms11061518.
- [260] A. F. Mohidin Batcha, D. M. R. Prasad, M. R. Khan, and H. Abdullah, “Biosynthesis of poly(3-hydroxybutyrate) (PHB) by Cupriavidus necator H16 from jatropha oil as carbon source,” *Bioprocess Biosyst Eng*, vol. 37, no. 5, pp. 943–951, May 2014, doi: 10.1007/s00449-013-1066-4.
- [261] M. D. Kim, M. Song, M. Jo, S. G. Shin, J. H. Khim, and S. Hwang, “Growth condition and bacterial community for maximum hydrolysis of suspended organic materials in anaerobic digestion of food waste-recycling wastewater,” *Appl Microbiol Biotechnol*, vol. 85, no. 5, pp. 1611–1618, Feb. 2010, doi: 10.1007/s00253-009-2316-x.
- [262] M. Miura, Y. Shimotori, H. Nakatani, A. Harada, and M. Aoyama, “Bioconversion of Birch Wood Hemicellulose Hydrolyzate to Xylitol,” *Appl Biochem Biotechnol*, vol. 176, no. 3, pp. 947–955, Jun. 2015, doi: 10.1007/s12010-015-1604-4.
- [263] C.-C. Chien *et al.*, “Functional Expression of phaCAB Genes from Cupriavidus taiwanensis Strain 184 in Escherichia coli for Polyhydroxybutyrate Production,” *Appl Biochem Biotechnol*, vol. 162, no. 8, pp. 2355–2364, Dec. 2010, doi: 10.1007/s12010-010-9008-y.
- [264] K. Mandon *et al.*, “Poly- β -Hydroxybutyrate Turnover in Azorhizobium caulinodans Is Required for Growth and Affects nifA Expression,” *J Bacteriol*, vol. 180, no. 19, pp. 5070–5076, Oct. 1998, doi: 10.1128/JB.180.19.5070-5076.1998.
- [265] B. A. Webb *et al.*, “Sinorhizobium meliloti chemotaxis to quaternary ammonium compounds is mediated by the chemoreceptor McpX,” *Mol Microbiol*, vol. 103, no. 2, pp. 333–346, Jan. 2017, doi: 10.1111/mmi.13561.
- [266] B. A. Webb, S. Hildreth, R. F. Helm, and B. E. Scharf, “Sinorhizobium meliloti Chemoreceptor McpU Mediates Chemotaxis toward Host Plant Exudates through Direct Proline Sensing,” *Appl Environ Microbiol*, vol. 80, no. 11, pp. 3404–3415, Jun. 2014, doi: 10.1128/AEM.00115-14.

- [267] L. D. Miller, C. K. Yost, M. F. Hynes, and G. Alexandre, “The major chemotaxis gene cluster of *Rhizobium leguminosarum* bv. *viciae* is essential for competitive nodulation,” *Mol Microbiol*, vol. 63, no. 2, pp. 348–362, 2007, doi: 10.1111/j.1365-2958.2006.05515.x.
- [268] T. Shidore, T. Dinse, J. Öhrlein, A. Becker, and B. Reinhold-Hurek, “Transcriptomic analysis of responses to exudates reveal genes required for rhizosphere competence of the endophyte *Azoarcus* sp. strain BH72,” *Environ Microbiol*, vol. 14, no. 10, pp. 2775–2787, Oct. 2012, doi: 10.1111/j.1462-2920.2012.02777.x.
- [269] B. Fan, L. C. Carvalhais, A. Becker, D. Fedoseyenko, N. von Wirén, and R. Borriss, “Transcriptomic profiling of *Bacillus amyloliquefaciens* FZB42 in response to maize root exudates,” *BMC Microbiol*, vol. 12, no. 1, p. 116, Dec. 2012, doi: 10.1186/1471-2180-12-116.
- [270] H. R. Schlaman, R. J. Okker, and B. J. Lugtenberg, “Regulation of nodulation gene expression by NodD in rhizobia,” *J Bacteriol*, vol. 174, no. 16, pp. 5177–5182, Aug. 1992, doi: 10.1128/jb.174.16.5177-5182.1992.
- [271] C. Amadou *et al.*, “Genome sequence of the β -rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia,” *Genome Res*, vol. 18, no. 9, pp. 1472–1483, 2008, doi: 10.1101/gr.076448.108.
- [272] B. Daubech *et al.*, “NoeM, a new nodulation gene involved in the biosynthesis of nod factors with an open-chain oxidized terminal residue and in the symbiosis with *mimosa pudica*,” *Molecular Plant-Microbe Interactions*, vol. 32, no. 12, pp. 1635–1648, 2019, doi: 10.1094/MPMI-06-19-0168-R.
- [273] A. Klonowska, R. Melkonian, L. Miché, P. Tisseyre, and L. Moulin, “Transcriptomic profiling of *Burkholderia phymatum* STM815, *Cupriavidus taiwanensis* LMG19424 and *Rhizobium mesoamericanum* STM3625 in response to *Mimosa pudica* root exudates illuminates the molecular basis of their nodulation competitiveness and symbiotic ev,” *BMC Genomics*, vol. 19, no. 1, 2018, doi: 10.1186/s12864-018-4487-2.
- [274] P. J. Rutten and P. S. Poole, “Oxygen regulatory mechanisms of nitrogen fixation in rhizobia,” 2019, pp. 325–389. doi: 10.1016/bs.ampbs.2019.08.001.
- [275] S. Ekici, G. Pawlik, E. Lohmeyer, H.-G. Koch, and F. Daldal, “Biogenesis of *cbb3*-type cytochrome c oxidase in *Rhodobacter capsulatus*,” *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, vol. 1817, no. 6, pp. 898–910, Jun. 2012, doi: 10.1016/j.bbabi.2011.10.011.
- [276] C. Angkawidjaja and S. Kanaya, “Family I.3 lipase: bacterial lipases secreted by the type I secretion system,” *Cellular and Molecular Life Sciences*, vol. 63, no. 23, pp. 2804–2817, Dec. 2006, doi: 10.1007/s00018-006-6172-x.
- [277] I. Linhartová *et al.*, “RTX proteins: a highly diverse family secreted by a common mechanism,” *FEMS Microbiol Rev*, vol. 34, no. 6, pp. 1076–1112, Nov. 2010, doi: 10.1111/j.1574-6976.2010.00231.x.
- [278] D. M. Russo *et al.*, “Proteins Exported via the PrsD-PrsE Type I Secretion System and the Acidic Exopolysaccharide Are Involved in Biofilm Formation by *Rhizobium leguminosarum*,” *J Bacteriol*, vol. 188, no. 12, pp. 4474–4486, Jun. 2006, doi: 10.1128/JB.00246-06.
- [279] S. A. Walker and J. A. Downie, “Entry of *Rhizobium leguminosarum* bv. *viciae* into Root Hairs Requires Minimal Nod Factor Specificity, but Subsequent Infection Thread Growth Requires *nodO* or *nodE*,” *Molecular Plant-Microbe Interactions®*, vol. 13, no. 7, pp. 754–762, Jul. 2000, doi: 10.1094/MPMI.2000.13.7.754.

- [280] K. V. Korotkov, M. Sandkvist, and W. G. J. Hol, “The type II secretion system: biogenesis, molecular architecture and mechanism,” *Nat Rev Microbiol*, vol. 10, no. 5, pp. 336–351, May 2012, doi: 10.1038/nrmicro2762.
- [281] B. González-Pedrajo and G. Dreyfus, “Sistemas de secreción de proteínas en las bacterias Gram-negativas: Biogénesis flagelar y translocación de factores de virulencia,” *Mensaje Bioquímico*, vol. XXVII. Depto. Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México. Cd. Universitaria, México, D.F., pp. 45–63, 2003.
- [282] S. Yang, F. Tang, M. Gao, H. B. Krishnan, and H. Zhu, “R gene-controlled host specificity in the legume–rhizobia symbiosis,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 43, pp. 18735–18740, Oct. 2010, doi: 10.1073/pnas.1011957107.
- [283] M. M. Saad, M. Crèvecoeur, C. Masson-Boivin, and X. Perret, “The Type 3 Protein Secretion System of *Cupriavidus taiwanensis* Strain LMG19424 Compromises Symbiosis with *Leucaena leucocephala*,” *Appl Environ Microbiol*, vol. 78, no. 20, pp. 7476–7479, Oct. 2012, doi: 10.1128/AEM.01691-12.
- [284] C. Clerissi *et al.*, “Parallels between experimental and natural evolution of legume symbionts,” *Nat Commun*, vol. 9, no. 1, p. 2264, Jun. 2018, doi: 10.1038/s41467-018-04778-5.
- [285] J. Zupan, C. A. Hackworth, J. Aguilar, D. Ward, and P. Zambryski, “VirB1* Promotes T-Pilus Formation in the *vir* -Type IV Secretion System of *Agrobacterium tumefaciens*,” *J Bacteriol*, vol. 189, no. 18, pp. 6551–6563, Sep. 2007, doi: 10.1128/JB.00480-07.
- [286] A. M. Hubber, J. T. Sullivan, and C. W. Ronson, “Symbiosis-Induced Cascade Regulation of the *Mesorhizobium loti* R7A VirB/D4 Type IV Secretion System,” *Molecular Plant-Microbe Interactions®*, vol. 20, no. 3, pp. 255–261, Mar. 2007, doi: 10.1094/MPMI-20-3-0255.
- [287] L. A. Lomovatskaya and A. S. Romanenko, “Secretion Systems of Bacterial Phytopathogens and Mutualists (Review),” *Appl Biochem Microbiol*, vol. 56, no. 2, pp. 115–129, Mar. 2020, doi: 10.1134/S0003683820020106.
- [288] C. M. Rojas, J. H. Ham, W.-L. Deng, J. J. Doyle, and A. Collmer, “HecA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of *Erwinia chrysanthemi* EC16 on *Nicotiana clevelandii* seedlings,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 20, pp. 13142–13147, Oct. 2002, doi: 10.1073/pnas.202358699.
- [289] B. S. Garavaglia, T. Zimaro, L. A. Abriata, J. Ottado, and N. Gottig, “XacFhaB adhesin, an important *Xanthomonas citri* ssp. *citri* virulence factor, is recognized as a pathogen-associated molecular pattern,” *Mol Plant Pathol*, vol. 17, no. 9, pp. 1344–1353, Dec. 2016, doi: 10.1111/mpp.12364.
- [290] K. Velmourougane, R. Prasanna, and A. K. Saxena, “Agriculturally important microbial biofilms: Present status and future prospects,” *J Basic Microbiol*, vol. 57, no. 7, pp. 548–573, Jul. 2017, doi: 10.1002/jobm.201700046.
- [291] J. Monjarás Feria and M. A. Valvano, “An Overview of Anti-Eukaryotic T6SS Effectors,” *Front Cell Infect Microbiol*, vol. 10, Oct. 2020, doi: 10.3389/fcimb.2020.584751.
- [292] S. B. de Campos, M. Lardi, A. Gandolfi, L. Eberl, and G. Pessi, “Mutations in Two *Paraburkholderia phymatum* Type VI Secretion Systems Cause Reduced Fitness in Interbacterial Competition,” *Front Microbiol*, vol. 8, Dec. 2017, doi: 10.3389/fmicb.2017.02473.

- [293] A. Salinero-Lanzarote *et al.*, “The Type VI secretion system of *Rhizobium etli* Mim1 has a positive effect in symbiosis,” *FEMS Microbiol Ecol*, vol. 95, no. 5, May 2019, doi: 10.1093/femsec/fiz054.
- [294] A. A. Alemneh, Y. Zhou, M. H. Ryder, and M. D. Denton, “Mechanisms in plant growth-promoting rhizobacteria that enhance legume–rhizobial symbioses,” *J Appl Microbiol*, vol. 129, no. 5, pp. 1133–1156, Nov. 2020, doi: 10.1111/jam.14754.
- [295] M. Snyder, S. Weissman, and M. Gerstein, “Personal phenotypes to go with personal genomes,” *Mol Syst Biol*, vol. 5, no. 1, Jan. 2009, doi: 10.1038/msb.2009.32.
- [296] R. Sorek and P. Cossart, “Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity,” *Nat Rev Genet*, vol. 11, no. 1, pp. 9–16, Jan. 2010, doi: 10.1038/nrg2695.
- [297] “www.agilent.com.”
- [298] P. J. A. Cock, C. J. Fields, N. Goto, M. L. Heuer, and P. M. Rice, “The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants,” *Nucleic Acids Res*, vol. 38, no. 6, pp. 1767–1771, Apr. 2010, doi: 10.1093/nar/gkp1137.
- [299] B. Ewing, L. Hillier, M. C. Wendl, and P. Green, “Base-Calling of Automated Sequencer Traces Using *Phred*. I. Accuracy Assessment,” *Genome Res*, vol. 8, no. 3, pp. 175–185, Mar. 1998, doi: 10.1101/gr.8.3.175.
- [300] S. Andrews, “FastQC: a quality control tool for high throughput sequence data.,” <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- [301] A. M. Bolger, M. Lohse, and B. Usadel, “Trimmomatic: a flexible trimmer for Illumina sequence data,” *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, Aug. 2014, doi: 10.1093/bioinformatics/btu170.
- [302] B. Langmead and S. L. Salzberg, “Fast gapped-read alignment with Bowtie 2,” *Nat Methods*, vol. 9, no. 4, pp. 357–359, Apr. 2012, doi: 10.1038/nmeth.1923.
- [303] Y. Liao, G. K. Smyth, and W. Shi, “featureCounts: an efficient general purpose program for assigning sequence reads to genomic features,” *Bioinformatics*, vol. 30, no. 7, pp. 923–930, Apr. 2014, doi: 10.1093/bioinformatics/btt656.
- [304] A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, and B. Wold, “Mapping and quantifying mammalian transcriptomes by RNA-Seq,” *Nat Methods*, vol. 5, no. 7, pp. 621–628, Jul. 2008, doi: 10.1038/nmeth.1226.
- [305] C. Trapnell *et al.*, “Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation,” *Nat Biotechnol*, vol. 28, no. 5, pp. 511–515, May 2010, doi: 10.1038/nbt.1621.
- [306] M. D. Robinson, D. J. McCarthy, and G. K. Smyth, “edgeR: a Bioconductor package for differential expression analysis of digital gene expression data,” *Bioinformatics*, vol. 26, no. 1, pp. 139–140, Jan. 2010, doi: 10.1093/bioinformatics/btp616.
- [307] Gene Ontology Consortium, “The Gene Ontology (GO) database and informatics resource,” *Nucleic Acids Res*, vol. 32, no. 90001, pp. 258D–261, Jan. 2004, doi: 10.1093/nar/gkh036.
- [308] M. Kanehisa, M. Furumichi, Y. Sato, M. Ishiguro-Watanabe, and M. Tanabe, “KEGG: integrating viruses and cellular organisms,” *Nucleic Acids Res*, vol. 49, no. D1, pp. D545–D551, Jan. 2021, doi: 10.1093/nar/gkaa970.
- [309] R. L. Tatusov, “The COG database: a tool for genome-scale analysis of protein functions and evolution,” *Nucleic Acids Res*, vol. 28, no. 1, pp. 33–36, Jan. 2000, doi: 10.1093/nar/28.1.33.

- [310] I. Jiménez-Guerrero *et al.*, “Transcriptomic Studies of the Effect of nod Gene-Inducing Molecules in Rhizobia: Different Weapons, One Purpose.,” *Genes (Basel)*, vol. 9, no. 1, Dec. 2017, doi: 10.3390/genes9010001.
- [311] R. T. Green, A. K. East, R. Karunakaran, J. A. Downie, and P. S. Poole, “Transcriptomic analysis of *Rhizobium leguminosarum* bacteroids in determinate and indeterminate nodules,” 2019, doi: 10.1099/mgen.0.000254.
- [312] J. L. Aguirre-Noyola, M. Rosenblueth, M. G. Santiago-Martínez, and E. Martínez-Romero, “Transcriptomic Responses of *Rhizobium phaseoli* to Root Exudates Reflect Its Capacity to Colonize Maize and Common Bean in an Intercropping System,” *Front Microbiol*, vol. 12, Oct. 2021, doi: 10.3389/fmicb.2021.740818.
- [313] T. Sambrook, J., Fritsch, E. R., & Maniatis, “Molecular Cloning: A Laboratory Manual (2nd ed.),” *Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.*, 1989.
- [314] E. Wood, “Molecular Cloning. A Laboratory Manual,” *Biochem Educ*, vol. 11, no. 2, p. 82, Apr. 1983, doi: 10.1016/0307-4412(83)90068-7.
- [315] W. M. Chen *et al.*, “*Ralstonia taiwanensis* sp. nov., isolated from root nodules of *Mimosa* species and sputum of a cystic fibrosis patient.,” *Int J Syst Evol Microbiol*, vol. 51, no. 5, pp. 1729–1735, Sep. 2001, doi: 10.1099/00207713-51-5-1729.
- [316] B. Kessler, V. de Lorenzo, and K. N. Timmis, “A general system to integratelacZ fusions into the chromosomes of gram-negative eubacteria: regulation of thePm promoter of theTOL plasmid studied with all controlling elements in monocopy,” *Mol Gen Genet*, vol. 233, no. 1–2, pp. 293–301, May 1992, doi: 10.1007/BF00587591.
- [317] M. Marchetti *et al.*, “Experimental Evolution of a Plant Pathogen into a Legume Symbiont,” *PLoS Biol*, vol. 8, no. 1, p. e1000280, Jan. 2010, doi: 10.1371/journal.pbio.1000280.
- [318] S. Cunnac, C. Boucher, and S. Genin, “Characterization of the ϵ -cis- ϵ -Acting Regulatory Element Controlling HrpB-Mediated Activation of the Type III Secretion System and Effector Genes in ϵ -*Ralstonia solanacearum*,” *J Bacteriol*, vol. 186, no. 8, pp. 2309 LP – 2318, Apr. 2004, [Online]. Available: <http://jb.asm.org/content/186/8/2309.abstract>
- [319] L.-K. Ng, I. Martin, M. Alfa, and M. Mulvey, “Multiplex PCR for the detection of tetracycline resistant genes,” *Mol Cell Probes*, vol. 15, no. 4, pp. 209–215, Aug. 2001, doi: 10.1006/mcpr.2001.0363.
- [320] J. H. Miller, *Experiments in molecular genetics*. Cold Spring Harbor Laboratory, 1972.
- [321] “QIAGEN CLC Genomics Workbench 23.0.1,” <https://digitalinsights.qiagen.com/>.
- [322] Y. Benjamini and Y. Hochberg, “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, Jan. 1995, doi: 10.1111/j.2517-6161.1995.tb02031.x.
- [323] J. Huerta-Cepas *et al.*, “Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper,” *Mol Biol Evol*, vol. 34, no. 8, pp. 2115–2122, Aug. 2017, doi: 10.1093/molbev/msx148.
- [324] Van Rossum G and Drake Jr FL., “Python reference manual,” *Centrum voor Wiskunde en Informatica Amsterdam*, 1995.
- [325] A. Untergasser *et al.*, “Primer3—new capabilities and interfaces,” *Nucleic Acids Res*, vol. 40, no. 15, pp. e115–e115, Aug. 2012, doi: 10.1093/nar/gks596.
- [326] M. Kubista *et al.*, “The real-time polymerase chain reaction,” *Mol Aspects Med*, vol. 27, no. 2–3, pp. 95–125, Apr. 2006, doi: 10.1016/j.mam.2005.12.007.

- [327] C. W. Balzarini, M. G., Gonzalez, L. A., Tablada, E. M., Casanoves, F., Di Rienzo, J. A., Robledo, “InfoStat Manual del Usuario,” no. December 2015, p. 334, 2008.
- [328] M. W. Pfaffl, “A new mathematical model for relative quantification in real-time RT-PCR,” *Nucleic Acids Res*, vol. 29, no. 9, pp. 45e–445, May 2001, doi: 10.1093/nar/29.9.e45.
- [329] J. Loh and G. Stacey, “Nodulation Gene Regulation in *Bradyrhizobium japonicum*: a Unique Integration of Global Regulatory Circuits,” *Appl Environ Microbiol*, vol. 69, no. 1, pp. 10–17, Jan. 2003, doi: 10.1128/AEM.69.1.10-17.2003.
- [330] L. ARAVIND, V. ANANTHARAMAN, S. BALAJI, M. BABU, and L. IYER, “The many faces of the helix-turn-helix domain: Transcription regulation and beyond,” *FEMS Microbiol Rev*, vol. 29, no. 2, pp. 231–262, Apr. 2005, doi: 10.1016/j.femsre.2004.12.008.
- [331] Z. Baharoglu and D. Mazel, “SOS, the formidable strategy of bacteria against aggressions,” *FEMS Microbiol Rev*, vol. 38, no. 6, pp. 1126–1145, Nov. 2014, doi: 10.1111/1574-6976.12077.
- [332] L. Lozano *et al.*, “Evolutionary Dynamics of Insertion Sequences in Relation to the Evolutionary Histories of the Chromosome and Symbiotic Plasmid Genes of *Rhizobium etli* Populations,” *Appl Environ Microbiol*, vol. 76, no. 19, pp. 6504–6513, Oct. 2010, doi: 10.1128/AEM.01001-10.
- [333] J. Huang, W. Yindeeyoungyeon, R. P. Garg, T. P. Denny, and M. A. Schell, “Joint Transcriptional Control of *xpsR*, the Unusual Signal Integrator of the *Ralstonia solanacearum* Virulence Gene Regulatory Network, by a Response Regulator and a LysR-Type Transcriptional Activator,” *J Bacteriol*, vol. 180, no. 10, pp. 2736–2743, May 1998, doi: 10.1128/JB.180.10.2736-2743.1998.
- [334] M. Janczarek, K. Rachwał, A. Marzec, J. Grządziel, and M. Palusińska-Szyszk, “Signal molecules and cell-surface components involved in early stages of the legume–rhizobium interactions,” *Applied Soil Ecology*, vol. 85, pp. 94–113, Jan. 2015, doi: 10.1016/J.APSOIL.2014.08.010.
- [335] B. D. Needham and M. S. Trent, “Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis,” *Nat Rev Microbiol*, vol. 11, no. 7, pp. 467–481, Jul. 2013, doi: 10.1038/nrmicro3047.
- [336] G. Alloing, K. Mandon, E. Boncompagni, F. Montrichard, and P. Frendo, “Involvement of Glutaredoxin and Thioredoxin Systems in the Nitrogen-Fixing Symbiosis between Legumes and Rhizobia,” *Antioxidants*, vol. 7, no. 12, p. 182, Dec. 2018, doi: 10.3390/antiox7120182.
- [337] W. G. Reeve, R. P. Tiwari, N. B. Kale, M. J. Dilworth, and A. R. Glenn, “ActP controls copper homeostasis in *Rhizobium leguminosarum* bv. *viciae* and *Sinorhizobium meliloti* preventing low pH-induced copper toxicity,” *Mol Microbiol*, vol. 43, no. 4, pp. 981–991, 2002, doi: 10.1046/j.1365-2958.2002.02791.x.
- [338] Y. Alguel *et al.*, “Crystal Structures of Multidrug Binding Protein TtgR in Complex with Antibiotics and Plant Antimicrobials,” *J Mol Biol*, vol. 369, no. 3, pp. 829–840, Jun. 2007, doi: 10.1016/j.jmb.2007.03.062.
- [339] L. Zhu, X. Xu, L. Wang, H. Dong, B. Yu, and Y. Ma, “NADP⁺-Preferring *d*-Lactate Dehydrogenase from *Sporolactobacillus inulinus*,” *Appl Environ Microbiol*, vol. 81, no. 18, pp. 6294–6301, Sep. 2015, doi: 10.1128/AEM.01871-15.
- [340] D. Harwani, P. Zangoui, and S. Mahadevan, “The *-Glucoside (bgl)* Operon of *Escherichia coli* Is Involved in the Regulation of *oppA*, Encoding an Oligopeptide Transporter,” *J Bacteriol*, vol. 194, no. 1, pp. 90–99, Jan. 2012, doi: 10.1128/JB.05837-11.

- [341] D. Egamberdieva, G. Renella, S. Wirth, and R. Islam, “Enzyme Activities in the Rhizosphere of Plants,” 2010, pp. 149–166. doi: 10.1007/978-3-642-14225-3_8.
- [342] D. E. Pettijohn, “Histone-like proteins and bacterial chromosome structure.,” *J Biol Chem*, vol. 263, no. 26, pp. 12793–6, Sep. 1988.
- [343] M. Del Campo, Y. Kaya, and J. Ofengand, “Identification and site of action of the remaining four putative pseudouridine synthases in *Escherichia coli*.,” *RNA*, vol. 7, no. 11, pp. 1603–15, Nov. 2001.
- [344] T. Muraki, M. Taki, Y. Hasegawa, H. Iwaki, and P. C. K. Lau, “Prokaryotic Homologs of the Eukaryotic 3-Hydroxyanthranilate 3,4-Dioxygenase and 2-Amino-3-Carboxymuconate-6-Semialdehyde Decarboxylase in the 2-Nitrobenzoate Degradation Pathway of *Pseudomonas fluorescens* Strain KU-7,” *Appl Environ Microbiol*, vol. 69, no. 3, pp. 1564–1572, Mar. 2003, doi: 10.1128/AEM.69.3.1564-1572.2003.
- [345] T. Tanabe, T. Funahashi, H. Nakao, S.-I. Miyoshi, S. Shinoda, and S. Yamamoto, “Identification and Characterization of Genes Required for Biosynthesis and Transport of the Siderophore Vibrioferrin in *Vibrio parahaemolyticus*,” *J Bacteriol*, vol. 185, no. 23, pp. 6938–6949, Dec. 2003, doi: 10.1128/JB.185.23.6938-6949.2003.
- [346] E. Eichhorn, J. R. van der Ploeg, and T. Leisinger, “Characterization of a Two-component Alkanesulfonate Monooxygenase from *Escherichia coli*,” *Journal of Biological Chemistry*, vol. 274, no. 38, pp. 26639–26646, Sep. 1999, doi: 10.1074/jbc.274.38.26639.
- [347] L. T. Rosa, M. E. Bianconi, G. H. Thomas, and D. J. Kelly, “Tripartite ATP-Independent Periplasmic (TRAP) Transporters and Tripartite Tricarboxylate Transporters (TTT): From Uptake to Pathogenicity,” *Front Cell Infect Microbiol*, vol. 8, Feb. 2018, doi: 10.3389/fcimb.2018.00033.
- [348] J. R. Nodwell and J. Greenblatt, “Recognition of boxA antiterminator RNA by the *E. coli* antitermination factors NusB and ribosomal protein S10,” *Cell*, vol. 72, no. 2, pp. 261–268, Jan. 1993, doi: 10.1016/0092-8674(93)90665-D.
- [349] S. Genin and T. P. Denny, “Pathogenomics of the *Ralstonia solanacearum* Species Complex,” *Annu Rev Phytopathol*, vol. 50, no. 1, pp. 67–89, Sep. 2012, doi: 10.1146/annurev-phyto-081211-173000.
- [350] G. Caetano-Anollés, L. G. Wall, A. T. De Micheli, E. M. Macchi, W. D. Bauer, and G. Favelukes, “Role of Motility and Chemotaxis in Efficiency of Nodulation by *Rhizobium meliloti*,” *Plant Physiol*, vol. 86, no. 4, pp. 1228–1235, Apr. 1988, doi: 10.1104/pp.86.4.1228.
- [351] K. Lang, A. Lindemann, F. Hauser, and M. Göttfert, “The genistein stimulon of *Bradyrhizobium japonicum*,” *Molecular Genetics and Genomics*, vol. 279, no. 3, pp. 203–211, Mar. 2008, doi: 10.1007/s00438-007-0280-7.
- [352] V. K. Ramachandran, A. K. East, R. Karunakaran, J. A. Downie, and P. S. Poole, “Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics,” *Genome Biol*, vol. 12, no. 10, 2011, doi: 10.1186/gb-2011-12-10-r106.
- [353] B. Huber, K. Riedel, M. Köthe, M. Givskov, S. Molin, and L. Eberl, “Genetic analysis of functions involved in the late stages of biofilm development in *Burkholderia cepacia* H111,” *Mol Microbiol*, vol. 46, no. 2, pp. 411–426, Oct. 2002, doi: 10.1046/j.1365-2958.2002.03182.x.
- [354] C. Xi, E. Schoeters, J. Vanderleyden, and J. Michiels, “Symbiosis-specific expression of *Rhizobium etli casA* encoding a secreted calmodulin-related protein,” *Proceedings of the*

- National Academy of Sciences*, vol. 97, no. 20, pp. 11114–11119, Sep. 2000, doi: 10.1073/pnas.210181097.
- [355] R. Timpl, D. Tisi, J. F. Talts, Z. Andac, T. Sasaki, and E. Hohenester, “Structure and function of laminin LG modules,” *Matrix Biology*, vol. 19, no. 4, pp. 309–317, Aug. 2000, doi: 10.1016/S0945-053X(00)00072-X.
- [356] S. Selvaraj, K. C. Saha, A. Chakraborty, S. N. Bhattacharyya, and A. Saha, “Toxicity of free and various aminocarboxylic ligands sequestered copper(II) ions to *Escherichia coli*,” *J Hazard Mater*, vol. 166, no. 2–3, pp. 1403–1409, Jul. 2009, doi: 10.1016/j.jhazmat.2008.12.060.
- [357] D. Corbett, S. Schuler, S. Glenn, P. W. Andrew, J. S. Cavet, and I. S. Roberts, “The combined actions of the copper-responsive repressor CsoR and copper-metallochaperone CopZ modulate CopA-mediated copper efflux in the intracellular pathogen *Listeria monocytogenes*,” *Mol Microbiol*, vol. 81, no. 2, pp. 457–472, Jul. 2011, doi: 10.1111/j.1365-2958.2011.07705.x.
- [358] T. S. de Castilho, T. B. Matias, K. P. Nicolini, and J. Nicolini, “Study of interaction between metal ions and quercetin,” *Food Science and Human Wellness*, vol. 7, no. 3, pp. 215–219, Sep. 2018, doi: 10.1016/j.fshw.2018.08.001.
- [359] Y. He *et al.*, “Chromium resistance characteristics of Cr(VI) resistance genes ChrA and ChrB in *Serratia* sp. S2,” *Ecotoxicol Environ Saf*, vol. 157, pp. 417–423, Aug. 2018, doi: 10.1016/j.ecoenv.2018.03.079.
- [360] S. S. Pandey, P. K. Patnana, R. Rai, and S. Chatterjee, “Xanthoferrin, the α - hydroxycarboxylate-type siderophore of *Xanthomonas campestris* pv. *campestris*, is required for optimum virulence and growth inside cabbage,” *Mol Plant Pathol*, vol. 18, no. 7, pp. 949–962, Sep. 2017, doi: 10.1111/mpp.12451.
- [361] N. Baichoo and J. D. Helmann, “Recognition of DNA by Fur: a Reinterpretation of the Fur Box Consensus Sequence,” *J Bacteriol*, vol. 184, no. 21, pp. 5826–5832, Nov. 2002, doi: 10.1128/JB.184.21.5826-5832.2002.
- [362] R. Platero, V. de Lorenzo, B. Garat, and E. Fabiano, “*Sinorhizobium meliloti* Fur-Like (Mur) Protein Binds a Fur Box-Like Sequence Present in the *mntA* Promoter in a Manganese-Responsive Manner,” *Appl Environ Microbiol*, vol. 73, no. 15, pp. 4832–4838, Aug. 2007, doi: 10.1128/AEM.00686-07.
- [363] M. R. O’Brian, “Perception and Homeostatic Control of Iron in the Rhizobia and Related Bacteria,” *Annu Rev Microbiol*, vol. 69, no. 1, pp. 229–245, Oct. 2015, doi: 10.1146/annurev-micro-091014-104432.
- [364] C. Li *et al.*, “Siderophore-Mediated Iron Acquisition Enhances Resistance to Oxidative and Aromatic Compound Stress in *Cupriavidus necator* JMP134,” *Appl Environ Microbiol*, vol. 85, no. 1, Jan. 2019, doi: 10.1128/AEM.01938-18.
- [365] F. Battistoni *et al.*, “Identification of an Iron-Regulated, Hemin-Binding Outer Membrane Protein in *Sinorhizobium meliloti*,” *Appl Environ Microbiol*, vol. 68, no. 12, pp. 5877–5881, Dec. 2002, doi: 10.1128/AEM.68.12.5877-5881.2002.
- [366] K. Eshelman, H. Yao, A. N. D. Punchi Hewage, J. J. Deay, J. R. Chandler, and M. Rivera, “Inhibiting the BfrB:Bfd interaction in *Pseudomonas aeruginosa* causes irreversible iron accumulation in bacterioferritin and iron deficiency in the bacterial cytosol,” *Metallomics*, vol. 9, no. 6, pp. 646–659, 2017, doi: 10.1039/C7MT00042A.

- [367] R. Santos, D. Herouart, A. Puppo, and D. Touati, “Critical protective role of bacterial superoxide dismutase in Rhizobium-legume symbiosis,” *Mol Microbiol*, vol. 38, no. 4, pp. 750–759, Nov. 2000, doi: 10.1046/j.1365-2958.2000.02178.x.
- [368] M. Mergeay and R. Van Houdt, “Adaptation to Xenobiotics and Toxic Compounds by Cupriavidus and Ralstonia with Special Reference to Cupriavidus metallidurans CH34 and Mobile Genetic Elements,” in *Biodegradative Bacteria*, Tokyo: Springer Japan, 2014, pp. 105–127. doi: 10.1007/978-4-431-54520-0_6.
- [369] X.-Q. Zhao, R.-C. Wang, X.-C. Lu, J.-J. Lu, J. Li, and H. Hu, “Tolerance and Biosorption of Heavy Metals by Cupriavidus metallidurans strain XXKD-1 Isolated from a Subsurface Laneway in the Qixiashan Pb-Zn Sulfide Minery in Eastern China,” *Geomicrobiol J*, vol. 29, no. 3, pp. 274–286, Apr. 2012, doi: 10.1080/01490451.2011.619637.
- [370] C.-C. Su, D. J. Rutherford, and E. W. Yu, “Characterization of the multidrug efflux regulator AcrR from Escherichia coli,” *Biochem Biophys Res Commun*, vol. 361, no. 1, pp. 85–90, Sep. 2007, doi: 10.1016/j.bbrc.2007.06.175.
- [371] K. Nonaka, H. Ohta, Y. Sato, and K. Hosokawa, “Utilization of Phenylpropanoids by Pseudomonas putida Soil Isolates and Its Probable Taxonomic Significance,” *Microbes Environ*, vol. 23, no. 4, pp. 360–364, 2008, doi: 10.1264/jsme2.ME08545.
- [372] U. Sauer and B. J. Eikmanns, “The PEP—pyruvate—oxaloacetate node as the switch point for carbon flux distribution in bacteria: We dedicate this paper to Rudolf K. Thauer, Director of the Max-Planck-Institute for Terrestrial Microbiology in Marburg, Germany, on the occasion of his 65th birthday,” *FEMS Microbiol Rev*, vol. 29, no. 4, pp. 765–794, Sep. 2005, doi: 10.1016/j.femsre.2004.11.002.
- [373] L. Walker, B. Lagunas, and M. L. Gifford, “Determinants of Host Range Specificity in Legume-Rhizobia Symbiosis,” *Front Microbiol*, vol. 11, no. November, pp. 1–13, 2020, doi: 10.3389/fmicb.2020.585749.
- [374] M. Sandy and A. Butler, “Microbial Iron Acquisition: Marine and Terrestrial Siderophores,” *Chem Rev*, vol. 109, no. 10, pp. 4580–4595, Oct. 2009, doi: 10.1021/cr9002787.
- [375] J. G. Bundy, M. P. Davey, and M. R. Viant, “Environmental metabolomics: a critical review and future perspectives,” *Metabolomics*, vol. 5, no. 1, pp. 3–21, Mar. 2009, doi: 10.1007/s11306-008-0152-0.
- [376] A. V. Aderemi, A. O. Ayeleso, O. O. Oyedapo, and E. Mukwevho, “Metabolomics: A Scoping Review of Its Role as a Tool for Disease Biomarker Discovery in Selected Non-Communicable Diseases,” *Metabolites*, vol. 11, no. 7, p. 418, Jun. 2021, doi: 10.3390/metabo11070418.
- [377] C. Li *et al.*, “Metabolomics in patients with psychosis: A systematic review,” *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, vol. 177, no. 6, pp. 580–588, Sep. 2018, doi: 10.1002/ajmg.b.32662.
- [378] G. Maker, T. Green, I. Mullaney, and R. Trengove, “Untargeted Metabolomic Analysis of Rat Neuroblastoma Cells as a Model System to Study the Biochemical Effects of the Acute Administration of Methamphetamine,” *Metabolites*, vol. 8, no. 2, p. 38, Jun. 2018, doi: 10.3390/metabo8020038.
- [379] L. D. Roberts, A. L. Souza, R. E. Gerszten, and C. B. Clish, “Targeted Metabolomics,” *Curr Protoc Mol Biol*, vol. 98, no. 1, Apr. 2012, doi: 10.1002/0471142727.mb3002s98.
- [380] A. C. Schrimpe-Rutledge, S. G. Codreanu, S. D. Sherrod, and J. A. McLean, “Untargeted Metabolomics Strategies—Challenges and Emerging Directions,” *J Am Soc Mass Spectrom*, vol. 27, no. 12, pp. 1897–1905, Dec. 2016, doi: 10.1007/s13361-016-1469-y.

- [381] W. B. Dunn *et al.*, “Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics,” *Metabolomics*, vol. 9, no. S1, pp. 44–66, Mar. 2013, doi: 10.1007/s11306-012-0434-4.
- [382] T. Fuhrer and N. Zamboni, “High-throughput discovery metabolomics,” *Curr Opin Biotechnol*, vol. 31, pp. 73–78, Feb. 2015, doi: 10.1016/j.copbio.2014.08.006.
- [383] J. E. Adaway, B. G. Keevil, and L. J. Owen, “Liquid chromatography tandem mass spectrometry in the clinical laboratory,” *Annals of Clinical Biochemistry: International Journal of Laboratory Medicine*, vol. 52, no. 1, pp. 18–38, Jan. 2015, doi: 10.1177/0004563214557678.
- [384] L. N. Williamson and M. G. Bartlett, “Quantitative liquid chromatography/time-of-flight mass spectrometry,” *Biomedical Chromatography*, vol. 21, no. 6, pp. 567–576, Jun. 2007, doi: 10.1002/bmc.844.
- [385] J. C. Schwartz, M. W. Senko, and J. E. P. Syka, “A two-dimensional quadrupole ion trap mass spectrometer,” *J Am Soc Mass Spectrom*, vol. 13, no. 6, pp. 659–669, Jun. 2002, doi: 10.1016/S1044-0305(02)00384-7.
- [386] T. Geiger, J. Cox, and M. Mann, “Proteomics on an Orbitrap Benchtop Mass Spectrometer Using All-ion Fragmentation,” *Molecular & Cellular Proteomics*, vol. 9, no. 10, pp. 2252–2261, Oct. 2010, doi: 10.1074/mcp.M110.001537.
- [387] M. Wang *et al.*, “Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking,” *Nat Biotechnol*, vol. 34, no. 8, pp. 828–837, Aug. 2016, doi: 10.1038/nbt.3597.
- [388] S. W. Gabrielson, “SciFinder,” *Journal of the Medical Library Association*, vol. 106, no. 4, p. 872, Oct. 2018, doi: 10.5195/JMLA.2018.515.
- [389] P. Shannon, “Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks,” *Genome Res*, vol. 13, no. 11, pp. 2498–2504, Nov. 2003, doi: 10.1101/gr.1239303.
- [390] S. P. Putri *et al.*, “Application of gas chromatography-mass spectrometry-based metabolomics in food science and technology,” *J Biosci Bioeng*, vol. 133, no. 5, pp. 425–435, May 2022, doi: 10.1016/j.jbiosc.2022.01.011.
- [391] O. A. Adebo *et al.*, “Application of gas chromatography–mass spectrometry (GC-MS)-based metabolomics for the study of fermented cereal and legume foods: A review,” *Int J Food Sci Technol*, vol. 56, no. 4, pp. 1514–1534, Apr. 2021, doi: 10.1111/ijfs.14794.
- [392] J. R. Everett, “Applications of Metabolic Phenotyping in Pharmaceutical Research and Development,” in *The Handbook of Metabolic Phenotyping*, Elsevier, 2019, pp. 407–447. doi: 10.1016/B978-0-12-812293-8.00014-1.
- [393] M. Libault, “Transcriptional Reprogramming of Legume Genomes: Perspective and Challenges Associated With Single-Cell and Single Cell-Type Approaches During Nodule Development,” *Front Plant Sci*, vol. 9, Nov. 2018, doi: 10.3389/fpls.2018.01600.
- [394] H. P. Bais, T. L. Weir, L. G. Perry, S. Gilroy, and J. M. Vivanco, “The role of root exudates in rhizosphere interactions with plants and other organisms,” *Annu Rev Plant Biol*, vol. 57, pp. 233–66, 2006, doi: 10.1146/annurev.arplant.57.032905.105159.
- [395] Y. Oppenheimer-Shaanan *et al.*, “A dynamic rhizosphere interplay between tree roots and soil bacteria under drought stress,” *Elife*, vol. 11, Jul. 2022, doi: 10.7554/eLife.79679.
- [396] V. Vives-Peris, C. de Ollas, A. Gómez-Cadenas, and R. M. Pérez-Clemente, “Root exudates: from plant to rhizosphere and beyond,” *Plant Cell Rep*, vol. 39, no. 1, pp. 3–17, Jan. 2020, doi: 10.1007/s00299-019-02447-5.

- [397] C. S. Mogal *et al.*, “UHPLC-MS/MS and QRT-PCR profiling of PGP agents and Rhizobium spp. of induced phytohormones for growth promotion in mungbean (var. Co4),” *Heliyon*, vol. 8, no. 5, p. e09532, May 2022, doi: 10.1016/j.heliyon.2022.e09532.
- [398] M. Lardi *et al.*, “Metabolomic Profiling of Bradyrhizobium diazoefficiens-Induced Root Nodules Reveals Both Host Plant-Specific and Developmental Signatures,” *Int J Mol Sci*, vol. 17, no. 6, p. 815, May 2016, doi: 10.3390/ijms17060815.
- [399] A. J. Ogden, M. Gargouri, J. Park, D. R. Gang, and M. L. Kahn, “Integrated analysis of zone-specific protein and metabolite profiles within nitrogen-fixing Medicago truncatula-Sinorhizobium medicae nodules,” *PLoS One*, vol. 12, no. 7, p. e0180894, Jul. 2017, doi: 10.1371/journal.pone.0180894.
- [400] B. Riviezzi *et al.*, “Comparative metabolite analysis of Delftia-Bradyrhizobium co-inoculated soybean plants using UHPLC-HRMS-based metabolomic profiling,” *Symbiosis*, vol. 85, no. 3, pp. 325–341, Dec. 2021, doi: 10.1007/s13199-021-00818-9.
- [401] A. L. Howison, J. Ewing M, Robson A, “External phosphate and calcium concentration but not the product of rhizobial nodulation genes, affect the attachment of Rhizobium meliloti to roots,” *Soil, Biol and Biochem*, vol. 25, no. 567–573, 1993.
- [402] M. Sisa, S. L. Bonnet, D. Ferreira, and J. H. Van der Westhuizen, “Photochemistry of Flavonoids,” *Molecules*, vol. 15, no. 8, pp. 5196–5245, Aug. 2010, doi: 10.3390/molecules15085196.
- [403] T. J. Mabry, K. R. Markham, and M. B. Thomas, *The Systematic Identification of Flavonoids*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1970. doi: 10.1007/978-3-642-88458-0.
- [404] E. A. Silva-Junior *et al.*, “Chemical Diversity in a Stingless Bee–Plant Symbiosis,” *ACS Omega*, vol. 4, no. 12, pp. 15208–15214, Sep. 2019, doi: 10.1021/acsomega.9b02096.
- [405] D. Khang, T. Dung, A. Elzaawely, and T. Xuan, “Phenolic Profiles and Antioxidant Activity of Germinated Legumes,” *Foods*, vol. 5, no. 4, p. 27, Apr. 2016, doi: 10.3390/foods5020027.
- [406] H. S. Gujral, P. Sharma, N. Gupta, and A. A. Wani, “Antioxidant properties of legumes and their morphological fractions as affected by cooking,” *Food Sci Biotechnol*, vol. 22, no. 1, pp. 187–194, Feb. 2013, doi: 10.1007/s10068-013-0026-8.
- [407] F. Pérez-Montañó *et al.*, “A transcriptomic analysis of the effect of genistein on Sinorhizobium fredii HH103 reveals novel rhizobial genes putatively involved in symbiosis,” *Sci Rep*, vol. 6, no. July, pp. 1–12, 2016, doi: 10.1038/srep31592.
- [408] H. Kobayashi, Y. N. Graven, W. J. Broughton, and X. Perret, “Flavonoids induce temporal shifts in gene-expression of nod-box controlled loci in Rhizobium sp. NGR234,” *Mol Microbiol*, vol. 51, no. 2, pp. 335–347, Feb. 2004, doi: 10.1046/j.1365-2958.2003.03841.x.
- [409] M. Janczarek, K. Rachwa??, A. Marzec, J. Grzadziel, and M. Palusi??ska-Szyszczyk, “Signal molecules and cell-surface components involved in early stages of the legume-rhizobium interactions,” *Applied Soil Ecology*, vol. 85, pp. 94–113, 2014, doi: 10.1016/j.apsoil.2014.08.010.
- [410] M. Lardi and G. Pessi, “Functional Genomics Approaches to Studying Symbioses between Legumes and Nitrogen-Fixing Rhizobia,” *High Throughput*, vol. 7, no. 2, p. 15, 2018, doi: 10.3390/ht7020015.
- [411] M. Shimamura *et al.*, “Phenolic Acids Induce Nod Factor Production in Lotus japonicus–Mesorhizobium Symbiosis,” *Microbes Environ*, vol. 37, no. 1, p. ME21094, 2022, doi: 10.1264/jsme2.ME21094.

- [412] Natalia V. Kostiuk 1, Maya B. Belyakova 2, Dzhulianna V. Leshchenko 2, Mikhail V. Miniaev 3, Margarita B. Petrova1, and Elena A. Kharitonova1, “Structural Characterization of the NodD Transcription Factor,” *American Journal of Bioinformatics Research*, vol. 3, no. 3, pp. 35–41, 2013.
- [413] S. Bag, A. Mondal, A. Majumder, S. K. Mondal, and A. Banik, “Flavonoid mediated selective cross-talk between plants and beneficial soil microbiome,” *Phytochemistry Reviews*, vol. 21, no. 5, pp. 1739–1760, Oct. 2022, doi: 10.1007/s11101-022-09806-3.
- [414] K. Novak, “Effect of exogenous flavonoids on nodulation of pea (*Pisum sativum* L.),” *J Exp Bot*, vol. 53, no. 375, pp. 1735–1745, Aug. 2002, doi: 10.1093/jxb/erf016.
- [415] N. H. Gomaa, M. O. Hassan, G. M. Fahmy, L. González, O. Hammouda, and A. M. Atteya, “Flavonoid profiling and nodulation of some legumes in response to the allelopathic stress of *Sonchus oleraceus* L.,” *Acta Bot Brasilica*, vol. 29, no. 4, pp. 553–560, Dec. 2015, doi: 10.1590/0102-33062015abb0153.
- [416] N. Mihailović, M. Bogdanović, G. Dražić, R. Filipović, and D. Radin, “Concentration-dependent influence of quercetin on nodulation process and the main characteristics of soybean inoculated with *Bradyrhizobium japonicum*,” *Plant Soil*, vol. 166, no. 2, pp. 243–246, Oct. 1994, [Online]. Available: <http://www.jstor.org/stable/42939687>
- [417] K. K. Compton and B. E. Scharf, “Rhizobial Chemoattractants, the Taste and Preferences of Legume Symbionts,” *Front Plant Sci*, vol. 12, May 2021, doi: 10.3389/fpls.2021.686465.
- [418] W. Malek, “Chemotaxis in *Rhizobium meliloti* strain L5.30,” *Arch Microbiol*, vol. 152, no. 6, pp. 611–612, Nov. 1989, doi: 10.1007/BF00425496.
- [419] L. Brechenmacher *et al.*, “Soybean metabolites regulated in root hairs in response to the symbiotic bacterium *Bradyrhizobium japonicum*,” *Plant Physiol*, vol. 153, no. 4, pp. 1808–1822, 2010, doi: 10.1104/pp.110.157800.
- [420] N. V. Chuiko, T. S. Antonyuk, and I. K. Kurdish, “The chemotactic response of *Bradyrhizobium japonicum* to various organic compounds,” *Microbiology (N Y)*, vol. 71, no. 4, pp. 391–396, 2002, doi: 10.1023/A:1019881124077.
- [421] B.-. J. Koo, D. C. Adriano, N. S. Bolan, and C. D. Barton, “ROOT EXUDATES AND MICROORGANISMS,” in *Encyclopedia of Soils in the Environment*, Elsevier, 2005, pp. 421–428. doi: 10.1016/B0-12-348530-4/00461-6.
- [422] K. Ma *et al.*, “Palmitic acid mediated change of rhizosphere and alleviation of *Fusarium* wilt disease in watermelon,” *Saudi J Biol Sci*, vol. 28, no. 6, pp. 3616–3623, Jun. 2021, doi: 10.1016/j.sjbs.2021.03.040.
- [423] P. de Werra, M. Péchy-Tarr, C. Keel, and M. Maurhofer, “Role of Gluconic Acid Production in the Regulation of Biocontrol Traits of *Pseudomonas fluorescens* CHA0,” *Appl Environ Microbiol*, vol. 75, no. 12, pp. 4162–4174, Jun. 2009, doi: 10.1128/AEM.00295-09.
- [424] S. M. Mandal, D. Chakraborty, and S. Dey, “Phenolic acids act as signaling molecules in plant-microbe symbioses,” *Plant Signal Behav*, vol. 5, no. 4, pp. 359–368, Apr. 2010, doi: 10.4161/psb.5.4.10871.
- [425] Y. Uraki and K. Koda, “Lignin,” in *Encyclopedia of Polymeric Nanomaterials*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 1073–1080. doi: 10.1007/978-3-642-29648-2_325.
- [426] F. Garabato *et al.*, “Especies nativas de Mimosa (Leguminosae) como reservorio de nuevos rizobios con posibles aplicaciones ambientales .”.

- [427] A. Clocchiatti, S. E. Hannula, M. van den Berg, M. P. J. Hundscheid, and W. de Boer, "Evaluation of Phenolic Root Exudates as Stimulants of Saptrophic Fungi in the Rhizosphere," *Front Microbiol*, vol. 12, Apr. 2021, doi: 10.3389/fmicb.2021.644046.
- [428] F. A. Loewus and P. P. N. Murthy, "myo-Inositol metabolism in plants," *Plant Science*, vol. 150, no. 1, pp. 1–19, Jan. 2000, doi: 10.1016/S0168-9452(99)00150-8.
- [429] R. H. Michell, "Inositol derivatives: evolution and functions," *Nat Rev Mol Cell Biol*, vol. 9, no. 2, pp. 151–161, Feb. 2008, doi: 10.1038/nrm2334.
- [430] P. R. A. Kohler, J. Y. Zheng, E. Schoffers, and S. Rossbach, "Inositol Catabolism, a Key Pathway in *Sinorhizobium meliloti* for Competitive Host Nodulation," *Appl Environ Microbiol*, vol. 76, no. 24, pp. 7972–7980, Dec. 2010, doi: 10.1128/AEM.01972-10.