



Caracterización metagenómica del moviloma plasmídico de la población bacteriana asociada a alfalfa.



Trabajo de Tesis Doctoral

2016

José Luis López

Dirección: Prof. Dr. Antonio Lagares



Universidad Nacional de La Plata

Facultad de Ciencias Exactas

Departamento de Ciencias Biológicas



El presente trabajo de Tesis para optar al grado de Doctor de la Facultad de Ciencias Exactas ha sido realizado en el Instituto de Biotecnología y Biología Molecular (IBBM) de la Facultad de Ciencias Exactas de la Universidad Nacional de La Plata – CCT – La Plata –CONICET, bajo la dirección del Profesor Dr. Antonio Lagares y la codirección del Profesor Dr. Mariano Pistorio

Abreviaturas

°C	Grados Celsius
ADN	Ácido Desoxirribonucleico
dNTPs	Desoxirribonucleótidos
DO	Densidad Óptica
g	Gramo
pb	Pares de bases
Kpb	Kilo pares de bases
Mpb	Mega pares de bases
l	Litro
M	Molar
mg	Miligramo
ml	Mililitro
mm	Milímetro
mM	Milimolar
PCR	Reacción en cadena de la polimerasa
pH	Potencial Hidrógeno
rpm	Revoluciones por minuto
µg	Microgramo
µl	Microlitro
µM	Micromolar
MPSme	Moviloma plasmídico de <i>Sinorhizobium meliloti</i>
MPAS	Moviloma plasmídico de los aislamientos asociados a semillas
COGs	Grupos de genes ortólogos (" <i>clusters of orthologous genes</i> ")
ORFs	Marcos de lectura abiertos (" <i>open reading frames</i> ")
PCV	Capacidades indicadoras de promoción del crecimiento vegetal
u.f.c	Unidades formadoras de colonias
Da	Dalton
UV	ultravioleta
min	minuto
PBS	<i>Phosphate buffer saline</i>
P/V	g/100 mL
SDS	Dodecilsulfato sódico

Índice

Resumen	1
CAPITULO I	3
Introducción.....	3
I.1. La importancia de los microorganismos asociados a las plantas.....	5
I.2. El microbioma en las plantas	5
I.3. Microbiota de la rizósfera	6
I.4. Microbiota de la filósfera	7
I.5. Microbiota de la endósfera	9
I.6. Microbiota asociada a semillas.....	11
I.7. La simbiosis fijadora de nitrógeno: <i>Sinorhizobium meliloti</i> -alfalfa.....	12
I.8. Microbioma en alfalfa	16
I.9. Movilomas plasmídicos bacterianos	16
I.10. Movilomas plasmídicos en <i>Sinorhizobium meliloti</i>	19
Objetivos.....	23
CAPITULO II	27
Materiales y métodos.....	27
II.1. Medios de cultivo utilizados	29
II.2 Medio mineral para irrigación de plantas.....	29
II.3. Aislamiento de bacterias asociadas a semillas de alfalfa.....	29
II.4. Obtención de bacterias en plántulas de alfalfa provenientes de semillas	30
II.5. Ensayos in vitro de capacidades indicadoras de promoción del crecimiento vegetal.....	30
II.6. Aislamiento de ADN total.....	31
II.7. Ensayos de PCR.....	31
II.8. Amplificación del gen 16s rRNA	31
II.9. Identificación de aislamientos por espectrometría de masas MALDI-TOF.....	32
II.10. Identificación de plásmidos de alto peso molecular.....	32
II.11. Purificación y secuenciamiento de los plásmidos crípticos.....	33
II.12. Identificación de plásmidos cerrados entre los ensamblados obtenidos del MPAS	34
II.13. Identificación de los cóntigos plasmídicos.....	34
II.14. Análisis de la composición funcional por COGs.....	35
II.15. Análisis de genes ortólogos por el software EDGAR.....	35
II.16. Análisis de la frecuencia de uso de codones	36

II.17. Búsquedas por similitud de secuencia.....	37
II.18. Anotación manual y asignación de grupos funcionales	37
CAPÍTULO III	39
Moviloma plasmídico de <i>Sinorhizobium meliloti</i>: secuenciamiento, caracterización funcional y vínculos evolutivos con el resto del genoma del rizobio.....	39
III.1. Introducción	41
III.2. Preparación y secuenciamiento masivo de plasmidoma críptico de una colección previamente caracterizada de aislamientos de <i>S. meliloti</i>	41
III.3. Comparación de las abundancias de COGs en el moviloma plasmídico críptico respecto de las presentes en el cromosoma y en los plásmidos simbióticos (pSyms). Variaciones composicionales, relaciones evolutivas	42
III.3. Análisis de genes ortólogos por EDGAR.....	59
III.4. Análisis de los genes presentes en el MPSme	62
CAPITULO IV.....	69
Caracterización del microbioma de la semilla de alfalfa.....	69
IV.1. Introducción	71
IV.2. Primeros ensayos. Semillas de <i>M. sativa</i> esterilizadas en superficie y germinadas en medio hidropónico estéril. Evidencias de la presencia de bacterias protegidas de la esterilización	71
IV.3. Aislamientos provenientes de semillas de alfalfa variedad Super Monarca.....	72
IV.4. Aislamientos provenientes de semillas de otras variedades de alfalfa	74
IV.5. Bacterias asociadas a plántulas crecidas de semillas esterilizadas.....	76
IV.6. Actividades asociadas a la promoción del crecimiento vegetal (PCV). Nuestros aislamientos.....	78
IV.7. Producción de bacteriocinas en los aislamientos de la colección	82
IV.8. Ensayos de inhibición de distintos hongos fitopatógenos	84
IV.9. Análisis de la diversidad plasmídica presente en la microbiota asociada a semilla	85
IV.10. Algunos comentarios acerca de los posibles roles biológicos conocidos para los géneros obtenidos en las semillas y plantas de alfalfa.....	87
CAPITULO V.....	91
Secuenciamiento del moviloma plasmídico críptico recuperado de bacterias de semillas de alfalfa.....	91
V.1. Acceso a los plásmidos crípticos presentes en el microbioma de semilla de alfalfa.....	93
V.2. Purificación y secuenciamiento del moviloma plasmídico en bacterias asociadas a la alfalfa acarreados en la semilla	95

V.3. Identificación de ensamblajes cerrados en el moviloma plasmídico proveniente de semilla de alfalfa.....	98
V.4. Contenido génico y predicción funcional de la información del moviloma plasmídico endófito	103
V.5. Evaluación de relaciones de ortología entre el MPSA y el MPS _{me}	104
V.6. Análisis de secuencias cromosomales presentes como subproducto de la secuenciación masiva de los plásmidos críticos.....	106
Conclusiones y perspectivas	111
Anexo	119
Referencias bibliográficas	137

Resumen

A través del tiempo las bacterias, como organismos unicelulares, han localizado eficientemente en sus pequeños genomas la información necesaria para afrontar el desafío de los cambios ambientales. En términos genéticos (informativos) las bacterias han concentrado la mayoría de las funciones basales (mantenimiento, *housekeeping*) en sus replicones más grandes (los cromosomas) dedicando compartimientos genéticos especiales y móviles (plásmidos, islas de ADN, transposones, integrones) para diversas respuestas adaptativas (a factores abióticos y bióticos). La partición del material genético en módulos, como elementos funcionales adicionales al cromosoma (“moviloma”, *mobilome*), facilitó la reducción del tamaño de los genomas permitiendo compartir genes (funciones) transitoriamente importantes (adaptativas) entre miembros separados de la propia comunidad. Una partición genética con redundancia previene, al mismo tiempo, la pérdida masiva de información genética cuando miembros individuales de la comunidad bacteriana mueren. Al mismo tiempo, variantes génicas nuevas son susceptibles de hacerse comunitarias a través del moviloma, excediendo la relevancia de las mismas para clones individuales debido a la posibilidad de transponer barreras de especie y de género. De este modo, el conjunto de plásmidos y otros elementos móviles en un ambiente dado representa un recurso genético central en el contexto de la evolución de las bacterias y de sus estrategias de manejo y uso colectivo de la información.

En nuestro laboratorio se estudia desde hace más de 30 años, y desde diferentes ángulos, la asociación entre *Sinorhizobium meliloti* y alfalfa como sistema modelo de una interacción simbiótica bacteria-planta en el suelo. Ante la posibilidad de contar con una colección de aislamientos de *S. meliloti* portadores de plásmidos bien caracterizados en su diversidad y capacidad conjugativa [1], en este trabajo de tesis hemos diseñado y puesto en práctica una estrategia para el aislamiento, caracterización, y análisis secuencial y funcional (a escala de megabases) de plásmidos crípticos de *S. meliloti*. En ese contexto, hemos luego investigado su rol en la evolución y conformación del genoma actual de los rizobios, y su relación con los plásmidos de otras bacterias asociadas a plantas. Como parte del trabajo hemos podido poner en evidencia vínculos de ortología de genes plasmídicos con genes de diferentes fracciones de los distintos replicones de *S. meliloti*, e inferir con varias herramientas (análisis de fracciones *core*, singletons, contenidos GC%, uso de codones) posibles relaciones de ancestralidad respecto de la adquisición de los mismos por parte *S. meliloti*. Hemos obtenido además evidencia que muestran al pSymA como el replicón con el contenido génico más vinculado al moviloma plasmídico críptico, y como el intercambiador más activo de genes metabólicos con el *pool* de genes móviles.

Por otra parte, teniendo en consideración que la micro-biota asociada a un nicho ecológico resulta de procesos permanentes de coevolución, en una segunda fase del trabajo aislamos y caracterizamos el moviloma plasmídico correspondiente a otras bacterias asociadas a semillas de alfalfa (simpátricas de *S. meliloti*, a partir de representantes de más de 10 géneros diferentes) e investigamos las características del mismo (genes, funciones) y su relación con el moviloma de los rizobios. Los resultados de los análisis realizados aportaron evidencia en apoyo del intercambio génico entre el moviloma de *S. meliloti* y el de la población de sus bacterias simpátricas, así como entre los plásmidos de estas últimas entre sí. La magnitud de información plasmídica aislada, secuenciada y analizada en esta tesis (megabases) colectada a

partir de una especie modelo de rizobio y de un conjunto de sus bacterias simpátricas representa el primer intento de abordar a escala comunitaria el análisis estructural de movilomas de un mismo nicho, incluyendo su contenido funcional, características comunes, y participación en la evolución de replicones más estables de las bacterias por donde transitan.

CAPITULO I
Introducción

I.1. La importancia de los microorganismos asociados a las plantas

La población humana creció 7 veces desde comienzos del siglo XIX [2]. Esto gradualmente condujo a que los recursos del planeta fueran sobreexplotados, con una pérdida masiva de biodiversidad, creciente cambio climático y perturbaciones en el ciclo del nitrógeno y fósforo [3]. Las mismas prácticas agrícolas que han incrementado los rendimientos agronómicos [4] para abastecer la demanda de alimentos de la creciente población mundial, han llevado a una distorsión masiva en el ciclo biogeoquímico del nitrógeno, donde la fijación de nitrógeno artificial por parte del hombre para su uso en la agricultura e industria ha superado ampliamente a la fijación biológica natural, implicando una sobrecarga de nitrógeno sobre todos los ecosistemas en forma de fertilizantes, ineficientemente utilizados, y desechos gaseosos como el óxido nitroso (N_2O), de altísima acción sobre el efecto invernadero, que ponen en riesgo el ecosistema global [5, 6]. La utilización de fertilizantes nitrogenados se estima en 10^{11} kg/año [7, 8], sin embargo cerca del 60% de los mismos no son adsorbidos por las plantas y la mayor parte de estos se filtran en las napas de agua. Un ejemplo de problema grave asociado a esto es el crecimiento desmedido de cianobacterias y algas en los mares por la filtración de este excedente de nitrógeno, afectando la cadena alimenticia marina [9]. Si bien los países desarrollados ya han logrado un incremento suficiente en el rendimiento de la producción de alimentos en base a estos fertilizantes contaminantes [10], los países en vías de desarrollo aún deben lograr un incremento semejante y es esperable que la utilización de los mismos fertilizantes sintéticos generen un daño aún mayor en el ecosistema. Sumado a esto, en un escenario optimista donde se duplique la producción de alimentos para el año 2050 [11], la misma no será suficiente para abastecer la demanda de una población creciente, que consume cada vez más productos de origen animal en sus dietas [12]. Otro de los problemas asociados al bajo rendimiento actual es el avance de la frontera agrícola y la deforestación sobre ambientes naturales, como las selvas en Brasil [13], Indonesia, Malasia, Paraguay, Bolivia, Zambia o Angola, como una solución transitoria para lograr este incremento en la producción pero que conlleva una pérdida estimada de 1,5 millones de km^2 desde el comienzo del siglo XXI [14]. En este contexto, es importante que seamos capaces de incrementar los rendimientos de la tierra que ya ha sido transformada en cultivable, y una de las formas más sostenibles de lograr esto es enfocando la investigación en las habilidades naturales de las plantas para incrementar su rendimiento. La forma de lograrlo se apoya en el estudio de los microorganismos asociados a las plantas y tiene el objetivo de generar alternativas ambientalmente sostenibles para el incremento de la producción global de alimentos.

I.2. El microbioma en las plantas

Actualmente se conoce que las plantas están colonizadas por un sorprendente número de microorganismos que incluso alcanzan densidades celulares mayores que las de la propia planta. Adicionalmente, la diversidad de genes microbianos sobrepasa enormemente el número de genes de la propia planta. Estos microorganismos pueden colonizar externamente la planta accediendo tanto desde la rizósfera (la interfase suelo-planta) como desde la filosfera (interfase aire-planta), por lo tanto la planta está en íntima conexión con la microbiota presente en el suelo y la microbiota circulante en el aire respectivamente [15-17]. Las plantas internamente pueden ser divididas en microambientes tales como la endorhiza (interior de las

raíces), la antósfera (flores), la espermosfera (semillas) y la carpósfera (frutas). Otra posible clasificación consiste en diferenciar entre los tejidos internos (la endosfera) y los tejidos externos (la ectósfera) [18]. A continuación se presentan los ecosistemas microbianos más frecuentemente estudiados en las plantas.

I.3. Microbiota de la rizósfera

Se le llama rizósfera a la zona de influencia de los exudados radiculares [19] y es considerada uno de los ecosistemas más complejos en la tierra [20-23]. En ella pueden encontrarse bacterias, hongos, oomycetes, nematodos, protozoos, algas, virus, arqueas y artrópodos [23-25]. La mayor parte de los miembros de este microbioma son parte de una compleja red alimenticia que utiliza la gran cantidad de nutrientes liberados por las raíces de la planta contenidos en exudados, células del borde, y mucílagos. Es conocido el rol de este ecosistema en la asimilación de carbono debido a la actividad microbiana. A diferencia de otros microambientes en plantas, la rizósfera se caracteriza por su alta abundancia microbiana [26] y diversas actividades bioquímicas [27]. La rizósfera fue propuesta como un escudo protector de las plantas frente a los patógenos presentes en el suelo [28].

Numerosos estudios han mostrado efectos de especificidad entre una determinada planta y la composición y abundancia de los microorganismos presentes en su rizósfera. Esta especificidad se observó tanto en cultivos agronómicos como en especies de plantas salvajes [19, 29-32].

Determinantes de la formación de comunidades microbianas en la rizósfera

La formación de las comunidades microbianas en la rizósfera está afectada por factores bióticos y abióticos. Muchos estudios han mostrado que la composición del suelo influye en la composición de las comunidades de hongos y bacterias en la rizósfera. Se ha observado que sus parámetros físico-químicos influyen en la fisiología de la planta modificando los exudados, que asimismo alteran la composición de la microbiota en la rizósfera [33-36]. En este sentido, la estructura de la comunidad microbiana resulta de una serie de factores: primero, el suelo y las semillas pueden ser considerados como fuentes de colonización microbiana [37], adicionalmente los parámetros físico-químicos del suelo, junto a los procesos biogeográficos, estructuran esta comunidad [38]. Además, la localización donde las plantas crecen determinan a qué microorganismos se encuentran expuestas y finalmente la especie de planta y su genotipo pueden determinar cuáles microorganismos pueden crecer y prosperar en la rizósfera. La especie de planta puede influenciar fuertemente la composición y actividad de la microbiota de la rizósfera, y las diferencias en la morfología de la raíz así como la cantidad y tipo de rizo-depósitos entre las plantas contribuyen en gran manera a este efecto específico de especie [39-44]. Ciertos metabolitos específicos liberados en la rizósfera pueden disparar múltiples respuestas en diferentes microorganismos de suelo. Por ejemplo, ciertos flavonoides producidos por plantas de soja pueden atraer no sólo a los simbioses como *Bradyrhizobium japonicum* sino también a patógenos como *Phytophthora sojae*. Los flavonoides también estimulan la germinación de esporas de micorrizas y la multiplicación de hifas. Un aspecto importante es que no sólo la especie de planta puede influenciar en la microbiota sino también el cultivar [36, 45-48]. Esto puede ser así debido a que los programas de crecimiento de granos se desarrollan generalmente sobre monocultivos en condiciones axénicas libres de patógenos

de suelo, por lo que minimizan la contribución del microbioma de la rizósfera al crecimiento de la planta y su sanidad. En este contexto, se ha mostrado que la domesticación que se ha llevado a cabo durante el cultivo moderno de plantas ha desfavoreciendo en algunos casos aquellas características de la planta que dan lugar y mantienen a microorganismos benéficos [49, 50].

Además de la influencia del tipo de suelo y del genotipo de las plantas, los microorganismos presentes en y sobre las semillas son una potencial fuente en la formación de la microbiota de rizósfera [51]. Éste es uno de los aspectos que menos se ha estudiado en la composición del rizobioma de la planta, es decir el llamado “efecto maternal” [52]. Estudios tempranos mostraron que la diversidad presente en las plántulas de pepino se parecía más a la presente en la tierra comparada con la presente en la superficie de las semillas [53]. Esto parece restarle importancia a dicho efecto, contrastando con los estudios que muestran que la microbiota puede transmitirse en el interior de las semillas, como lo muestra un trabajo sobre variedades de maíz salvaje y moderno, donde parece existir una microbiota conservada de endófitos en semillas a lo largo del tiempo [54]. El reclutamiento de microbiota desde las semillas a la rizósfera sugiere que las plantas pueden transmitir microorganismos específicos (quizás sólo una pequeña fracción de las que colonizan la planta) de una generación a la otra. Esta idea de acarreamiento de una microbiota que luego formará parte estable en el rizobioma tiene implicancias importantes en la co-evolución de las interacciones planta-microorganismo en los ecosistemas naturales.

I.4. Microbiota de la filósfera

La filósfera comprende la parte aérea de las plantas y está dominada por las hojas. El hábitat de la superficie de las hojas es vasto, estimándose por modelado de vegetación [55] que el área global de hojas es de aproximadamente mil millones de km^2 de superficie superior e inferior de hojas, lo que significa una superficie dos veces mayor que la superficie de tierra en el planeta. Asumiendo que existen 10^6 - 10^7 bacterias por cm^2 en la superficie de la hoja [15], la población presente en el total de hojas podría ser de cerca de 10^{23} bacterias. En cuanto a los cambios ambientales que experimentan, el ciclo de vida de las plantas se completa cada año en las plantas anuales, mientras que las perennes forman y dividen hojas a lo largo del año. Por lo tanto, la filósfera puede ser considerada como un ambiente de vida media corta en oposición a la rizósfera. Los colonizadores exitosos de la filósfera serán aquellos que se multipliquen y ocupen nuevos nichos mientras las hojas se expanden. Los microorganismos epífitos están expuestos a la atmósfera, se ven sujetos al ciclo diario y son influenciados directamente por la luz del sol e indirectamente por el metabolismo de la planta. Por otra parte la cutícula hidrofóbica que recubre la superficie de las hojas reduce la evaporación de agua así como la filtración de metabolitos de la planta, por lo que puede considerarse como un ambiente oligotrófico. A pesar de estos aspectos unificadores existen grandes heterogeneidades tanto en la macroescala (cambios climáticos estacionales, localización de las hojas en la planta y la composición de la cutícula) como en la microescala (presencia de venas, estomas y estructuras como tricomas e hidátodos que alteran la distribución de nutrientes [56, 57]. Sumado a estas variaciones nutricionales, los microorganismos filósfericos están expuestos a las variaciones en la radiación UV del sol y en la disponibilidad de agua, por lo que la distribución y multiplicación de microorganismos se produce de manera irregular en las

hojas [58]. Además, estos microorganismos se enfrentan frecuentemente con una serie de compuestos antibacterianos producidos por las mismas plantas, bacterias y hongos, o bien pueden ser expuestos a una respuesta de defensa en las plantas al momento de ingresar por los apoplastos. En este último caso esta respuesta podrá ser contrarrestada a menudo mediante la secreción de efectores por parte de los microorganismos [59], siendo este tipo de respuesta determinante cuando el modo de vida epífito representa sólo una etapa en el ciclo de vida en una bacteria que luego es endófito o patógena.

Origen de las comunidades filoféricas

La observación de que las poblaciones bacterianas se reestablecen en la filósfera de las plantas anuales con patrones similares a los observados en los años previos [60], sugiere la presencia de reservorios de estas bacterias en las plantas. Aunque una fracción de microorganismos en la filósfera puede explicarse debido a colonizadores provenientes del aire, el estudio de los perfiles del gen 16S rRNA bacterianos sugiere que existe un reclutamiento temprano de una comunidad núcleo que coloniza la planta desde las semillas, el suelo u otras plantas, el cual explicaría la consistencia en la comunidad bacteriana dominante entre miembros individuales de plantas anuales [60].

Resistencia a estreses en la filósfera

La filósfera está expuesta al oxígeno producido en la fotosíntesis y a la luz UV durante el día. En consecuencia los microorganismos colonizadores son expuestos a daños mediados por ROS (respuesta oxidativa sistémica) en sus ácidos nucleicos, lípidos y proteínas. Algunos mecanismos de protección que mostraron experimentalmente ser eficientes en mejorar el desempeño de las bacterias filoféricas son: producción de pigmentos, activación de mecanismos de reparación del ADN (por ej. fotoliasas), catalasas, superóxido dismutasas y proteínas de protección al ADN durante la escasés nutricional.

Otro factor ambiental presente en la filósfera es la desecación [61]. Se conocen al menos dos mecanismos por los cuales las bacterias contrarrestan estos factores de estrés: la formación de agregados y la excreción de compuestos bioactivos. La importancia en la formación de agregados se mostró mediante el análisis de la supervivencia de células de *Pseudomonas syringae* comparando células individuales y en forma de agregados, frente a condiciones de desecación y humedad [62]. Los agregados involucran la formación de exopolisacáridos que ayudan a mantener una capa hidratada rodeando a la bacteria [63-65]. Asimismo la excreción de biosurfactantes aumenta la humificación y son comúnmente producidos por las bacterias filoféricas [64, 65]. Las fluctuaciones en la disponibilidad de agua también implican estreses osmóticos que son contrarrestados mediante la producción de solutos compatibles, como colina o trehalosa, tanto por síntesis *de novo* como por importación de osmoprotectantes derivados de plantas [66, 67].

Adaptación metabólica en la filósfera

En contraste con la rizósfera, se sostiene que la filósfera es un ambiente escaso en nutrientes [15], por lo que se especula que las bacterias son capaces de concentrar nutrientes gracias al incremento de humificación debido a la producción de biosurfactantes [65] o bien mediante la

producción de ácido indol acético (AIA), que promueve el crecimiento vegetal generando en el proceso adelgazamiento de la pared celular de la planta y la consecuente liberación de sacáridos [15, 68]. Esta hipótesis sostiene que la producción de AIA en bacterias filósfericas mejora el desempeño de los microorganismos colonizadores a la vez que favorece el crecimiento de la planta [69]. Otras fuentes de carbono presentes en la filósfera son los isoprenos y compuestos volátiles monocarbonados [70]. Por ejemplo se sabe que el metanol producido por las plantas favorece el crecimiento de *Methylobacterium extorquens* y levaduras metilotróficas como *Candida boidinii* [71-73]. Ciertos derivados del metanol presentes en la filósfera [74] favorecen la presencia del género *Hyphomicrobium* [75]. Se ha reportado la presencia de numerosos transportadores del tipo ABC, involucrados en el transporte de azúcares y aminoácidos en *Pseudomonas* spp., y del tipo TonB (TBDTs) en *Sphingomonas* spp. [76]. Esto sugiere que estas especies sobreexpresan aquellos transportadores que puedan incrementar la concentración de ciertos sustratos presentes en baja concentración en la filósfera. La utilización de amonio [77] y la fijación biológica de nitrógeno en la filósfera se ha reportado como una estrategia para evitar la limitación del nitrógeno [78]. Asimismo, la producción de transportadores de otros macronutrientes esenciales como fosfatos, sulfatos y hierro se han reportado como formas de adaptación a esta escasés nutricional [79]. Se ha indicado que la producción de sideróforos por parte de las bacterias filósfericas es proporcional a la producción de taninos por la planta huésped, debido a la capacidad de quelar el hierro por parte de estos últimos y como parte de la respuesta de la bacteria ante esta insuficiencia del hierro disponible [80]. Estudios recientes mostraron que las bacterias fotótrofas que habitan la filósfera producen bacteriorodopsinas para generar energía a partir de la luz solar [81, 82].

El estudio de cambios de expresión por análisis de transcriptómica y proteómica en las bacterias filósfericas puede proveer indicios sobre los mecanismos de adaptación a los diferentes estreses a los que se ven expuestas en la filósfera. Un ejemplo de esto lo muestra un estudio que reveló al regulador de respuesta PhyR como un elemento clave en la adaptación al calor, desecación, daño oxidativo, exposición a luz UV y estrés osmótico, el cual mostró regular positivamente genes involucrando la expresión de chaperonas, síntesis de osmoprotectante trehalosa y biosíntesis de carotenoides, entre otras respuestas en *M. extorquens* [83, 84] y *Sphingomonas* spp. [85]. En *Pseudomonas syringae* el cofactor sigma 22 mostró ser importante para la adaptación a la filósfera [63].

I.5. Microbiota de la endósfera

Los microorganismos son considerados endófitos cuando colonizan del interior de las plantas en la mayor parte de su ciclo de vida, no generan patogénesis y frecuentemente son aislados de tejidos vegetales desinfectados superficialmente. Los endófitos fueron ignorados históricamente y considerados contaminantes posiblemente debido en parte a su baja abundancia en comparación con las bacterias rizosféricas [86, 87]. Hoy en día se considera que de las más de 300 mil especies de plantas conocidas, cada una de ellas es huésped de al menos un endófito [88]. En este sentido existe un universo enorme de biodiversidad e interacciones endófitos-planta por abordar. Los endófitos colonizan frecuentemente los mismos nichos que los patógenos, lo cual los hace buenos candidatos de cara al biocontrol de microorganismos [89] [90], insectos [88] y nematodos [86] fitopatógenos. En algunos casos los endófitos

pueden favorecer la germinación, promover el enraizamiento de la plántula en condiciones adversas [91] y mejorar el crecimiento de la planta [92]. Algunos endófitos bacterianos han mostrado prevenir el desarrollo de la patogénesis a través de la síntesis *de novo* de novedosos compuestos antifúngicos, mostrando un rol importante como biocontroladores en las plantas y su rol privilegiado en el manejo integrado de plagas. Por otro lado, el estudio de metabolitos producidos por microorganismos endófitos es de sumo interés por ser una fuente potencial de novedosas moléculas para tratamientos de enfermedades en humanos y otros animales, tal como ha sido el caso del taxol [88]. Junto con la producción de nuevos químicos, muchos endófitos han mostrado una capacidad natural de degradación de xenobióticos. La habilidad de algunos endófitos de mostrar resistencia a metales pesados/antimicrobianos y de degradar distintos compuestos orgánicos probablemente tenga relación con su exposición natural a diversos compuestos, en su interacción con el nicho suelo/planta [93-98].

Aislamiento y diversidad de endófitos bacterianos

El nicho endófito ofrece protección del ambiente a aquellas bacterias que puedan colonizar y establecerse *in planta*. Estas bacterias generalmente ocupan los espacios intracelulares y han sido aisladas de todos los compartimentos de la planta, incluyendo de las semillas [99]. Las mismas se han aislado tanto de plantas monocotiledóneas como dicotiledóneas, incluyendo tanto especies leñosas (como pera y roble) como herbáceas (como azúcar y maíz). Los estudios clásicos de diversidad de bacterias endófitas se han enfocado en la caracterización de aislamientos obtenidos a partir de tejidos internos luego de una desinfección de las superficies de las plantas con hipoclorito de sodio o agentes similares [100].

Frecuentemente, las comunidades bacterianas que habitan tallos, raíces y hojas de una gran variedad de plantas han sido estudiadas por técnicas basadas en el gen ARNr 16S tales como RFLP, DGGE o clonado y secuenciación de fragmentos de este gen. Ciertos estudios señalan que cinco taxones bacterianos muestran los niveles más prometedores de colonización y persistencia: *Cellulomonas*, *Clavibacter*, *Curtobacterium*, *Pseudomonas* y *Microbacterium* [101, 102].

Endófitos como promotores del crecimiento vegetal

Los endófitos pueden promover el crecimiento vegetal por mecanismos similares a los frecuentemente estudiados en rizobacterias. Estos incluyen solubilización de fosfatos [103, 104], producción de ácido indol acético [105] y producción de sideróforos [106]. Éstos también pueden proveer con vitaminas esenciales para las plantas [107]. Otros atributos adaptativos han sido relacionados a los endófitos, tales como ajuste osmótico, regulación estomatal, modificación de la morfología radicular, mejora en la incorporación de nutrientes y aceleración de la acumulación y metabolismo del nitrógeno [108, 109].

Endófitos y biocontrol

Los endófitos poseen la capacidad de aminorar o prevenir los efectos deletéreos de ciertos organismos patógenos. Sus efectos benéficos parecen ocurrir por mecanismos similares a los descritos en las bacterias rizosféricas [108-110]. Algunas enfermedades de origen fúngico, bacteriano o viral, e incluso en algunos casos daños causados por insectos y nematodos,

pueden ser reducidos mediante la inoculación con endófitos [111-114]. Se cree que ciertos endófitos inducen un fenómeno conocido como resistencia inducida sistémica (ISR), el cual es fenotípicamente similar a la resistencia sistémica adquirida (SAR) y genera una respuesta sistémica de la planta que es efectiva contra distintos patógenos virales, bacterianos o fúngicos así como contra distintos herbívoros [115].

Productos naturales de bacterias endófitas

Muchos endófitos son miembros de géneros bacterianos comunes de suelo, tales como *Pseudomonas*, *Burkholderia* y *Bacillus* [116]. Estos géneros son bien conocidos por su producción de metabolitos secundarios diversos tales como antibióticos, compuestos orgánicos volátiles, antifúngicos, antivirales o agentes insecticidas. Mientras un amplio rango de compuestos biológicamente activos han sido aislados de microorganismos endófitos, todavía siguen siendo una fuente inexplorada de nuevos productos naturales. Aunque la mayor parte de la atención se ha centrado en la producción de antimicrobianos de origen fúngico, cierto número de compuestos de bajo peso molecular, activos a bajas concentraciones con actividad en humanos, animales y microorganismos patógenos de plantas se han aislado de endófitos.

Tabla I.1: Algunos metabolitos de interés agronómico o biomédico aislados de distintos microorganismos endófitos.

Organismo	Planta asociada	Agente activo	Actividad	Referencia
<i>Taxomyces andreanae</i>	<i>Taxus brevifolia</i>	Taxol	Anticancerígeno	[117]
<i>Pseudomonas viridiflava</i>	gramineas	Ecomicina B y C	Antimicrobiano	[118]
<i>Streptomyces griseus</i>	<i>Kandelia candel</i>	p-ácidos aminoacetofenólicos	Antimicrobiano	[119]
<i>Streptomyces NRRL 30562</i>	<i>Kennedia nigricans</i>	Munumbicinas	Antibiótico	[120]
<i>Streptomyces NRRL 30566</i>	<i>Grevillea pteridifolia</i>	Kakadumicina	Antibiótico	[121]
<i>Serratia marcescens</i>	<i>Rhyncholacis penicillata</i>	Oocidina A	Antifúngico	[88]
<i>Paenibacillus polymyxa</i>	Trigo	Fusaricidina A-D	Antifúngico	[122]
<i>Cytospora sp.</i>	<i>Quercus sp. 103</i>	Ácido citónico A y D	Antiviral	[123]
<i>Streptomyces sp.</i>	<i>Monstera sp.</i>	Coronamicina	Antimalárico, antifúngico	[124]

I.6. Microbiota asociada a semillas

Dos de las preguntas más frecuentes en conexión con las bacterias endófitas aisladas en las plantas es cuál es su origen y cómo ingresan en los tejidos de las plantas en la naturaleza. Las bacterias endófitas pueden provenir tanto de las semillas [125-130], como del suelo rizosférico [86] o de la filósfera [131].

Poco se ha explorado la presencia de bacterias en el interior de las semillas, que pueden ser halladas al esterilizar la superficie de las mismas con distintos agentes desinfectantes y su posible rol en la colonización de la planta. Algunos trabajos recientes comienzan a develar que las bacterias colonizan tejidos germinativos [132], son acarreadas por las plantas y se observa su transmisión vertical a través de sucesivas generaciones [52], posiblemente *vía* colonización

de los tejidos adultos y acarreamiento en el interior de las semillas. Un trabajo reciente ha comparado la diversidad bacteriana presente entre una variedad ancestral y una domesticada de plantas del género *Zea*, cultivadas en diferentes suelos a partir de semillas sin esterilizar, y el mismo sugiere que la microbiota presente en los tallos y raíces de ambas variedades es coincidente en su mayor medida a la diversidad acarreada en sus semillas y en menor medida a una microbiota presente en los distintos suelos. Esto sugiere la idea de que la diversidad bacteriana que colonizará la endosfera de la planta es resguardada y acarreada en la semilla, dejando la influencia de colonización de la microbiota presente en el suelo en un segundo plano [133].

Además, al explorar la microbiota en variedades diferentes de semillas del mismo género, el mismo grupo ha mostrado que ciertos géneros están conservados en el interior de las semillas de distintas variedades del género *Zea* ancestrales y domesticadas [54], e incluso que algunos de los aislamientos de las semillas de variedades ancestrales fueron capaces de inhibir hongos fitopatógenos, mostrando ser un reservorio natural de microorganismos protectivos [134]. En tal contexto es razonable asumir que años de co-evolución entre una planta y su propio microbioma hayan servido para seleccionar aquellas especies (o clones) que favorecen, por sobre otras, el desarrollo de la planta que las porta.

1.7. La simbiosis fijadora de nitrógeno: *Sinorhizobium meliloti*-alfalfa

Las bacterias agrupadas en las familias Rhizobiaceae, Phyllobacteriaceae y Bradyrhizobiaceae, habitan el suelo en condiciones de crecimiento libre y se asocian en simbiosis con las raíces de las legumbres como organismos fijadores de nitrógeno. Se conocen a la fecha más de 60 especies dentro de estos rizobios fijadores pertenecientes al orden de los Rhizobiales, agrupados actualmente en 7 géneros: *Ensifer* (también conocido como *Sinorhizobium*), *Mesorhizobium*, *Rhizobium*, *Bradyrhizobium*, *Azorhizobium*, *Allorhizobium* y *Agrobacterium*.

Sinorhizobium meliloti es un rizobio capaz de establecer asociaciones simbióticas con plantas de los géneros *Medicago*, *Melilotus* y *Trigonella*. Esta asociación es el resultado de un complejo diálogo molecular entre los simbioses, que se diferencian a lo largo de la interacción para dar lugar a un nuevo órgano en las raíces de las plantas, el nódulo fijador de nitrógeno.

Intercambio de señales

El diálogo comienza cuando compuestos aromáticos (flavonoides) presentes en los exudados de la planta sirven como una primera señal para que las bacterias rizosféricas produzcan lipoquitooligosacáridos, denominados factores Nod [135]. Para lograr esta activación, los flavonoides se unen a reguladores transcripcionales ya expresados en la bacteria (NodD), los cuales entonces inducen la transcripción de los restantes genes *nod* de rizobios. [135, 136]. Los factores Nod activan múltiples respuestas en la planta, preparándola para recibir a la bacteria [137]. Tanto los factores Nod como los exopolisacáridos simbióticos producidos por las bacterias inducen a la planta a formar hilos de infección, túbulos finos llenos de bacterias que penetran el tejido cortical de la planta y conducen a las bacterias hacia las células “blanco”. Una de las primeras respuestas de la planta al factor Nod es el incremento de los niveles intracelulares de calcio en los pelos radiculares (“spiking”)[137] y alteraciones al citoesqueleto del pelo radicular [138-141]. Luego de estas respuestas, sucede el enrollamiento de los pelos

radiculares, que atrapan a las bacterias rizobiales en el pelo radicular curvo colonizado [142, 143]. En simultáneo los factores Nod estimulan a las células del cortex radicular a retomar la mitosis en un proceso que depende de la interrupción del transporte de auxinas en la planta causada por los rizobios y por los flavonoides, y que dará lugar al primordio del nódulo y a las células que recibirán a los rizobios [138, 144, 145].

Desarrollo del hilo de infección

En todas las interacciones rizobio-leguminosa, las bacterias deben penetrar al cortex radicular antes que puedan fijar nitrógeno [135]. Esto lo logran a partir de la producción de hilos de infección por parte de la planta, previamente siendo atrapadas por la punta del pelo radicular curvo. Se han encontrado varios mutantes que son incapaces de colonizar el pelo de infección pero no se conoce aún el mecanismo responsable. [146-148]. Las bacterias que han sido atrapadas por el hilo de infección y pueden producir exopolisacáridos (EPS) simbióticos y factores Nod, producen el crecimiento hacia adentro de la membrana celular del hilo de infección, y pueden invadir así los tejidos internos [143]. La invasión ocurre incluso si el factor Nod y los exopolisacáridos son aportados por distintas *S. meliloti* atrapadas juntas en el mismo hilo de infección [149]. El crecimiento del hilo de infección hacia adentro supone un cambio de polaridad en el crecimiento [143] y es en esa zona del crecimiento del hilo donde se produce la división celular bacteriana [150].

S. meliloti produce succinoglicano y galactoglicano, involucrados en la formación del hilo de infección [151, 152]. Los mutantes de los genes involucrados en la producción de succinoglicano (*exoY*) no pueden formar hilos de infección, quedándose arrestados formando una microcolonia en la punta del hilo curvo, mientras que los mutantes en genes involucrados en las modificaciones de los exopolisacáridos (*exoH*) forman hilos de infección abortivos [153], lo cual refleja la importancia de los EPS en la formación del hilo de infección. Por otro lado cuando los factores Nod no están en su forma estructural completa o la percepción de los mismos es interrumpida, también se generan fenotipos de producción de hilos de infección abortivos [154-156]. Una vez que las bacterias alcanzan la base de las células del hilo de infección deben generar nuevos hilos a través de las siguientes capas celulares, hasta llegar al nódulo en formación [144]. En el mismo se generará una zona de meristema central que persistirá a lo largo de la vida del nódulo [138, 144]. En el caso de *M. truncatula* y *M. sativa*, estos meristemas son persistentes y los nódulos se denominan indeterminados. En el caso de *L. japonicus* los nódulos son determinados, donde todas las células se dividen en sincronía, se diferencian y senescen [157]. En los nódulos indeterminados, el hilo de infección cruza la zona meristemática por donde la oleada mitótica ya ha pasado [157] para alojarse en las células capaces de ser infectadas y las bacterias se diferencian a un fenotipo fijador de nitrógeno, llamado bacteroide. Este cambio implica que la bacteria desarrolle la multiploidía debido a múltiples rondas de endoreduplicación sin citocinesis [144, 158]. Las ventajas de la multiploidía involucran una alta tasa de transcripción y alta actividad metabólica, lo cual se cree favorecería el proceso de fijación de nitrógeno [159].

Colonización intracelular

Las células de la corteza interior de la planta internalizan a las bacterias entrantes en compartimentos membranosos que maduran a estructuras llamadas simbiosomas. Cuando la

bacteria encuentra su tejido blanco en el cortex interior de la raíz, puede ser internalizada por una célula cortical y establecer un nicho en dicha célula. Cada bacteria es endocitada en un compartimento individual formado por membrana derivada del hilo de infección. Esta estructura formada por una bacteria individual y la membrana endocítica es llamada simbiosoma [160]. En los nódulos indeterminados, las bacterias en su membrana se dividen en sincronía hasta diferenciarse a bacteroides fijadores de nitrógeno [161].

Diferenciación y supervivencia

Una vez que la bacteria ha sido englobada dentro de la membrana de la célula huésped, debe sobrevivir dentro del compartimento del simbiosoma y diferenciarse a la forma bacteroide. Tanto factores de la bacteria como de la planta son necesarios. Uno de los mecanismos de defensa más importantes de las bacterias Gram-negativas contra el ambiente extracelular es el lipopolisacárido (LPS), que es un componente mayoritario de la membrana externa, el cual consiste de un lípido A anclado a la membrana, una cadena de polisacárido y una unidad repetida de antígeno O [162]. El gen *bacA* en *S. meliloti* es requerido para la producción del lípido A en su estructura correcta que le permite la supervivencia en la célula huésped. Un mutante en el gen *lpsB*, produce un núcleo de LPS con una composición alterada de azúcares [162] y, aunque puede invadir la raíz, es envuelta en la membrana del simbiosoma y comienza a elongarse a la forma bacteroide, no puede completar el programa a bacteroide y comenzar a fijar N₂ [162]. Otros genes que mutados impiden una fijación normal de nitrógeno son: el *sitA*, el cual es un transportador de manganeso (el mutante no puede crecer en medio mínimo sin manganeso adicional [163], cuyo mutante puede invadir la planta con eficiencia reducida pero sus bacteroides senescen y no pueden fijar nitrógeno. El mutante en el cluster *pha* presenta un fenotipo parecido, tratándose de un transportador de potasio activo en condiciones alcalinas, lo que sugiere que los mutantes que fallan en la homeostasis de ciertos iones no pueden persistir en el ambiente intracelular [164]. Los mutantes en reguladores globales de respuesta a estrés como *rpoH1* también son deficientes en la simbiosis, en este caso el gen codifica el factor sigma 32 y parece invadir normalmente pero forma bacteroides que senescen prematuramente [165]. En contraste, el mutante en el gen *relA* es defectivo en múltiples etapas en la formación del nódulo. Este mutante carece de la respuesta al estrés por falta de nutrientes [166]. Otros requerimientos descritos para la correcta simbiosis son la ribonucleótido reductasa dependiente de vitamina B12, la cual en rizobios está adaptada para su correcto funcionamiento independiente de la concentración de oxígeno. Este enzima es esencial para la abundante producción de dNTPs requerida en el proceso de endoreplicación. El transporte eficiente de fosfatos ha mostrado ser relevante debido a que los mutantes en los sistemas *phoCDET* y *pstCAB* forman nódulos pequeños que son incapaces de fijar nitrógeno [167, 168].

Adaptación metabólica en el desarrollo nodular

Una vez que las bacterias han sido englobadas por una membrana completamente funcional, incluidas en un ambiente anaerobio y se han diferenciado a bacteroide, la expresión de los genes de nitrogenasa y enzimas respiratorias de condiciones microaeróbicas comienza y puede fijarse nitrógeno [169]. Entre los componentes regulatorios sensores de la concentración de oxígeno encontramos a FixL, FixJ, NifA, sigma 54 y FixK [169]. Estos reguladores son los

responsables de muchos de los cambios en la expresión de las proteínas que se han detectado durante la diferenciación a bacteroide [170]. En general estos cambios incluyen una regulación negativa de la mayoría de los procesos metabólicos y un aumento concomitante de los productos génicos involucrados en la fijación de nitrógeno y respiración [170]. La respiración provee a la nitrogenasa con 16 moléculas de ATP y 8 electrones que se estiman son requeridos para reducir 1 molécula de N_2 a 2 moléculas de NH_4^+ [171]. El amonio que es liberado por los bacteroides luego se incorporada por canales de amonio en la membrana peribacteroide [172, 173] y es asimilado principalmente por las enzimas glutamina y asparagina sintetetas [174, 175]. Por otro lado, se requiere un flujo de carbono constante para la diferenciación a bacteroide y la fijación de nitrógeno. Los gránulos de polihidroxibutirato (PHB) que son producidos por *S. meliloti* durante el proceso de invasión, son degradados durante la diferenciación a bacteroide y se usan como una fuente de carbono preferencial [176]. Sin embargo, los mutantes en la síntesis y degradación de PHB también pueden desarrollar nódulos normales, aunque son menos competitivos en infecciones mixtas [176, 177]. El carbono necesario también es provisto por la planta en forma de malato (adquirido mediante el sistema de importación bacteriano Dct) [171, 178]. Se requiere la conversión de malato a piruvato para una correcta fijación de nitrógeno [179]. Para ello se sugiere que existe una producción de acetil-CoA desde malato por la enzima málica y la piruvato deshidrogenasa, de manera de canalizar carbono al ciclo de Krebs [171]. Algunas proteínas en plantas se ha demostrado asimismo importantes para la fijación. Por ejemplo, las leghemoglobinas que le dan al nódulo un color rojo característico son necesarias para establecer el ambiente microaeróbico en los nódulos [180]. Otra proteína requerida es la sacarosa-sintasa (SS) que cataliza la conversión de sacarosa a UDP-glucosa y fructosa que finalmente se metaboliza a malato y se transporta al bacteroide [171, 181]. La formación de clusters hierro sulfuro en las subunidades de nitrogenasa requieren asimismo el transporte de sulfato (Sst1) hacia las membranas del simbiosoma [182, 183].

Evasión de la respuesta inmune. Componentes de los rizobios

Durante la invasión y la simbiosis, los rizobios son capaces de evadir la respuesta inmune innata de la planta. Los epitopes que generalmente son reconocidos por las plantas para montar una respuesta inmune sobre las bacterias invasoras pueden estar ausentes en el caso de los rizobios. Tal es el caso del epitope Flg22 de la flagelina [184]. Asimismo las legumbres, a diferencia de otras plantas, son incapaces de montar una respuesta a epitopes frecuentemente reconocidos tales como el péptido Elf18 (del factor de elongación Tu) o Csp15 (de la proteína de respuesta a estrés térmico) [185]. Estas y otras adaptaciones de las respuestas inmunes parecen haberse adquirido durante la evolución para permitir la simbiosis. Las defensas naturales de la planta también se cree que están involucradas en la autolimitación a la formación de nódulos [186]. Por otro lado, se postula que los exopolisacáridos producidos por las bacterias son responsables de la atenuación de la respuesta inmune de la planta [187]. En cuanto al LPS, frecuentemente causante de respuesta oxidativa en las plantas, en el caso de *S. meliloti* presenta un tipo de lípido A que evita este tipo de respuestas en sus plantas huéspedes *M. sativa* y *M. truncátula* [188, 189].

I.8. Microbioma en alfalfa

No son muchos los estudios que han indagado acerca de otros microorganismos presentes en la planta de alfalfa, más allá del extenso trabajo en el estudio de su bacteria simbiote natural, *Sinorhizobium meliloti*. Uno de los primeros trabajos realizados para caracterizar a bacterias que acompañan a *S. meliloti* en alfalfa mostró la presencia de *Erwinia herbicola* en semillas de 9 variedades de alfalfa, y su influencia sobre la capacidad de nodulación de diferentes cepas de *S. meliloti* [190]. En cuanto a la presencia de otras bacterias presentes en los nódulos, se han hallado distintas especies incluyendo *Bacillus megaterium*, *Brevibacillus chosinensis* y *Microbacterium trichothecenolyticum*. Dichos aislamientos mostraron generar mayor número de nódulos al coinocularse con *S. meliloti* pero sin mostrar cambios en el tamaño de las plantas al comparar con las inoculadas sólo con *S. meliloti*. Asimismo, estos aislamientos no rizobiales mostraron promover el crecimiento de la planta al ser inoculados en ausencia de *S. meliloti* [191].

Uno de los estudios más recientes del microbioma de alfalfa fue realizado en base a análisis de perfiles por T-RFLP, en plantas crecidas en condiciones de mesocosmos en suelos arenosos no calcareos, donde se muestra que las comunidades presentes en tallos y hojas de alfalfa son diferentes a las de nódulos y suelos [192]. En el mismo, se utilizaron muestras de suelo, nódulos, tallos y hojas de las plantas crecidas durante 2 años y se compararon los perfiles mediante escalado multidimensional no métrico (N-MSD). Los resultados mostraron que las comunidades presentes en suelo y nódulos diferían significativamente de las de tallos-hojas. Las bibliotecas generadas a partir de las muestras de suelo, nódulos y de tallos-hojas, mostraron que la mayor diversidad de géneros se encontró en el suelo, y la menor diversidad en nódulos, tal como se esperaba. Sólo 4 clases de bacterias se encontraron en nódulos y 7 clases tanto en suelo como en tallos-hojas. Al analizar los géneros de la clase *Alfaproteobacteria* se encontró que en nódulos sólo se hallaba el género *Sinorhizobium* mientras que en suelo las familias Rhizobiaceae, Bradyrhizobiaceae, Methylocystaceae, Hypomicrobiaceae y Caulobacteraceae estaban presentes. En el caso de tallos-hojas, las familias Rhizobiaceae, Aurantimonadaceae y Methylobacteriaceae sumado a familias del orden Sphingomonadales se hicieron presentes. Aunque el género *Sinorhizobium* resultaba estar ausente tanto en las muestras de suelo como en las de tallos-hojas, fue detectada su presencia en las de tallos-hojas por medio de qPCR mediante el uso de los genes *rpoE1* y *nodC*. En este trabajo tampoco pudieron identificar perfiles de T-RFLP en semillas esterilizadas con HgCl₂, probablemente debido al bajo título presente de microorganismos asociados a las semillas o a este método de esterilización utilizado.

I.9. Movilomas plasmídicos bacterianos

Mientras que los organismos unicelulares evolucionan en ambientes diversos y con condiciones cambiantes, las células que forman parte de organismos multicelulares ejercen funciones más especializadas gracias a la protección que reciben desde redes homeostáticas que amortiguan las variaciones del entorno. En este contexto evolutivo, las bacterias concentrado eficientemente en sus pequeños genomas la información necesaria para afrontar el desafío de los cambios ambientales [193-195]. En términos genéticos (informativos), y con pocas excepciones, las bacterias han concentrado la mayoría de sus funciones basales (housekeeping) en sus replicones más grandes (los cromosomas) mientras que ciertos

compartimientos genéticos móviles (plásmidos, islas de ADN, transposones, integrones) se observa que concentran genes relacionados a respuestas adaptativas [196, 197]. La partición del material genético en módulos, como elementos funcionales adicionales al cromosoma (“moviloma”, mobilome), facilitó la reducción del tamaño de los genomas permitiendo compartir genes (funciones) transitoriamente importantes (adaptativas) entre miembros separados de la propia comunidad. Una partición genética con redundancia previene, al mismo tiempo, la pérdida masiva de información genética cuando miembros individuales de la comunidad bacteriana mueren. El conjunto de plásmidos y otros elementos móviles en un ambiente dado representa el recurso genético principal donde las bacterias codifican y concentran sus respuestas moleculares a los estímulos ambientales (bióticos y abióticos). En este contexto, la obtención y caracterización en escala ómica de la información genética asociada a plásmidos en particular, y al moviloma en general, permite el acceso a nuevas proteínas y actividades enzimáticas diseñadas evolutivamente para acomodar/adaptar a los microorganismos al ambiente y a sus hospederos eucarióticos [198].

Plásmidos: características estructurales, informacionales y de movilización. Su rol como mediadores de adaptación.

Una gran proporción de la biodiversidad informacional y funcional en las bacterias está presente en plásmidos, esto es en replicones extracromosomales (en general circulares) y más pequeños que el cromosoma celular. Los plásmidos deben replicarse, controlar su número de copias, y asegurar su herencia en cada división celular. Los plásmidos están presentes en todas las ramas del árbol evolutivo bacteriano y se han encontrado en todas las comunidades bacterianas estudiadas hasta la fecha, incluyendo el suelo, ríos, y ambientes marinos. En cuanto a su complejidad, ciertos (mega) plásmidos pueden ser del tamaño de cromosomas pequeños y pueden contener varios replicones compatibles co-integrados. En contraste con esto, los aislamientos bacterianos naturales contienen a menudo plásmidos pequeños que abarcan solamente genes de la replicación, algunos genes asociados a actividades no esenciales en condiciones de crecimiento en el laboratorio y genes de función aún desconocida. Los plásmidos se cree son, en general, vehículos portadores de información ligada a necesidades funcionales transitorias vinculadas a procesos no codificados en el cromosoma bacteriano. Así, éstos pueden intercambiar DNA con el cromosoma de la bacteria huésped, y ser transferidos especialmente *via* conjugación a otras bacterias como mediadores de transferencia horizontal de genes (HGT, *Horizontal Gene Transfer*) [199]. Es un hecho bien reconocido que la habilidad de las bacterias para explotar nuevos entornos y responder a nuevas presiones de selección se sustenta principalmente en base a mecanismos de adquisición de nuevos genes por transferencia horizontal, más que por acumulación de mutaciones puntuales y modificación secuencial de genes [200-202]. Probablemente, el número de soluciones que la naturaleza ha encontrado para dar espacio a la expresión de nuevas habilidades es de distribución muy pequeña, y la transferencia horizontal de información (experiencia evolutiva) ocurre más rápidamente que la generación de nuevos caminos originados por acumulación de mutaciones puntuales y modificación secuencial de genes [200-203]. Así, la HGT entre bacterias es considerada como una de las principales fuerzas directoras que gobiernan la evolución de las estructuras genómicas procariontas [204, 205].

Metagenómica y plásmidos

La diversidad biológica y molecular presente en las comunidades bacterianas naturales es muy alta, con la mayor parte de ella no caracterizada, y el conjunto microbiano aun no explorado sobrepasa por mucho la información disponible de los microorganismos procariotas depositados en las diversas colecciones de cultivos [206, 207]. En paralelo a la creciente disponibilidad de secuencias de genomas completos bacterianos, la cantidad de secuencias plasmídicas completas en los últimos años se ha incrementado desde las 1490 secuencias completas hacia el año 2008 a casi 5 mil secuencias completas en Noviembre de 2015, habiendo sido aproximadamente la mitad de ellas secuenciadas como parte de proyectos de genomas bacterianos completos y la otra mitad incorporada como plásmidos completos en proyectos metagenómicos o en los cuales sus cromosomas no pudieron completarse. En lo que hace al ambiente suelo en particular, la diversidad de los microorganismos se ha explotado durante muchos años basado en el cultivo *in vitro* y el aislamiento y caracterización de las diferentes especies microbianas. Esto ha limitado *per se* la explotación de los recursos asociados a microorganismos no cultivables, que hasta hace algunos años representaban un depósito génico inexplorado. Las aproximaciones metagenómicas, independientes de la necesidad de cultivo, saltean este inconveniente y han comenzado a usarse con éxito cierto [208-211].

La exploración de plásmidos presentes en bacterias no cultivables ha sido posible a través de la recuperación de replicones conjugativos provenientes de las mismas, que codifican resistencias específicas de interés. En estas estrategias, ciertas bacterias marcadas son utilizadas como recipientes en experimentos de conjugación realizados con bacterias endógenas de varios ambientes. Los plásmidos aislados son caracterizados detalladamente con respecto a perfiles de resistencia, y agrupados según su grupo de incompatibilidad y a características de la transferencia. Es posible analizar a partir de muestras de ADN total de distintas comunidades la presencia y proporción de plásmidos pertenecientes a distintos grupos de incompatibilidad tales como IncP-1, IncP7 e IncP-9 [212]. Otras aproximaciones actuales al estudio metagenómico de los plásmidos incluyen: la preparación de gradientes de ultracentrifugación por cloruro de cesio de muestras de ADN de comunidades [211], la derivación bioinformática de genes/cóntigos relacionados a plásmidos como herramienta para la minería de secuencias plasmídicas en proyectos de metagenómica [213], la captura ayudada por transposones [214] y la degradación de ADN lineal mediante el uso de exonucleasas seguida por reacción de amplificación por desplazamiento múltiple (MDA), mediante el uso de la polimerasa viral phi29 [215].

En el caso de la purificación por gradientes de CsCl, las ventajas de este método incluyen la obtención de plásmidos de alto peso molecular intactos y en gran cantidad, siendo en este sentido una técnica superior a las demás. En el caso de la predicción bioinformática, el enfoque requiere que los cóntigos posean algún gen atribuible a un plásmido, por lo que es aplicable para analizar sólo cóntigos grandes o que presentan genes tales relacionados a la replicación, conjugación o genes de mantenimiento plasmídico. Además, su dependencia de la existencia de esos genes plasmídicos depositados en las bases de datos lo limita a secuencias genómicas donde esos genes estén presentes. La estrategia de TRACA ("*transposon aided capture*") mediada por transposones depende de la estabilidad del plásmido en *E. coli*, con lo cual su aplicación es limitada por fuera del rango de plásmidos que son mantenidos en dicha

especie receptora. Por último, la aproximación dependiente de la digestión de ADN lineal tiene la desventaja de que los plásmidos grandes son frecuentemente nickeados durante los procesos de purificación, con lo que luego de la digestión se encuentran frecuentemente sólo plásmidos de tamaños pequeños, aunque la disminución de la contaminación cromosomal es notable y se logra una buena purificación de ADN plasmídico de tamaños menores a 10 Kpb.

I.10. Movilomas plasmidicos en *Sinorhizobium meliloti*

Una característica común de los rizobios es que su genoma presenta un gran número de plásmidos que codifican un diverso número de funciones [216, 217]. Los plásmidos que llevan información necesaria para la simbiosis son llamados plásmidos simbióticos o pSyms [218-220]. Además de estos elementos simbióticos los rizobios presentan otros plásmidos, referidos como plásmidos no-simbióticos, crípticos o accesorios, que no son indispensables para la simbiosis o la vida libre en el laboratorio y cuyos roles biológicos aún no han sido esclarecidos. El hecho de que este tipo de plásmidos se pierda, se incorpore o modifique su número de copias o su contenido genético en la célula, es consistente con el rol que los plásmidos muestran como vectores transientes que posibilitan cambios adaptativos en las poblaciones bacterianas [221].

En muchos rizobios, el ADN plasmídico puede constituir un porcentaje sustancialmente alto del ADN bacteriano, tal como sucede en *S. meliloti* donde la contribución de los mega (plásmidos) llega a ser casi un 45% del total de su genoma. En la cepa de *S. meliloti* 1021, el genoma consiste de un cromosoma de 3,65 Mpb y dos plásmidos simbióticos pSymA y pSymB, cuyos tamaños son de 1,36 y 1,68 Mpb respectivamente. En otras cepas, otros plásmidos no simbióticos han sido descritos adicionalmente [1, 222-226], representando un porcentaje mayor de la información del genoma contenida en los plásmidos.

Megaplásmidos simbióticos

El contenido génico de los megaplásmidos simbióticos se conoce desde el secuenciamiento de la cepa *S. meliloti* 1021 [227, 228]. El megaplásmido pSymA codifica genes requeridos para la nodulación (genes *nod*), genes codificantes de la nitrogenasa (*nif*) y genes requeridos para la fijación de nitrógeno (*fix*). Además, pSymA contiene genes involucrados en el metabolismo, transporte del carbono y del nitrógeno así como de resistencia y respuesta a estreses que le otorgan a *S. meliloti* ventaja adaptativas en su nicho especializado [227]. Su secuenciamiento también mostró que el pSymA podría haber sido adquirido via transferencia horizontal a *S. meliloti* debido al menor contenido de GC (60,4%) en comparación con sus otros replicones (el cromosoma y el megaplásmido pSymB).

El megaplásmido pSymB codifica genes de adquisición de nutrientes así como enzimas involucradas en biosíntesis de polisacáridos y actividades catabólicas. El pSymB también muestra características comunes al cromosoma como lo es la presencia de genes esenciales del metabolismo (housekeeping), tales como el tRNA correspondiente a arginina, el gen *minCDE* involucrado en la división celular y la ruta completa de biosíntesis de la asparagina. En contraste, no se encontró presente en el pSymA ningún gen esencial para la vida libre. En un estudio de genómica comparativa realizado por hibridización de distintos aislamientos de *S. meliloti* frente a un microarray de la cepa *S. meliloti* 1021 realizado por Giuntini y

colaboradores [229], dichos autores concluyeron que la mayor cantidad de polimorfismos fue encontrada presente en el pSymA. Este megaplásmido mostró además una gran densidad de elementos transponibles, secuencias de inserción y elementos relacionados a fagos. En general, dichos elementos tienden a acumularse en lugares donde no existen genes esenciales. Toda esta información, sumada a la observación de que el megaplásmido pSymA puede ser curado mientras que el megaplásmido pSymB no, sugiere que el pSymA es un replicón que conserva características no esenciales semejantes a las codificadas en los plásmidos mientras que el pSymB es un replicón que ya ha adquirido un rol esencial del tipo cromosomal. La cepa curada en pSymA no mostró capacidades diferentes de crecimiento en medios ricos comparada con la cepa salvaje pero no es capaz de crecer en medios definidos con algunas fuentes de carbono presentes como únicas fuentes de carbono y energía [230]. Los genes presentes en pSymB se observaron más distribuidos entre diferentes taxones en comparación con los otros replicones de *S. meliloti* [231]. En este contexto, el pSymB es considerado como un replicón que tiene un rol en la diferenciación intra-especie a través de selección génica positiva, mientras que el pSymA tendría un rol en la adquisición de nuevas funciones [231]. Otro estudio sobre 3 cepas distintas de *S. meliloti* (1021, AK83 y BL225C) [232] mostró que el pSymB presenta una gran resistencia a los rearrreglos genómicos, mostrando una sintenia casi perfecta. Sin embargo, el pSymA mostró un bajo grado de sintenia y una mayor tasa de rearrreglos. El pSymA de la cepa *S. meliloti* SM11 presenta una diversidad mayor de funciones en comparación con los otros pSymA analizados, incluyendo regiones relacionadas a diferentes plásmidos de *Sinorhizobium medicae* WSM419 [233].

Plásmidos no-simbióticos en *S. meliloti*

Los plásmidos en los rizobios se consideran crípticos, accesorios o no simbióticos cuando no son necesarios para establecer una simbiosis completa [217]. El tamaño de los mismos es muy variable (entre 7 Kpb y 260 Kpb). Algunas regiones de los mismos pueden estar involucrados en la aparición de rearrreglos genéticos debido a la presencia de secuencias de inserción, o cointegraciones de otros plásmidos [234, 235]. La presencia de secuencias repetidas en los plásmidos crípticos sugiere que la recombinación de los mismos con otros plásmidos, simbióticos y no simbióticos, en los rizobios es frecuente [236]. La información disponible indica que los plásmidos crípticos proveen a los rizobios de funciones con ventajas adaptativas, que influyen en el *fitness* ambiental de los mismos, por lo cual se asume que los cambios fenotípicos derivados de la mutación de sus genes son difíciles de detectar en el laboratorio [237]. En este sentido, en ensayos de liberación al ambiente de cepas modificadas genéticamente se han detectado cepas salvajes de *S. meliloti* capaces de desplazar y prevalecer en los nódulos, las cuales mostraron contener plásmidos crípticos responsables de dichas diferencias [238]. La observación de ciertas ventajas en la simbiosis atribuidas a diferentes plásmidos crípticos [239] o incluso la presencia de genes homólogos a los implicados en la nodulación [240] parece indicar cierta ventaja no esencial en el proceso simbiótico para este tipo de plásmidos, que por otro lado no parece ser el caso en otros plásmidos crípticos que parecen afectar el proceso simbiótico [241]. Es posible, entonces, que el rol de este compartimiento génico accesorio en las funciones adaptativas de la bacteria no sea unívoco, pudiendo a veces sugerir una ventaja adaptativa frente a la simbiosis y en otros casos una desventaja o quizás una ventaja en otro tipo de condiciones aún no esclarecidas.

Con respecto a los estudios de diversidad plasmídica realizados en *S. meliloti*, la metodología más conveniente para la identificación de plásmidos crípticos consiste en el método de *lisis in situ* en gel de agarosa [242], mediante el cual los plásmidos pueden observarse porque mantienen una mayor integridad en comparación con otros métodos. Entre los trabajos de diversidad de plásmidos crípticos en *S. meliloti*, una colección de aislamientos llevada a cabo en nuestro laboratorio fue la base de un estudio en donde la diversidad de perfiles plasmídicos fue representativa de los plásmidos presentes en dicha población bacteriana. La misma fue recuperada de 25 suelos diferentes de Argentina y en este trabajo se indagó acerca de la correlación entre la constitución genómica de los rizobios y el contenido de los plásmidos crípticos presentes en la misma [1]. El análisis por geles de Eckhardt sobre 64 aislamientos mostró que los mismos presentaban 22 perfiles plasmídicos distintos con al menos 38 bandas discretas en los geles. Para evaluar si la diversidad de plásmidos presentes era representativa de la población muestreada, se utilizó un análisis acumulativo por índice de Shannon-Weaver [243] y el mismo mostró que se alcanzó un *plateau* en los resultados observados con un número de aislamientos ($n > 49$). Además el mismo trabajo mostró que los perfiles plasmídicos individuales no se correlacionaban con los marcadores cromosomales de sus huéspedes, sugiriendo una gran actividad de transferencia de plásmidos entre cepas distintas. Dicha colección de aislamientos de *S. meliloti* constituye el punto de partida de este trabajo de tesis doctoral.

Replicación, partición y mantenimiento en plásmidos no simbióticos de rizobios

Aunque no se conoce la contribución ni el rol adaptativo de este tipo de plásmidos en los rizobios y la mayoría de sus proteínas son de función desconocida, se sabe que se replican y mantienen de manera muy estable a las sucesivas generaciones. Los genes que dirigen la replicación y partición en los plásmidos (simbióticos y no simbióticos) en los rizobios pertenecen a la familia RepABC [244, 245]. La misma familia de proteínas, organizadas mayoritariamente en operones, es la responsable de la replicación y partición de casi todos los plásmidos y segundos cromosomas de las bacterias pertenecientes al orden Rhizobiales, y de la mayoría de este tipo de replicones en alfaproteobacterias [245].

Conjugación y evidencias de la transferencia de genes en los plásmidos no simbióticos

Tanto los mecanismos de transferencia de genes (HGT) como la recombinación y la conversión génica parecen ser los mecanismos que moldean la estructura genómica en los rizobios. Dentro de la gran cantidad de plásmidos no simbióticos presentes en los rizobios muchos de ellos han mostrado ser autotransmisibles o movilizables lo que sugiere su rol como vehículos de adaptación, evolución y diversificación en estas especies. En el trabajo realizado previamente en nuestro laboratorio, se estudió cuán activos en movilidad eran los plásmidos crípticos presentes en nuestra colección, representando los 22 perfiles distintos (22 OTUs), y se observó que el 14% de los mismos eran movilizables, indicando autotransmisibilidad o movilización por parte de otro plásmido "*helper*". Asimismo, el 29% de los aislamientos que representaban a un OTU diferente pudieron retransmitir al plásmido modelo pSmeLPU88b a una tercer cepa receptora, indicando cierta ubicuidad en las funciones "*helper*" necesarias en los aislamientos de *S. meliloti* [1]. La evaluación de los mecanismos moleculares involucrados en la movilización de estos plásmidos presenta gran relevancia en la comprensión de los

mecanismos mediante los cuales los plásmidos pueden intercambiar funciones adaptativas en una misma especie y su rol en la configuración genómica de los replicones.

Objetivos

OBJETIVO GENERAL

Los replicones plasmídicos crípticos forman parte del compartimiento génico móvil mediante el cual la comunidad bacteriana potencialmente intercambia (comparte) funciones adaptativas para el ambiente en que se desarrolla. Teniendo en cuenta nuestro interés en la biología y genómica de rizobios como bacterias modelos de suelo simbiotes de plantas, en el presente trabajo de tesis desarrollaremos una aproximación metagenómica para la exploración de la información génica presente en los plásmidos crípticos de diferentes aislamientos de *S. meliloti* y de otras bacterias simpátricas asociadas a plantas de alfalfa. Empleando como modelo de análisis una bacteria de suelo de importancia agrícola, el presente trabajo de tesis se propone avanzar en la caracterización estructural y funcional de los movilomas plasmídicos.

OBJETIVOS ESPECÍFICOS

- Obtención y purificación de los plásmidos crípticos de una colección de aislamientos de *S. meliloti* previamente caracterizada en nuestro laboratorio en términos de los plásmidos presentes, y las características de transferencia conjugativa de los mismos.
- Construcción de una colección de aislamientos de bacterias de la microbiota asociada a semillas y plantas de alfalfa. Análisis de su diversidad. Caracterización del contenido plasmídico presente en los aislamientos. Preparación y purificación de sus plásmidos crípticos.
- Secuenciamiento a saturación (a escala de megabases) del DNA de los replicones crípticos presentes en *S. meliloti* y en la biota simpátrica asociada a alfalfa. Anotación y análisis bioinformático como herramienta para la predicción de funciones (*data mining*).
- Evaluación de relaciones de ortología y búsqueda de vínculos evolutivos entre el moviloma plasmídico de *S. meliloti* y los replicones crípticos presentes en el rizobio (cromosoma, pSymB, pSymA).
- Comparación de los movilomas plasmídicos secuenciados de *S. meliloti* y de sus bacterias simpátricas. Análisis de relaciones de ortología hacia la evaluación de eventos de transferencia lateral entre ambos movilomas.
- Consolidación de una imagen (modelo) que provea evidencia estructural, funcional, y de los vínculos del moviloma críptico de *S. meliloti* como vector de la evolución genómica del rizobio y como mediador de la interacción con otros genomas bacterianos que comparten el mismo hábitat.

CAPITULO II

Materiales y métodos

II.1. Medios de cultivo utilizados

Composición de los medios de cultivo y condiciones de crecimiento bacteriano

Para el crecimiento de cepas de los aislamientos provenientes de semillas y plántulas de alfalfa se utilizó el medio de cultivo LB-Glucosa [246], cuya composición es la siguiente: Glucosa 1g/l, NaCl 5g/l, Triptona 10g/l, Extracto de levadura 5g/l.

Para el caso de *Sinorhizobium meliloti* se utilizó el medio de cultivo TY [247], cuya composición es: CaCl₂·2H₂O 0,9 g/l, Triptona 5 g/l, Extracto de levadura 3 g/l.

Los cultivos se incubaron por el tiempo requerido a 28 °C con agitación (180 rpm) o a 28 °C en incubadora (en el caso de las placas con medio sólido).

Para la preparación de medios sólidos se utilizó agar en una concentración de 15 g de agar por cada litro de medio de cultivo.

La estimación del número de bacterias se realizó por medición fotométrica de la densidad óptica a 600 nm (una DO₆₀₀ = 0.1 corresponde a aproximadamente a 2×10⁷ UFC/ml).

II.2 Medio mineral para irrigación de plantas

Composición del medio mineral Fåhraeus para irrigación de plantas

Las plántulas crecidas a partir de las semillas esterilizadas superficialmente fueron irrigadas con Medio Mineral de Fåhraeus cuya composición es (g/l): CaCl₂·2H₂O 0.114, MgSO₄·7H₂O 0.120, Citrato Férrico 0.005, KH₂PO₄ 0.1, Na₂HPO₄ 0.15. Dicha solución puede prepararse 1000 veces más concentrada (1000X) y es esterilizada en autoclave.

Una vez diluída, a la solución final se le adiciona una solución de micronutrientes, la cual es asimismo preparada 1000X y filtrada independientemente en filtros de 0,22 micrones.

Solución Micronutrientes (1000X) (g/l): KCl 3.73, H₃BO₃ 1.55, SO₄Mn·H₂O 0.85, CuSO₄·5H₂O 0.13, ZnSO₄·7H₂O 0.58, (NH₄)₆Mo₇O₂₄·4H₂O 0.01.

II.3. Aislamiento de bacterias asociadas a semillas de alfalfa

Seis diferentes variedades de semillas de *Medicago sativa* (alfalfa) que presentan distintas características agronómicas y antigüedad fueron provistas por el Instituto Nacional de Tecnología Agropecuaria (INTA-Castelar). Para el aislamiento de bacterias endófitas, réplicas de entre 1 y 2 gramos de semillas fueron esterilizadas superficialmente con H₂SO₄ concentrado (98% P/P) durante 10 min, lavadas con abundante agua destilada estéril en baño de hielo para evitar el calentamiento por neutralización y neutralizadas con solución *buffer* fosfato estéril (PBS: 123 mM NaCl, 9.8 mM Na₂HPO₄, 2.4 mM KH₂PO₄, pH 7.4). Alternativamente, las semillas fueron esterilizadas con solución de hipoclorito de sodio (11 gCl/L) durante 10 minutos y lavadas con abundante agua destilada estéril. En ambos casos, 500 µL del agua sobrante luego del lavado final fue plaqueado en 5 placas de medio LB sólido como control de esterilidad externo de las semillas. Posteriormente, éstas fueron incubadas en 25 mL de PBS por 4 horas para permitir la entrada de agua y la liberación gradual de los exudados. Luego de transcurrido ese tiempo, 100 µL del sobrenadante fueron plaqueados en medio LB e incubados a 28 °C durante 4 días hasta que todas las colonias visualmente distintas fueron aisladas en medio de cultivo LB. Dichos aislamientos fueron reaislados sucesivamente en LB hasta la obtención de

un cultivo homogéneo y luego fueron criopreservadas a -20 y -80 °C en medio LB suplementado con 50% y 25% de glicerol, respectivamente.

Las variedades de semillas utilizadas en los ensayos fueron gentilmente provistas por el Ing. Alejandro Peticari. Las mismas son variedades experimentales estudiadas en el INTA-Castelar y presentan distintas características agronómicas, detalladas en la Tabla A8 del Anexo.

II.4. Obtención de bacterias en plántulas de alfalfa provenientes de semillas

Las plántulas de alfalfa fueron obtenidas a partir de semillas esterilizadas en superficie según se detalla en la sección II.3, luego germinadas en medio agar agua y crecidas agrupadas de a 4 plantas en tubos de ensayo cónicos tipo Falcon de 50mL estériles, conteniendo vermiculita esterilizada como soporte y tapones de algodón que permiten el intercambio gaseoso y a la vez evitan la contaminación externa a lo largo del experimento. Las plantas fueron incubadas en cámara de plantas (Binder™, Modelo KBW 400) con un ciclo programado de 16 horas luz (22 °C) y 8 horas de oscuridad (16 °C) y fueron regularmente irrigadas con solución Fåhraeus estéril en cámara de flujo laminar. Luego de 30 días de incubación las plántulas fueron cosechadas. Cada plántula fue machacada en mortero estéril con 1 mL de solución PBS y se plaquearon 100 µL del sobrenadante (de la muestra directa o de diluciones seriadas 1:10) en medio LB sólido. Una vez crecidas las colonias a 28 °C durante 4 días, las mismas fueron reaisladas sucesivas veces para obtener un cultivo puro, y criopreservadas.

Preparación de la vermiculita en nuestros ensayos

La vermiculita se lavó con abundante agua destilada, fue secada en estufa a 60 °C y fraccionada en tubos cónicos tipo Falcon de 50 mL cerrados con tapones de algodón-gasa para evitar contaminaciones exteriores durante los ensayos. Estos tubos ya preparados fueron esterilizados 4 veces en autoclave. En cada ensayo dos tubos fueron abiertos en flujo laminar y su vermiculita fue embebida en 10 mL de PBS estéril y luego plaqueada en medio LB sólido e incubada toda la noche a 28 °C, de manera de comprobar que no existe microbiota remanente en la vermiculita con capacidad de crecimiento en el medio de cultivo que utilizamos en nuestros ensayos. El mismo procedimiento se realizó sobre dos tubos con vermiculita, pero sin contener semillas, al finalizar los ensayos luego de los 30 días de incubación en la cámara de plantas, de manera de corroborar que la contaminación externa está ausente bajo este sistema de crecimiento aislado de la microbiota del aire en el ambiente.

II.5. Ensayos in vitro de capacidades indicadoras de promoción del crecimiento vegetal (PCV)

Con el objetivo de explorar si los aislamientos presentaban alguna actividad de promoción del crecimiento vegetal (PCV) *in vitro*, diferentes pruebas bioquímicas fueron evaluadas, tales como solubilización de fosfatos [248], producción de quitinasas [249], sideróforos [250], celulasas [251], ácido indol acético [252], proteasas [253]. Adicionalmente, se realizaron ensayos de biocontrol contra algunos hongos fitopatógenos de géneros tales como *Rhizopus*, *Alternaria* y *Fusarium* [254], y ensayos de producción de bacteriocinas según Parret y colaboradores [255].

II.6. Aislamiento de ADN total

Para la preparación de ADN total de los aislamientos bacterianos, se tomaron bacterias crecidas en medio sólido (aproximadamente una punta de tip) o líquido (sedimento de 500 μL de un cultivo de $\text{DO}_{600}=0,8$, centrifugado a 10000 g durante 5 minutos), y se resuspendieron en 100 μL de N-lauril sarcosina al 0,1% en TE (50 mM Tris, 20 mM EDTA, pH=7,0). Luego se centrifugaron (10.000xg 2 minutos) y se lavaron con solución 1M de NaCl. A continuación, se centrifugó a (10.000xg 2 minutos) y el sedimento se resuspendió en 250 μL de TE (50 mM Tris, 20 mM EDTA, pH=7,0) suplementado con sacarosa al 20%. Se agregaron 250 μL de lisozima 5 mg/ml en TE (50 mM Tris, 20 mM EDTA, pH=7,0). Esta mezcla se incubó 1 hora a 37 °C. Posteriormente se agregaron 50 μL de una solución de Pronasa 10 mg/ml y 5% de N-lauril sarcosina en TE (50 mM Tris, 20 mM EDTA, pH=7,0), se mezcló por inversión y se incubó a 37 °C durante 1-2 horas. Transcurrido este tiempo, se agregaron 70 μL de una solución de acetato de sodio 3M, 200 μL de fenol equilibrado con Tris a pH=8 y 100 μL de cloroformo. Se agitó suavemente y se centrifugó durante 15 minutos a 10.000xg. La fase acuosa se pasó a un nuevo tubo y se precipitó el ADN por agregado de 1 volumen de isopropanol o 3 volúmenes de etanol. Se centrifugó 15 minutos a 10.000xg. El sedimento se lavó con etanol 70%, se dejó secar al aire. Finalmente, se resuspendió en el volumen deseado de agua (en general 25-50 μL) y se añadió 1 μL de una solución 1mg/ml de RNasa.

II.7. Ensayos de PCR

Los ensayos de PCR fueron realizados en ciclador Bioer Thermal Cycler. Las mezclas de reacción utilizadas se prepararon según protocolo propuesto por el proveedor de la enzima. La concentración final de los componentes fue: oligonucleótidos (dNTPs) 0,2 μM ; Mg^{+2} 1,5 – 3,0 mM, cebadores (“*primers*”) 0,2 – 1 μM , Taq ADN polimerasa 1 unidad enzimática. El tampón de reacción (“*buffer*”) es de composición desconocida y se utiliza según las indicaciones del fabricante. Los moldes de ADN para las amplificaciones fueron obtenidos utilizando alguna de las siguientes metodologías: procedimiento de “*hervidita*” o bien preparación de ADN total.

La amplificación de los productos fue visualizada en geles de agarosa (0,8 a 1%) teñidos con bromuro de etidio a una concentración final de 1,0 $\mu\text{g}/\text{ml}$. Las reacciones se realizaron en 10 μL finales utilizando 2,5 μL de muestra.

II.8. Amplificación del gen 16s rRNA

Como material de partida para las PCR se utilizaron lisados celulares obtenidos tomando material bacteriano crecido en medio sólido (punta de tip) y resuspendiendo en 50 μL de agua destilada apta para PCR y luego sometiénolo a un baño de agua 15 minutos a 100 °C para provocar la lisis celular. Posteriormente se centrifuga el tubo a 10000 rpm durante 10 minutos y se toma 1 μL del sobrenadante, se realiza una dilución 1:10 en agua apta para PCR del mismo y se utiliza 1 μL de la dilución como material molde de la PCR. Cuando las PCR no dieron una banda nítida como se esperaba, alternativamente se utilizó como molde de reacción una extracción de ADN total preparada como se detalla en la sección II.6.

Se amplificaron fragmentos del gen 16s rDNA mediante el uso de los *primers* fD1/rD1 [256] generando amplicones de 1.48 kpb de tamaño, y los *primers* Y1/Y2 [257] que generan

amplicones de entre 550 y 600 pb. Dichos productos fueron secuenciados mediante el uso de los mismos *primers* fD1 o Y1, dependiendo el caso, con la colaboración de la plataforma de secuenciamiento del INTA-Castelar. El análisis de las lecturas obtenidas por BLASTn permitió estimar la taxonomía de los aislamientos obtenidos a nivel de género.

Tabla II.1. Cebadores ("*primers*") utilizados para la amplificación de las regiones parciales del gen 16S rRNA.

Denominación	Secuencia	Referencia
Y1	5'-TGGCTCAGAACGAACGCTGGCGGC-3'	[257]
Y2	5'-CCCACTGCTGCCTCCCGTAGGAGT-3'	[257]
fD1	5'-AGAGTTTGATCCTGGCTCAG-3'	[256]
rD1	5'-AAGGAGGTGATCCAGCC-3'	[256]

II.9. Identificación de aislamientos por espectrometría de masas MALDI-TOF

La identificación bacteriana por MS MALDI-TOF fue realizada mediante el empleo del *software* MALDI *Biotyper* 3.1 (de la empresa Bruker Daltonics)[258]. La preparación de las muestras para su posterior identificación fue llevada a cabo de acuerdo al método de extracción de proteínas por lisis celular descrito por Bruker Daltonics. Inicialmente, colonias bacterianas de 12-24 hs de crecimiento en medio LB fueron resuspendidas en etanol 80%, centrifugadas a 10.000 rpm por 2 min y finalmente resuspendidas en 20 μ l de ácido fórmico 70%:acetonitrilo 100% (1:1). Para el análisis de las muestras 1 μ l de suspensión celular fue adicionado sobre cada uno de los *spot* de una placa MALDI (384 *spots*) que, inmediatamente después de secarse, fue recubierta con 1 μ l de solución saturada de matriz (10 mg/mL) consistente de ácido α -ciano-4 hidroxí cinámico (HCCA) resuspendido en solución orgánica (50% acetonitrilo + 2,5% tri-fluoro acético). Los espectros de masas fueron obtenidos manualmente en un rango de 2 a 20 KDa en modo lineal positivo utilizando el *software* FlexControl (3.0) acoplado a un espectrómetro de masas modelo Ultraflex (Bruker Daltonics). Se utilizó el calibrante BTS (*bacterial test standard*) (Bruker Daltonics) que consiste en proteínas ribosomales de la cepa *E. coli* DH5 α junto a las proteínas RNasa A y mioglobina, presentando masas moleculares correspondientes a 13.683 y 16.954 Da, respectivamente. La clasificación a nivel de género o especie de los espectros obtenidos fue realizada de acuerdo a los valores de puntuación (*scores*) del *software* MALDI *Biotyper*: un valor mayor a 2 indica una identificación a nivel de especie, un valor entre 1,7 y 1,9 indica una identificación a nivel de género y un valor menor a 1,7, una identificación no confiable. La base de datos del *software* comercial fue nutrida con la adición de las nuevas cepas identificadas en este trabajo, las cuales pudieron ser identificadas por secuenciación del gen 16S rRNA y cuyos espectros de masas promedios fueron obtenidos sumando 20 espectros de masas independientes que presentaron intensidades variables de entre 6×10^3 a 10^4 unidades arbitrarias [259].

II.10. Identificación de plásmidos de alto peso molecular

Los perfiles de plásmidos fueron obtenidos por el método de lisis *in situ* de Eckhardt [242] con las modificaciones realizadas por Wheatcroft [260]. Las bacterias fueron cultivadas en medio TY líquido. Alícuotas de 1 ml de cultivos crecidos hasta una $DO_{600} = 0,4$ fueron centrifugadas y las células fueron lavadas con 1 ml de solución 0,3% (P/V) de sarcosilato de sodio. Las células

de cada muestra fueron centrifugadas nuevamente, resuspendidas en 15 μ l de solución de ficoll 20% (P/V), y mantenidas en agua hielo hasta el momento de sembrar el gel de lisis/separación (0,75% de agarosa con buffer TBE (Tris 0,09 M; ácido bórico 0,09 M; EDTA 0,002 M)). Previo a la siembra de las muestras en cada fosa del gel se introdujeron 25 μ l de una solución de SDS 10% (P/V) con xylene cyanol, y la electroforesis se corrió a polaridad invertida hasta que el frente retrocedió un centímetro. Se detuvo la corrida y se procedió a la siembra. Para ello se agregaron 15 μ l de solución de lisis (sacarosa 8% ficoll 2% Tris 50 mM EDTA 25 mM, RNasa 0,4 mg/ml lisozima 1mg/ml) a cada una de las muestras, se mezclaron en forma suave e inmediatamente se sembraron 15 μ l de la mezcla por calle. Se corrió con la polaridad normal a 30 V 30 min y luego a 110 V 2 h, luego se lavó con agua 30 min y se tiñó con bromuro de etidio (0,5 mg/ml).

II.11. Purificación y secuenciamiento de los plásmidos crípticos

Colección de aislamientos de *Sinorhizobium meliloti* utilizados

Dieciocho aislamientos de nuestra colección de *Sinorhizobium meliloti* fueron seleccionados para la purificación de ADN plasmídico. Cada uno de estos aislamientos presenta un perfil plasmídico diferente [1] y se estima que un total de 38 bandas correspondientes a diferentes plásmidos de distintos tamaños (entre 20 y 200 kpb) están presentes en esta colección, tal como se visualiza por geles de lisis *in situ* [242]. A su vez, esta colección representa a la diversidad de plásmidos colectados en los sitios de muestreo, tal como se observa por el análisis del índice de Shannon [1]. Los aislamientos que pudieron ser exitosamente purificados y finalmente secuenciados fueron los siguientes: LPU 9, LPU 198, LPU 5, LPU 87, LPU 86, LPU 196, LPU 258, LPU 26, LPU 88, LPU 57, LPU 178, LPU 118, LPU 116, LPU 243, LPU 15, LPU 110, LPU 121, LPU 122 y LPU 244.

Enriquecimiento de los plásmidos crípticos del MPSme mediante gradientes de cloruro de cesio y secuenciamiento

Los aislamientos de *S. meliloti* se crecieron individualmente a 28 °C en medio TY y sus plásmidos fueron purificados por medio de la ultra centrifugación en gradientes de CsCl, según el método descrito por Jouanin y colaboradores [261], de cara al enriquecimiento de los plásmidos de alto peso molecular no simbióticos (crípticos). A partir de este procedimiento se logra un incremento notable de la representación de los plásmidos crípticos en la muestra, respecto al ADN cromosómico o al perteneciente a megaplásmidos.

Cada preparación de ADN fue digerida con la enzima de restricción *EcoRI* a 37 °C durante 4 horas, luego corrida en un gel de agarosa 0,8 % p/v, teñida con bromuro de etidio y revelada a luz UV (260 nm). La calidad del enriquecimiento plasmídico fue evaluada en base a la minimización del fondo cromosómico presente. Una muestra de cantidades equivalentes en masa de ADN de cada aislamiento fue preparada para su posterior secuenciamiento, partiendo de concentraciones medidas en el espectrofotómetro NanoDrop™. Dicha muestra fue secuenciada utilizando la tecnología de Ion Torrent™ [262]. Adicionalmente, y debido a la presencia de errores de secuenciación sobre secuencias homopoliméricas, la misma muestra se secuenció utilizando la plataforma Illumina™. Dicho secuenciamiento fue realizado en el

CeBiTec, Universidad de Bielefeld, Alemania. Las lecturas fueron ensambladas utilizando Newbler v.2.6, de Roche™.

Enriquecimiento de los plásmidos crípticos del MPAS mediante gradientes de cloruro de cesio y secuenciamiento

Los aislamientos provenientes de semillas y plántulas de alfalfa (que se describen en la sección V.2) fueron crecidos en LB y sus plásmidos fueron purificados por gradientes de CsCl [261]. Para su secuenciamiento se preparó una mezcla de iguales cantidades en masa, partiendo de concentraciones medidas en el espectrofotómetro NanoDrop™.

El secuenciamiento de esta muestra de plásmidos pertenecientes a aislamientos asociados a las semillas de alfalfa (MPAS) se realizó mediante la construcción de una biblioteca *mate-pair*. La misma consistió en la ruptura del ADN plasmídico en fragmentos de largo promedio de 4.8 kpb, su purificación y la posterior unión de adaptadores que permiten circularizar los fragmentos y sirven de secuencias blanco de los cebadores que se utilizan en el secuenciamiento. Una vez circularizado el ADN se fragmenta nuevamente y se procede al secuenciamiento *pair-end* de Illumina utilizando las secuencias blanco localizadas en los adaptadores mencionados. Las lecturas obtenidas mediante esta biblioteca consisten en fragmentos de ADN separados por una distancia promedio de 4.8 kpb que permite ensamblar cóntigos y unirlos evitando los problemas de ensamblado asociados a secuencias repetitivas largas, frecuentes en genomas bacterianos. La construcción de la biblioteca y el secuenciamiento fueron realizados en el CeBiTec, Universidad de Bielefeld, Alemania.

II.12. Identificación de plásmidos cerrados entre los ensamblados obtenidos del MPAS

La identificación de los plásmidos cerrados se realizó mediante dos *softwares* de edición y finalización de genomas, Consed y BACCardI [263, 264] en el CeBiTec, Universidad de Bielefeld.

II.13. Identificación de los cóntigos plasmídicos

Considerando que el enriquecimiento por gradientes de CsCl genera una muestra con contaminantes de ADN cromosómico y de megaplásmidos, se estableció un método para discriminar cuáles cóntigos fueron ensamblados por lecturas provenientes de los plásmidos crípticos enriquecidos.

i) Se calculó el número de lecturas por largo de contig (L/C) para aquellos cóntigos que presentaban una cobertura completa y un alto porcentaje de identidad (99-100%) al alinearlos contra cromosomas o plásmidos simbióticos de aislamientos previamente secuenciados de *S. meliloti*. Estos cóntigos contaminantes presentan una media de L/C y desviaciones estándar bajas (ver Tabla A1, Anexo). Contrariamente, aquellos cóntigos que presentan una alta cobertura L/C, mayor a 2 desviaciones estándar de la media de los contaminantes, fueron considerados provenientes de plásmidos crípticos.

ii) Se realizó un filtrado de los cóntigos basado en alineamientos BLASTn contra los cromosomas y megaplásmidos secuenciados de las cepas *Sinorhizobium meliloti* BL225C, AK83, 1021, SM11, Rm41 y GR4, tomando en cuenta la cobertura y el porcentaje de identidad del

alineamiento. Aquellos cóntigos con más de 95% de cobertura y más de 90% de identidad contra estas cepas, fueron removidos de la colección. En todos los casos se corroboró que estos cóntigos tuvieran asimismo un bajo valor de L/C y no fueran, en principio, posibles transferencias de ADN cromosomal hacia los plásmidos crípticos.

iii) Para finalizar, los cóntigos remanentes con bajos valores de L/C, y que escaparon al filtrado realizado en ii) debido a baja cobertura en el alineamiento o bajo % de identidad con los genomas secuenciados, fueron considerados posibles cóntigos pertenecientes a ADN cromosómico o megaplasmídico. Dichos cóntigos fueron excluidos de los análisis, pudiendo tratarse tanto de secuencias cromosomales o megaplasmídicas exclusivas de las nuevas cepas, o bien de ADN de plásmido críptico que haya resultado estar en baja representatividad en la muestra secuenciada.

Luego de los sucesivos pasos de filtrado, los cóntigos mayores a 1000 pb fueron importados al sistema de anotación GenDB [265], que permite la predicción de marcos de lectura abiertos y la anotación sobre los mismos. La predicción de ORFs sobre los cóntigos pequeños que obtuvimos en el ensamblado muchas veces generó ORFs más cortos debido a la interrupción del cóntigo.

II.14. Análisis de la composición funcional por COGs

El análisis de la variación en la composición funcional basada en grupos de genes ortólogos COGs (*“cluster of orthologous genes”*) de la sección III.3 fue llevada a cabo sobre los cromosomas y megaplásmidos (pSymA y pSymB) de diferentes cepas completamente secuenciadas de *S. meliloti*, y sobre diferentes plásmidos pertenecientes a la familia *Rhizobiaceae*, los cuales se encuentran listados en la Tabla A2 del Anexo. Los ORFs de dichos cromosomas, megaplásmidos y plásmidos completamente secuenciados fueron descargados de la base de datos GenBank, y mediante el uso del servidor WebMGA [266] se identificaron aquellos genes que pudieron ser asignados a alguna clase COG. A partir de dicho análisis pudimos obtener el número de ORFs que, en cada replicón, era asignado a cada clase de COGs (de la A a la Z). En base a estos valores y al número total de ORFs asignables a cualquier clase COG, se calcularon los porcentajes de cada clase COG representada en el total de ORFs de dicho replicón.

Las matrices construidas por las distintas composiciones de COGs calculadas en cada replicón fueron la información de partida con la que efectuamos los análisis de componentes principales (PCA), el cual fue realizado *online* mediante el *software* “MetaboAnalist”, representando los tres primeros componentes de variación [267].

El análisis jerárquico multivariado de clusters (o “mapa de calor”) se realizó mediante el mismo *software* partiendo de la misma matriz de datos de entrada.

II.15. Análisis de genes homólogos por el software EDGAR

El *software* EDGAR [268] fue diseñado para buscar homólogos entre diferentes genomas/replicones/metacóntigos (definiendo a estos últimos como la combinación de genes no redundantes de un grupo dado de replicones/genomas). Este *software* está ideado para

recuperar genes homólogos tomados en el sentido filogenético como aquellos que presentan un origen ancestral común. Los homólogos que divergen y devienen ortólogos debido a un evento de especiación [269] son aquellos que conservan su función en diferentes organismos [270]. Para hallar homólogos, el método de búsqueda por BBH (“*bi-directional best hits*”) que utiliza EDGAR permite búsquedas masivas de genes de manera más eficiente con respecto a otros métodos con la limitación de no poder tomar en consideración los parálogos (homólogos producto de una duplicación, frecuentemente presentes en un mismo genoma) [271]. El *software* emplea búsquedas por BLASTp utilizando una matriz de sustitución BLOSUM62, realizando BBH como se mencionó previamente, y tomando como valores de corte, para determinar si dos ORFs son homólogos, a aquellos determinados entre dos genomas por el análisis de SRV (*score ratio values*), empleado por Lerat y colaboradores [272], con modificaciones para adaptar el criterio de elección de valores de corte a géneros de variadas distancias filogenéticas [268].

Mediante el uso de EDGAR, en el capítulo III se buscaron genes “singletones”, es decir aquellos presentes en una cepa y ausentes en las restantes analizadas. Este análisis se realizó para cada cromosoma, pSymA y pSymB de las cepas completamente secuenciadas de *S. meliloti*, listadas en la Tabla A2 del Anexo. Asimismo se buscaron genes homólogos presentes en todas las cepas de *S. meliloti* para cada uno de los replicones antes mencionados (genes “core” en cromosoma, pSymA y pSymB). Por otro lado, el *software* se utilizó para obtener conjuntos de genes homólogos entre todos los marcos de lectura abiertos del MPS_{me} como conjunto de referencia y el conjunto de ORFs presentes en el pSymA, pSymB y cromosoma de *S. meliloti*. Como parte del análisis del conjunto de genes del “core” se calcularon las curvas de la función de decaimiento de genes (“core development plot”) para el pSymA y pSymB mediante el uso de EDGAR.

Por último, se utilizó el algoritmo de EDGAR para obtener y contar ortólogos entre el MPS_{me} y distintos plásmidos (agrupados como metacóntigos) y cromosomas de especies bacterianas relacionadas, por un lado (sección III.3), y entre los ensamblados obtenidos del moviloma plasmídico de bacterias asociadas a la semilla (MPAS) y el MPS_{me}, además de distintos conjuntos de plásmidos agrupados en metacóntigos (sección V.5).

Los resultados de estos últimos análisis mencionados fueron expresados como el Número de Ortólogos identificados en un replicón dado (o metacóntigo) por cada Cien Genes (NOCG)(densidad de ortólogos). Estas estimaciones fueron utilizadas para rastrear aquellos replicones que contribuyen más a la composición presente en el MPS_{me}, tanto de *S. meliloti*, como en otras especies. Los diferentes metacóntigos que utilizamos para buscar ortólogos contra el MPS_{me} fueron definidos como se adjuntan en la Tabla A2 del Anexo.

II.16. Análisis de la frecuencia de uso de codones

Para el cálculo de la frecuencia relativa del uso de codones (RSCU, *relative synonymous codon usage* [273]), se utilizaron cada uno de los conjuntos de genes homólogos obtenidos mediante EDGAR (ver sección II.15) y mediante el uso del *software* GCUA [274] se realizó el cálculo del RSCU acumulativo. Posteriormente se generó una matriz conteniendo 59 de los 64 codones posibles (se exceptúan los codones de terminación UGA, UAG, UAA, el codón codificante de la

Metionina, AUG, y el codón codificante del Triptófano, UGG, los cuales son invariables). A partir de estas frecuencias de codones se generaron matrices y se analizaron por análisis de componentes principales (PCA) mediante el uso del *software* XLSTAT, versión 2/07/2013 (Addinisoft™).

II.17. Búsquedas por similitud de secuencia

Las búsquedas de similitud nucleotídica y aminoacídica se realizaron utilizando las secuencias nucleotídicas completas de los ensamblados o contigs en el caso de BLASTn y las secuencias aminoacídicas de los ORFs en el caso de las búsquedas por BLASTp y tBLASTn. Se utilizó como valor de "expect" 10^{-5} .

II.18. Anotación manual y asignación de grupos funcionales

Los ORFs que fueron predichos y anotados automáticamente mediante la plataforma GenDB [265] en los proyectos de secuenciación de los movilomas plasmídicos MPS_{me} y MPAS, y los presentes en los diferentes conjuntos de genes homólogos obtenidos por EDGAR [268], fueron revisados manualmente y a cada ORF se le asignó una letra indicando la pertenencia a un determinado grupo funcional, basándonos en las predicciones funcionales, respetando el siguiente criterio:

H: proteínas hipotéticas conservadas, dicha categoría incluye a todo ORF que presenta al menos algún homólogo en los genomas secuenciados y depositados en la base de datos del GenBank, detectado a través de las búsquedas secuenciales y parámetros establecidos en la plataforma GenDB [265], pero para el cual aún se desconoce la función bioquímica asociada a alguno de sus homólogos o basada en otro tipo de predicciones bioinformáticas (HMM-HMM, etc). Las anotaciones automáticas asociadas a estos ORFs presentan las descripciones "hypothetical protein", "conserved protein", "putative membrane protein" o "putative secreted protein".

U: predicciones de ORFs que no presentan homólogos en los genomas secuenciados y depositados en la base de datos del GenBank, detectado a través de las búsquedas secuenciales y parámetros establecidos en la plataforma GenDB [265]. Para estos genes aún se desconoce la función bioquímica asociada asignable por alguna predicción bioinformática.

T: ORFs con predicciones funcionales asociadas al transporte de moléculas.

S: se incluyen ORFs con predicciones funcionales relacionadas a la regulación transcripcional, traduccional, traducción de señales, sistemas de dos componentes, etc.

P: incluye todos los ORFs con predicciones relacionadas a la replicación, partición y segregación del ADN.

A: ORFs con funciones predichas relacionadas al metabolismo del ADN tales como: metilación y restricción, helicasas, ADN o ARN polimerasas, ligasas, nucleasas.

MGE: ORFs con funciones predichas relacionadas a la conjugación, transposición, secuencias de inserción y genes relacionados a fagos.

M: ORFs con predicciones funcionales relacionadas a diversas funciones metabólicas que incluyen las rutas biosintéticas o catalíticas de las bases nitrogenadas, aminoácidos, ácidos grasos, rutas centrales como glicólisis, ciclo de Krebs y pentosas fosfato, resistencia a antibióticos y a compuestos xenobióticos, respuestas oxidativas o de estrés, peptidoglicanos, lipopolisacáridos, metabolismo del nitrógeno y relacionado al proceso simbiótico, degradación de macromoléculas, movilidad bacteriana, resistencia a metales, chaperonas y predicciones metabólicas generales no asignables a una ruta en particular.

X: En caso del análisis de los conjuntos de genes homólogos al *MPS_{me}* y de fracciones de genes singletons (ver capítulo III) se utiliza alternativamente esta letra para indicar ORFs con anotaciones funcionales asociadas a funciones de fagos, transposasas o secuencias de inserción.

CAPÍTULO III

**Moviloma plasmídico de *Sinorhizobium meliloti*:
secuenciamiento, caracterización funcional y vínculos
evolutivos con el resto del genoma del rizobio**

III.1. Introducción

Teniendo en cuenta que los plásmidos crípticos contienen una considerable cantidad y diversidad de información génica en los rizobios, los cuales muestran un genoma especialmente particionado (cromosoma, megaplásmidos, plásmidos crípticos), en este primer capítulo nos propusimos comenzar a investigar cuál podía ser el rol tanto adaptativo (qué funciones están codificadas y qué ventajas ofrecen) como evolutivo (qué rol juegan en la estructuración de la información génica en estas bacterias) que juegan estos replicones en la biología de *S. meliloti*. Para ello, y haciendo uso de una colección previamente estudiada en nuestro laboratorio [1], abordamos su estudio con un enfoque genómico, el cual nos permitió acceder en gran escala al repertorio de plásmidos crípticos de *S. meliloti* como un modelo para estudiar aspectos de ancestralidad de sus genes y evaluar su perfil de funciones génicas, de cara a evaluar hipótesis sobre los roles biológicos que presentan los movilomas plasmídicos en las bacterias.

III.2. Preparación y secuenciamiento masivo de plasmidoma críptico de una colección previamente caracterizada de aislamientos de *S. meliloti*

Nuestro laboratorio ha caracterizado previamente una colección de aislamientos de *S. meliloti* que se conformó en base a su diversidad de perfiles plasmídicos [1]. Con el objetivo de secuenciar el conjunto de plásmidos crípticos presentes en esta colección de 18 aislamientos de *S. meliloti*, se realizaron purificaciones de ADN seguidas de gradientes de CsCl a fin de obtener ADN enriquecido en los plásmidos crípticos de alto peso molecular no simbióticos (ver Materiales y métodos, sección II.11). Una muestra enriquecida en ADN plasmídico de cada uno de los aislamientos fue secuenciada utilizando las plataformas Ion Torrent (PGM) e Illumina (MiSeq), resultando en un total de más de 4 millones de lecturas con ambas plataformas, y un total de cerca de 1 Gpb de secuencia obtenida. El ensamblado de las lecturas resultó en un total de 5.763 cóntigos (mayores a 500 pb) y 9,66 Mpb de secuencia no redundante con una cobertura promedio de 103 lectura/kilobase secuenciada. Para el curado de las secuencias que nos permite sustraer la información cromosomal contaminante y conservar la proveniente de los plásmidos crípticos, tomamos al azar 30 cóntigos diferentes que contenían genes cromosomales conocidos de *S. meliloti*, calculamos el número promedio de lecturas/kilobase sumado a dos desviaciones estándar como criterio de corte, y excluimos todos aquellos cóntigos que mostraron una menor cobertura que dicho valor. En promedio, los cóntigos cromosomales que descartamos mostraron una cobertura de 49 ± 13 lecturas/kilobase, lo cual contrastó con la mayor cobertura observada en los cóntigos de origen plasmídico (usualmente mayor que 400 lecturas/kilobase)(ver Anexo, tabla A1). Un análisis similar fue realizado para los genes de los pSym con resultados de cobertura similares a los obtenidos para el cromosoma, posiblemente debido a su gran tamaño, muy bajos números de copias, y bajo nivel de contaminación en nuestra preparación de plásmidos. La información filtrada nos permitió identificar 337 cóntigos mayores de 1 kpb y que en conjunto totalizaron 1,46 Mpb de secuencia no redundante de ADN plasmídico. El N50 de esta colección de cóntigos (calculado como el largo del cóntigo tal que la suma del largo del conjunto de cóntigos menores a él iguale o supere a mitad de la secuencia no redundante) fue de 8,1 kpb, siendo el cóntigo más largo de 44 kpb. La búsqueda en la secuencia curada de marcos de lectura abiertos (ORFs por su sigla en inglés, *Open Reading Frames*) utilizando la plataforma GenDB y luego un curado

manual, predijo la presencia de unas 2.174 secuencias codificantes, incluyendo ORFs truncados al final de los cóntigos (ver Materiales y métodos, sección II.13). Considerando el total de la secuencia curada y el número de cepas a partir de las que realizamos las preparaciones, resultó una cantidad (promedio) de al menos 80 kpb de ADN en plásmidos crípticos/cepa, con un aporte en promedio de 120 genes/cepa.

El análisis de la secuencia del moviloma plasmídico críptico mostró que el mismo tiene un contenido porcentual GC (GC%) de 58,26%, valor que es consistente con los observados en plásmidos crípticos individuales previamente secuenciados en *S. meliloti* (ver por ejemplo los replicones con códigos de acceso HE995406.1, AF327371.1, DQ145546.1, EF066650.1 y JQ665880.1). En contraste con esta observación, y según ha sido reportado previamente, los cromosomas y megaplásmidos pSymB (“crómidos”, según Harrison y colaboradores [275]) presentan un contenido GC% de 62,7% y de 62,4%, respectivamente [228, 276]; y los megaplásmidos pSymA de 60,4 % [227] algo más cercano al observado en el moviloma plasmídico aquí presentado.

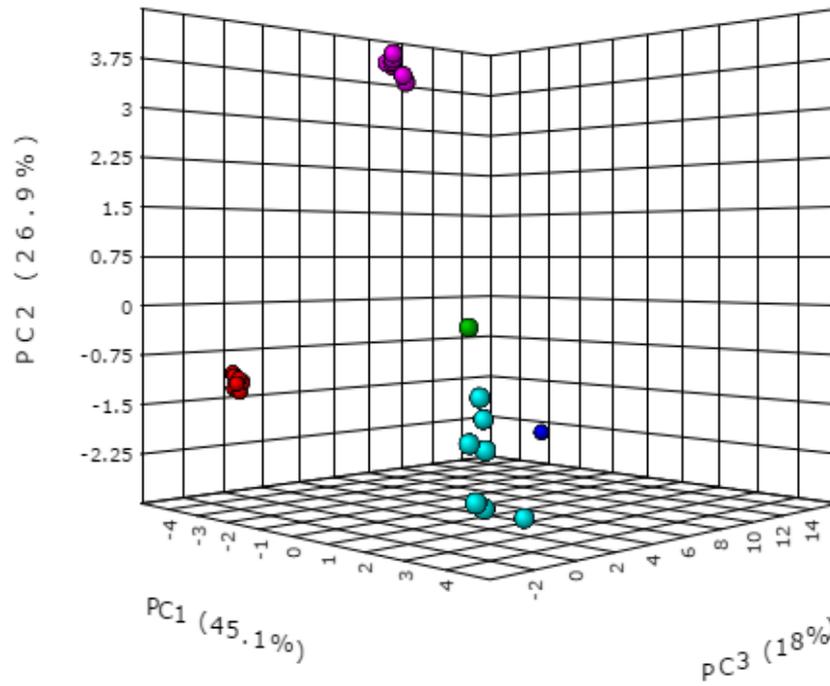
III.3. Comparación de las abundancias de COGs en el moviloma plasmídico críptico respecto de las presentes en el cromosoma y en los plásmidos simbióticos (pSyms).

Variaciones composicionales, relaciones evolutivas

Una de las formas más sencillas de evaluar el tipo de funciones codificadas en un dado replicón es analizar el tipo de proporción de COGs [277], que corresponden al conjunto de genes presentes. Así, con el objeto de comparar las abundancias relativas de genes que codifican funciones celulares diferentes en los distintos replicones de *S. meliloti*, se calcularon las proporciones de clases COGs (A a Z) presentes en los cromosomas, pSymAs y pSymBs de todas las cepas secuenciadas a la fecha, así como para el moviloma plasmídico críptico que aquí hemos secuenciado. El cálculo de las proporciones de los distintos COGs (ver Materiales y métodos, sección II.14) en cada replicón (compartimiento genómico) fue utilizado como variable de entrada para realizar el análisis de componentes principales (PCA por su sigla en inglés, *Principal Component Analysis*) que se muestra en la Figura III.1.A. Los resultados indican que los diferentes replicones (cromosomas, pSymA y pSymB) y el moviloma plasmídico críptico mapean en diferentes regiones en el espacio de tres dimensiones definido por los componentes principales (PC) 1, 2 y 3; que en dicho análisis representaron 90% de la variación total acumulada según se aprecia en la Figura III.1.B (curva verde).

La mayor variación composicional entre los pSymAs comparada con la de los pSymBs y la de los cromosomas, se refleja en una dispersión más amplia de las localizaciones de los pSymA en el espacio PC1-PC2-PC3 (Figura III.1, esferas celestes). Los megaplásmidos pSymBs, a los que se refiere frecuentemente como “crómidos” [275], mostraron una dispersión composicional más acotada (esferas violetas) y comparable a la dispersión observada para los cromosomas de *S. meliloti* (esferas rojas). En concordancia con esto, la distancia media calculada para cada tipo de replicón a su posición promedio en el espacio PCA fue de 0,05; 0,12 y 0,95 para los cromosomas, pSymBs y pSymAs respectivamente.

A



B

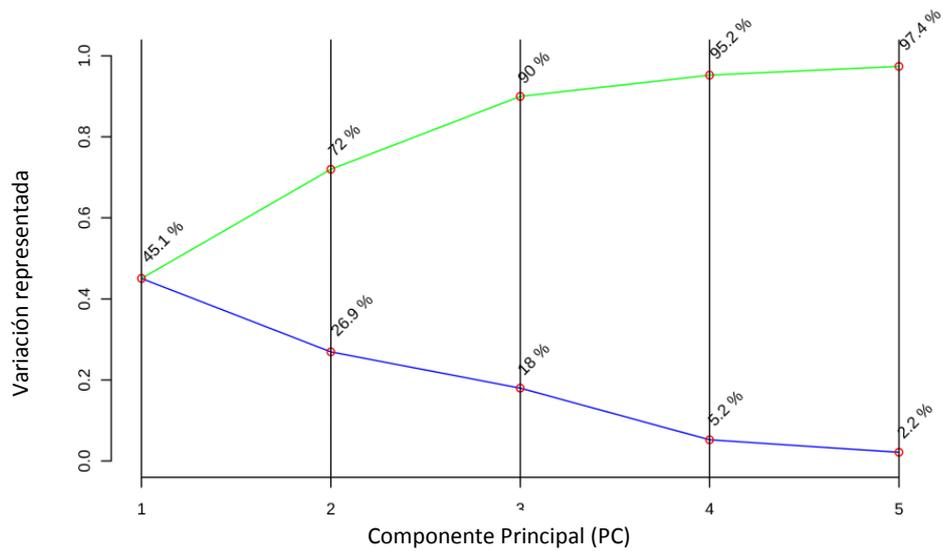


Figura III.1. Análisis de componentes principales utilizando como variables de entrada la proporción de COGs existente en el conjunto de los genes del cromosoma, de los pSyms y del moviloma plasmídico críptico de *S. meliloti*. Parte A. Localización de cada uno de los replicones en el espacio de los componentes principales 1, 2 y 3. El color de los círculos indica los siguientes replicones: rojo, cromosoma; verde: moviloma plasmídico de la familia Rhizobiaceae (secuencias de GenBank); azul, moviloma plasmídico críptico de *S. meliloti* (esta tesis); celeste, pSymA; violota, pSymB. Parte B. Detalle de la proporción de la variación representada por los primeros 5 componentes principales (curva azul), y de sus valores acumulados (curva verde).

El análisis de las abundancias de COGs sirvió para investigar las características funcionales generales de los replicones y para explorar las relaciones entre ellos, independientemente de la presencia o ausencia de marcadores individuales específicos. La Figura III.2 presenta un análisis jerárquico multivariado de clusters para explorar la ocurrencia de patrones específicos de clases COG (de la A a la Z como las *variables del análisis*, que se listan a la derecha del gráfico) asociadas a cada tipo de compartimiento genómico (las *muestras*, que corresponden a las columnas listadas en la parte superior). Los resultados muestran la predominancia de clases específicas de COGs asociadas a cada tipo de replicón, con COGs del *cluster* B₃ (dominada, como se espera, por diversas funciones “housekeeping” junto con funciones de movilidad), *cluster* A (transducción de señales, transporte y metabolismo de iones inorgánicos y producción y conversión de energía) y *cluster* B₁ (mecanismos de defensa, transporte y metabolismo de carbohidratos y biogénesis de pared celular, membrana y envoltura) como las predominantes en los cromosomas, pSymAs y pSymBs, respectivamente. El moviloma plasmídico de Rhizobiaceae se agrupó con el pSymA (composiciones similares del *cluster* A) y se separó del moviloma plasmídico de *S. meliloti*, con este último emergiendo como un compartimiento genómico funcionalmente distinto con predominio de funciones asociadas al tráfico y secreción (clase COG “U”), y funciones -como se espera- relacionadas al mantenimiento y estabilidad de plásmidos (clases COG “L” y “D”)(todas del *cluster* B₂).

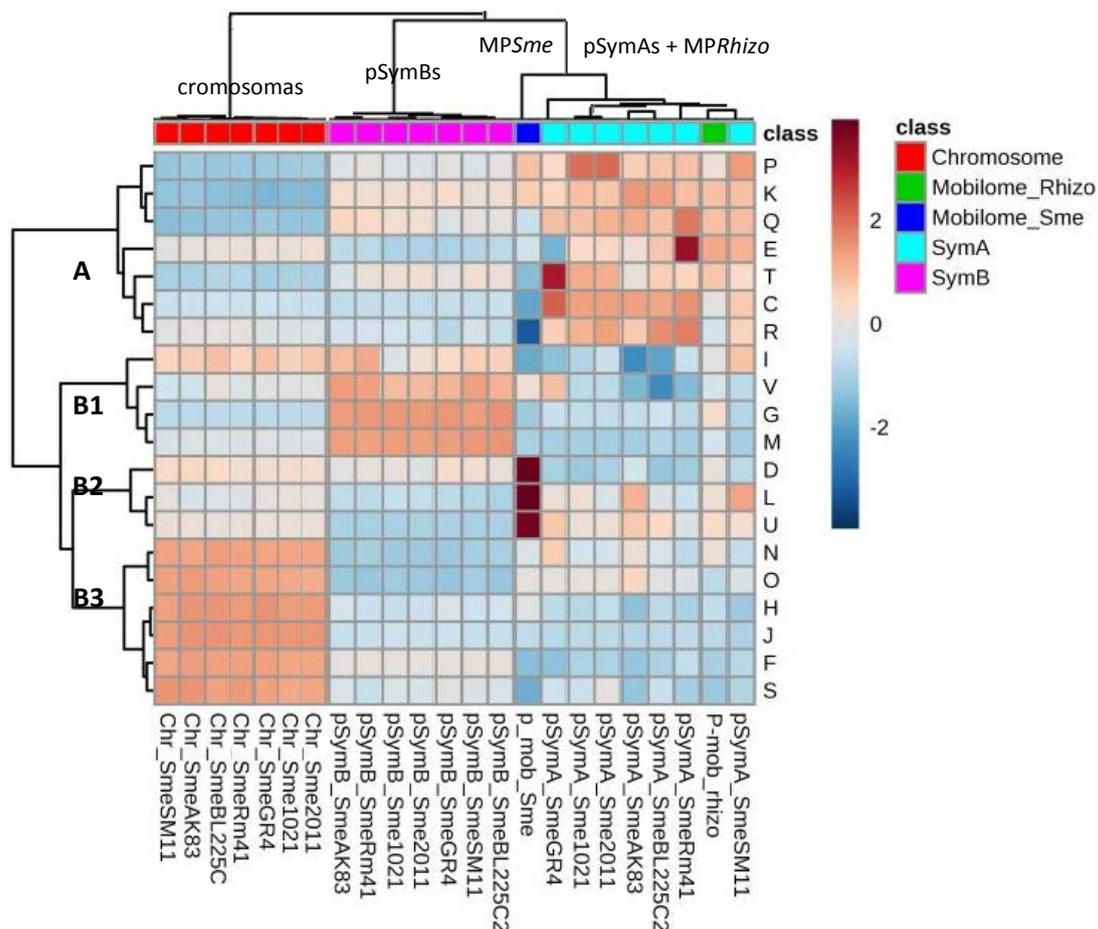


Figura III.2. Análisis jerárquico multivariado de clusters empleando como variables las proporciones de clases COG en los diferentes replicones de *S. meliloti*. MPRhizo, moviloma plasmídico de la familia Rhizobiaceas (secuencias disponibles en GenBank); MPSme, moviloma plasmídico críptico de *S. meliloti*.

En comparación con la heterogeneidad de las clases COG entre los diferentes pSymA (diferentes colores conservados a lo largo de las columnas contiguas de estos megaplásmidos), una composición COG altamente conservada –menos variable- fue observada tanto para el cromosoma y para el pSymB. Es interesante destacar, sin embargo, que de acuerdo al perfil funcional los pSymBs se agruparon dentro del *cluster* de los replicones plasmídicos con los pSymAs, el MPRhizo y el MPSme.

Evidencias de transferencia lateral de genes entre el moviloma plasmídico críptico y los pSym(s) y el cromosoma de *S. meliloti*

El análisis presentado en la sección anterior mostró que los diferentes compartimientos genómicos de *S. meliloti* portan –con diversos grados de variación- perfiles funcionales que les son distintivos. Hemos visto, además, que cada uno de los replicones presenta diferente grado de variación en su contenido génico siguiendo un orden decreciente pSymAs > pSymBs ≈ cromosomas. A partir de esta última observación, exploramos en más detalle las fuentes de variaciones entre los contenidos génicos de un mismo tipo de replicón. Así, a partir de las secuencias genómicas completas de diferentes cepas de *S. meliloti* disponibles en el GenBank calculamos el contenido de singletons por cada millar de genes presentes en cada uno de los replicones. En concordancia con la mayor diversidad de clases COG observada entre los pSymA(s) comparados con la diversidad observada entre los pSymB(s) o entre los cromosomas (Figura III.1.A); los pSymA(s), según puede observarse en los datos que se presentan en la Figura III.3, fueron también los replicones con mayor densidad de singletons (esto es, genes presentes en el replicón de una cepa pero no en los homólogos de las otras).

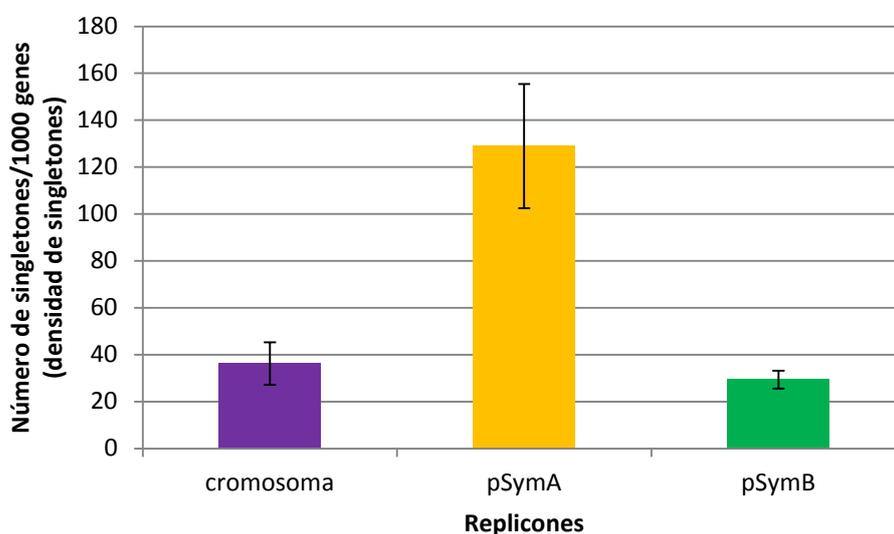


Figura III.3. Valores medios de la cantidad de singletons/1000 genes presentes en los cromosomas y en los megaplásmidos pSymAs y pSymBs de *S. meliloti*, calculados a partir del análisis de los genomas de siete cepas diferentes disponibles en el GenBank. Las barras de error indican las desviaciones estándar de la media.

Como se sugiere para los plásmidos del grupo *Bacillus cereus* [278], es concebible que los singletons representen a los genes más recientemente adquiridos durante la evolución de los replicones. Por lo tanto, la información de la Figura III.3 sugiere que la incorporación de genes

por transferencia horizontal se produce de manera más frecuente en el pSymA, comparado con el pSymB y los cromosomas. A partir de los resultados obtenidos existen dos fuertes evidencias que sugieren al moviloma plasmídico como posible fuente de varios de los singletons observados. En primer lugar el contenido GC% del moviloma plasmídico críptico de *S. meliloti* es similar al contenido GC% de los singletons independientemente del replicón que estemos analizando, y diferente al contenido GC% de los genomas core de cualquiera de los replicones (genomas core refiere en este caso al conjunto génico que se conserva en todos los replicones homólogos del conjunto de cepas analizadas)(Figura III.4).

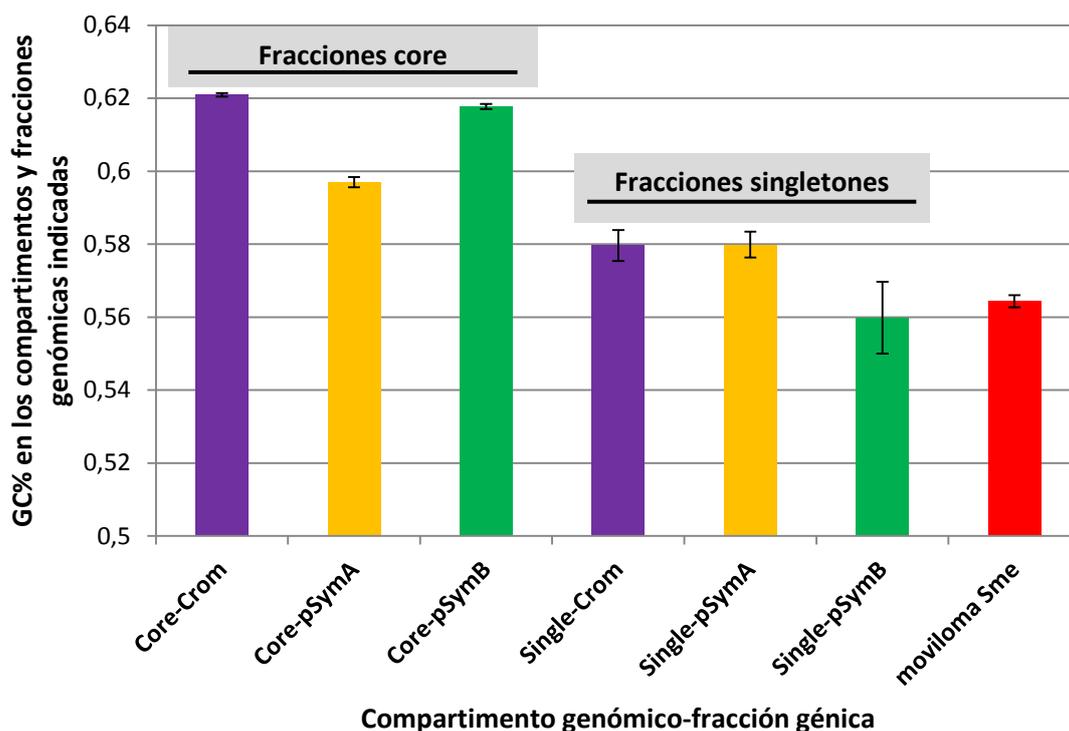


Figura III.4: Valores medios del contenido G+C porcentual (GC%) correspondiente a diferentes compartimentos (cromosomas, megaplásmidos) y fracciones (core, singletons) genómicas de *S. meliloti*, comparados con el GC% del moviloma plasmídico críptico del mismo rizobio (barra roja a la derecha). Las barras de error indican desviaciones estándar de la media.

La segunda evidencia que apoya la ocurrencia pasada de transferencia génica desde el moviloma plasmídico críptico a los pSymAs se presenta en la Figura III.5 en la que se evalúa a través de un PCA el uso de codones en los diferentes compartimentos genómicos completos (cromosomas, pSymAs, y pSymBs), y en varias de sus fracciones génicas (genes core, singletons). Según puede apreciarse, mientras el uso de codones en cada genoma core es cercano al correspondiente a su replicón completo (cuadrados y círculos con un mismo color); el uso de codones de la fracción singletons de cada replicón (triángulos) es más cercano al observado en el moviloma plasmídico críptico (círculo rojo). Por lo tanto hay una relación clara entre el contenido GC y el uso de codones entre los singletons y el moviloma plasmídico críptico. Resultó también evidente que mientras los usos de codones de los genomas core de los pSymBs (cuadrados verdes) son cercanos a los de los cromosomas (cuadrados azules), los usos de codones de los genomas core de los pSymAs (cuadrados amarillos) se desplazan hacia el uso de codones correspondiente al moviloma plasmídico críptico (círculo rojo).

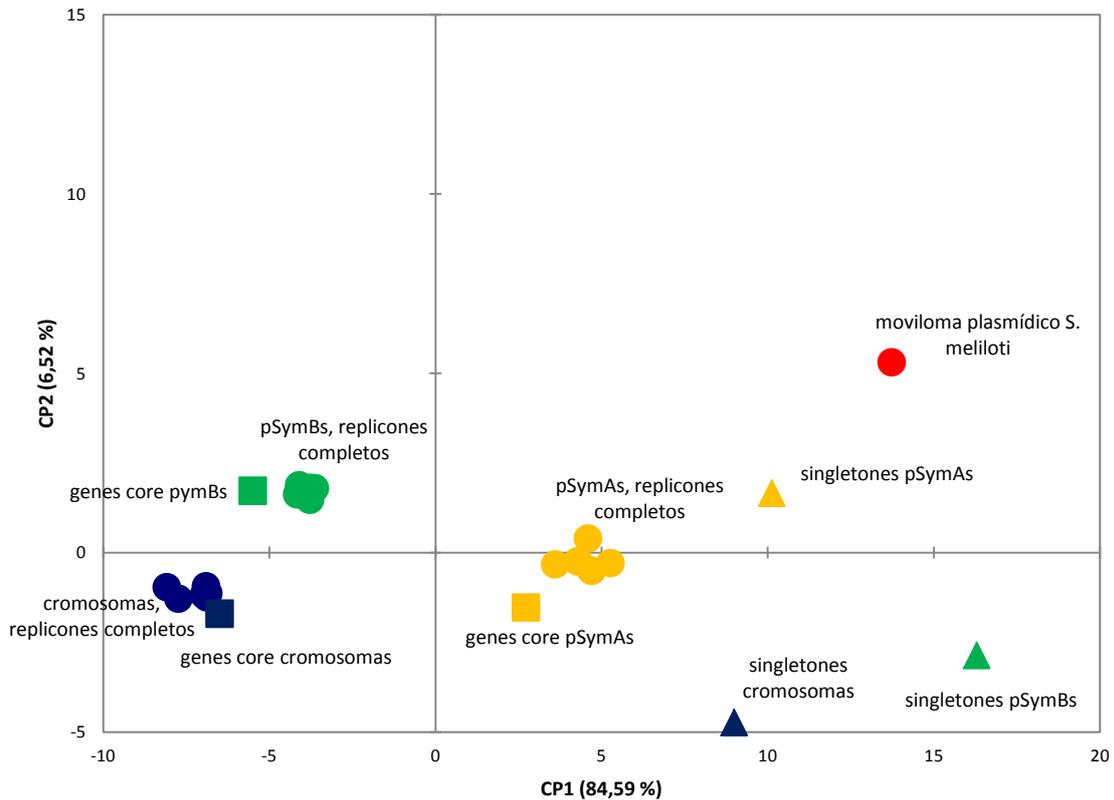


Figura III.5. Análisis de componentes principales utilizando como variables de entrada la proporción de uso de codones excluyendo AUG (único para Met), UGG (único para Trp); y UAA, UAG, UGA (todos de terminación). Los círculos azules, verdes y amarillos corresponden a los replicones completos cromosomales, pSymBs y pSymAs, respectivamente. Los cuadrados y los triángulos utilizan el mismo código de colores y corresponden a los genomas core y singletons, respectivamente.

Las evidencias derivadas de los contenidos GC% y del uso de codones en los distintos replicones y fracciones genómicas de *S. meliloti* apoyan ambas fuertemente la transferencia génica pretérita desde plásmidos no simbióticos a los tres replicones de este rizobio (cromosoma y pSyms), evidenciada hoy en miembros de las fracciones genómicas de los singletons.

Considerando la evidencia actual aceptada que sostiene que el uso de codones es moldeado durante la evolución a nivel celular [279, 280], surge la pregunta referida a ¿por qué los diferentes replicones presentan un uso de codones claramente distintivo? Probablemente, la diferencia en el uso de codones entre los genes core del pSymA(s), pSymB(s) y cromosomas refleje las diferencias originales entre el uso de codones en el rizobio pre-simbiótico (el cromosoma ancestral) y las de las bacterias ancestrales, dadoras de los replicones pSym. De acuerdo a esta posibilidad, la mayor parte de los genes core conservados entre los pSymA(s) y los pSymB(s) deberían haberse incorporado a sus replicones correspondientes antes de la adquisición de estos megaplásmidos por el precursor pre-simbiótico de *S. meliloti*. La función de decaimiento de los genes core al incrementar el número de pSyms incluidos en el análisis (Figura III.6) sirve para inferir que los precursores ancestrales del pSymA y pSymB en el momento de su adquisición podrían haber tenido en el orden de 500 y 1200 genes de los

actuales presentes en el core, respectivamente (es decir, un 40% y 80% de su contenido génico actual).

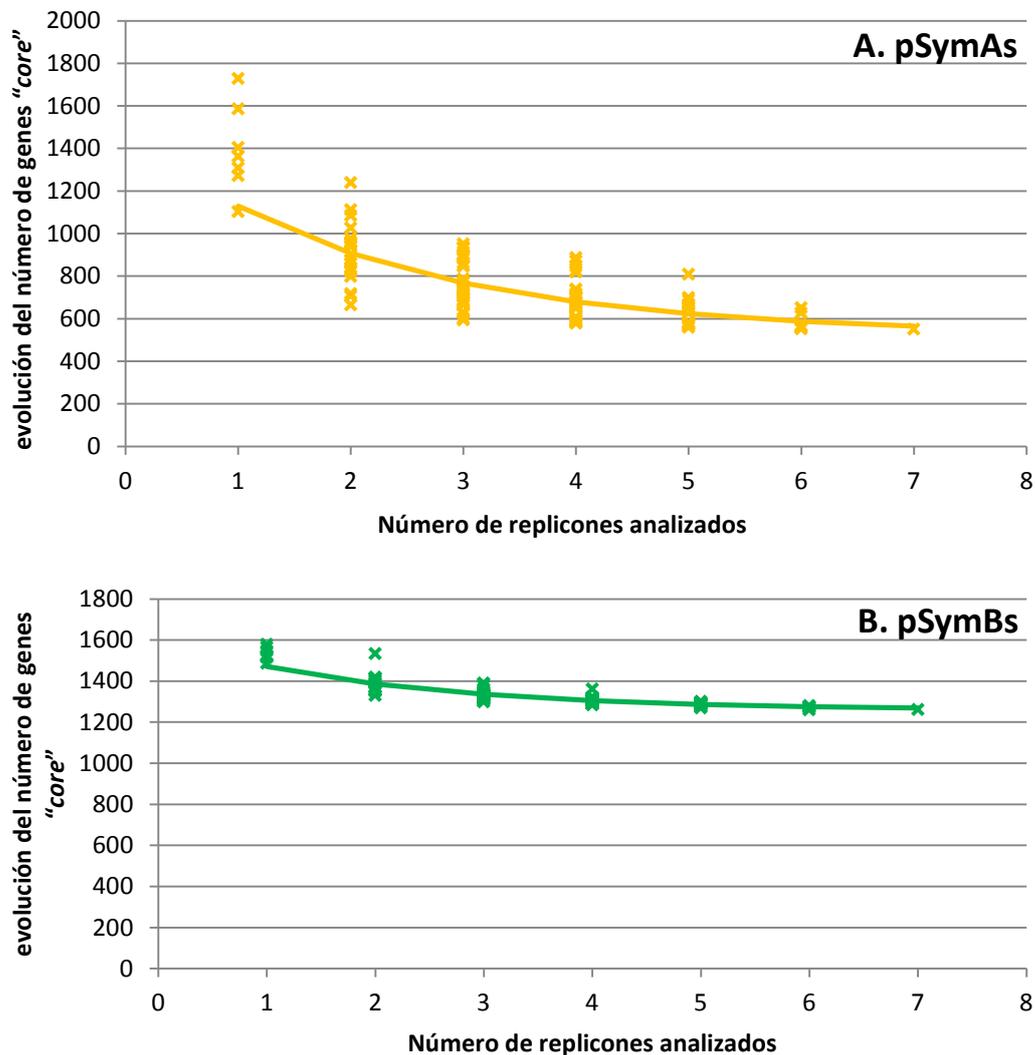


Figura III.6: Función de decaimiento de los genes core del pSymA (parte A) y pSymB (parte B) calculada con el *software* EDGAR (ver Materiales y Métodos, sección II.11) al adicionar n replicones (n de 1 a 7) en sus distintas combinaciones representadas por cruces en la figura. Las funciones que describen las curvas de las figuras son las siguientes: $Y=953 * \exp(-X/2,171) + 527,843$ (pSymA), y $Y=354,861 * \exp(-x/1,941) + 1259,758$ (pSymB).

Homólogos del genoma core y de los singletons de *S. meliloti* en el moviloma plasmídico críptico

El análisis comparativo de los genes core y singletons de los pSym y de los cromosomas respecto de sus homólogos en el moviloma plasmídico, sirvió para investigar cómo varió el uso de codones de dichos marcadores dentro de *S. meliloti* por un lado, respecto de cómo lo hizo siendo parte del moviloma plasmídico en circulación por el otro. La posibilidad de contar con homólogos al MPS_{me} de las fracciones core y singletons de cada replicón nos permitió comparar el uso de codones de los genes en su versión actual localizada en el cromosoma,

pSymA o pSymB (en el rizobio), respecto del uso de codones de sus genes homólogos localizados en el moviloma plasmídico. En la Figura III.7 se presenta un PCA en el que se comparan los usos de codones de los genes core y singletons de diferentes replicones (cromosomas y pSyms) para los que encontramos homólogos dentro del moviloma. Se observa en primer lugar que los homólogos al moviloma que están presente en cada fracción core (los **homólogos al moviloma en el core del pSymA, pSymB y cromosoma se indican con diamantes** amarillos, verdes y violetas, respectivamente) tienen cada uno un uso de codones que es similar a sus correspondientes **conjuntos totales de genes core (indicados con cuadrados** amarillos, verdes y violetas, respectivamente). En contraste con esta observación, los **homólogos de los mismos genes core pero ahora en su versión localizada en el moviloma plasmídico** (asteriscos amarillos, verdes y violetas) tienen –en promedio– un uso de codones que es similar en todos los casos al del moviloma completo (círculo rojo). Si la mayoría de estos homólogos se hubieran transferido recientemente entre alguna de las fracciones core (tanto de los megaplásmidos como de los cromosomas) y el moviloma, los mismos aún deberían presentar perfiles de uso de codones similares a los de su replicón de origen, y este no es el caso según lo observado en la Figura III.7. El uso de codones en los genes del core homólogos al moviloma dependiendo de su localización (rizobio / moviloma), sugiere fuertemente que esta fracción génica ha sido adquirida ancestralmente por el rizobio evolucionando al uso de codones actual de manera similar al del resto de los genes core de cada replicón, y distanciándose de la evolución del uso de codones de sus homólogos del moviloma (diamantes vs. asteriscos). Aunque aún no está claro cómo fueron los mecanismos que originaron los megaplásmidos simbióticos primordiales (reducción por pérdida de genes de replicones de mayor tamaño, transferencia de genes desde un moviloma ancestral hasta constituir estos megaplásmidos), los homólogos core al moviloma situados en los megaplásmidos deben haber evolucionado en huéspedes diferentes antes de entrar en *S. meliloti*, tal como se refleja por su propia frecuencia de uso de codones. **Es así probable que los genes core de los pSymA y pSymB homólogos al moviloma hayan sido incorporados a estos replicones antes de la adquisición de estos replicones por parte de los rizobios.** Además, el hecho de que la diferencia en el uso de codones (distancias en el PCA de la Figura III.7) entre los genes del MPS_{me} y sus homólogos core en los replicones de *S. meliloti* muestre incrementarse desde el pSymA al cromosoma≈pSymB (distancia desde el asterisco a los diamantes en amarillo, verde y azul; respectivamente) es un fuerte indicativo de una relación evolutiva más reciente del moviloma con el pSymA, y mucho más ancestral con los genes core homólogos del pSymB y el cromosoma.

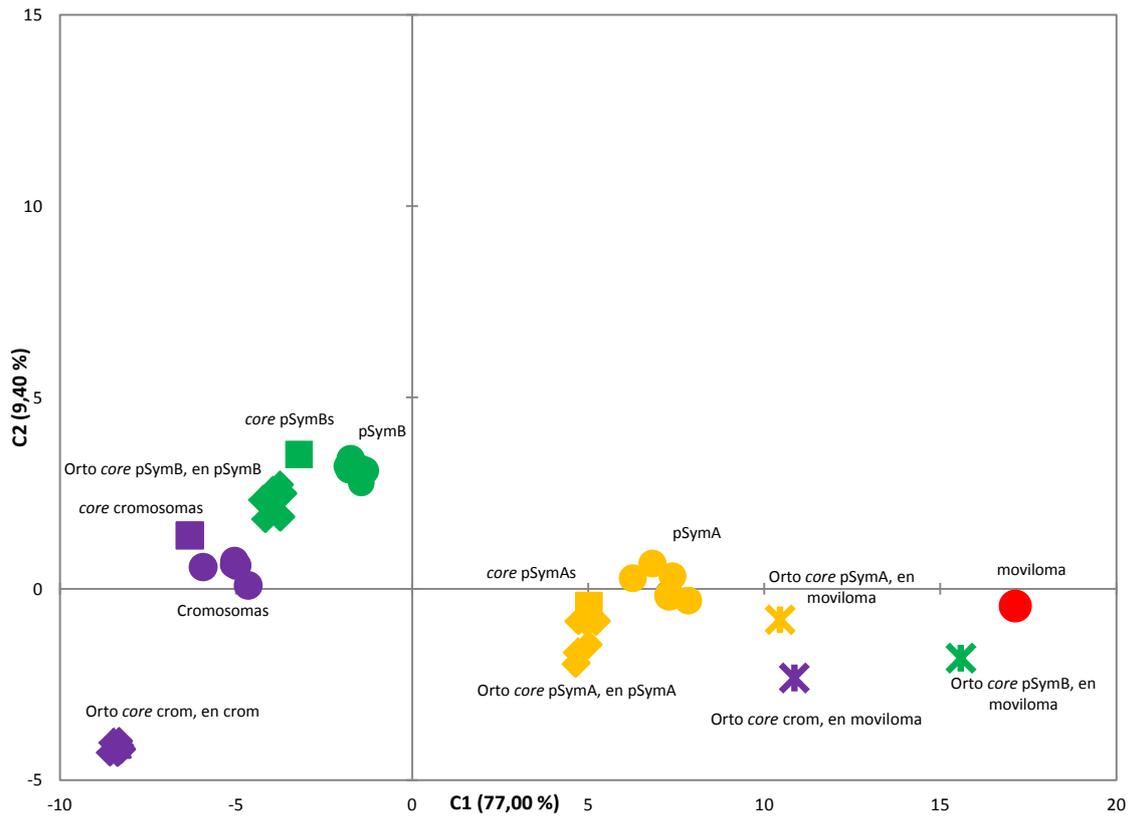


Figura III.7. Análisis de componentes principales utilizando como variables de entrada la proporción de uso de codones de los genes core de los diferentes replicones de *S. meliloti* (cromosomas y pSyms) para los que se encontraron homólogos dentro del moviloma plasmídico críptico. Los replicones y el tipo de fracción analizada se indica según los siguientes códigos de colores y símbolos, respectivamente: **Replicones.** violeta: cromosoma; amarillo: pSymA; verde: pSymB; rojo: moviloma; **Fracciones.** círculos: replicones completos, cuadrados: core completo para el replicón indicado, diamantes: genes core homólogos al moviloma, localizados en el replicón indicado; asteriscos: genes core homólogos al moviloma, localizados en el moviloma.

Otro aspecto interesante surgió de la diferencia observada entre el uso de codones de los homólogos al MPS_{me} en el core de cada replicón (diamantes), y el conjunto completo de genes del core (cuadrados). Mientras que en el caso de los megaplásmidos esta diferencia fue pequeña (uso de codones similar), resultó llamativo el claro corrimiento de los homólogos al moviloma presentes en el core cromosomal (diamantes violetas) respecto de su conjunto core completo (cuadrado violeta). La posible existencia de un subconjunto de genes homólogos con el cromosoma que se desvía notablemente con respecto al promedio de los genes conservados en todos los cromosomas, nos llevó a pensar que podría existir una relación de ancestralidad aún mayor para los homólogos al MPS_{me} presentes en el core cromosomal. De manera de indagar sobre cómo la ancestralidad podía vincularse a cambios en el uso de codones en los distintos subconjuntos de genes presentes en el core cromosomal, se generaron subconjuntos dentro del core que fueran asimismo homólogos a cromosomas de especies bacterianas con diferente relación taxonómica a *S. meliloti* (ver Tabla A2 del Anexo). La representación por PCA de estos subconjuntos de genes en la Figura III.8 mostró que a medida que los subconjuntos de homólogos al core cromosomal del rizobio incluían genes presentes en cromosomas cada vez más alejados filogenéticamente (coincidente con un menor número de genes homólogos en

los subconjuntos), se obtenía una localización de dichos genes en el espacio PC1, PC2 (diamantes naranjas) cada vez más alejada del conjunto del core total de *S. meliloti* (cuadrado violeta). En particular, la localización de los diferentes grupos de homólogos al core cromosomal de *S. meliloti* siguió un trayecto hacia valores cada vez más bajos de los componentes principales 1 y 2 al crecer la distancia filogenética entre los rizobios y las bacterias de cuyos cromosomas se infirieron los homólogos (en el gráfico puntos desde el core ancestral 1 al 7). Esta observación fue indicativa por un lado de que existe una composición preferencial en el uso de codones hacia donde progresan los genes más ancestrales en un cromosoma, y que este efecto de ancestralidad puede detectarse en un análisis de este tipo. Los resultados de la Figura III.8 muestran por otra parte que: a) que el trayecto seguido por el uso de codones incluyó el punto correspondiente a los genes cromosomales core de *S. meliloti* homólogos del moviloma, y b) que es probable (por la posición respecto de los puntos de la curva) que los genes homólogos con los plásmidos críticos que están alojados en el core cromosomal presenten un uso de codones de mayor ancestralidad que el resto del conjunto de los genes del core. Tales observaciones tomadas en conjunto sugieren muy fuertemente que la fracción de genes del core cromosomal que son homólogos con genes del moviloma corresponden posiblemente a genes que fueron adquiridos/compartido con un moviloma ancestral en los estadios primordiales de la constitución del cromosoma como replicón.

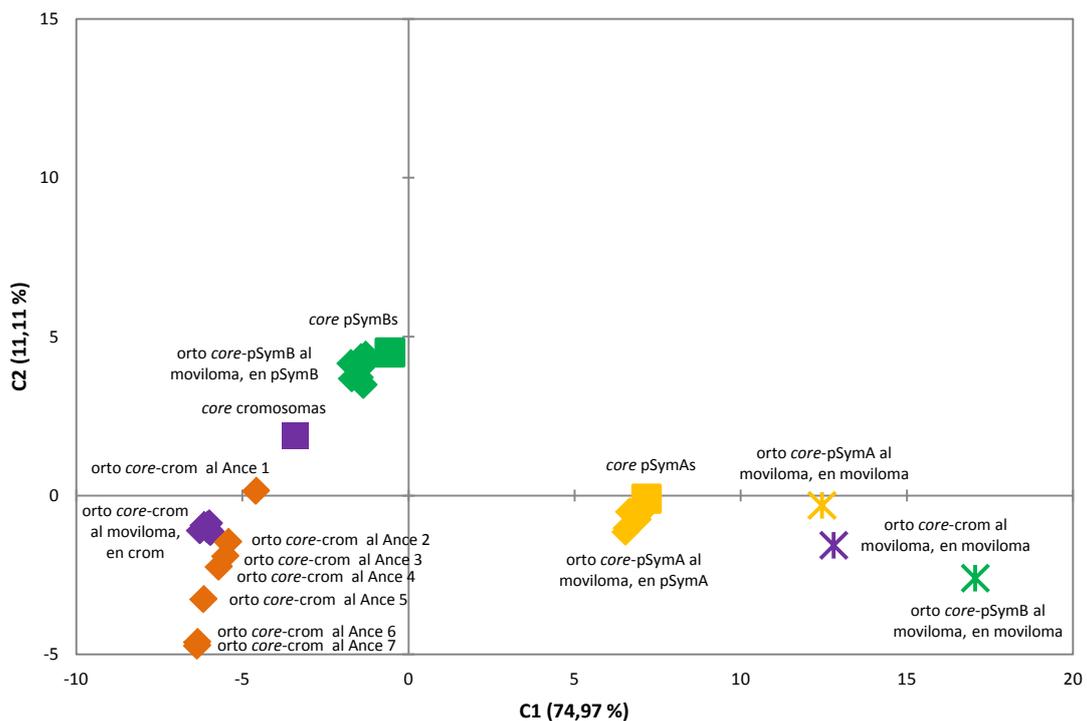


Figura III.8. Análisis de componentes principales utilizando como variables de entrada la proporción de uso de codones de los genes core del cromosoma de *S. meliloti* que son homólogos con genes cromosomales de bacterias que presentan con distinto grado de vínculo filogenético con el rizobio. Los replicones y el tipo de fracción analizada se indica según los siguientes códigos de colores y símbolos, respectivamente: **Replicones.** violeta: cromosoma; amarillo: pSymA; verde: pSymB; naranja: homólogos presentes en el core cromosomal a conjuntos de cromosomas de distintas bacterias (Ance1-Ance7). **Fracciones.** cuadrados: core completo para el replicón indicado, diamantes: genes core homólogos a cromosomas particulares (1 a 7) o al moviloma de plásmidos críticos de *S. meliloti*.

El análisis de las funciones presentes en este subconjunto de genes, parte del core cromosomal y homólogos al moviloma, (diamantes violetas “orto core-crom al moviloma, en crom” en la Figura III.9, Tabla A4 del Anexo) nos indica que la mayoría de ellos corresponden a genes del metabolismo celular (45%, M), seguido por genes de funciones aún desconocidas (21%, H), transporte (21%, T), reguladores transcripcionales (9%, R) y algunas pocas funciones de metabolismo del ADN (3%, A) y de origen fágico (1%, F).

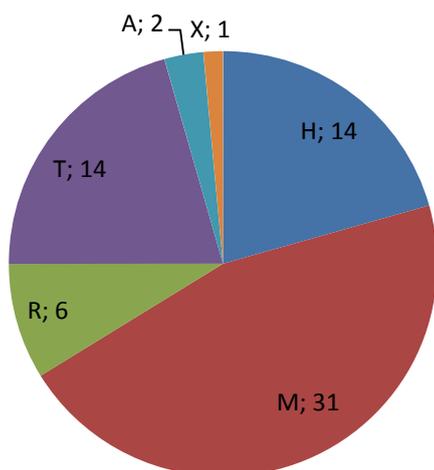


Figura III.9: Categorías funcionales presentes en el conjunto de genes ancestrales (ver texto) presentes en el core cromosomal de *S. meliloti* y ortólogos al MPSme. H: funciones de función desconocida, M: metabolismo, R: regulación, T: transporte, A: metabolismo del ADN, X: proteínas asociadas a fagos.

Estas funciones del core cromosomal debieron haber sido transferidas desde plásmidos ancestrales (moviloma primitivo) al cromosoma de un proto rizobio, para luego ajustar su uso de codones al observado hoy para esta fracción génica distinguible del resto de los genes core y de los demás genes cromosomales. Es posible que la incorporación de nuevas secuencias al moviloma plasmídico críptico con el secuenciamiento y análisis de nuevas cepas, de lugar a la identificación de nuevos miembros de este conjunto temprano de genes cromosomales. El conjunto de genes identificados representa hasta el momento sólo el 2% de la fracción core del cromosoma.

Para explorar cuántos de estos genes son asimismo parte de otros cromosomas de bacterias relacionadas, se buscaron ortólogos de este grupo en los cromosomas core de *Agrobacterium* y *Agromonas* (“Ance 2” en Figura III.8, Tabla A3 del Anexo). Estas funciones debieron ser de origen común a los plásmidos crípticos y el cromosoma ancestral a estas tres especies. El análisis precedente mostró que 26% de los genes del core cromosomal de *S. meliloti* con homólogos en el moviloma tiene también ortólogos en las fracciones core cromosomales de *Agrobacterium/Agromonas*, y que dicho conjunto está asociado a funciones metabólicas relacionadas con la biosíntesis de fenilalanina, glutamina, lipopolisacáridos, tioredoxina, biotina y chaperonas, enzimas de degradación de xenobióticos y del metabolismo del carbono, así como proteínas de transporte de azúcares, aminoácidos y poliaminas. Finalmente, y con el propósito de caracterizar un subconjunto de genes aún más ancestral, analizamos los ortólogos al MPSme de un conjunto más amplio de cromosomas de bacterias de la familia *Rhizobiaceae* (conjunto “Ance 7” en la Figura III.8, Tabla A3 del Anexo). El análisis arrojó un

total de 9 genes que se listan en la Tabla III.1. Los mismos son comunes a las distintas especies de la familia, están presente en el moviloma críptico y en la fracción core del cromosoma de *S. meliloti*, y presentan un comportamiento de uso de codones propio de genes adquiridos tempranamente en la filogenia de la familia.

Tabla III.1. Genes ortólogos entre el MPS_{Sme} y diferentes cromosomas de *Rhizobiaceae*.

denominación del OFR	predicción bioinformática de la función	categoría	Homólogo en plásmido críptico (1)	Homólogo en plásmido simbiótico (1)
<i>S. meliloti</i> mobilome final 147	K01946 biotin carboxylase	M	pRM41A (<i>Sme</i>), pSINME01 (<i>Sme</i>)	No
<i>S. meliloti</i> mobilome final 395	K09972 general L-amino acid transport system ATP-binding protein	T	pRetNIM1c (<i>R.etli</i>), pRtrCIAT899b (<i>R.tro</i>)	No
<i>S. meliloti</i> mobilome final 696	K01627 2-dehydro-3-deoxyphosphoactonate aldolase (KDO 8-P synthase)	M	No	No
<i>S. meliloti</i> mobilome final 1258	K01915 glutamine synthetase	M	pHRC017 (<i>Sme</i>), pSINME01 (<i>Sme</i>)	No
<i>S. meliloti</i> mobilome final 1296	acetylglutamate kinase (EC:2.7.2.8)	M	pSmeSM11a (<i>Sme</i>)	No
<i>S. meliloti</i> mobilome final 1360	K00979 3-deoxy-manno-octulosonate cytidyltransferase (CMP-KDO synthetase)	M	No	No
<i>S. meliloti</i> mobilome final 1366	K00384 thioredoxin reductase (NADPH)	M	pNGR234b (<i>Sfre</i>), pSYMA (<i>Sme</i>)	core-pSymA (<i>Sme</i>)
<i>S. meliloti</i> mobilome final 1420	hypothetical protein	H	pHRB800 (<i>Sme</i>), pRmeGR4b (<i>Sme</i>), pHRC017 (<i>Sme</i>), pNGR234b (<i>Sfre</i>)	core-pSymA (<i>Sme</i>)
<i>S. meliloti</i> mobilome final 1528	Probable cold shock protein y4cH	M	pHRB800 (<i>Sme</i>), pRmeGR4b (<i>Sme</i>), pHRC017 (<i>Sme</i>)	core-pSymA (<i>Sme</i>)

(1) Empleando tBLASTn se exploró la presencia de potenciales homólogos en otros plásmidos de *S. meliloti*. Se evaluó asimismo su presencia en el pSymA de todas las cepas y en el core del pSymA. Se indica la pertenencia del plásmido críptico como: *Sme* (*S. meliloti*), *R.tro* (*Rhizobium tropici*), *R.etli* (*Rhizobium etli*), *Sfre* (*Sinorhizobium fredii*).

Los genes de la primera columna de la Tabla III.1 pertenecientes al MPS_{Sme} presentaron en la mayoría de los casos genes homólogos en alguno de los plásmidos crípticos ya secuenciados de *S. meliloti*. Tal observación muestra para este conjunto génico la existencia frecuente de redundancia genómica. Algunos de estos genes mostraron asimismo homólogos en el core del pSymA, lo cual pone en evidencia la redundancia de los mismos también en el megaplásmido simbiótico. La evidencia de algunos de estos genes presentes en plásmidos de otras especies relacionadas muestra la existencia de genes metabólicos ancestrales que, estando presente en el plasmidoma críptico de *S. meliloti* y en el core cromosomal de *Rhizobiaceae*, son además parte de un moviloma plasmídico más amplio que circula –como era esperable– por varios miembros de la familia. Por el análisis antes expuesto de la frecuencia de uso de codones, estos genes no son copias recientes en los plásmidos crípticos de las versiones cromosomales, sino por el contrario son sus homólogos que mantuvieron un camino evolutivo independiente en los plásmidos crípticos. Es interesante preguntarse entonces qué dinámicas evolutivas mantienen a estos genes presentes en un compartimento por demás “plástico”, si asumimos que los plásmidos crípticos codifican información que no es esencial para su hospedador, y que estos genes son prescindibles para la supervivencia del mismo plásmido críptico. Es interesante preguntarse si los genes metabólicos ancestrales presentes actualmente en los plásmidos crípticos complementan procesos metabólicos ya presentes en la célula mediante la adición de copias del gen, o bien si los plásmidos crípticos son reminiscencias de y suministros

actuales de genes en la construcción de los distintos replicones. ¿Cómo logran estos genes metabólicos mantenerse intactos en replicones plásticos, no esenciales, evadiendo su delección por mecanismos de recombinación, inserción o por acumulación de mutaciones puntuales que lleven a la pérdida de su función a lo largo de la evolución si no contribuyen con un rol esencial en algún contexto fisiológico?

Homólogos al MPSme presentes en la fracción singletones

El análisis de los genes homólogos al MPSme que están presentes en la fracción singletones del pSymA, pSymB y del cromosoma arrojó resultados que contrastaron de modo marcado con los obtenidos al evaluar los genes homólogos de las fracciones core de los mismos replicones (sección anterior). Al tratarse de genes singletones, los mismos pueden ser considerados como el producto ya sea de genes adquiridos recientemente, o bien como la resultante de la pérdida de genes particulares en un clon ancestral que dio luego origen a 6 de las 7 cepas analizadas de *S. meliloti*. El análisis que realizamos a continuación apoya la primera de las posibilidades anteriores.

La posibilidad de contar con ambas versiones de los genes homólogos, esto es los localizados por un lado en el MPSme (en los plásmidos crípticos), y por el otro en cada replicón; nos permitió comparar el uso de codones en uno y otro compartimento genómico. Según esperaríamos si se tratase de una transferencia reciente de los singletones desde el MPSme a cada uno de los replicones del rizobio, el uso de codones de ambas versiones de homólogos se muestra en efecto similar según puede apreciarse en el PCA de la Figura III.9 (pequeña distancia entre asteriscos y triángulos de cada color), a diferencia de lo que describimos previamente para los homólogos del core al MPSme (Figura III.7, distancia entre diamantes y asteriscos de cada color). Además, el hecho de que el uso de codones (de los diferentes pares de homólogos en general) se encuentre cercano al del conjunto de los genes del MPSme indica que la transferencia tuvo que haberse dado desde los plásmidos crípticos hacia el pSymA, pSymB y cromosoma, respectivamente; y no a la inversa. Si los singletones se hubiesen transferido recientemente en sentido opuesto, hacia los plásmidos crípticos; el uso de codones de ambas versiones de los genes hubiese sido similar (y no lo es) al promedio de los genes del replicón de origen (pSymA, pSymB o cromosoma) (triángulos y asteriscos hubiesen estado cerca de los círculos). Los resultados de la Figura III.10 apoyan la hipótesis de una adquisición reciente de singletones desde el MPSme hacia los rizobios, donde además la mayoría de los genes recibidos (singletones) aún no han modificado sustancialmente su uso de codones hacia el de su nueva posición genómica, preservando aún un uso de codones similar al del MPSme que se consolida como la posible fuente de dichos marcadores. De todos modos, y más allá de estas consideraciones, nótese que la posición de los triángulos desplazados hacia el uso de codones de sus propios replicones completos (círculos) es posiblemente indicativo del comienzo de una incipiente adaptación de los nuevos singletones al uso de codones de sus nuevos replicones (efecto que no se observa para el caso de los singletones homólogos al MPSme presentes en el pSymB). El análisis precedente apoya con fortaleza la hipótesis que muestra a los singletones homólogos al MPSme como productos de la transferencia más reciente de genes desde los plásmidos crípticos a los diferentes replicones en *S. meliloti*, con signos leves de adaptación a un nuevo uso de codones.

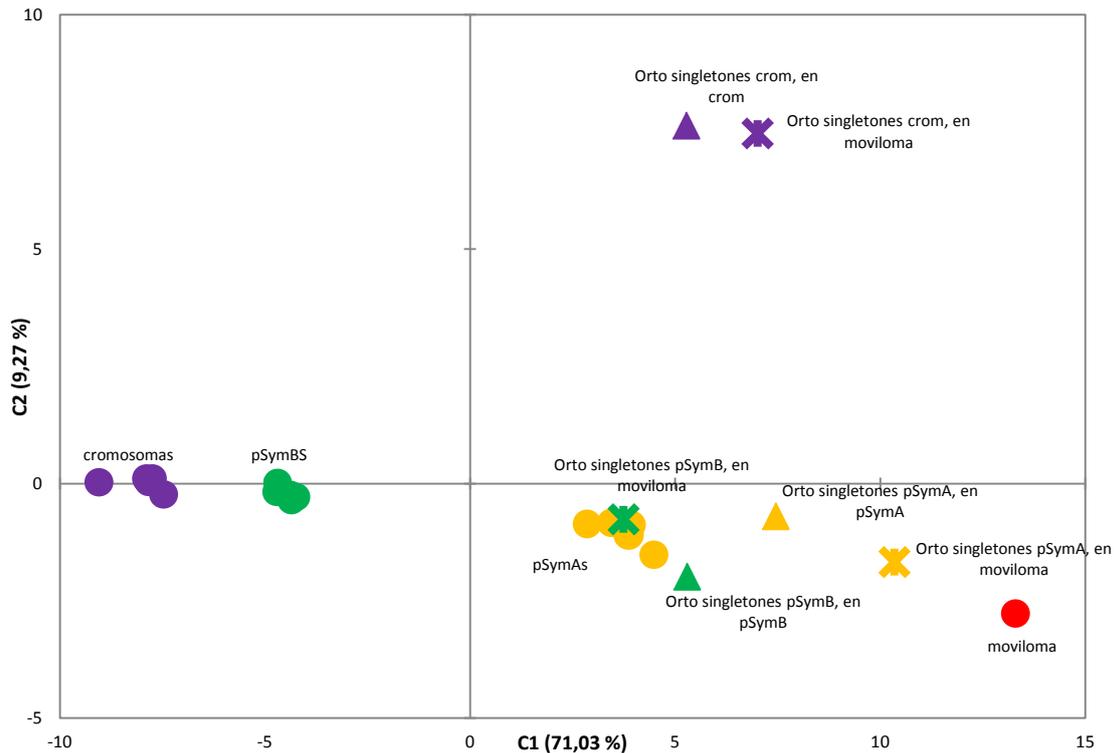


Figura III.10. Análisis de componentes principales utilizando como variables de entrada la proporción de uso de codones de los genes singletones de los diferentes replicones de *S. meliloti* (cromosomas y pSyms) para los que se encontraron homólogos dentro del moviloma plasmídico críptico. Los replicones y el tipo de fracción analizada se indica según los siguientes códigos de colores y símbolos, respectivamente: **Replicones.** violeta: cromosoma; amarillo: pSymA; verde: pSymB; rojo: moviloma; **Fracciones.** círculos: replicones completos, triángulos: genes singletones homólogos al moviloma, localizados en el replicón indicado; asteriscos: genes singletones homólogos al moviloma, localizados en el moviloma.

Al observar las funciones codificadas por los singletones del cromosoma que son homólogos a genes del MPS_{me}, observamos que la mayor proporción de ellos está relacionada a distintos mecanismos de movilidad de genes (tales como transposasas, recombinasas, metilasas, proteínas relacionadas a fagos) y a genes de función desconocida (proteínas hipotéticas con homólogos secuenciados)(Figura III.11.A3). Tal perfil funcional contrasta con la dominancia de genes que codifican para funciones metabólicas, de transporte y regulatorias en los genes core cromosomales homólogos al MPS_{me} (Figura III.9 y Figura III.11.A1) según mostramos en la sección anterior. De hecho, la dominancia de genes metabólicos en las fracciones core con homólogos en el MPS_{me} parece ser una característica general, independientemente del replicón que se considere (Figura III.11.A1, B1, y C1). Además, el mismo cambio composicional mencionado en el cromosoma puede apreciarse entre los homólogos al MPS_{me} en las fracciones core y singleton del pSymB (Figuras III.11.B1 y III.11.B3) Contrastando con la similitud entre los singletones del cromosoma y los del pSymB, los singletones del pSymA homólogos al MPS_{me} (Figura III.11.C3) mostraron una proporción mucho más alta de genes asociados a funciones metabólicas. Es interesante que una marcada proporción de genes metabólicos en el pSymA es también propia de la fracción de genes homólogos al MPS_{me} distribuida de manera variable (fracción VD)(Figura III.11.C2). La fracción VD corresponde a aquella representada por genes compartidos entre al menos 2, y en nuestro caso nunca por más de 6 cepas.

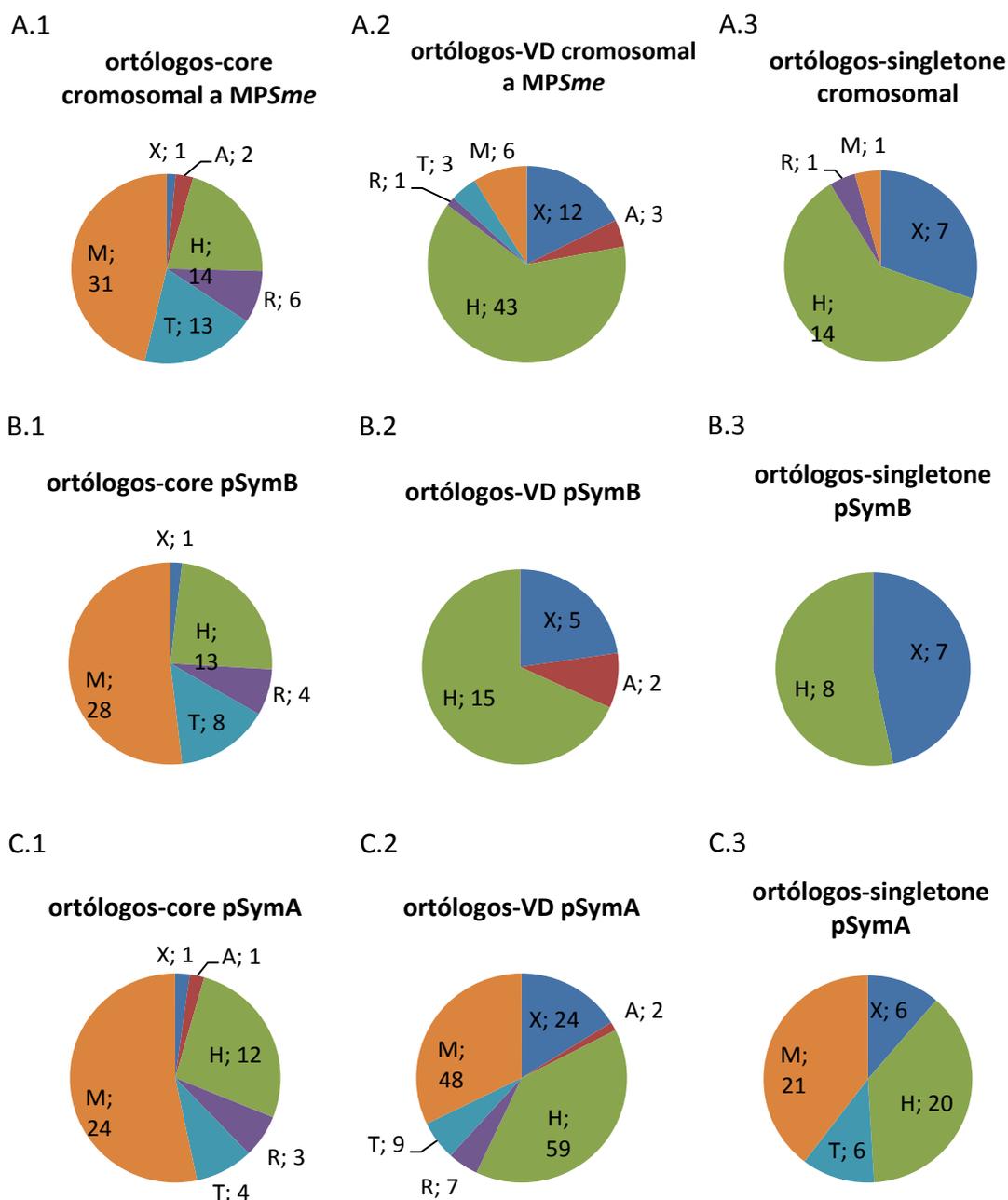


Figura III.11. Predicción informática de las funciones asociadas a los homólogos al MPSme presentes en la fracción core (columna 1), en la distribuida de manera variable (VD)(columna 2) y en la fracción singletons (columna 3). Dichos datos se presentan para los replicones cromosomales (fila A), para el pSymB (fila B) y para el pSymA (fila C). Se indica el número de genes homólogos al MPSme pertenecientes a cada grupo funcional. La predicción de funciones se realizó sobre los grupos de homólogos obtenidos con EDGAR [268] y realizando una clasificación manual en grupos funcionales (ver sección II.18). Las letras dentro de las gráficas corresponden a las siguientes categorías. M: metabolismo; H: función desconocida; X: funciones de fagos, transposasas, secuencias de inserción; T: transporte; R: regulación; A: modificación del ADN.

Al extender el mismo análisis sobre **la totalidad de los genes singletons** en los 3 replicones (los que presentan homólogos al MPSme y los que no, Figura III.12) se observa una diferencia similar en la composición funcional entre el pSymA por un lado y el pSymB y cromosoma por otro: mientras que el pSymA muestra una mayor proporción de genes metabólicos (M), de transporte (T) y regulación (R), los pSymB y cromosomas muestran una mayor proporción de

proteínas hipotéticas (H) y genes móviles y de fagos (X). Esto nos lleva a preguntarnos por qué el pSymA sigue incorporando genes que codifican una mayor diversidad de funciones metabólicas, mientras que el pSymB y cromosoma parecen estar más sesgados a la incorporación de genes desconocidos o relacionados a elementos móviles.

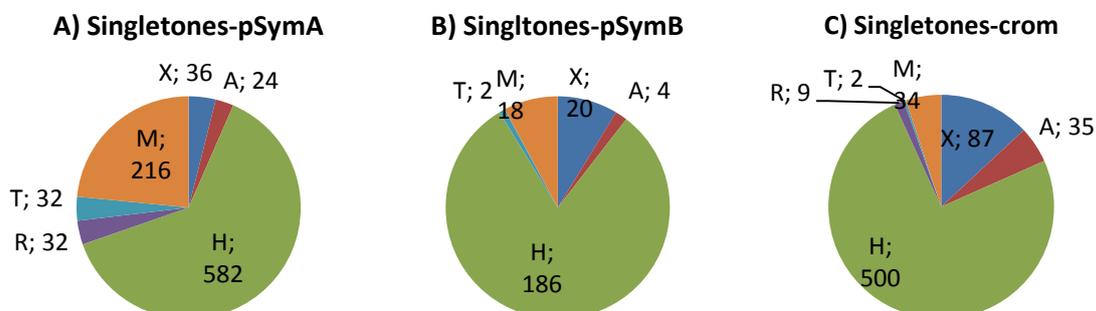


Figura III.12. Predicción informática de las funciones asociadas a los homólogos al *MPSme* presentes en la fracción singletons completa. Dichos datos se presentan para los replicones pSymAs (A), para los pSymBs (B) y cromosomales (C). Se indica el número de genes homólogos al *MPSme* pertenecientes a cada grupo funcional. La predicción de funciones se realizó sobre los grupos de homólogos obtenidos con EDGAR [268] y realizando una clasificación manual en grupos funcionales (ver sección II.18). Las letras dentro de las gráficas corresponden a las siguientes categorías. M: metabolismo; H: función desconocida; X: funciones de fagos, transposasas, secuencias de inserción; T: transporte; R: regulación; A: modificación del ADN.

Hasta aquí el conjunto de evidencias obtenidas indica por un lado que varios genes tempranamente compartidos entre los replicones de *S. meliloti* y los plásmidos crípticos persistieron a lo largo del proceso evolutivo, y son hoy identificables como homólogos del moviloma en las fracciones core de los diferentes replicones. Varios de dichos genes están relacionados a funciones metabólicas. Por otra parte, los genes de adquisición más reciente que corresponden a homólogos al *MPSme* de la fracción singletons de los cromosomas y pSymB codifican principalmente para transposasas, intrones, y una proporción destacada de genes aún de función aun desconocida, mientras que los recientemente adquiridos por el pSymA muestran una diversidad funcional similar a la que se observa en los homólogos al *MPSme* en su fracción core.

Es posible que sólo el pSymA muestre una plasticidad tal que permita la adquisición de una mayor diversidad de funciones génicas por no alterar la viabilidad de la bacteria (una menor presión de selección debido a un rol más secundario en vida libre), o bien que exista una mayor probabilidad intrínseca por parte del mismo de incorporar secuencias provenientes de plásmidos (mayor densidad de secuencias recombinogénicas). Será interesante explorar si esta misma observación se da entre los plásmidos simbióticos de otras bacterias noduladoras y sus respectivos plásmidos crípticos.

La fracción de genes singletons que han ingresado en un replicón (mediante integración de elementos, de plásmidos o de fagos) transitará por un proceso de selección luego del cual quedarán incorporados al repertorio más estable sólo aquellos genes que contribuyan a la adaptación (*fitness*) de la especie a un determinado nicho. Tal hipótesis se ve reforzada a partir del análisis del conjunto completo de genes en la fracción singleton (del que los homólogos al

MPSme que hemos analizado representan sólo una pequeña fracción), donde encontramos una distribución predominante de funciones relacionadas a fagos, transposasas y genes de función desconocida como parte de la fracción singletones (Figura III.12). Será interesante analizar en qué grado los plásmidos crípticos o los bacteriofagos colaboran en la adquisición de los genes por la fracción singleton en cada replicón de *S. meliloti*. En relación a ellos, el moviloma que hemos analizado en este trabajo, que consiste en aproximadamente 1,46 Mb de secuencia no redundante y ha mostrado portar homólogos de 6,5%, 6,2%, y 3,6% de los genes singletones presentes en los cromosomas, pSymB, y pSymA de *S. meliloti*. Debe tenerse en cuenta que por un lado el tamaño del moviloma plasmídico críptico accesible a la especie incluye secuencias que desconocemos y a las que no hemos accedido por limitaciones de muestreo (geográficas y de volumen de aislamientos analizados). Posiblemente sea esta la razón que impone la mayor limitación a trazar en el moviloma plasmídico el origen de la gran mayoría de los singletones. La mejora en la cobertura del MPSme modificará sin dudas la proporción de singletones en condiciones de ser analizados como posibles provenientes del plasmidoma críptico accesible a la especie. Por otro lado, debe tenerse en cuenta que muchos genes que son hoy parte de las fracciones singletones en los distintos replicones), pueden no ser ya parte del moviloma críptico circulante, circunstancia que los dejará como genes huérfanos de homólogos de sus ancestros plasmídicos. De todos modos, este argumento parece adquirir mayor peso para genes huérfanos de las fracciones core que han sido adquiridos en tiempos ancestrales desde movilomas crípticos con configuraciones seguramente muy diferentes.

Más allá de todas las consideraciones precedentes para rastrear el origen de los singletones, la mayor densidad de genes singletones observada para el pSymA respecto de los otros replicones (Figura III.3), el tipo de funciones asociadas a ellos (Figura III.12) y su uso de codones próximo al observado para el MPSme son todas evidencias que indican que el pSymA coevoluciona con los plásmidos crípticos de manera mucho más íntima que los otros dos replicones de *S. meliloti*. Desde el punto de vista funcional es interesante asimismo mencionar que para el caso particular del pSymA, los genes que aparecen como adquiridos más recientemente desde los plásmidos crípticos (los singletones homólogos al MPSme, Figura III.3 C3) muestran un perfil funcional parecido al de las fracciones core ortólogas al MPSme presentes en los distintos replicones (Figura III.3 A1, B1, C1). En este contexto es interesante explorar qué características en la dinámica de transferencia de genes o en la importancia en la vida de la bacteria tiene en particular el pSymA, que le permite un genoma core más reducido, y una mayor proporción de genes variables o únicos, es decir un carácter más plástico con respecto a los otros replicones en *S. meliloti*.

III.3. Análisis de genes ortólogos por EDGAR

Estimación cuantitativa de ortólogos al comparar diferentes plásmidos y cromosomas bacterianos contra el MPS_{me}

Con el propósito de investigar relaciones de ancestralidad entre los genes presentes en el MPS_{me} y diferentes replicones de especies bacterianas de la familia Rhizobiaceae y de otras bacterias relacionadas, utilizando el *software* EDGAR [268], realizamos una búsqueda de ortólogos según se muestra en la Figura III.13. El tamaño de cada grupo compartido de ortólogos al MPS_{me} fue expresado como el número de ortólogos identificados por cada cien genes en cada replicón (o metacóntigo artificial) analizado (índice NOCG en las ordenadas del gráfico), como una medida de la densidad de ortólogos. Como era esperable, el NOCG resultó ser mayor para los diferentes plásmidos (lado derecho del gráfico, barras verdes) que para los diferentes cromosomas (lado izquierdo, barras azules) en consonancia con un vínculo más cercano entre el MPS_{me} y los compartimentos plásmidicos de las distintas especies. Además, la búsqueda de ortólogos al MPS_{me} con plásmidos de otras especies mostró NOCG mayores para aquellos géneros que pertenecen a la clase Alphaproteobacteria (NOCGs > 10%, y hasta 40%) comparado con Beta- y Gammaproteobacteria (NOCGs < 10%). Notablemente, ambos pSyms de *S. meliloti* mostraron valores de NOCG intermedios comparados con los de cromosomas y plásmidos en las distintas especies (ambos megaplásmidos indicados con barras amarillas), donde pSymA aparece como el replicón más cercano al moviloma críptico en términos de contenido de ortólogos, en consistencia con las evidencias de las secciones anterior en que analizamos su contenido de singletones y uso de codones.

A partir del análisis precedente resultó además particularmente interesante la alta densidad de ortólogos compartidos entre los plásmidos de *S. meliloti* y de especies como *Ochrobactrum* y *Sinorhizobium fredii*, sumado a los plásmidos de otras especies asociadas con diferentes plantas. Es posible entonces que el grupo de ortólogos que hemos reconocido sea la consecuencia visible de un intercambio activo de genes (más o menos ancestral) entre los géneros indicados, que podrá indagarse si ha sido favorecido por relaciones simpátricas entre algunas de estas bacterias. No debe descuidarse, de todos modos, que la cercanía del MPS_{me} a algunos de estos géneros y no a otros puede estar aún sesgada por la cantidad de secuencias disponibles en el GenBank para unos y otros, atento a que la información de secuencia disponible para los distintos organismos no es la misma.

Para profundizar en la similitud que muestran los genes del MPSme en relación a otros plásmidos o cromosomas secuenciados, realizamos una búsqueda por tBLASTn contra la base de datos de GenBank haciendo uso de todos los marcos de lectura abiertos (ORFs) predichos bioinformáticamente y observamos que una gran cantidad de los mismos codifican proteínas ya presentes en distintos plásmidos de aislamientos de *S. meliloti*, *S. medicae*, *S. fredii*, *Agrobacterium*, *Mesorhizobium*, *Rhizobium gallicum* y especies aisladas de suelo como *Polymorphum gilvum* (Figura III.14). Esto indica que algunos plásmidos en nuestra colección presentan genes ya secuenciados en otros plásmidos de *S. meliloti*, lo cual indica cierto grado de redundancia funcional entre aislamientos obtenidos de sitios geográficos distintos. Por otro lado también es indicativo de que las otras especies presentan movilomas plasmídicos que presentan algún vínculo de transferencia de genes con mayor o menor grado de ancestralidad, que será interesante caracterizar.

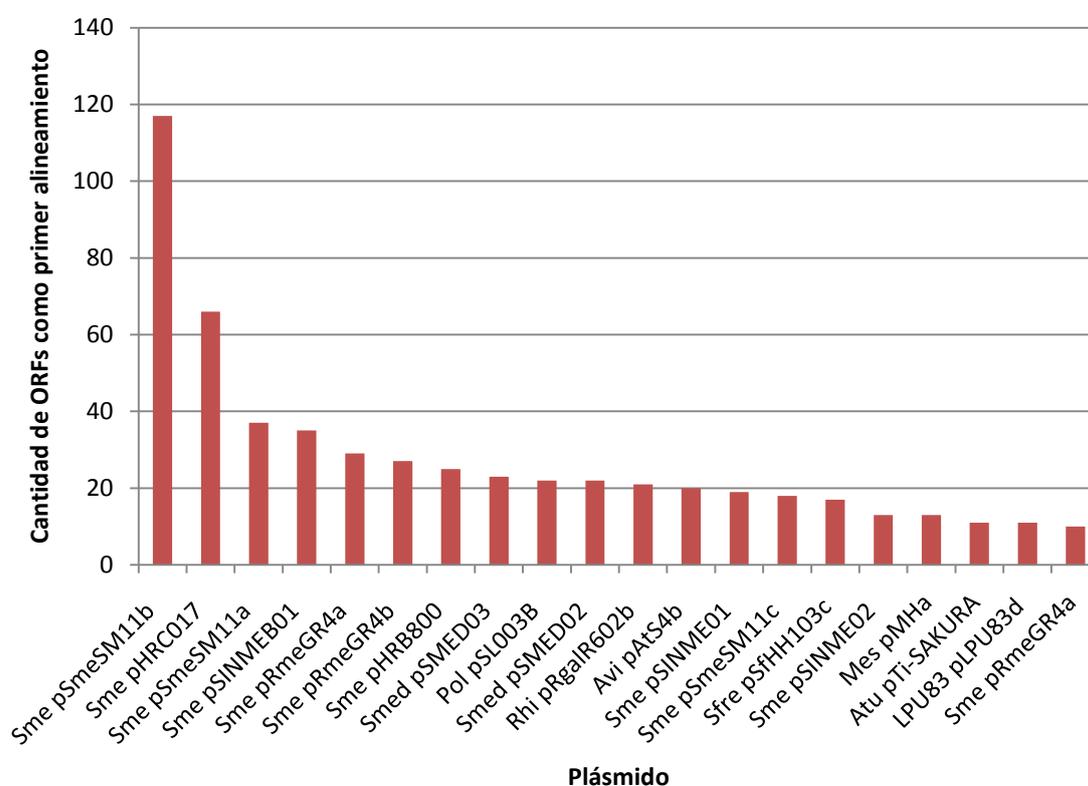


Figura III.14: Cantidad de ORFs que presentaron como primer alineamiento un plásmido depositado en GenBank mediante alineamientos por tBLASTn (Expect 10^{-5}), a partir de cada secuencia aminoacídica de ORFs. Se representan sólo aquellos plásmidos con los que alinearon más de 10 ORFs. Se indica la especie que contiene a cada plásmido como: Sme, *Sinorhizobium meliloti*; Smed, *Sinorhizobium medicae*; Pol, *Polymorphum gilvum*; Rhi, *Rhizobium gallicum*; Avi, *Agrobacterium vitis* S4; Sfre, *Sinorhizobium fredii*; Mes, *Mesorhizobium huakuii* 7653R; Atu, *Agrobacterium tumefaciens*; LPU83, *Rhizobium* sp. LPU83.

Por otra parte, los ORFs que no mostraron como primer alineamiento algún plásmido, sí mostraron similitud con distintos cromosomas como los de algunas cepas de *S. meliloti* y otras especies noduladoras como *Mesorhizobium loti*, *S. fredii*, *S. medicae*, *Mesorhizobium australicum*, *Ensifer adhaerens* y otras bacterias poco estudiadas como *Pannonibacter*, aislada de caña de azúcar [281]. Es interesante estudiar estos genes como posibles candidatos de transferencias génicas entre cromosomas y plásmidos crípticos de *S. meliloti* y otras especies.

Un caso interesante puede ser el los genes con similitud a la isla simbiótica de *Mesorhizobium loti*, cuya presencia en los plásmidos puede estar relacionada a una mejora en la interacción rizobio-leguminosa.

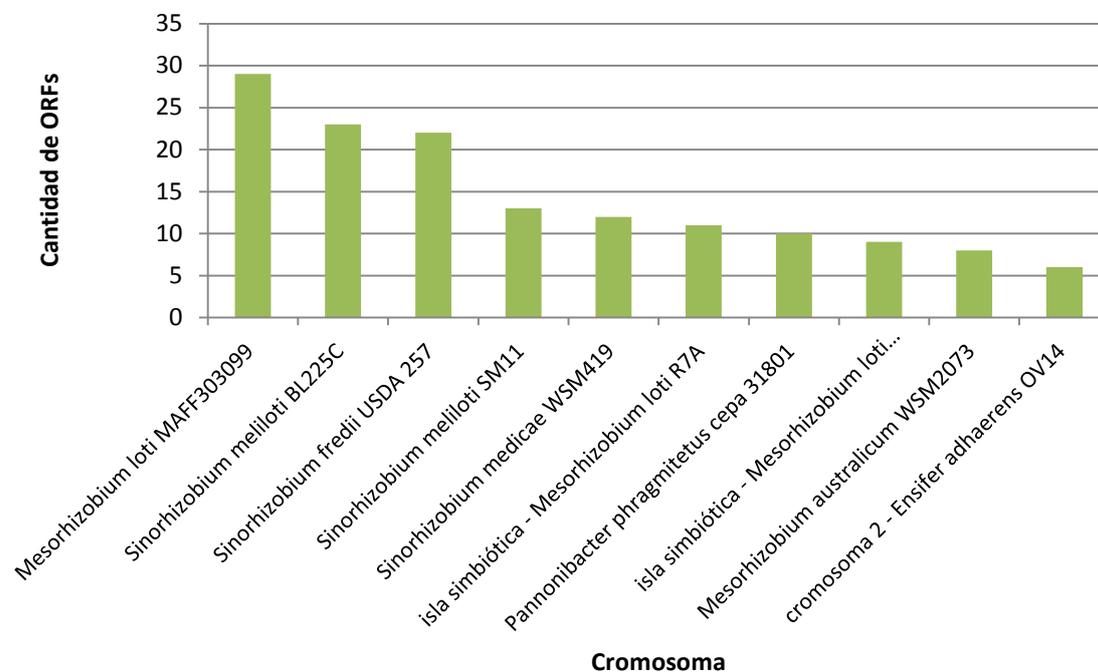


Figura III.15: Cantidad de ORFs que presentaron como primer alineamiento una secuencia cromosomal depositada en GenBank mediante alineamientos por tBLASTn (Expect 10^{-5}), a partir de cada secuencia aminoacídica de ORF. Se representan los 10 cromosomas con mayor número de ORFs alineados.

III.4. Análisis de los genes presentes en el MPSme

Distribución por categorías funcionales

Un análisis inicial de los ORFs presentes en todos los cóntigos pertenecientes al plasmidoma de *S. meliloti* (filtrados según se indicó en la primera sección de este capítulo) mostró en primer lugar, como era esperable, que muchos de ellos incluían genes asociados a funciones frecuentemente halladas en replicones plasmídicos tales como genes de replicación, estabilización, movilización y transposición entre otros. Según ya hemos mencionado, mediante la plataforma GenDB [265], se identificaron y anotaron 2.174 ORFs en el conjunto de cóntigos plasmídicos. Estos ORFs fueron manualmente curados, anotados y clasificados de acuerdo a su función en distintas categorías, tal como se muestra en la Figura III.16, ver Materiales y métodos, sección II.18. En la misma, puede observarse que el grupo de funciones mayoritarias está representado por proteínas hipotéticas conservadas (H) (31%). En esta categoría incluimos los ORFs que mostraron al menos un homólogo en otro genoma y que son anotadas automáticamente ya sea como "hypothetical protein", "conserved protein", "putative membrane protein" o "putative secreted protein". Tal grupo de ORFs no posee homólogos asignables bioinformáticamente que presente alguna función bioquímica conocida, por lo que es un grupo de genes de especial importancia a la hora de identificar nuevas

funciones, ya que puede tratarse de genes de amplia distribución filogenética sin evidencias experimentales aún respecto de su función. En términos de abundancia siguieron los grupos funcionales de: metabolismo (21%), elementos genéticos móviles (19%), el de genes hipotéticos “ORFans” (10%)(sin homólogos conservados detectados bioinformáticamente y de función desconocida), el de ORFs asociados a actividades de transporte (6%), ORFs con productos asociados a regulación transcripcional, los sistemas de dos componentes e histidina quinasas (5%), los ORFs de replicación, partición y segregación (4%) y finalmente los ORFs de metabolismo del ADN (4%).

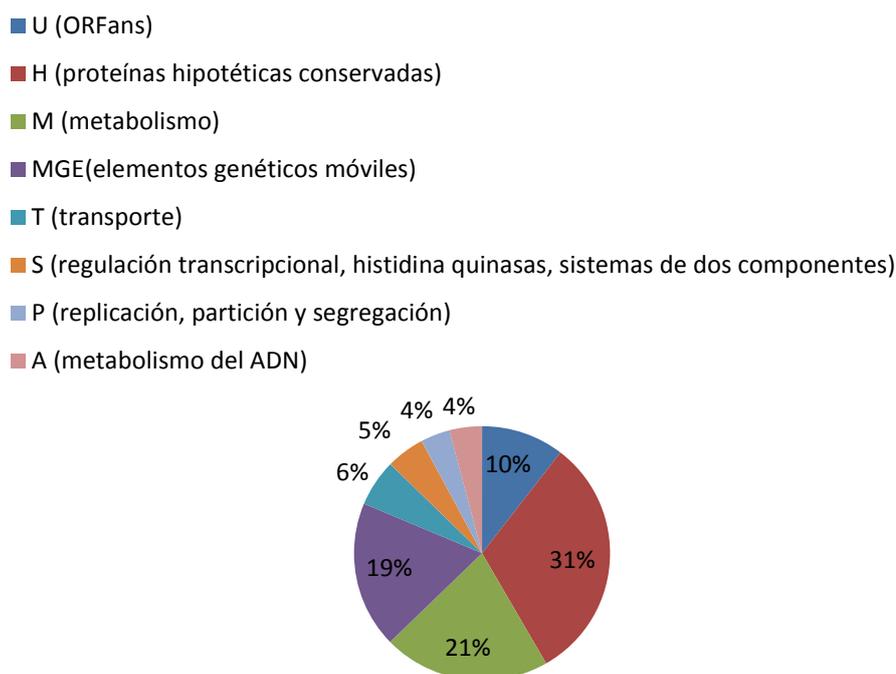


Figura III.16. Composición funcional de los ORFs predichos en el MPSme. El análisis de los cóntigos del MPSme analizados con la plataforma GenDB [265] permitió reconocer, curar, y anotar 2.174 ORFs que pudieron ser clasificados en los grupos funcionales que se muestran en la figura.

Elementos genéticos móviles (MGE)

Esta categoría incluye a todos los ORFs relacionados a la transposición (7,7%), conjugación (7,8%) y metabolismo de fagos (2,7%). Estos ORFs son en su mayor parte los responsables de los mecanismos involucrados en la transferencia horizontal de genes (HGT) y se encuentran típicamente en plásmidos. Es bien conocido que la habilidad de explotar un nuevo ambiente por parte de las bacterias para enfrentar nuevas presiones de selección y adaptarse a los diferentes nichos ecológicos depende principalmente en la adquisición de nuevos genes a través de mecanismos de transferencia horizontal de genes [282-284], por lo que estos mecanismos de HGT en conjunto con los procesos de transposición y recombinación son los principales mecanismos que operan en la evolución procariota [205, 285]).

Genes relacionados a la conjugación

Los plásmidos frecuentemente codifican módulos de genes asociados a su movilidad por conjugación, ya sea de manera autotransmisible, o a través de su movilización por la presencia de funciones de ayuda (“helper”) desde otro plásmido que provee los genes de formación del pili conjugativo (Mpf) [286]. En nuestro grupo hemos caracterizado previamente las propiedades de movilización conjugativa de la colección de aislamientos de *S. meliloti* cuyo moviloma hemos secuenciado, habiendo identificado 3 aislamientos autotransmisibles [222]. Adicionalmente, se hallaron 5 aislamientos que podían re-transferir el plásmido críptico movilizable previamente caracterizado pSmeLPU88b a una tercer cepa receptora, sugiriendo la presencia de funciones “helper” Mpf en estos aislamientos. En el presente secuenciamiento del plasmidoma de esta misma colección, hemos detectado la presencia de 138 ORFs asociados al proceso de conjugación, los cuales se detallan en la Tabla III.2. De acuerdo a la clasificación inicialmente propuesta por Ding y colaboradores [287], y luego ampliada [288], encontramos genes conjugativos de los tipos I y algunos que podrían pertenecer a sistemas del tipo IV, según se describe a continuación. Dentro de módulos de conjugación del tipo I, encontramos 4 regiones completas Dtr y 3 regiones completas Mpf en contigios de diferente largo. Los sistemas de tipo I completos en nuestros contigios constaron de los genes *traG*, *traD*, *traC*, *traA*, *traF*, *traB* y *traH*, con un arreglo similar al descrito previamente [287] para la región Dtr. Además, dos de los cuatro sistemas completos Dtr de tipo I se hallaron asociados a los ORFs codificantes para los genes regulatorios *traR* y *traM*, mientras que los dos restantes no presentaron estas regiones asociadas al mismo contigio (sino que presentaron en su lugar otros genes no relacionados). El resto de estos ORFs se encontraron en su mayoría parcialmente unidos o bien aislados en contigios más pequeños. Esto se debió a la falta de lecturas de mayor largo para completar el ensamblado de contigios más extensos dentro de las regiones de conjugación, no pudiendo determinar por esto de qué sistema conjugativo forman parte. En forma adicional se hallaron 3 regiones Mpf completas [287], ensambladas en contigios. Estas regiones presentaron los ORFs codificantes de los genes *trbI*, *trbH*, *trbG*, *trbF*, *trbL*, *trbK*, *trbJ*, *trbE*, *trbD*, *trbC*, *trbB*, *tral* y en dos de los casos los genes regulatorios *traR* y *traM*, que ocasionalmente se encuentran asociados a las regiones Mpf. En estos últimos dos casos los genes *traR* y *traM* se encuentran seguidos del operón de replicación *repABC*. Cabe destacar que en ningún caso se observó una configuración sinténica diferente a la descrita por Ding y colaboradores [287] en los contigios donde se hallaron ORFs conjugativos parcialmente unidos en sistemas de tipo I.

En total pudimos identificar 12 ORFs codificantes para posibles relaxasas de Dtr, pero el número de los ORFs codificantes de los genes restantes de los distintos sistemas de conjugación es variable (Tabla III.2), aunque nos permite suponer que, si hubiese como es frecuente una única relaxasa/plásmido, el número de Mpfs de tipo I o II puede estar presente en hasta 6 plásmidos, debido al número máximo de ORFs de referencia identificados para cada uno de ellos. Por otro lado se encontraron genes *mobC* y *mobZ* que permiten suponer la presencia de sistemas de conjugación del tipo IV. Nuestro grupo ya había reportado por primera vez las relaxasas de regiones Dtr tipo IV (compuesto de los genes *mobZ*, *mobC* y *parA*) y la presencia de las mismas dentro del moviloma que ahora hemos secuenciado [288]. Más allá de la pertenencia de las 12 relaxasas a los Dtrs mencionados, las mismas mostraron diferentes grados de similitud con otras de plásmidos crípticos de las siguientes especies:

Agrobacterium tumefaciens, *Nitrobacter hamburguensis*, *Sinorhizobium fredii*, *Agrobacterium rhizogenes*, *Sinorhizobium fredii*, *Sinorhizobium meliloti* y *Agrobacterium vitis*.

Tabla III.2: Genes identificados en el MPSme asociados a diferentes funciones de conjugación plasmídica.

Módulo Mpf	83 (genes)	Módulo Dtr	47 (genes)
<i>trbI</i>	6	<i>traG</i>	9
<i>trbH</i>	5	<i>traD</i>	4
<i>trbG</i>	5	<i>traC</i>	5
<i>trbF</i>	4	<i>traA</i>	12
<i>trbL</i>	5	<i>traF</i>	4
<i>trbK</i>	3	<i>traB</i>	5
<i>trbJ</i>	5	<i>traH</i>	4
<i>trbE</i>	6	<i>traR</i>	2
<i>trbD</i>	4	<i>traM</i>	2
<i>trbC</i>	6		
<i>trbB</i>	5	ORFs del sistema tipo IV	8
<i>virB1</i>	1		
<i>virB2</i>	2	<i>mobC</i>	4
<i>virB3</i>	2	<i>mobZ</i>	4
<i>virB4</i>	3		
<i>virB5</i>	4		
<i>virB6</i>	6		
<i>virB8</i>	3		
<i>tral</i>	3		
<i>traR</i>	2		
<i>traM</i>	3		

Proteínas de replicación

Los plásmidos de rizobios muestran un alto grado de estabilidad que asegura su propagación a las sucesivas generaciones a pesar de estar presentes en un número bajo de copias por célula. Tal característica es una propiedad derivada de los mecanismos de replicación y partición que frecuentemente se hallan en estos replicones. Los genes de replicación que están presentes en los plásmidos crípticos de rizobios son en su mayoría miembros de la familia *repABC* [244], la cual es además el sistema de replicación más frecuentemente observado en la clase Alphaproteobacteria. Dicha región de replicación incluye el gen *repC* (que codifica para una replicasa, proteína que participa en el inicio de la replicación) y los genes *repA* y *repB* (que codifican proteínas asociadas a la correcta segregación plasmídica). En nuestro plasmidoma hemos hallado 14 ORFs con similitud de secuencia a genes *repA*, 13 ORFs similares a *repB* y 14 ORFs similares a *repC*. Estos últimos con alta similitud de secuencia a genes *repC* presentes en los plásmidos pSmeSM11b de *Sinorhizobium meliloti*, pSFHH103 de *Sinorhizobium fredii*, pAtS4c de *Agrobacterium vitis* y pRL7 y pRL8 de *Rhizobium leguminosarum*. De entre los cóntigos que presentan estos genes, 9 de ellos contienen los operones *repABC* completos. La ausencia de otras proteínas de replicación en la muestra corrobora que el sistema de replicación *repABC* es el predominante en los plásmidos crípticos de *Sinorhizobium meliloti*. De entre los 14 módulos *repABC* hallados, 4 de ellos están ubicados de forma lindante a los módulos conjugativos Mpf.

Genes metabólicos presentes en el moviloma de Sme

Al analizar los 260 genes que clasificamos como potencialmente asociados a funciones metabólicas observamos, según mostramos en la Figura III.17, que la mayor parte de los ORFs codifican enzimas que no pueden ubicarse en ninguna ruta metabólica en particular (44%); ya sea porque se conoce sólo alguna predicción enzimática general del ORF (tal como podría ser “hidrolasa”, “fosfatasa”, “alquil-transferasa”), o bien porque la enzima asociada podría participar en distintas rutas metabólicas.

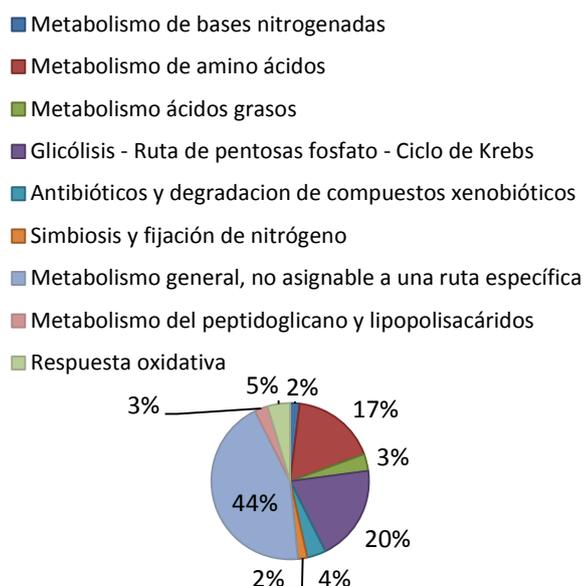


Figura III.17. Distribución de funciones predichas asociadas a los ORFs vinculados a funciones metabólicas en el MPSme.

El siguiente grupo de ORFs metabólicos en abundancia está representado por enzimas del metabolismo central (20%), ya sean de utilización de distintas fuentes de carbono, glicólisis, ruta de las pentosas fosfato, ciclo de Krebs y fosforilación oxidativa. Este grupo es interesante en cuanto a la hipótesis de la presencia de genes de utilización de azúcares como xilosa, manosa, o del sistema de transporte de azúcares (fosfotransferasa) como parte de una mejora en el aprovechamiento de distintas fuentes de carbono y la presencia de reacciones intermediarias en las rutas del metabolismo central que podrían dar indicios de un posible rol de los plásmidos crípticos en la modulación metabólica celular fina o mejora del rendimiento (“*fitness*”) celular en condiciones de vida libre, respuesta a estreses o simbiosis.

El grupo de ORFs del metabolismo de aminoácidos (17%) comprende enzimas involucradas en la biosíntesis y degradación de distintos aminoácidos y, si bien no fue posible recrear ninguna ruta completa a partir de los genes presentes en el moviloma, deberá considerarse la posibilidad que los mismos mejoren la eficiencia en el metabolismo de alguno de los aminoácidos, y/o provean conexiones metabólicas que pueden ser benéficas para la competencia de las bacterias frente a otras que no posean dichos genes.

Entre los genes llamados de “respuesta oxidativa” (5%) encontramos enzimas como alquil-hidroperoxido-reductasas, que han sido descritas como enzimas que, en mutantes de

catalasa que no mostraban daño oxidativo durante el crecimiento aeróbico, bioconvierten el peróxido de hidrógeno [290]. También en relación al estrés oxidativo otras enzimas halladas comprenden a la familia de las tioredoxinas y tioredoxinas reductasas responsables de proteger a la célula de las especies reactivas de oxígeno [291]; y glutatión transferasas responsables de la detoxificación de distintos compuestos xenobióticos [292]. Otra familia de enzimas halladas en el moviloma fueron metionina sulfóxido reductasas, reponsables de la transformación de residuos de sulfóxidos de metionina a residuos de metionina en proteínas expuestas a daño oxidativo [293]. Se encontraron además genes de proteasas Lon que tienen relevancia en la degradación de proteínas desnaturalizadas bajo situaciones de estrés celular [294].

Entre los “antibióticos y degradación de compuestos xenobióticos” (4%) ubicamos genes involucrados en la biosíntesis de penicilinas, poliquétido sintetasas y gramidicina. Por otro lado también hallamos enzimas de resistencia a compuestos lactámicos, bleomicina, microcistina y virginiamicina. Esto ha puesto en evidencia la presencia de parte de un arsenal adaptativo de producción y degradación de compuestos con actividad antimicrobiana que forman parte de la química de la competencia y adaptación a distintos nichos, y que son comúnmente encontrados en bacterias de suelo, rizosféricas y endófitas en plantas [295].

Dentro los genes pertenecientes al metabolismo de ácidos grasos, bases nitrogenadas, peptidoglicano y lipopolisacáridos, hallamos unas pocas enzimas involucradas en la biosíntesis y degradación de los mismos, pero no encontramos rutas metabólicas completas.

Por último, en el pequeño grupo de genes relacionados al metabolismo del nitrógeno hemos podido detectar la presencia de los genes *bacA*, *nodQ*, y otros relacionados a la producción de exopolisacáridos (tales como *exoB* y otros de síntesis de succinoglicano). Dichos genes tienen relevancia en el establecimiento del proceso de fijación biológica de nitrógeno [296] y por ello son particularmente clasificados, de cara a evaluar un posible rol en la simbiosis fijadora de nitrógeno. Desconocemos cómo dichos genes se han integrado al moviloma, y el significado de su presencia en los plásmidos crípticos.

Funciones de transporte

Dentro de las funciones relacionadas al transporte (6%), encontramos genes codificantes de proteínas transportadoras de la familia ABC, con especificidad para transporte de distintos compuestos tales como hierro, cisteína, oligopeptidos, azúcares como maltosa, galactosa, rafinosa, manitol, alfa-glucósido, celobiosa, quitobiosa, arabinooligosacárido y manosa. En ninguno de estos casos se encuentran, sin embargo, los 3 ORFs que darían como resultado el transportador completo. También se encontraron genes de transportadores de metales tales como *czcA* (cobalto, zinc y cadmio). Encontramos asimismo algunos ejemplos de proteínas pertenecientes a los sistemas de secreción de tipo III (genes *yscR*, *yscT*, *yscU*, *yscN* y *yscQ*) descritos en el eflujo de efectores proteicos relacionados a patogénesis. En el caso de los sistemas de tipo IV relacionados a la conjugación, estos fueron descritos anteriormente (ORFs del sistema Mpf).

En este capítulo hemos comenzado a explorar algunas ideas que nos representan a los plásmidos crípticos como una fuente de información génica pasada y reciente que modificó la información codificada en los otros replicones de los sinorizobios. En esta relación informacional, los cromosomas, pSymAs y pSymBs parecen haber intercambiado genes ancestrales con los plásmidos crípticos para moldear sus estructuras informacionales. En cuanto a los genes recientes, los resultados nos muestran por un lado que la mayoría de estas transferencias son desde los plásmidos crípticos hacia los demás replicones y que el tipo de funciones que persisten en estos últimos sugieren un vínculo más estrecho entre los pSymAs y los plásmidos crípticos, que será interesante explorar en mayor detalle. Estos análisis también nos permitieron saber que existen relaciones de ancestralidad o de transferencia de genes con plásmidos de otras especies que también forman parte del microbioma de diferentes plantas. Para continuar con los posibles alcances que esta transferencia de información génica puede tener entre diferentes bacterias y *S. meliloti*, en el siguiente capítulo abordamos la caracterización de la microbiota que cohabita la planta de alfalfa, de manera de acceder a la diversidad microbiana con la que *S. meliloti* podría interactuar así como al compartimiento plasmídico que dichas especies pudieran contener.

CAPITULO IV

Caracterización del microbioma de la semilla de alfalfa

IV.1. Introducción

Según ya hemos observado en nuestro laboratorio [1], y hemos visto en el capítulo precedente; los plásmidos de *S. meliloti* son parte de un compartimento genético muy móvil, con límites de intercambio fuera de la especie que no han sido aún caracterizados. La disponibilidad de más de 1,4 Mb de secuencia de ADN de plásmidos crípticos (no-pSym) de *S. meliloti*, nos permite abordar ahora la búsqueda de genes homólogos en plásmidos de bacterias de otros géneros que se asocien a las plantas de alfalfa en simpatria con los propios rizobios, y explorar así el grado de ubicuidad de los genes del moviloma. Como puede observarse en la literatura, las plantas son profusamente colonizadas en su superficie (filósfera, rizósfera [16, 23]) y también en su interior (endosfera [86]), razón por la cual el estudio de los microbiomas resulta esencial para la comprensión de las bases fisiológicas y evolución de los sistemas planta-bacterias. Si bien la composición de los microbiomas (géneros/especies microbianas) asociados a diferentes plantas (y sus partes específicas) han sido particularmente estudiados en los últimos años, es aún incipiente la información sobre la ecología microbiana sobre/dentro del huésped, sobre sus cambios e interacciones, y sobre flujo horizontal de información como motor de diversidad y evolución. Como parte de la adaptación a diversos nichos presentes en la planta, se ha observado que la composición de esta microbiota es muy diversa y que su metabolismo forma parte de las respuestas que el conjunto planta-microbioma presenta frente a las condiciones ambientales a las que se ve expuesta. Atento a esta consideración, es de esperar que el moviloma plasmídico circulante entre los miembros de este microbioma contenga información asociada a respuestas adaptativas que dichas bacterias utilizan de manera transiente, y que muy probablemente comparten a partir de sus elementos móviles. El estudio de los plásmidos a nivel genómico permite explorar qué tipo de información se encuentra codificada en ellos, qué respuestas biológicas pueden estar ligadas a esta fuente de información móvil, y finalmente cuáles son los límites de la transferencia de esta información entre las especies que una planta y su entorno alojan (en sus distintos nichos).

En el contexto precedente del rol de los microbiomas en la fisiología de las plantas, de su composición cambiante, y del rol del flujo génico en la evolución natural de las comunidades bacterianas allí alojadas, en el presente capítulo hemos abordado el aislamiento e identificación de bacterias naturalmente asociadas a alfalfa con el propósito de investigar sus contenidos plasmídicos, el conjunto de genes presentes y las posibles funciones asociadas, así como el grado de vínculo (ortología) con los genes del moviloma ya caracterizada del rizobio simbiótico simpátrico, *S. meliloti*.

IV.2. Primeros ensayos. Semillas de *M. sativa* esterilizadas en superficie y germinadas en medio hidropónico estéril. Evidencias de la presencia de bacterias protegidas de la esterilización

En un primer acercamiento nos propusimos la obtención de bacterias asociadas a plantas crecidas en el laboratorio, en tierra proveniente de regiones alfareras del país de modo de acceder a parte de la microbiota que coloniza la planta naturalmente en condiciones de campo. Sin embargo, durante los ensayos iniciales de control utilizando sólo semillas

esterilizadas en superficie (con ácido sulfúrico o con hipoclorito de sodio, ver Materiales y métodos, sección II.3) y crecidas en medio para plantas con vermiculita y en esterilidad, observamos que las plantas estaban siempre colonizadas con bacterias, tanto en sus raíces como en la parte aérea. La observación reiterada en sucesivos ensayos nos llevó a pensar que la carga bacteriana podría provenir de las propias semillas, de una zona protegida del efecto de los agentes químicos enérgicos que utilizamos durante la esterilización superficial. En la sección siguiente describiré la estrategia utilizada para el aislamiento y caracterización de las bacterias asociadas a las semillas y resistentes a la esterilización superficial.

IV.3. Aislamientos provenientes de semillas de alfalfa variedad Super Monarca.

Tipificación

Para explorar los microorganismos provenientes de las semillas de alfalfa, las mismas fueron esterilizadas superficialmente del mismo modo que en los primeros ensayos control y sus exudados colectados luego de 4 horas de permanencia en medio PBS fueron plaqueados en distintos medios nutritivos, resultando en la sucesiva obtención y colección de aislamientos bacterianos que visualmente presentaron morfologías muy variadas (Figura IV.1). La repetición de los ensayos permitió coleccionar más de 60 aislamientos, encontrándose títulos del orden de 10^2 - 10^3 u.f.c./g de semilla.

La identificación de los aislamientos por las técnicas de MALDI-TOF (ver Materiales y métodos, sección II.9) y el secuenciamiento parcial del gen de la subunidad 16S ribosomal (ARNr 16S) nos permitió agruparlos y clasificarlos a nivel de género. En la Tabla IV.1 se muestran los 28 géneros obtenidos a partir de semillas de alfalfa en la variedad Super Monarca ("S"), que fue utilizada como la variedad de referencia en este trabajo de tesis.

Si bien el gen ARNr 16S resulta ser el marcador más utilizado para la identificación en bacterias, su utilización requiere su amplificación por PCR y la posterior purificación y secuenciamiento del amplicón para cada aislamiento obtenido. Aunque este método fue el de elección al momento de identificar los géneros aislados que presentaban diferencias morfológicas claras, no nos resultó adecuado a la hora de realizar múltiples ensayos cuantitativos sobre la totalidad de la diversidad bacteriana cultivable, siendo un método relativamente lento y costoso para la tipificación bacteriana en comparación al MALDI-TOF, el cual permite un tratamiento rápido de los aislamientos cultivables por el cual pudimos, una vez consolidada una biblioteca interna de los aislamientos provenientes de semillas y plantas, comenzar a explorar si la diversidad observada proveniente de las semillas se mantenía constante en sucesivas repeticiones de los ensayos.

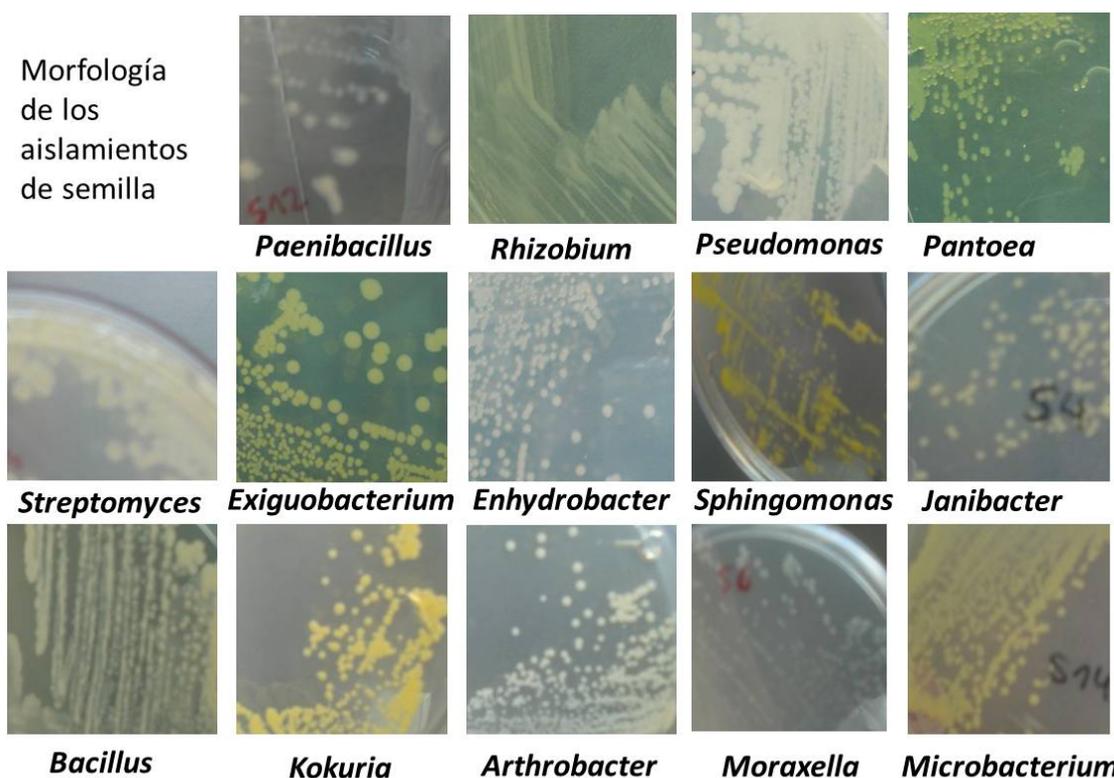


Figura IV.1. Genotipificación y morfología de algunos géneros aislados a partir de exudados de semilla de *Medicago sativa*, variedad Super Monarca esterilizadas superficialmente.

Con el objetivo de comprender si esta diversidad asociada a las semillas era reproducible en los distintos ensayos para luego tratar de establecer los límites de esta diversidad dentro de una misma variedad de semilla, en nuestro laboratorio hemos comenzado a realizar ensayos tendientes a establecer los límites de la diversidad bacteriana obtenida, la cual muestra una gran variación de géneros y de cantidades relativas de aislamientos para los diferentes géneros en cada ensayo, como se muestra en los 4 ensayos presentados en la Figura IV.2, donde cada ensayo independiente presenta diferencias tanto en abundancia como en presencia de los diferentes géneros identificados. Será importante avanzar en la cobertura de la diversidad observada de cara a establecer los límites de estas variaciones y la completa caracterización del microbioma acompañante de semilla.

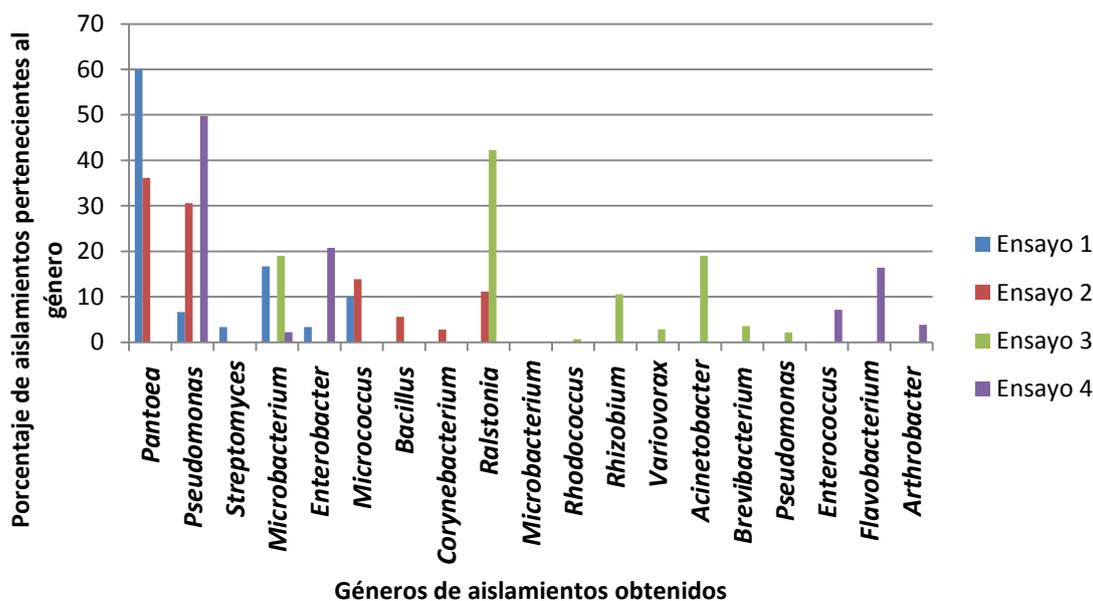


Figura IV.2: Ensayos cuantitativos de la diversidad bacteriana presente en las semillas de alfalfa, variedad S. Se indica la cantidad de aislamientos pertenecientes a cada género particular, identificados por EM-MALDI-TOF en 4 ensayos independientes.

IV.4. Aislamientos provenientes de semillas de otras variedades de alfalfa

Para evaluar si la presencia de bacterias asociadas a semillas y tolerantes a la esterilización superficial era un fenómeno propio de la variedad S de *Medicago sativa*, el mismo tipo de análisis fue extendido a otras variedades de semillas de alfalfa. También en la Tabla IV.1 pueden observarse los géneros bacterianos identificados en las 4 nuevas variedades de alfalfa ensayadas (D, G, B y F, ver Materiales y métodos sección II.3, tabla A8 del Anexo). Los hallazgos mostraron géneros previamente identificados en la variedad S, apoyando la hipótesis de una comunidad bacteriana conservada entre las distintas variedades de semillas, tal como se observó para semillas del género *Zea* [54]. En nuestro caso, los géneros *Bacillus*, *Kokuria* y *Arthrobacter* se encontraron presentes en la mayoría de las variedades estudiadas. Deberá analizarse la extensión de la ubicuidad observada y si dicho fenómeno es producto de una coevolución semilla-microbioma donde han existido fenómenos de selección.

Por otro lado, hallamos géneros asociados sólo a semillas de otras variedades y ausentes en la variedad S, tal como es el caso de los géneros *Curtobacterium*, *Planococcus*, *Frigobacterium*, *Roseomonas*, *Staphylococcus* y *Terribacillus*, *Sporosarcina*, *Sphingobium*, *Agrococcus*, *Azorcus*, *Brevibacillus* y *Solibacillus*. Esta observación plantea la posibilidad de que, además de la presencia de géneros conservados/frecuentes a la que ya hemos hecho referencia, existan dentro de las semillas bacterias menos ubicuas, como vía de evaluación y eventual incorporación de nuevos miembros con actividades diferentes a la comunidad del microbioma. No debe de todos modos descartarse la posibilidad alternativa que las semillas estén mucho más colonizadas que lo que se evidencia por las bacterias liberadas en 4 h de ensayo (pueden contener bacterias que no se liberan), estando como alternativa todas la variedades colonizadas por una diversidad similar de géneros bacterianos.

Tabla IV.1. Géneros bacterianos aislados de semillas de seis variedades distintas de alfalfa (S,D,G,B,F,C).

<u>Géneros</u>	<u>Variedades</u>	<u>Géneros</u>	<u>Variedades</u>
<i>Bacillus</i>	S,D,G,B,F	<i>Moraxella</i>	S
<i>Arthrobacter</i>	S,D,G,B,C	<i>Sphingomonas</i>	S,G
<i>Microbacterium</i>	S,G,B	<i>Acinetobacter</i>	S
<i>Kocuria</i>	S,B,C,F,G	<i>Pantoea</i>	S,G,B
<i>Janibacter</i>	S,G	<i>Staphylococcus</i>	S, D
<i>Paenibacillus</i>	S,B,G	<i>Planococcus</i>	D,G
<i>Brevundimonas</i>	S	<i>Curtobacterium</i>	G
<i>Terracoccus</i>	S	<i>Roseomonas</i>	G
<i>Luteococcus</i>	S	<i>Frigoribacterium</i>	B
<i>Streptomyces</i>	S	<i>Terribacillus</i>	B
<i>Enhydrobacter</i>	S	<i>Exiguobacterium</i>	S, G
<i>Sporosarcina</i>	D,G	<i>Solibacillus</i>	D
<i>Brevibacillus</i>	F	<i>Rhizobium</i>	S
<i>Pseudomonas</i>	S	<i>Flavobacterium</i>	S
<i>Enterococcus</i>	S	<i>Brevibacterium</i>	S
<i>Varioborax</i>	S	<i>Rhodococcus</i>	S
<i>Ralstonia</i>	S	<i>Corynebacterium</i>	S
<i>Micrococcus</i>	S	<i>Enterobacter</i>	S
<i>Agrococcus</i>	G	<i>Sphingobium</i>	G
<i>Azoarcus</i>	G		

Un recorrido por la bibliografía nos indica que la mayoría de estos géneros ya han sido reportados en semillas esterilizadas superficialmente de distintas plantas. *Bacillus*, *Acinetobacter*, *Streptomyces*, *Flavobacterium*, *Corynebacterium*, *Pseudomonas* y *Micrococcus* ya habían sido encontrados en semillas y óvulos de distintas especies de plantas [128]. Representantes del género *Rhizobium* por su parte, fueron hallados previamente en semillas de maíz [52, 54, 297]. Asimismo, mientras bacterias de los géneros *Pantoea*, *Pseudomonas* y *Staphylococcus* fueron aisladas de semillas de variedades del género *Zea* [54], *Brevundimonas* y *Enhydrobacter* fueron aislados específicamente de semillas de *Zea Mays* [297]. Cabe señalar por último que bacterias de los géneros *Kocuria* y *Luteococcus* [298]; *Arthrobacter*, *Paenibacillus*, y *Sphingomonas* [299]; así como *Curtobacterium*, *Frigobacterium*, *Microbacterium* [52] y *Exiguobacterium* pudieron ser todas aisladas de semillas de plantas de arroz. Existe por su parte un reporte en que se informa el aislamiento de *Moraxella* de semillas de *Abutilon theophrasti* [300].

En el caso de *Roseomonas*, si bien hasta la fecha no ha sido reportada estar presente en semillas, se la ha aislado recientemente a partir de suelo rizosférico [301] asociado a cultivos de arroz. En el caso de *Janibacter*, este género se ha aislado recientemente en suelos de pre y

pos-cosecha de alfalfa, mostrando además propiedades de promoción del crecimiento [302]. El suelo rizosférico también ha mostrado albergar al género *Terribacillus* [303].

Es interesante resaltar que tanto los géneros *Terribacillus* como *Planococcus* han sido aislados en suelos salinos o ambientes de fuerte estrés osmótico, característica común a la semilla [304-306]. *Terracoccus* por su parte ha sido aislado en plantas de tomate [307] aunque aún no se había reportado su presencia en semillas. Los resultados que hemos mostrado hasta aquí presentan al menos dos características destacadas: a) la presencia en semillas de alfalfa de bacterias con una marcada diversidad, descritas en su mayoría previamente en sistemas diversos, y b) la localización en las semillas en un compartimento que las hace muy resistentes a tratamientos exhaustivos con agentes químicos deshidratantes u oxidantes. Mostraremos en la sección siguiente los análisis para evaluar la presencia de bacterias dentro de plantas provenientes de semillas esterilizadas en superficie, y el análisis de los aislamientos recuperados.

IV.5. Bacterias asociadas a plántulas crecidas de semillas esterilizadas

Una de las preguntas que surge a la obtención de bacterias asociadas a las semillas es si éstas se encuentran allí de manera transitoria como parte del desarrollo de su ciclo biológico en el suelo o ambiente circundante a las semillas, o bien si además las mismas han adquirido la capacidad de colonizar tempranamente la planta desde la semilla y han perdurado en los estadios posteriores del desarrollo. Para indagar este aspecto, las semillas esterilizadas superficialmente fueron germinadas y luego crecidas en condiciones de asepsia en tubos estériles con vermiculita, y finalmente molidas en mortero con solución mineral y plaqueadas en medio nutritivo LB - Glucosa. La aparición de colonias que observamos fue prueba evidente de que (al menos algunas de) las bacterias contenidas en las semillas colonizan tejidos como tallos, raíces y hojas. Usando el procedimiento precedente varios aislamientos fueron recuperados de plántulas de distintas variedades de alfalfa, los que fueron luego identificados taxonómicamente y comparados con los aislamientos provenientes de semilla que mostramos en la Tabla IV.1. En la Tabla IV.2 se presenta la lista de géneros que pudieron ser identificados en plantas de alfalfa crecidas a partir de semillas de seis variedades diferentes esterilizadas superficialmente.

Este ensayo mostró que distintos aislamientos provenientes de las semillas tienen la capacidad de colonizar la plántula en crecimiento, poniendo en evidencia que las semillas constituyen un reservorio natural de microbios capaces de colonizar y establecerse en los estadios posteriores del desarrollo de la planta. En la Tabla IV.2 se observa que 4 de los 21 géneros obtenidos en plántulas de las variedades examinadas no se encontraron en nuestros ensayos colonizando las semillas de alguna de las distintas variedades. De todos modos, es importante observar que, atento al modo en que hemos realizado los ensayos (semillas esterilizadas en superficie y creciendo en ambiente estéril), todos los aislamientos de las plántulas deben necesariamente provenir de la semilla. Esto significa que los aislamientos de dichos géneros no pudieron ser obtenidos cultivando los exudados, probablemente debido a: 1) que bacterias de los mismos no se liberan al medio durante las 4 horas en que son incubadas las semillas, o 2) dado que los experimentos a partir de los cuales se obtuvieron bacterias desde las semillas fueron independientes de los que se utilizaron para obtener aislamientos de las plántulas, y que el

número de aislamientos obtenidos a partir de las semillas es bajo, es posible que el número de ensayos y la cantidad de aislamientos obtenidos de semillas no haya sido la suficiente para cubrir toda la diversidad presente en las mismas, es decir, que la cobertura de la diversidad presente en las semillas haya sido baja. Algunos de estos géneros encontrados sólo en plántulas y no en semillas fueron previamente aislados a partir de semillas de distintas plantas como en el caso de *Stenotrophomonas* [299] y *Novosphingobium* [308].

Tabla IV.2. Géneros bacterianos correspondientes a los aislamientos recuperados de plántulas de alfalfa de 6 variedades diferentes (S,D,B,C,G,F) crecidas a partir de semillas esterilizadas superficialmente según se indica en Materiales y métodos, sección II.4. En la tabla se indica además los casos en bacterias de un mismo género que habían sido ya obtenidas de exudados de semillas colectados a las 4 h posteriores a la esterilización (ver sección IV.2 a IV.4 en este capítulo).

<u>Géneros aislados de plántulas</u>	<u>Variedades</u>	<u>Aislados en semillas</u>
<i>Microbacterium</i>	S,D,B,C	Sí
<i>Pantoea</i>	S,G,C,D	Sí
<i>Stenotrophomonas sp</i>	S,G,F,C	No
<i>Acinetobacter</i>	D,G,F	Sí
<i>Arthrobacter</i>	S,C,F,G	Sí
<i>Paenibacillus</i>	S,C,B	Sí
<i>Kocuria</i>	F,C	Sí
<i>Rhizobium</i>	S,C	Sí
<i>Exiguobacterium</i>	S	Sí
<i>Corynebacterium sp</i>	S	Sí
<i>Flavobacterium sp</i>	S	Sí
<i>Sphingobium</i>	S	Sí
<i>Moraxella</i>	S	Sí
<i>Novosphingobium</i>	S	No
<i>Pseudomonas</i>	S,B	Sí
<i>Staphylococcus</i>	B,C	Sí
<i>Oerskovia</i>	B	No
<i>Micrococcus</i>	D	Sí
<i>Brevibacillus</i>	C	Sí
<i>Bacillus</i>	S	Sí
<i>Sphingobacterium</i>	B	No

En contraste con las observaciones anteriores, aislamientos de 17 de los 21 géneros (81 %) recuperados de las plántulas sí mostraron estar presentes en los exudados de semillas. A su vez, los aislamientos de estos géneros representan sólo a 17 de los 39 (44%) géneros asociados a las semillas de las distintas variedades. Este constituye un subconjunto de géneros que pudieron colonizar la plántula de manera detectable, indicando que posiblemente sólo una fracción de la diversidad presente en las semillas sea capaz de colonizar la plántula. La hipótesis de una transmisión vertical de las bacterias presentes en las semillas está apoyada por los trabajos publicados en semillas del género *Zea* [54], de arroz [52] y de uva [132], y nos conduce a pensar en un rol biológico de larga data (y de coevolución con la planta) para estos microorganismos, en comparación con aquellos colonizadores que acceden a la planta desde el medio exterior en distintas etapas de la vida de la misma (colonización horizontal). Muy posiblemente las bacterias asociadas y protegidas por las semillas han adquirido un mecanismo de perpetuación en el hospedador por medio de la colonización de los tejidos reproductivos de la planta para transmitirse a las sucesivas generaciones. La evaluación de

esta hipótesis es plausible cultivando la planta en condiciones de asepsia a lo largo de varias generaciones, y observando cómo es la propagación de la flora bacteriana en las semillas neoformadas. En este contexto, la diversidad bacteriana asociada a semillas debe considerarse como un bioinoculante consorcial natural, cuyas actividades en relación a la sanidad y al crecimiento de las plantas debe ser explorado en detalle. En un esquema como el precedente puede ocurrir: a) que la microbiota asociada a semillas conste de un núcleo de géneros conservados capaces de colonizar todas las variedades, y de subconjuntos de géneros que se asocien específicamente con variedades particulares o grupos de ellas (efecto de la interacción específica variedad-microorganismo), o bien b) que, en el otro extremo, que no exista ningún sesgo hacia asociaciones de variedades particulares con perfiles específicos de biodiversidad. Será necesario el análisis de un número mayor de variedades explorando la biodiversidad con buen grado de penetración (de semillas y plantas de cada variedad) para evaluar cuál de los escenarios anteriores describen al microbioma de las semillas de *M. sativa*.

Este trabajo representa el primer estudio de diversidad de la microbiota proveniente de las semillas en alfalfa. No obstante, ya existía un trabajo sobre la diversidad presente en nódulos, tallo y hojas de plantas crecidas en condiciones de mesocosmos, por Galardini y colaboradores [232]. En el mismo, se estudió la diversidad al nivel de familia mediante la técnica de T-RFLP. En aquella ocasión se pudieron detectar miembros de las familias Methylobacteriaceae, Rhizobiaceae y Sphingomonadaceae en tallos y hojas de alfalfa. En nuestros ensayos hemos detectado algunos géneros presentes en las mismas familias, entre los aislamientos de plántulas crecidas de semillas estériles. Tal es el caso de los géneros *Roseomonas*, *Rhizobium*, *Sphingomonas*, *Sphingobium* y *Novosphingobium*. Por otra parte, la presencia de *Erwinia herbicola* (*Pantoea*) había sido reportada por Handelsman y colaboradores en alfalfa [190].

IV.6. Actividades asociadas a la promoción del crecimiento vegetal (PCV). Nuestros aislamientos

Más allá de la mayor o menor especificidad en la conformación de los microbiomas respecto de la especie o variedad de semilla que se considere (ver sección anterior), la presencia de bacterias asociadas a semillas debe haber resultado en la selección de microorganismos que, de modo directo o indirecto, medien actividades que sean solidarias con la sanidad y el normal crecimiento de la planta. Atendiendo a esta posibilidad, y a la experiencia de coevolución acumulada en años, los aislamientos que se propagan verticalmente a través de semillas constituyen muy buenos candidatos (y posiblemente bioseguros) de cara a la búsqueda de cepas con potencial uso como biofertilizantes promotores del crecimiento vegetal (PCV). El desarrollo de bioinoculantes como posible reemplazo de los agroquímicos en los cultivos es una tendencia creciente destinada a sostener la producción y disminuir el daño ecológico que los modos de producción de alimentos generan actualmente sobre el planeta [309]. Los microorganismos PCV suelen clasificarse en “biofertilizantes”, cuando aumentan la biodisponibilidad de un nutriente en el suelo o bien potencian la capacidad de una planta a incorporarlo; “fitoestimuladores”, cuando tienen la capacidad de producir cambios en los reguladores del crecimiento vegetal (tales como ácido indol acético, ácido giberélico, citoquinas y etileno) o “biopesticidas/biocontroladores”, cuando tienen la capacidad de inhibir agentes patógenos mediante la producción de bacteriocinas, antibióticos, antifúngicos, sideróforos o bioinsecticidas [310]. Muchos microorganismos han sido estudiados y aplicados

como bioinoculantes durante décadas (sea por mejorar el crecimiento y/o la nutrición vegetal, o como biocontroladores), incluyendo bacterias y hongos. Aunque resulta en una mirada limitada, la forma de evaluar la potencialidad de microorganismos como PCV se realiza a través de pruebas de tamizado (*screening*) en las que se ensayan -con pruebas simples *in vitro*- actividades bioquímicas que frecuentemente están relacionadas con la promoción del crecimiento de plantas y el biocontrol (producción de hormonas vegetales, producción de sideróforos, producción de enzimas específicas, producción de compuestos antibacterianos y compuestos antifúngicos, etc.). Dichas actividades pueden realizarse en el laboratorio y son un primer indicativo de si un aislamiento puede presentar o no capacidades reales como PCV. La estrategia está apoyada en la experiencia de que una buena proporción de los microorganismos que presentan dichas actividades *in vitro* muestran asimismo capacidades PCV en ensayos *in vivo* con plantas [310, 311]. Frecuentemente, un mismo aislamiento puede presentar varias actividades *in vitro* ligadas potencialmente a la PCV, lo cual es ventajoso en el desarrollo de bioinoculantes. Cabe resaltar, de todos modos, que el hecho que una bacteria presente actividades *in vitro* compatibles con la promoción de crecimiento de plantas no significa que lo hará *in planta* de modo que pueda ser catalogado efectivamente como microorganismo PCV. La actividad efectiva de un microorganismo como PCV en suelo depende de factores abióticos (edáficos y climáticos) y bióticos (predadores microbianos, efectos consorciales, compatibilidad con las comunidades microbianas presentes) complejos, que no se reflejan en las pruebas simples de evaluación de caracteres de PCV. Debe tenerse en cuenta, por otro lado, que el limitado conocimiento respecto de mecanismos menos convencionales de promover el crecimiento de plantas (ej. inducción de cambios en la tolerancia a estreses) condiciona fuertemente el uso de las pruebas de tamizado acotándolo a la búsqueda de actividades bioquímicas vinculadas sólo a los efectos de promoción más conocidos. De todos modos, en este contexto y con las consideraciones anteriores, las pruebas de tamizado han sido en muchos casos útiles para a la selección de aislamientos de uso práctico como PCV.

Caracterización de los aislamientos de nuestra colección recuperados de semillas y plántulas de *M. sativa*

Como una primera aproximación para explorar si los aislamientos naturalmente asociados a las semillas de alfalfa son capaces de generar alguna mejora en el crecimiento de la planta se realizaron diferentes ensayos *in vitro* para evaluar distintas actividades asociadas en muchos casos, como hemos dicho, con la PCV.

Una de las actividades que exploramos fue la producción de componentes con actividad de sideróforos como responsables de la quelación del hierro, nutriente escaso en suelos neutros o básicos debido a la baja solubilidad de la forma oxidada (Fe^{+3}) de ese elemento. La planta genera en la rizósfera un ambiente ácido que favorece la solubilización del hierro y la capacidad de producir los sideróforos puede darles ventajas a los microorganismos durante la colonización de la rizósfera [312]. Además, la quelación del hierro por parte de microorganismos benéficos puede evitar que ciertas bacterias y hongos patógenos colonicen la rizósfera [313]. Entre nuestros aislamientos, más de la mitad (55%) mostró un fenotipo productor de sideróforos (ver Tabla IV.3 columna 4). Los géneros con esta capacidad que pudimos identificar están listados en la Tabla IV.4. Varios de los géneros que hemos

encontrado ya han sido estudiados en la producción de sideróforos, determinando estructuras químicas y mecanismos de acción para los mismos. Tal es así en el caso de los géneros *Arthrobacter* [314], *Bacillus* [315], *Corynebacterium* [316], *Pseudomonas* [317], *Rhizobium* [318], *Paenibacillus* [319]. También se ha reportado producción de sideróforos en géneros como *Microbacterium* [320], *Moraxella* [321], *Exiguobacterium* [322], *Pantoea* [323], *Janibacter* [302], *Staphylococcus* [324], *Kokuria* [325], *Terribacillus* [303], *Stenotrophomonas* [326], y *Novosphingobium* [327]. Para el género *Curtobacterium* no se ha reportado producción de sideróforos en la bibliografía.

Otro mecanismo muy estudiado en relación a las capacidades de PCV es la solubilización de fosfatos insolubles por parte de los microorganismos. Se considera que dicho mecanismo provee la mayor fuente de fósforo para las plantas (el fósforo es el segundo nutriente limitante después del nitrógeno), tanto en la naturaleza como en los agroecosistemas. En nuestra colección, 19 % de los aislamientos fue capaz de solubilizar fosfatos como se muestra en la tabla IV.3, columna 3. Los aislamientos que presentaron mejor capacidad de solubilización de fosfatos fueron los pertenecientes al género *Pantoea* (ver tabla IV.4), frecuentemente citado en la bibliografía debido a esta capacidad [328]. Es interesante notar que los aislamientos obtenidos de exudados de semilla fueron más propensos a producir proteasas y celulasas que aquellos que se aislaron de plántulas. Lo contrario sucedió en el caso de la solubilización de fosfatos, atento a que los aislamientos con esa actividad están representados en gran medida por miembros del género *Pantoea* que fueron colectados casi en su totalidad a partir de plántulas y que además representaron el 50% del total de los aislamientos con esta capacidad. Los géneros *Pseudomonas*, *Bacillus*, *Exiguobacterium*, si bien no han mostrado tan buenos resultados, han producido halos de hidrólisis moderados, y ya han sido descritos como solubilizadores de fosfatos [329-331]. Los otros géneros que presentaron dicha capacidad en nuestra colección (*Brevundimonas*, *Microbacterium* y *Staphylococcus*) no se han reportado en la literatura hasta la fecha como solubilizadores de fosfatos.

Un tercer fenotipo evaluado en los aislamientos fue la producción de la hormona vegetal ácido indol acético (AIA). Los cambios fisiológicos relacionados a la producción de esta auxina, la más importante en la planta, están relacionados al alargamiento y división celular, diferenciación y respuestas a la luz y la gravedad [332]. El crecimiento radicular y el desarrollo de raíces laterales que conduce a una mayor área superficial, conllevan a una mejor asimilación de nutrientes por parte de la planta y por ende a un mayor crecimiento. Su producción ha sido tempranamente estudiada en bacterias patógenas como *Pseudomonas savastanoi*, *Agrobacterium tumefaciens* y *Agrobacterium rhizogenes* [333, 334], pero también en bacterias que establecen interacciones beneficiosas con plantas como *Azospirillum sp.* [335], *Pseudomonas*, *Xanthomonas* [336], *Rhizobium* [337] y *Bradyrhizobium* [338]. Según el tipo de interacción, la producción de AIA puede resultar tanto en la fitoestimulación como en la patogénesis para evadir la respuesta defensiva de la planta [339]. Las bacterias productoras de AIA pueden variar respecto de la ruta biosintética que utilizan para la producción de la hormona [340]. Las mismas rutas identificadas en bacterias están presentes en las plantas, pero su utilización es diferente en los dos casos. Cinco de las rutas biosintéticas conocidas utilizan al triptófano como su precursor y son llamadas rutas dependientes de triptófano. Entre estas, la ruta del indol-3-acetamida (IAM) es mayormente utilizada en bacterias y su presencia en plantas ha sido sólo recientemente demostrada [341]. La ruta del indol-3-piruvato se cree que es la mayoritaria en plantas, aunque la completa elucidación de las enzimas involucradas

no ha sido realizada en plantas a la fecha, aunque sí en bacterias. Tanto las rutas de la triptamina así como la del indol-3-acetonitrilo (IAN) han sido detectadas en bacterias como en plantas, mientras que la ruta de la oxidasa de cadena lateral del triptófano (TSO) ha sido detectada sólo en bacterias al día de hoy. La producción de AIA en bacterias está relacionada a la presencia de estreses, donde en particular en bacterias asociadas a plantas se cree que la producción de AIA está relacionada a estreses como el nutricional o de potencial matricial (principalmente en bacterias epífitas), o el de pH (principalmente en bacterias rizosféricas). La presencia de exudados o extractos de plantas ha mostrado también ser un regulador positivo en la producción de AIA como en el caso de *Xanthomonas axonopodis* [342]. En el caso de *Rhodococcus fascians* la producción de AIA fue estimulada, de manera más compleja, sólo por la presencia de extractos de plantas que habían sido infectadas con la propia bacteria [343]. En el caso de *Pantoea agglomerans* se ha visto incrementada la producción de AIA en contacto con la superficie de la planta [344] y en otro estudio sobre el efecto de la superficie de plantas en la producción de AIA se ha observado una heterogeneidad en la expresión de los genes biosintéticos asociada a las diferencias en el microambiente presentes en dichas superficies [345]. En cuanto a las condiciones de laboratorio, el agregado de triptófano (precursor de la mayor parte de las rutas y presente en los exudados en las plantas) se ha mostrado que puede inducir la producción de AIA en distintas bacterias y es utilizado en los ensayos para investigar la capacidad *in vitro* de producción de AIA. Se ha mostrado en *Azospirillum* que, mientras el antranilato (precursor del triptófano) inhibe la producción de AIA, la presencia de triptófano la induce, inhibiendo al mismo tiempo la producción de antranilato [346]. En los aislamientos de nuestra colección, la producción de AIA fue detectada en 33 % de ellos sin la adición de triptófano, y en 45 % de los aislamientos cuando fue agregado triptófano al medio de cultivo. Los aislamientos que pudimos identificar como productores de AIA pertenecen todos a excepción de uno, a géneros con miembros en los que se ha reportado esta característica, y que fueron aislados frecuentemente en asociación con distintas plantas. Tal es el caso de aislamientos de los géneros *Pantoea* [347], *Rhizobium*, *Pseudomonas* [348], *Bacillus* [349], *Microbacterium* [349], *Arthobacter* [350], y *Kokuria* [351]. En el caso del género *Terribacillus*, si bien su presencia ha sido detectada como endófito de *Tridax procumbens* [304], no ha sido a la fecha reportado como productor de AIA.

Si bien en la Tabla IV.3 se presentan las incidencias de cada una de las actividades de interés dentro de la colección de aislamientos, en la Tabla A5 del Anexo se muestra el perfil completo de actividades para cada uno de los aislamientos, de modo que pueda apreciarse que algunos de ellos son capaces de expresar varias características positivas frecuentemente asociadas a la PCV, hecho que los posiciona como los candidatos de preferencia a la hora de escoger candidatos para evaluar efectos *in planta*.

Tabla IV.3. Resultados de los ensayos de evaluación *in vitro* de caracteres asociados a la PCV en aislamientos de semillas y plantas de alfalfa.

	Proteasas	Celulasas	Fosfatos	Sideróforos	AIA LB	AIA LB c/ Trp
Aislamientos analizados totales	115	118	126	110	107	51
Aislamientos positivos totales	41	54	24	60	35	23
% positivos	36	46	19	55	33	45
Analizados de semillas	75	76	88	69	66	17
Analizados de plántula	40	41	38	41	41	34
% de positivos de semillas	41,3	53,9	11,4	56,5	37,9	35,3
% de positivos de plántulas	25	32	37	51	24	50

Tabla IV.4. Lista de los diferentes géneros bacterianos identificados en semillas y/o plántulas de alfalfa indicando el tipo de actividad/es asociada/s a cada uno de ellos. El número de aislamientos con la actividad indicada para cada género particular se muestra entre paréntesis.

Proteasas	Celulasas	Fosfatos	Sideróforos	AIA	AIA cTrip
<i>Exiguobacterium</i> (1)	<i>Microbacterium</i> (4)	<i>Exiguobacterium</i> (1)	<i>Microbacterium</i> (3)	<i>Microbacterium</i> (2)	<i>Microbacterium</i> (4)
<i>Stenotrophomonas</i> (2)	<i>Pantoea</i> (3)	<i>Pantoea</i> (8)	<i>Moraxella</i> (1)	<i>Rhizobium</i> (4)	<i>Pantoea</i> (3)
<i>Arthrobacter</i> (2)	<i>Moraxella</i> (1)	<i>Pseudomonas</i> (2)	<i>Exiguobacterium</i> (1)	<i>Pseudomonas</i> (2)	<i>Rhizobium</i> (1)
<i>Janibacter</i> (2)	<i>Exiguobacterium</i> (1)	<i>Brevundimonas</i> (1)	<i>Pseudomonas</i> (1)	<i>Pantoea</i> (1)	<i>Pseudomonas</i> (2)
<i>Bacillus</i> (3)	<i>Corynebacterium</i> (1)	<i>Staphylococcus</i> (2)	<i>Corynebacterium</i> (1)	<i>Arthrobacter</i> (2)	<i>Staphylococcus</i> (2)
<i>Microbacterium</i> (2)	<i>Arthrobacter</i> (5)	<i>Microbacterium</i> (1)	<i>Arthrobacter</i> (6)	<i>Kocuria</i> (2)	<i>Paenibacillus</i> (1)
<i>Paenibacillus</i> (1)	<i>Pseudomonas</i> (2)	<i>Bacillus</i> (1)	<i>Pantoea</i> (8)	<i>Bacillus</i> (1)	<i>Kocuria</i> (1)
<i>Novosphingobium</i> (1)	<i>Janibacter</i> (2)		<i>Brevundimonas</i> (1)	<i>Terribacillus</i> (1)	
<i>Sporosarcina</i> (1)	<i>Bacillus</i> (3)		<i>Janibacter</i> (2)		
<i>Curtobacterium</i> (1)	<i>Kocuria</i> (2)		<i>Bacillus</i> (2)		
<i>Terribacillus</i> (1)	<i>Stenotrophomonas</i> (1)		<i>Rhizobium</i> (3)		
<i>Kocuria</i> (1)	<i>Paenibacillus</i> (2)		<i>Staphylococcus</i> (2)		
<i>Micrococcus</i> (1)	<i>Staphylococcus</i> (1)		<i>Paenibacillus</i> (2)		
	<i>Micrococcus</i> (1)		<i>Kocuria</i> (2)		
	<i>Roseomonas</i> (1)		<i>Terribacillus</i> (1)		
	<i>Curtobacterium</i> (1)		<i>Stenotrophomonas</i> (1)		
			<i>Novosphingobium</i> (1)		
			<i>Curtobacterium</i> (1)		

IV.7. Producción de bacteriocinas en los aislamientos de la colección

Los microorganismos producen una extraordinaria diversidad de mecanismos de defensa antimicrobianos que incluyen antibióticos, subproductos metabólicos, enzimas líticas, exotoxinas y bacteriocinas. Es claro que diferentes arsenales antimicrobianos se encuentran ampliamente distribuidos entre las especies bacterianas, de arqueas y fúngicas analizadas a la fecha, aunque no se sabe exactamente qué rol tienen los mismos en la configuración del

establecimiento de las diferentes comunidades microbianas. En el contexto del presente trabajo de caracterización del microbioma de *M. sativa*, exploramos la capacidad de los aislamientos de generar compuestos antimicrobianos con acción hacia otras bacterias presentes en tierra, semillas o plantas de alfalfa.

Entre los mecanismos de antibiosis, el de producción de bacteriocinas fue el que elegimos analizar en un primer acercamiento a las interacciones microbianas negativas, de competición, en la planta de alfalfa. Las bacteriocinas son proteínas o péptidos con efecto antibacteriano que tienen la propiedad, a diferencia de los antibióticos por ejemplo, de tener un rango de acción relativamente estrecho, afectando generalmente a unas pocas especies bacterianas cercanas taxonómicamente, a menudo sobre especies o cepas cercanas a la misma que las produce. Tal característica lo hace un sistema de competición muy atractivo dentro de las plantas en tanto que pueden eventualmente operar para que bacterias específicas puedan competir con miembros particulares de la comunidad sin afectar el resto de la composición del microbioma. El conocimiento de mecanismos como el anterior es esencial para el diseño de bioinoculantes que, portadores de actividades benéficas, puedan hacerse un lugar dentro de nichos específicos de las plantas donde puedan expresar actividades de interés sin ser excluidos. Al mismo tiempo, la posibilidad creciente de elaborar inoculantes consorciales, deberá contemplar compatibilidades biológicas que tengan en cuenta estos aspectos. Las distintas familias de bacteriocinas incluyen proteínas que difieren en sus tamaños, blancos bacterianos, modos de acción y mecanismos de inmunidad. Distintas familias han sido caracterizadas particularmente en ciertas especies, encontrándose presentes preferencialmente en plásmidos como en el caso de las colicinas estudiadas en *E. coli* [352], o bien en cromosomas como en el caso de las piocinas estudiadas en *Pseudomonas* [353]. Los géneros Gram-positivos, por su parte, producen diversas familias de bacteriocinas, encontrándose entre las más estudiadas aquellas producidas por bacterias lácticas [354]. Una de las diferencias más importantes de las bacteriocinas producidas por Gram-positivas es su mayor rango de acción, actuando sobre diferentes géneros e incluso sobre bacterias Gram-negativas. Una de las aplicaciones agronómicas propuestas para las bacteriocinas es su utilización en la defensa de la planta contra cepas de bacterias patógenas. Se ha sugerido para este fin el clonado de múltiples genes de bacteriocinas en plantas o bien la utilización de aislamientos productores de bacteriocinas que sean capaces de ser utilizables a la vez como biofertilizantes y como antagonistas de bacterias patógenas de interés [355].

Para comenzar a explorar si existen mecanismos de antibiosis entre los aislamientos endófitos de semillas y los inoculantes de *S. meliloti* utilizados en la actualidad, la producción de bacteriocinas fue evaluada sobre un césped de la cepa de referencia *S. meliloti* 2011, observando los resultados presentados en la tabla IV.5. Se observa en la misma que la inhibición es producida por algunos aislamientos de los géneros *Bacillus*, *Paenibacillus* y *Stenotrophomonas*. Es importante resaltar que un estudio de actividad bacteriocina en semillas coinoculadas es necesario para asegurarnos que dichos aislamientos no son incompatibles con *S. meliloti*. Cabe destacar que otros aislamientos de los mismos géneros no mostraron actividad bacteriocina *in vitro*.

Tabla IV.5. Aislamientos productores de bacteriocinas *in vitro* contra un césped de la cepa *S. meliloti* 2011 y resumen de los aislamientos compatibles como producto de este análisis. Se indica en la columna 4 el número de aislamientos que serían compatibles con *S. meliloti*, debido a no presentar inhibición frente a la misma, y sus capacidades PCV entre un total de 56 aislamientos analizados.

Aislamiento	Género	Inhibición	PCV <i>in vitro</i> en aislamientos compatibles	
SS7	<i>Bacillus</i>	++	Total analizados	56
SS20	<i>Panibacillus</i>	+	Positivos no compatibles	8
SP5	<i>Stenotropomas</i>	+	Negativos compatibles	48
SP13	<i>Pseudomonas</i>	+	Proteasas	17
BS8	<i>Bacillus</i>	+	Celulasas	20
DS11	<i>Bacillus</i>	+	Solubilización de fósforos	14
GS4	<i>Bacillus</i>	+	Sideróforos	33
GS21	<i>Microbacterium</i>	+	Ácido indol acético	25

Por otra parte, se evaluaron unos 49 aislamientos que no presentaron actividades bacteriocinas contra *S. meliloti*, entre los cuales algunos muestran ser buenos candidatos para ensayar sus propiedades PCV *in vivo* coinoculando con *S. meliloti*.

Asimismo se evaluó la producción de bacteriocinas con los distintos aislamientos de la colección de cara a conocer si existen incompatibilidades entre los mismos que afecte el desarrollo de posibles inoculantes consorciales, y a la vez conocer si la producción de bacteriocinas en diferentes nichos de la planta puede ser un factor de competencia que impida a las bacterias susceptibles colonizar el mismo. Estos resultados se presentan en la Tabla A6 del Anexo.

IV.8. Ensayos de inhibición de distintos hongos fitopatógenos

Los aislamientos endófitos de nuestra colección fueron evaluados en su capacidad de inhibición de algunos hongos fitopatógenos de alfalfa y de otros cultivos. Cuatro hongos fueron evaluados *in vitro* contra cerca de 90 aislamientos de nuestra colección. En la Tabla IV.6 se presentan las características de algunos de los aislamientos con características destacadas respecto de sus capacidades de control biológico y de actividades de PCV caracterizadas en la sección precedente.

La elección de los hongos blanco se realizó teniendo en cuenta ya sea su efecto sobre alfalfa, o sobre otros cultivos de mucha relevancia agronómica local o regional. *Sclerotinia sclerotium* produce la pudredumbre blanca, que afecta notablemente a la papa en zonas tropicales, pero también afecta la soja, maní, girasol, canola, alfalfa, chauchas y a hortalizas como lechuga, tomate, apio, coliflor y repollo [356]. Por su parte, *Alternaria alternata* se ha reportado como fitopatógena en más de 380 especies de plantas; y *Fusarium solani* es un hongo fitopatógeno de alfalfa, maíz, banana, batata, café, cítricos, soja, [357, 358]. Finalmente, *Fusarium graminearum* afecta principalmente a la cebada y al trigo, pero también a otros cultivos como centeno, triticale, arroz, avena y maíz, además de colonizar otras plantas sin necesariamente producir síntomas [359]. Este hongo también puede producir mortalidad y defectos

reproductivos en animales que se alimentan con él, incluyendo infecciones en humanos [360]. En nuestra colección se hallaron 8 aislamientos que incluyeron los géneros *Bacillus*, *Paenibacillus*, *Microbacterium*, *Pseudomonas*, *Pantoea* y uno aún no identificado (NI), que mostraron efectos inhibitorios sobre los hongos fitopatógenos *Sclerotinia sclerotium*, *Alternaria alternata* y *Fusarium solani*. La mayoría de estos aislamientos presentaron antagonismo contra los 3 hongos, lo cual es interesante para el control de varios hongos utilizando un mismo inoculante. Además, el hecho que varios de los aislamientos anteriores pertenezcan a géneros formadores de esporas los hace buenos candidatos para la formulación de bioinoculantes más estables. Ningún aislamiento de los 90 ensayados presentó inhibición contra *Fusarium graminearum*.

Tabla IV.6. Actividades asociadas al control biológico y a la PCV en aislamientos escogidos de la colección de bacterias aisladas de semilla.

Ais	Género	SS	AA	FS	Pro	Cel	Fos	Sid	AIA	Bac
SS7	<i>Bacillus</i>	+	+	NE	+	-	-	+	-	++
SS20	<i>Paenibacillus</i>	+	NE	+	-	-	-	-	-	+
SP10	<i>Microbacterium</i>	+	+	+	+	+	+	+	-	-
SP11	<i>Pseudomonas</i>	-	++	-	-	+	+	+	+	-
SP23	NI	+	NE	+	NE	NE	NE	NE	NE	-
BS14	<i>Bacillus</i>	+	NE	+	+	+	-	+	+	-
GS1	<i>Pantoea</i>	-	NE	+	NE	NE	NE	NE	NE	-
GS38	<i>Bacillus</i>	+	NE	+	NE	NE	NE	NE	NE	-

SS: *Sclerotinia sclerotium*, **AA:** *Alternaria alternata*, **FS:** *Fusarium solani*, **Pro:** proteasas, **Cel:** celulasas, **Fos:** solubilización de fosfatos, **Sid:** Producción de sideróforos, **AIA:** Producción de ácido indol acético, **Bac:** Producción de bacteriocinas contra *S. meliloti*, **Ais:** aislamiento. **NE:** no ensayado.

IV.9. Análisis de la diversidad plasmídica presente en la microbiota asociada a semilla

Como parte del desarrollo de esta tesis centrada en el acceso a la información génica codificada en plásmidos de bacterias asociadas a la planta de alfalfa fue necesaria la obtención de aislamientos que pudieran portar plásmidos. En esta búsqueda inicial, cuyos resultados mostramos a lo largo de este capítulo, pudimos observar que la planta de la alfalfa era colonizada inesperadamente desde la semilla y por esta razón nos fue relevante caracterizar la diversidad asociada a la semilla y acceder a la fracción de la misma que podía colonizar la plántula durante el primer mes de desarrollo de la planta en condiciones de asepsia. El paso siguiente fue el de analizar si los aislamientos colectados tanto en semillas como en plántula presentaban o no evidencias de poseer plásmidos en sus genomas. Dicha caracterización la

llevamos a cabo realizando geles de lisis *in situ* por geles de Eckhardt [242], mediante los cuales determinamos que 28 aislamientos entre 130 analizados (21,5 %) mostraron algún perfil plasmídico por este método. La Tabla IV.7 resume estos aislamientos, sus géneros asociados, la cantidad de bandas observadas en los geles de lisis y si fueron obtenidos desde exudados de semillas o de plántulas.

Tabla IV.7. Aislamientos de la colección que presentan bandas plasmídicas por geles de lisis *in situ*. Referencias: AIS, denominación del aislamiento; Nro de bandas: cantidad de bandas observadas en geles de lisis.

AIS	Género	Nro de bandas	nicho	AIS	Género	Nro de bandas	nicho
SP7	<i>Rhizobium</i>	8	plántula	SP13	<i>Pseudomonas</i>	1	plántula
SP17	<i>Rhizobium</i>	6	plántula	SP15	<i>Stenotrophomonas</i>	1	plántula
SS8	<i>Pantoea</i>	3	semilla	GS2	<i>Curtobacterium</i>	1	semilla
SP9	<i>Rhizobium</i>	4	plántula	GS27	<i>Curtobacterium</i>	1	semilla
GP2	<i>Pantoea</i>	3	plántula	BS6	<i>Bacillus</i>	1	semilla
SP6	<i>Pantoea</i>	3	plántula	DP3	<i>Micrococcus</i>	1	plántula
SP18	<i>Rhizobium</i>	4	plántula	SP20	<i>Novosphingobium</i>	2	plántula
BS7	<i>Bacillus</i>	3	semilla	DS2	<i>Sporosarcina</i>	1	semilla
GP1	<i>Pantoea</i>	3	plántula	DS8	<i>Planococcus</i>	1	semilla
SP11	<i>Pseudomonas</i>	2	plántula	GS6	<i>Bacillus</i>	1	semilla
SP12	<i>Rhizobium</i>	2	plántula	GS7	<i>Bacillus</i>	1	semilla
GS9	<i>Curtobacterium</i>	1	semilla	GS11	NI	1	semilla
SP21	<i>Pantoea</i>	3	plántula	GS20	<i>Roseomonas</i>	1	semilla
SS12	<i>Moraxella</i>	1	semilla	SP22	<i>Pantoea</i>	3	plántula

Es interesante destacar que algunos de estos aislamientos mostraron ser del mismo género y presentar a su vez diferentes perfiles plasmídicos, como es el caso de *Rhizobium*. Por otro lado, los perfiles observados para los distintos géneros fueron siempre diferentes y en algunos casos la dificultad de obtención de un gel con buena definición, como en el caso de los aislamientos de *Pantoea*, permitió observar algunas bandas en común pero no pudimos determinar si los mismos presentaron diferencias en los perfiles. Como resultado, la elección de aislamientos para la obtención de las muestras plasmídicas representativas en nuestra colección se realizó eligiendo aquellos aislamientos que pudieron mostrar con claridad diferentes perfiles plasmídicos. La purificación y secuenciación de dicha muestra será detallada en el capítulo V.

IV.10. Algunos comentarios acerca de los posibles roles biológicos conocidos para los géneros obtenidos en las semillas y plantas de alfalfa.

¿Interacción beneficiosa o patogénica con la planta? Rango de hospedadores. Interacciones bióticas complejas

El análisis de la bibliografía relacionada a los diferentes géneros bacterianos obtenidos aquí en plantas y semillas de alfalfa muestra a un conjunto de ellos ya descritos anteriormente en plantas, y unos pocos géneros que no habían sido identificados como colonizadores de las mismas. Asimismo, se han reportado ejemplos de géneros que, además de colonizar plantas, se encuentran distribuidos en otros ambientes lo que puede indicar que distintas especies del género se han adaptado a distintos nichos ecológicos [361], o bien que una misma especie esté adaptada a la vez a distintos ambientes y aquí sólo hemos evidenciado su rol dentro de nichos en la planta. Así, algunos aislamientos pertenecientes a los géneros aquí reportados ya habían sido identificados como integrantes de microbiomas de otras plantas, insectos u otros animales. La búsqueda de la mayoría de los géneros bacterianos mostrados en este trabajo arroja como resultado que los mismos han sido aislados de endosfera, filósfera, rizósfera o semillas de diversas plantas, en variadas zonas geográficas y en plantas tanto salvajes como agrícolas. Un ejemplo se muestra en una reciente revisión realizada por Bredow y colaboradores [362], donde a partir de 7 estudios de diversidad bacteriana en la filósfera de 8 cultivos de relevancia económica mundial (café, azúcar, caña, habas, maíz, soja, tomate y uva), mediante el secuenciamiento parcial del gen ARNr 16S, se observa la presencia de los géneros *Pantoea*, *Stenotrophomonas*, *Acinetobacter*, *Paenibacillus*, *Kocuria*, *Rhizobium*, *Flavobacterium*, *Novosphingobium*, *Pseudomonas*, *Staphylococcus*, *Micrococcus*, *Bacillus*, *Microbacterium*, *Brevundimonas*, *Sphingomonas*, *Curtobacterium*, además de otros géneros no identificados en alfalfa en este trabajo de tesis.

Se describen a continuación las características más relevantes de algunos de los géneros asociados a semillas y plantas de alfalfa que hemos aislado y comenzado a caracterizar en este trabajo de tesis.

***Pantoea*.** Distintas especies del género *Pantoea* han sido aisladas como microorganismos oportunistas en humanos, como fitopatógenas y endófitas en plantas, asociadas a diversos insectos y también en agua y suelo [363]. Existe un ejemplo de una relación entre una especie patogénica del género *Pantoea* que utiliza un insecto como vector para infectar a una planta [364]. Es evidente que la capacidad de adaptación de las especies incluidas en este género es muy amplia y es tentador suponer que algunas especies del género transiten tanto plantas como insectos, pudiendo establecer relaciones simbióticas y/o patogénicas con los distintos huéspedes. Tal característica las hace especialmente interesantes para analizar en detalle cuáles son los determinantes de compatibilidad/incompatibilidad para el acceso/rechazo a/de ciertas asociaciones con un fondo genético conservado. Es también necesario evaluar qué implicancia clínica tienen los aislamientos que pudieran ser aplicables biotecnológicamente.

***Stenotrophomonas*.** Las especies de este género también muestran una gran plasticidad adaptativa, reflejándose esto en las diferencias genómicas que existen entre los aislamientos secuenciados (ver comentario precedente en relación a *Pantoea*). Estas especies han sido aisladas principalmente asociadas a plantas y presentan excelentes capacidades PCV,

producción de enzimas de interés industrial, actividades de biorremediación, producción de diversas biomoléculas de interés y capacidades de biocontrol [365, 366]. Esta especie contiene múltiples ejemplos donde se ha probado actividad antifúngica y de patogenicidad contra nematodos fitopatógenicos, hecho que debe tenerse en cuenta a la hora de entender desde lo biológico y evolutivo la presencia de estas bacterias en plantas. Más allá de estas consideraciones, existen diferentes aislamientos clínicos oportunistas multiresistentes de *Stenotrophomonas*, y si bien se cree que el principal hábitat del género es el suelo y las plantas, los mecanismos que se sospecha son responsables de su impacto clínico parecen ser los mismos que dichas bacterias utilizan para asociarse a las plantas [366]; circunstancia por la que su utilización biotecnológica debe ser evaluada cuidadosamente.

Paenibacillus. Los miembros de este género se han estudiado durante largo tiempo las propiedades de PCV en diferentes aislamientos, las de fijación de nitrógeno, producción de metabolitos y enzimas de interés industrial, así como capacidades antifúngicas y de biocontrol de insectos por la presencia de las proteínas insecticidas Cry en algunos de sus aislamientos [367-370]. En nuestra colección, este género ha mostrado producir celulasas, proteasas, AIA y sideróforos. La capacidad de formación de endosporas les da a los integrantes de este género (así como a los de *Bacillus*) una ventaja para la producción de inoculantes estables en el tiempo. Su presencia se ha reportado asociada a microbiomas de insectos como la termita [371], lo que sugiere explorar eventuales relaciones entre el microbioma de plantas y el de un insecto que la digiere. Se ha estudiado la enfermedad producida sobre escarabajos japoneses por una especie de *Paenibacillus* que se comercializa como biocontrolador en EUA [372], mostrando una vez más la existencia de una bacteria simbiote distribuida en gran variedad de plantas, que afecta a un insecto que se alimenta de las raíces de la misma. La misma especie ha sido aislada como endófito de pino e inoculada exitosamente en maíz, donde mostró desarrollar capacidades de PCV [373]. El género *Paenibacillus* también presenta multiresistencia [374], aunque existe poca bibliografía respecto a su aislamiento en el ambiente clínico [375]. La necesidad de cuidar también en este género la posible ubicuidad de mecanismos de patogenicidad, extensibles aun a vertebrados está ejemplificado en el reciente aislamiento de *Paenibacillus* causante de enfermedad en serpientes vectores de schistosomiasis [376]. El uso eventual de aislamientos como el precedente en estrategias de biocontrol (como de serpientes y schistosomiasis) resalta nuevamente la necesidad de avanzar en profundidad en la caracterización de los mecanismos moleculares que participan en el establecimiento de asociaciones mutualistas o patogénicas para un mismo género de bacterias.

Kokuria. El género *Kokuria* se utiliza en industria como control de esterilización, dado que es muy resistente a diferentes factores físicos, encontrándose presente en ambientes extremos [377] y formando parte incluso de la microbiota resistente a la radiación en zonas como Chernobyl [378]. El género se ha encontrado en el suelo [379], sedimentos marinos [380], frecuentemente en aislamientos clínicos [381, 382] y en rizósfera de distintas plantas por lo que se han caracterizado en sus propiedades de PCV [383].

Rhizobium. El género *Rhizobium* ha sido especialmente estudiado en relación a su capacidad de fijar nitrógeno en asociación con leguminosas (ver Introducción de esta tesis), hallándose en suelo y plantas, sobre todo asociado a la rizósfera, aunque también como endófitos en

plantas adultas y asociadas a las semillas, tal como hemos reportado en este trabajo. Sus especies [384] han sido empleadas como biofertilizantes en agricultura desde hace más de 120 años por su capacidad fijadora de nitrógeno atmosférico, y utilizadas como uno de los modelos más estudiados de interacción planta-bacteria. Sus capacidades PCV han sido mucho menos estudiadas, y menos aún la posible significancia de la presencia de estas bacterias dentro de las semillas.

Flavobacterium. Especies del género *Flavobacterium* han sido aisladas de suelo [385], ambientes acuáticos y plantas presentes en ellos [386, 387], rizósfera y filósfera de diversas plantas terrestres, mostrando además capacidades PCV y de biocontrol [388]. Si bien algunos aislamientos presentan severa patogenicidad contra distintos peces [389], los genomas de los aislamientos de plantas terrestres se diferencian de los presentes en los aislamientos patogénicos por su mayor proporción de genes de utilización de xilosa, arabinosa y pectinas; y por la presencia exclusiva de genes de degradación de hemicelulosas que sugieren su adaptación al nicho terrestre en base a estas modificaciones genómicas [390].

Novosphingobium. Las especies de este género presentan adaptaciones genómicas a distintos ambientes [391], habiendo sido aisladas en suelos y plantas, incluyendo aislamientos con capacidad de degradación de compuestos aromáticos [392]. También como parte de la microbiota de *Homalodisca vitripennis*, insecto plaga que además es vector de la especie fitopatogena *Xylella fastidiosa* [393].

Exiguobacterium. En el caso del género *Exiguobacterium*, el mismo se ha aislado de nódulos de Fenugreek (*Trigonella foenum-graecum*) y ha mostrado producir distintas capacidades PCV *in vitro* (solubilización de fósforo, producción de sideróforos, AIA e inhibición de distintos hongos fitopatogénicos), pero sobre todo ha mostrado ser capaz de potenciar el crecimiento de la planta (largo de tallo y raíces) cuando es co-inoculada con *S. meliloti* [322].

Consideraciones finales.

Más allá del rol de fitopromotores o fitopatogénicos de las bacterias presentes en semillas y de su relación con variantes del mismo género asociadas a representantes del reino animal, será importante analizar su origen (ancestralidad), eventual función(es), y vínculos con otros microbiomas debido a que su presencia parece ser un hecho muy ubicuo en distintas especies vegetales. Será importante analizar a los distintos miembros del microbioma de las semillas, las variaciones existentes entre ellos, el significado para la planta, los mecanismos de transmisión vertical, y su vínculo con bacterias ambientales del mismo género pero asociadas a otros organismos que comparten interacciones bióticas con plantas (insectos, vertebrados). El avance en estas cuestiones a nivel biológico, genómico, y bioquímico, es esperable que permita identificar aquellos determinantes que han sustentado la transición de géneros específicos entre ambientes (muy) diferentes, haciendo de los microbiomas grupos dinámicos y en permanente evolución condicionados por el propio huésped, y por los (micro y macro) organismos que con él conviven.

En este capítulo hemos explorado la diversidad bacteriana asociada a las semillas de alfalfa, tanto aquella presente en su interior (protegidas de agentes externos) como aquella que es capaz de colonizar la plántula a partir de estas semillas. Estos resultados mostraron que las semillas son fuentes naturales de futuros colonizadores de la planta y sugiere la idea de que esta asociación sea un mecanismo que plantas y bacterias han adquirido de manera de seleccionar y transmitir una diversa microbiota hacia las sucesivas generaciones. En este sentido, el estudio de las capacidades de promoción del crecimiento vegetal y la exploración de las especies capaces de perdurar a lo largo de las generaciones puede revelarnos qué tipo de beneficios otorgarían a las plantas aquellas bacterias incluídas en dicho vínculo coevolutivo. Este conocimiento es importante para el diseño de bioinoculantes más efectivos, porque permite indagar aquellas características que son necesarias para una asociación más estable y duradera, así como indagar sobre las composiciones consorciales más estables y beneficiosas desde el punto de vista del crecimiento de la planta.

Con el objetivo de profundizar el estudio de las capacidades adaptativas que las especies que conforman el microbioma pueden intercambiar entre sí, quisimos explorar aquella información genética codificada en los plásmidos de manera de saber cómo los mismos contribuyen con distintas actividades génicas y de cara a examinar posibles relaciones entre los plásmidos de *S. meliloti* analizados en el Capítulo III y los presentes en las otras bacterias que cohabitan la planta de alfafa. En el próximo capítulo abordaremos el secuenciamiento de una muestra de los plásmidos previamente mostrados como un conjunto representante de la información plasmídica presente en el microbioma proveniente de la semilla de alfalfa.

CAPITULO V

Secuenciamiento del moviloma plasmídico críptico recuperado de bacterias de semillas de alfalfa

V.1. Acceso a los plásmidos crípticos presentes en el microbioma de semilla de alfalfa

La propuesta central del desarrollo de esta tesis fue la de explorar desde el punto de vista estructural y funcional el moviloma plasmídico presente en los rizobios simbioses de alfalfa como describimos en el Capítulo III, y analizar luego el moviloma de las bacterias simpátricas (asociadas a la misma raíz) con el propósito de analizar su contenido génico y eventual vínculo (intercambio) con el moviloma de los rizobios. Con este propósito, en el capítulo anterior describimos parte del microbioma que coloniza a la planta de alfalfa, aquel que proviene de la semilla. Como hemos dicho, parte de este microbioma (algunos géneros/especies) es posible que esté íntimamente ligado a las semillas transmitiéndose verticalmente a través de las sucesivas generaciones. Ya que es esperable que dichos aislamientos mantengan una relación de coevolución muy cercana y ancestral con la planta y con los rizobios, decidimos analizar el moviloma plasmídico en ese microbioma. En un primer paso hacia el estudio del moviloma plasmídico de estas comunidades, se realizó una caracterización de los perfiles plasmídicos por geles de lisis *in situ* [242], en los distintos aislamientos colectados que presentamos en el Capítulo IV precedente. El análisis de 130 aislamientos recuperados de las semillas pertenecientes a los géneros y especies descritas en la Tabla IV.7 del capítulo anterior nos permitió reconocer 28 aislamientos portadores de plásmidos, entre los cuales pudimos identificar 18 perfiles diferentes (Figura V.1) en aislamientos de 12 géneros diferentes (Tabla V.1).

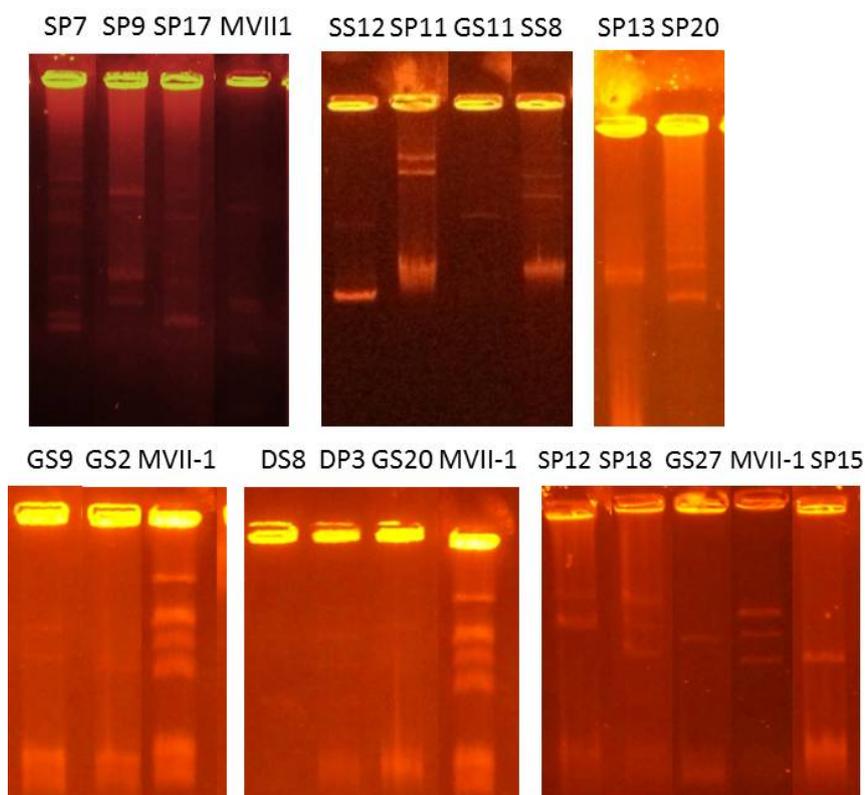


Figura V.1. Perfiles de plásmidos correspondientes a 18 aislamientos recuperados de la colección de bacterias asociadas a semillas y plantas de alfalfa. Se muestra en algunos casos el perfil plasmídico de la cepa de referencia *S. meliloti* MVII-1, que presenta bandas correspondientes a plásmidos de los siguientes tamaños: 1500 kpb, 156 kpb, 135 kpb, 81 kpb y 54 kpb [219].

Los aislamientos con mayor riqueza en perfiles plasmídicos fueron los pertenecientes al género *Rhizobium* con 5 perfiles de plásmidos diferentes y entre 3 a 8 bandas cada uno que corresponden a replicones de tamaños entre aproximadamente 70 kpb a 1700 kpb. Los plásmidos visualizados para los aislamientos de *Pantoea* mostraron poseer perfiles con bandas conservadas correspondientes a plásmidos de aproximadamente 600, 170 y 80 kpb. Otros géneros que mostraron perfiles plasmídicos fueron *Pseudomonas*, *Moraxella*, *Stenotrophomonas*, *Micrococcus*, *Curtobacterium*, *Roseomonas*, *Arthrobacter*, *Planococcus* y *Novosphigobium*, los cuales presentaron en general entre 1 y 2 bandas plasmídicas. La búsqueda en GenBank nos indica que para el caso de los géneros *Curtobacterium* y *Roseomonas* no se habían identificado plásmidos hasta la fecha (Noviembre 2015), con poca información también disponible para los géneros *Stenotrophomonas*, *Micrococcus* y *Planococcus* (ver Tabla V.1).

Tabla V.1. Lista de géneros presentes en semillas de alfalfa en los que hemos encontrado plásmidos, con la descripción del número de plásmidos que han sido completamente secuenciados y depositados en el GenBank para esos mismos géneros.

Género	Nº de plásmidos completos actualmente disponibles	Aislados de plantas
<i>Rhizobium</i>	125	Sí
<i>Pantoea</i>	82	Sí
<i>Pseudomonas</i>	166	Sí
<i>Moraxella</i>	14	No
<i>Stenotrophomonas</i>	6	Sí
<i>Micrococcus</i>	9	No
<i>Curtobacterium</i>	0	-
<i>Roseomonas</i>	0	-
<i>Arthrobacter</i>	57	Sí
<i>Planococcus</i>	10	No
<i>Novosphigobium</i>	23	No

Según hemos visto en el capítulo precedente son diversos los géneros asociados a semillas que colonizan las plantas en crecimiento, multiplicándose luego para dar forma al microbioma de la misma (semillas, rizósfera, filósfera, y compartimento endofítico), que competirán y se sumarán a los que puedan provenir de la microflora colonizante del suelo de implantación. A la sólo vista de las bacterias encontradas en semilla hemos podido detectar una diversidad importante de plásmidos contenidos en géneros diversos, dando sustento a la posibilidad real de un intercambio horizontal de información tanto entre ellos como con los rizobios simbióticos cuyo contenido plasmídico ha sido descrito en detalle en el capítulo III de esta tesis y sus activas propiedades de movilidad en estudios previos de nuestro laboratorio [1]. Con el propósito de buscar evidencias de posibles eventos de transferencia, y si existiesen de describir las características del mismo (entre qué bacterias, límites, relaciones de ortología), en las secciones siguientes abordamos el análisis de las secuencias plasmídicas del microbioma de

semillas. El análisis del contenido génico plasmídico y su caracterización funcional serán además abordados con el propósito de explorar y reconocer nuevos aportes del moviloma a la dinámica de la comunidad microbiana y a su relación con la planta, el resto de la biota y el entorno.

V.2. Purificación y secuenciación del moviloma plasmídico en bacterias asociadas a la alfalfa acarreados en la semilla

Como paso inicial en el estudio del moviloma plasmídico que acompaña esta diversidad de especies bacterianas, se cultivaron los 18 aislamientos que habían mostrado distintos perfiles plasmídicos y posteriormente se enriquecieron dichos plásmidos en gradientes de cloruro de cesio del mismo modo que describimos para la preparación de los plásmidos de *S. meliloti*. Los aislamientos que pudieron ser exitosamente enriquecidos en los gradientes, pertenecieron a los géneros detallados en la Tabla V.2.

Tabla V.2. Aislamientos de la colección cuyos plásmidos fueron enriquecidos y secuenciados en este estudio.

#AIS	Género	# de bandas	nicho	#AIS	Género	# de bandas	nicho
SP7	<i>Rhizobium</i>	8	plántula	SP13	<i>Pseudomonas</i>	2	plántula
SP17	<i>Rhizobium</i>	6	plántula	SP15	<i>Stenotrophomonas</i>	1	plántula
SS8	<i>Pantoea</i>	4	semilla	GS2	<i>Curtobacterium</i>	2	semilla
SP9	<i>Rhizobium</i>	4	plántula	GS27	<i>Curtobacterium</i>	1	semilla
SP18	<i>Rhizobium</i>	4	plántula	DP3	<i>Micrococcus</i>	3	plántula
SP11	<i>Pseudomonas</i>	2	plántula	SP20	<i>Novosphingobium</i>	3	plántula
SP12	<i>Rhizobium</i>	2	plántula	DS8	<i>Planococcus</i>	1	semilla
GS9	<i>Curtobacterium</i>	2	semilla	GS11	NI	1	semilla
SS12	<i>Moraxella</i>	1	semilla	GS20	<i>Roseomonas</i>	1	semilla

Luego de la purificación, con las muestras obtenidas se preparó una mezcla con cantidades equivalentes en masa de cada una. Dicha mezcla fue secuenciada mediante la tecnología Illumina *mate pair* (ver Materiales y métodos, sección II.11). El ensamblado de las lecturas generó 2.210 cóntigos con un largo medio de 2.072 pb, una moda de 300 pb y una mediana de 449 pb (el contigo más largo fue de 127.120 pb). Cerca del 90% de los cóntigos tuvieron tamaños menores que 4.334 pb. La distribución descendente de tamaños resultó en un N50 de 11.901 pb.

Del análisis del largo de los cóntigos versus la cantidad de lecturas que los formaron, que se presenta en la Figura V.2.A., se observa que para cóntigos diferentes pero de un mismo largo (misma posición en el eje x) el número de lecturas fue variable (diversos valores en el eje y). Además, si se analiza la **densidad de lecturas (cobertura)**, esto es el número de lecturas normalizadas al largo de cada contigo (Figura V.2.B), se observa que la misma es variable e

independiente del largo de los cóntigos. Dicha observación refleja que la mezcla secuenciada consistió de fragmentos de replicones presentes en distintas concentraciones, como era esperable teniendo en cuenta que: a) las preparaciones de plásmidos que hemos realizado provienen de mezclas de varias cepas en proporciones que no fueron exactamente equivalentes, b) las cepas contienen plásmidos de diferente número de copias (y tamaños), y c) las diferentes preparaciones con esas características fueron mezcladas en cantidades equivalentes en masa sin relación con los moles de cada plásmido presente en cada conjunto que se incorporó a la mezcla. La mezcla en cantidades equivalentes en masa de cada preparación (mezcla de varias cepas) permitió la secuenciación de una cantidad equivalente de ADN plasmídico desde cada una de ellas.

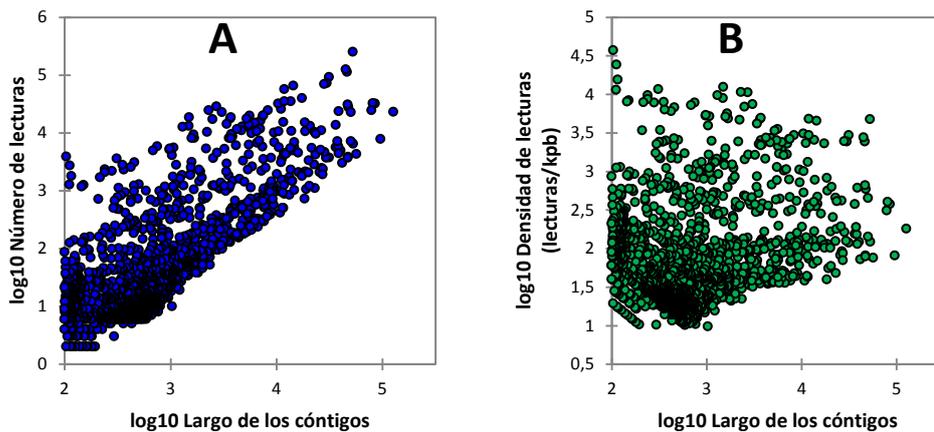


Figura V.2. Número (parte A) y densidad de lecturas (cobertura)(parte B) de la secuenciación en relación al largo de los cóntigos. Los números están expresados como logaritmos en base 10.

El secuenciamiento realizado a partir de la biblioteca “*mate-pair*”, que aporta información posicional respecto de dos lecturas de secuencias (distancia aproximada a la que se encuentran dentro del genoma) permitió la unión de muchos de los cóntigos en ensambles de mayor tamaño (ver más abajo), superando secuencias repetitivas que son frecuentemente la mayor limitación en el ensamblado de genomas bacterianos [394]. La biblioteca “*mate-pair*” nos permitió obtener una mejora notoria con respecto al secuenciamiento previo de la mezcla de ADN plasmídico de *S. meliloti*. Como resultado final se obtuvieron 133 **ensambles (“scaffolds”)** mayores a 2,0 kpb y de hasta 335,5 kpb, que incluyeron a 465 de los 2.210 cóntigos totales obtenidos, es decir 21% de ellos.

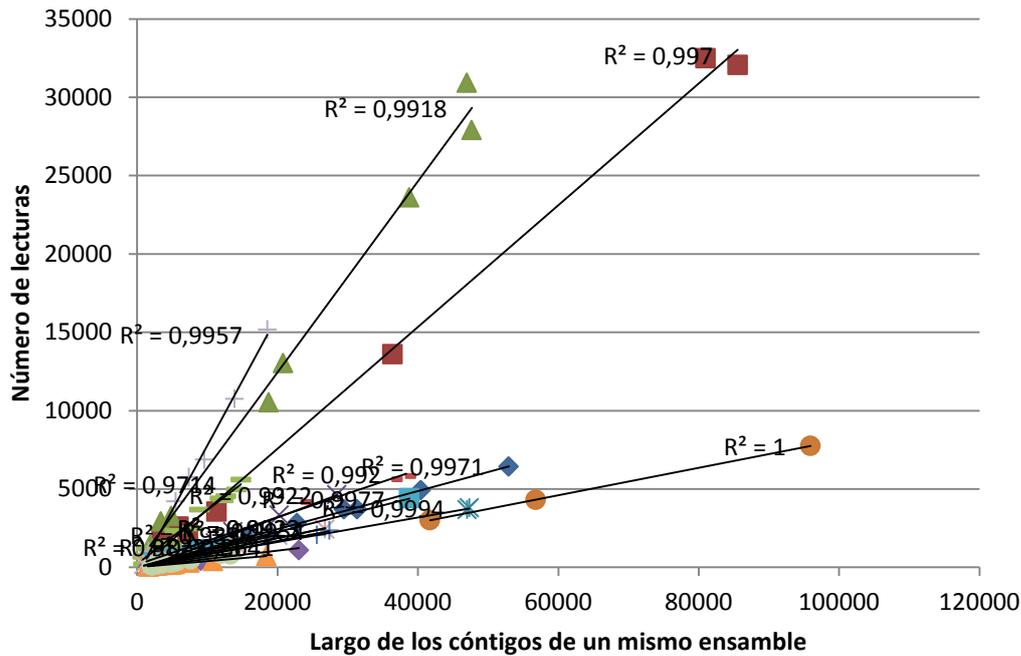


Figura V.3: Número de lecturas en función del largo de los cóntigos (pb) pertenecientes a 16 ensamblados (“scaffolds”) distintos. Se emplearon símbolos diferentes para ensamblados diferentes, y símbolos iguales para los distintos cóntigos dentro de un mismo ensamble. Cada ensamble muestra una pendiente característica (cobertura). Se representan las tendencias lineales con su coeficiente de correlación para cada ensamble.

Los ensamblados (“scaffolds”) obtenidos nos permitieron analizar en mayor detalle la cobertura de cada uno de los cóntigos que formaban parte de cada uno de ellos por separado, según se muestra en la Figura V.3. Las rectas con coeficientes de correlación cercanos a la unidad indican que los cóntigos generados tuvieron una cobertura homogénea a lo largo de cada ensamble, lo esperable para cóntigos que son parte, efectivamente, de un mismo replicón dentro de la mezcla secuenciada, más allá de la representatividad que ese plásmido tenga dentro de la misma. La representatividad de cada ensamble en la muestra viene dada por la pendiente de las rectas (cobertura). A pesar de que la cobertura entre los distintos ensamblados en la muestra varía en más de 100 veces (Tabla A7 del Anexo), el enriquecimiento de los plásmidos en la muestra y la cobertura total del secuenciamiento fue suficiente para obtener ensamblados de decenas de kilobases incluso para los plásmidos menos enriquecidos en la muestra (ej. aún la recta de círculos marrones, con poca pendiente respecto del resto). Las pendientes relativas de las rectas dependen de los diferentes factores que determinaron la mayor o menor presencia de un dado replicón plasmídico en la mezcla de secuenciación, que han sido mencionados en la sección precedente

Con el propósito de verificar el grado en que nuestra preparación en gradientes de CsCl había conseguido enriquecerse en ADN plasmídico respecto de cromosomal y megaplasmídico, analizamos la cobertura alcanzada por unos y otros según se muestra en la Figura V.4. Consistente con el procesamiento de una muestra enriquecida en plásmidos, y de manera similar a lo ocurrido con el secuenciamiento que presentamos en el Capítulo 3 para los plásmidos de *S. meliloti*, la cobertura para los ensamblados con alta similitud a secuencias cromosomales y de megaplásmidos (C y M, en este caso similares por BLASTn a bacterias depositadas en GenBank, pertenecientes a los distintos géneros aquí identificados por ARNr

16s o biotipado por MALDI-TOF), mostraron coberturas menores que las de los ensamblados con similitud a diferentes plásmidos (P). Asimismo, varios ensamblados con gran similitud a plásmidos presentaron bajas coberturas, debido a su baja representación en la muestra final. Esta última observación indica que si bien la elevada densidad de lecturas/kpb en relación a la media del secuenciamiento es un buen indicador para el reconocimiento de secuencias plasmídicas en la muestra, varias secuencias plasmídicas no fueron reconocidas con este único criterio. Atendiendo a ello, además de la cobertura para identificar secuencias plasmídicas los ensamblados se analizaron por BLASTn contra la base nr de GenBank clasificando fragmentos como (potencialmente) plasmídicos a aquellos que presentaron similitud con plásmidos ya reportados y que a su vez no alinearon con secuencias cromosomales. Como producto de este proceso de filtrado manual, 3,34 Mpb de secuencia no redundante fue obtenido en el total de los plásmidos crípticos presentes en la microbiota endófito proveniente de la semilla de alfalfa.

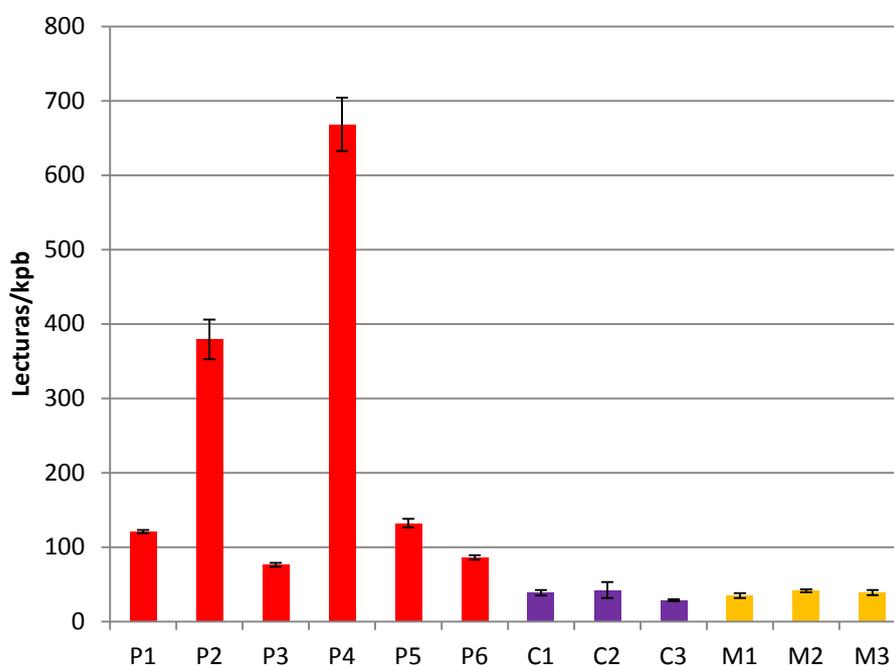


Figura V.4. Densidad de lecturas (cobertura) alcanzada por diferentes ensamblados (*scaffolds*) asignados a plásmidos (P_n), cromosomas (C_n) o megaplásmidos (M_n) por búsqueda de similitud BLASTn. Se tomaron al azar 6 ensamblados plasmídicos, 3 cromosomales y 3 provenientes de megaplásmidos. Las barras de error indican desviaciones estándar de la media.

V.3. Identificación de ensamblados cerrados en el moviloma plasmídico proveniente de semilla de alfalfa.

A partir de los ensamblados obtenidos pudo establecerse que 15 de ellos correspondían a plásmidos cerrados que se presentan en la Tabla V.3 (Materiales y métodos, sección II.12). Los mismos mostraron distintos grados de similitud nucleotídica con replicones plasmídicos ya secuenciados de distintas especies pero ninguno superó una cobertura de alineamiento mayor que 55%, lo cual nos muestra que los plásmidos secuenciados son diferentes a los depositados a la fecha en las bases de datos. Con el objetivo de identificar aquellos aislamientos de nuestra colección que contienen a dichos plásmidos cerrados, se realizaron ensayos de PCR con

cebadores específicos sobre el conjunto de los aislamientos cuyos plásmidos secuenciamos. Los aislamientos que alojaron a los diferentes plásmidos cerrados pudieron ser identificados y se describen en la columna 7 de la Tabla V.3.

Tabla V.3. Listado y descripción general (tamaño, huésped, similitud por BLASTn con otros replicones) de los plásmidos que pudieron ser cerrados [ensambles ("scaffolds") circulares] luego del secuenciamiento.

PLÁSMIDO	NC	TAMAÑO	1º SIMILITUD POR BLASTn	COB	%I	GÉNERO PORTADOR (IDENTIFICACIÓN)
pBo1	3	194.465	<i>Pseudomonas putida</i> ND6, plásmido pND6-2	55%	98%	<i>Pseudomonas</i>
pBo2	14	117.971	<i>Arthrobacter aureescens</i> TC1, plásmido TC2	4%	81%	<i>Arthrobacter</i>
pBo3	6	112.593	<i>Sinorhizobium meliloti</i> , plásmido pHRC017	4%	84%	<i>Rhizobium</i>
pBo4	17	111.177	<i>Methylobacterium oryzae</i> CBMB20, plásmido pMOC1	1%	78%	<i>Roseomonas</i>
pBo5	1	50.361	<i>Psychrobacter</i> sp. PRwf-1	2%	88%	<i>Moraxella</i>
pBo6	1	79.080	<i>Pantoea vagans</i> C9-1, plásmido pPag2	13%	91%	<i>Pantoea</i>
pBo7	2	57.475	<i>Sphingomonas</i> sp. NS2, plásmido 201	9%	79%	<i>Novosphingobium</i>
pBo8	9	57.311	<i>Agrobacterium tumefaciens</i> , plásmido pTi-SAKURA	3%	99%	<i>Rhizobium</i>
pBo9	4	56.353	<i>Agrobacterium tumefaciens</i> Ti, plásmido pTiBo542	10%	93%	<i>Rhizobium</i>
pBo10	1	52.694	<i>Arthrobacter chlorophenolicus</i> A6	7%	84%	sin identificar
pBo11	6	51.706	sin similitud			<i>Roseomonas</i>
pBo12	6	51.621	<i>Starkeya novella</i> DSM 506	6%	74%	<i>Roseomonas</i>
pBo13	6	38.685	<i>Micrococcus</i> sp. A1, plásmido pLMA1	31%	99%	<i>Micrococcus</i>
pBo14	4	26.348	<i>Sphingopyxis</i> sp. Kp5.2, plásmido pSfKp5.2	38%	99%	<i>Novosphingobium</i>
pBo15	4	18.892	<i>Agrobacterium</i> sp. H13-3, cromosoma lineal	20%	78%	r <i>Rhizobium</i>

NC: Número de cóntigos que forman el plásmido. COB: Cobertura del alineamiento entre el plásmido y la secuencia depositada en GenBank. %I: porcentaje de identidad entre el plásmido y la secuencia depositada en GenBank.

Algunos de los plásmidos cerrados mostraron un buen alineamiento por BLASTn y alta cobertura con otros plásmidos presentes genomas bacterianos ya depositadas en GenBank. En varios casos los mejores alineamientos correspondieron a genomas bacterianos coincidentes o relacionados (duplas *Sinorhizobium-Rhizobium*, *Methylobacterium-Roseomonas*, *Psychrobacter-Moraxella*, *Sphingomonas-Novosphingobium*, *Starkeya-Roseomonas*) con el de los aislamientos en que hemos encontrado los plasmidos secuenciados (comparar columnas 4 y 7 de la Tabla V.3). La evidencia revela que los mismos han transitado e intercambiado información con representantes de generos/especies filogenéticamente relacionadas. En ningún caso, sin embargo, la cobertura de las regiones similares superó 55% del largo de ninguno de nuestros plásmidos (Tabla V.3, columna 5) poniendo en evidencia el carácter variable de los mismos. En consonancia con esta idea, el alineamiento con otros plásmidos o replicones en distintas regiones de un mismo ensamble reveló el carácter modular, propio de este tipo de replicones, en los plásmidos que hemos analizado.

Más allá de las observaciones precedentes, la mayor parte de las secuencias no mostró similitud con otros replicones conocidos, lo que pone en evidencia que la fracción de

diversidad plasmídica que se ha explorado aún es pequeña y que existe en cada una de estas especies una cantidad importante de información codificada en plásmidos que es aún desconocida. En el caso de algunos plásmidos contenidos en aislamientos de los géneros *Arthrobacter*, *Rhizobium*, *Roseomonas* y *Moraxella*, casi la totalidad de las secuencias han sido nuevas.

Para profundizar el análisis anterior realizamos a continuación búsquedas por similitud BLASTp con cada uno de los marcos de lectura abiertos predichos para los ensamblados. Dado que la tasa de sustitución aminoacídica es mucho menor que la nucleotídica a lo largo de la evolución genómica, las búsquedas por BLASTp son más sensibles en exploraciones taxonómicas. En la Tabla V.4 se muestra la similitud de los productos traducidos codificados en los plásmidos (cerrados) del microbioma de semillas de alfalfa respecto de proteínas codificadas en otros plásmidos de diferentes bacterias (columnas 4, 5 y 6).

Tabla V.4. Búsquedas de similitud de los productos traducidos codificados en los plásmidos (cerrados) del microbioma de semillas de alfalfa.

Plás-mido	Nº ORFs	1º SIMILITUD POR BLASTn	1º similitud por BLASTp	2º similitud por BLASTp	3º similitud por BLASTp	AISLAMIENTO PORTADOR (IDENTIFICACIÓN)
pBo1	222	<i>Pseudomonas putida</i> ND6, plásmido pND6-2	<i>Pseudomonas</i> (100%)	NC	NC	<i>Pseudomonas</i>
pBo2	136	<i>Arthrobacter aurescens</i> TC1, plásmido TC2	<i>Arthrobacter</i> (61,2%)	<i>Kokuria</i> (3,3%)	<i>Rhodococcus</i> (2,5%)	<i>Arthrobacter</i>
pBo3	107	<i>Sinorhizobium meliloti</i> , plásmido pHRC017	<i>Rhizobium</i> + <i>Agro</i> (63,5 %)	<i>Neorhizobium</i> (7,3%)	<i>Ochrobactrum</i> (4,2%)	<i>Rhizobium</i>
pBo4	128	<i>Methylobacterium oryzae</i> CBMB20, plásmido pMOC1	<i>Roseomonas</i> + <i>Acetobacteraceae</i> (35,9%)	<i>Skermanella</i> (9%)	<i>Methylobacterium</i> (7,7%)	<i>Roseomonas</i>
pBo5	55	<i>Psychrobacter</i> sp. PRwf-1	<i>Enhydrobacter</i> (37,7%)	<i>Paenibacillus</i> (32,1%)	<i>Actinobacillus</i> (7,5%)	<i>Moraxella</i>
pBo6	63	<i>Pantoea vagans</i> C9-1, plásmido pPag2	<i>Pantoea</i> (70,7%)	<i>Yersinia</i> (6,9%)	<i>Serratia</i> (3,4%)	<i>Pantoea</i>
pBo7	65	<i>Sphingomonas</i> sp. NS2, plásmido 201	<i>Sphingobium</i> (47,4%)	<i>Sphingomonas</i> (26,3%)	<i>Sphingopyxis</i> (3,5%)	<i>Novosphingobium</i>
pBo8	64	<i>Agrobacterium tumefaciens</i> plasmid pTi-SAKURA	<i>Rhizobium</i> + <i>Agro</i> (38,2 %)	<i>Sinorhizobium</i> (9,1%)	<i>Labrenzia</i> (9,1%)	<i>Rhizobium</i>
pBo9	63	<i>Agrobacterium tumefaciens</i> Ti, plásmido pTiBo542	<i>Rhizobium</i> + <i>Agro</i> (75,9%)	<i>Sinorhizobium</i> (7,4%)	<i>Shinella</i> (1,9%)	<i>Rhizobium</i>
pBo10	58	<i>Arthrobacter chlorophenolicus</i> A6	<i>Arthrobacter</i> (56,4%)	<i>Nesterenkonia</i> (10,3%)	<i>Microbacterium</i> (5,1%)	sin identificar
pBo11	83	sin similitud	<i>Roseomonas</i> + <i>Acetobacteraceae</i> (83,6%)	<i>Methylobacterium</i> (3,3%)	<i>Aureimonas</i> (1,6%)	<i>Roseomonas</i>
pBo12	56	<i>Starkeya novella</i> DSM 506	<i>Roseomonas</i> + <i>Acetobacteraceae</i> (95,7%)	<i>Bradyrhizobium</i> (13%)	<i>Varioborax</i> (8,7%)	<i>Roseomonas</i>
pBo13	52	<i>Micrococcus</i> sp. A1 plasmid pLMA1	<i>Micrococcus</i> (53,8%)	<i>Actinobacteria</i> (17,9%)	<i>Pseudonocardia</i> (12,8%)	<i>Micrococcus</i>
pBo14	37	<i>Sphingopyxis</i> sp. Kp5.2 plasmid pSfKp5.2	<i>Sphingobium</i> (37,5%)	<i>Sphingomonas</i> (20,8%)	<i>Sphingomonadaceae</i> (20,8%)	<i>Novosphingobium</i>
pBo15	21	<i>Agrobacterium</i> sp. H13-3 linear chromosome	<i>Agrobacterium</i> (100%)	NC	NC	<i>Rhizobium</i>

ORFs: marcos de lectura abiertos. Números entre paréntesis indican porcentajes de ORFs que presentan su mejor alineamiento por BLASTp contra una proteína secuenciada en dicho género. NC: No corresponde (valores de expect $\leq 10^{-5}$).

En primer lugar pudo verse que varios ensamblajes que por BLASTn dieron resultados positivos sólo sobre algunas regiones de su secuencia (regiones de similitud cortas en relación al tamaño total del plásmido, Tabla V.3, columna 5), pudieron ahora, por BLASTp de sus marcos de lectura abiertos, ser vinculados con mucha mejor cobertura a replicones ya reportados del banco de datos.

Los resultados, que confirman y extienden los obtenidos por BLASTn, muestran que el género con mayor similitud (columna 4) coincidió en todos los casos con el identificado para el aislamiento que porta el plásmido en nuestra colección. Los géneros que se posicionan en el segundo y tercer lugar en las búsquedas por BLASTp (columnas 5 y 6) mostraron en general mucho menos porcentaje de similitud evidenciando claramente la preferencia de huésped en los que los plásmidos han circulado y moldeado su estructura presente. En ese contexto, casi la totalidad de los marcos de lectura abiertos que dieron resultado positivo en los alineamientos ($\text{expect} \leq 10^{-5}$) lo hicieron con secuencias aminoacídicas de taxonomía conocida a nivel de género (o en su defecto a nivel de familia) sugiriendo, como hemos dicho, el tipo de bacterias en las que los plásmidos han estado presente.

Mientras que pocos plásmidos mostraron todos sus marcadores pertenecientes a un solo género (pBo1 y pBo15), algunos de los plásmidos mostraron marcadores pertenecientes a una misma familia taxonómica, como en los casos de *Arthrobacter*, *Kokuria* y *Nesterenkonia* (familia **Micrococcaceae**); *Rhizobium/Agrobacterium* y *Sinorhizobium* (familia **Rhizobiaceae**); *Sphingobium*, *Sphingomonas* y *Sphingopyxis* (familia **Sphingomonadaceae**); y *Pantoea*, *Yersinia* y *Serratia* (familia **Enterobacteriaceae**). En relación a esta observación es importante destacar que el límite de relación parece extenderse a nivel de orden/clase para varios de estos ensamblajes, con lo que resulta evidente un escenario donde los plásmidos transitan preferencialmente un contexto de género o familia, pero donde también aparece la presencia de genes de grupos taxonómicos más alejados como evidencia (ya sea de ancestralidad o) de transferencia horizontal de genes entre especies con diferente grado de vínculo taxonómico. En la Figura V.5. se presenta a modo de ejemplo el carácter mosaico de uno de los plásmidos de la Tabla V.4.

Los resultados presentados muestran, para la comunidad aquí estudiada, cómo los plásmidos representan al conjunto de información al que distintas especies (con distinto grado de cercanía) tienen acceso vía transferencia horizontal de genes. Los resultados del análisis sugieren que existe un conjunto preferencial de géneros frecuentados por un plásmido en particular, donde: a) uno de ellos se corresponde con el del huésped del que hemos recuperado el plásmido bajo análisis en nuestro microbioma de las semillas de alfalfa, y b) el resto de los géneros con secuencias similares se corresponde en varios casos con bacterias presentes dentro del mismo microbioma, **sugiriendo muy fuertemente a la simpatria (con la planta como concentrador) como un factor positivo en favor de la HGT**. Ejemplos de ellos son pares *Arthrobacter* y *Microbacterium*, *Arthrobacter* y *Kocuria*, *Rhizobium* y *Sinorhizobium*, *Enhydrobacter* y *Paenibacillus* que se presentan en las Tablas V.3 y V.4 y en las descripciones de las especies bacterianas que conforman el moviloma realizada en el Capítulo 4 precedente.

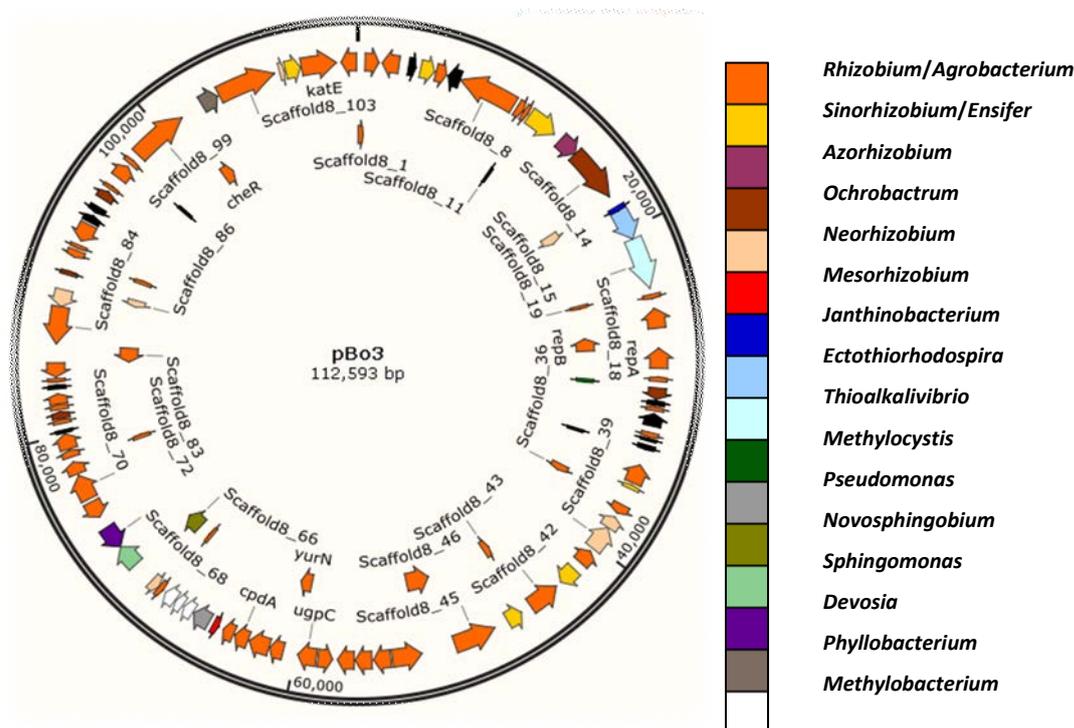


Figura V.5: Esquema del plásmido pBo3. Se muestran los marcos de lectura abiertos (ORFs) con cada color representando el género correspondiente al primer alineamiento obtenido en búsquedas BLASTp (valores de expect $\leq 10^{-5}$). En negro se muestran aquellos ORFs que no mostraron un alineamiento con dichos parámetros de búsqueda.

Ensamblados abiertos

La mayor parte de los ensamblados que obtuvimos no pudieron circularizarse en plásmidos cerrados. Esto es normal en secuenciamientos genómicos bacterianos, y en nuestro caso puede deberse a la presencia de secuencias repetitivas más largas que los insertos de la biblioteca *matepair* realizada, a limitaciones en la cobertura de la misma o bien debido a problemas de ensamblado. Como se muestra en la Tabla V.5, los ensamblados lineales, al igual que los plásmidos cerrados anteriormente presentados, mostraron diversos tamaños, coberturas de secuenciamiento y grados de similitud con otros replicones ya secuenciados.

Tabla V.5. Búsquedas de similitud de los productos traducidos codificados en los ensamblados lineales del microbioma de semillas de alfalfa. Las búsquedas por BLASTp arrojaron indicios fuertes respecto de los huéspedes por los que los plásmidos han circulado y además colectado información génica.

ENSAMBLE	nº ORFs	1º SIMILITUD POR BLASTn	1º similitud por BLASTp	2º similitud por BLASTp	3º similitud por BLASTp
scaffold00002	238	<i>Sphingobium</i> sp. plásmido SYK-6 (23%)	<i>Sphingomonadaceae</i> (43%)	<i>Novosphingobium</i> (32,2%)	<i>Sphingobium</i> (12,2%)
scaffold00004	196	<i>Methylobacterium populii</i> BJ001, plásmido pMPOP02 (2%)	<i>Sphingobium</i> (43,1%)	<i>Sphingomonas</i> (32,8%)	<i>Sphingopyxis</i> (4,6%)
scaffold00005	129	<i>Agrobacterium vitis</i> S4, plásmido pAtS4b (31%)	<i>Agrobacterium</i> (41,4%)	<i>Rhizobium</i> (27,3%)	<i>Microvirga</i> (3,1%)
scaffold00007	119	<i>Sinorhizobium</i> sp. M14, plásmido pSinA (43%)	<i>Rhizobium</i> (34,2%)	<i>Sinorhizobium</i> (14,9%)	<i>Agrobacterium</i> (13,2%)
scaffold00010	104	<i>Ochrobactrum anthropi</i> strain OAB, plásmido 1 (11%)	<i>Agrobacterium</i> (30,1%)	<i>Rhizobium</i> (27,2%)	<i>Aminobacter</i> (6,8%)
scaffold00012	83	<i>Agrobacterium vitis</i> S4, plásmido pAtS4a (7%)	<i>Rhizobium</i> (44,2%)	<i>Agrobacterium</i> (13%)	<i>Rhodobacteraceae</i> (3,9%)
scaffold00014	80	<i>Rhizobium</i> sp. LPU83, plásmido pLPU83a (40%)	<i>Rhizobium</i> (50%)	<i>Agrobacterium</i> (11,8%)	<i>Sinorhizobium</i> (10,5%)

ORFs: marcos de lectura abiertos. Entre paréntesis se indica la cobertura del alineamiento BLASTn en la columna 3 y el porcentaje de ORFs asignados a cada género por BLASTp, en columnas 4, 5 y 6.

V.4. Contenido génico y predicción funcional de la información del moviloma plasmídico endófito

Como se realizó con el moviloma plasmídico de *Sinorhizobium meliloti* (MPS_{me}), los marcos de lectura abiertos detectados por la anotación automática de la plataforma GenDB fueron reanotados manualmente, con los que luego se confeccionó una lista de funciones por categorías en todos los ensamblajes y cóntigos que conforman el moviloma plasmídico de bacterias provenientes de semillas de alfalfa (MPSA). La Figura V.6 muestra dicha distribución funcional comparada con la del MPS_{me} previamente secuenciado.

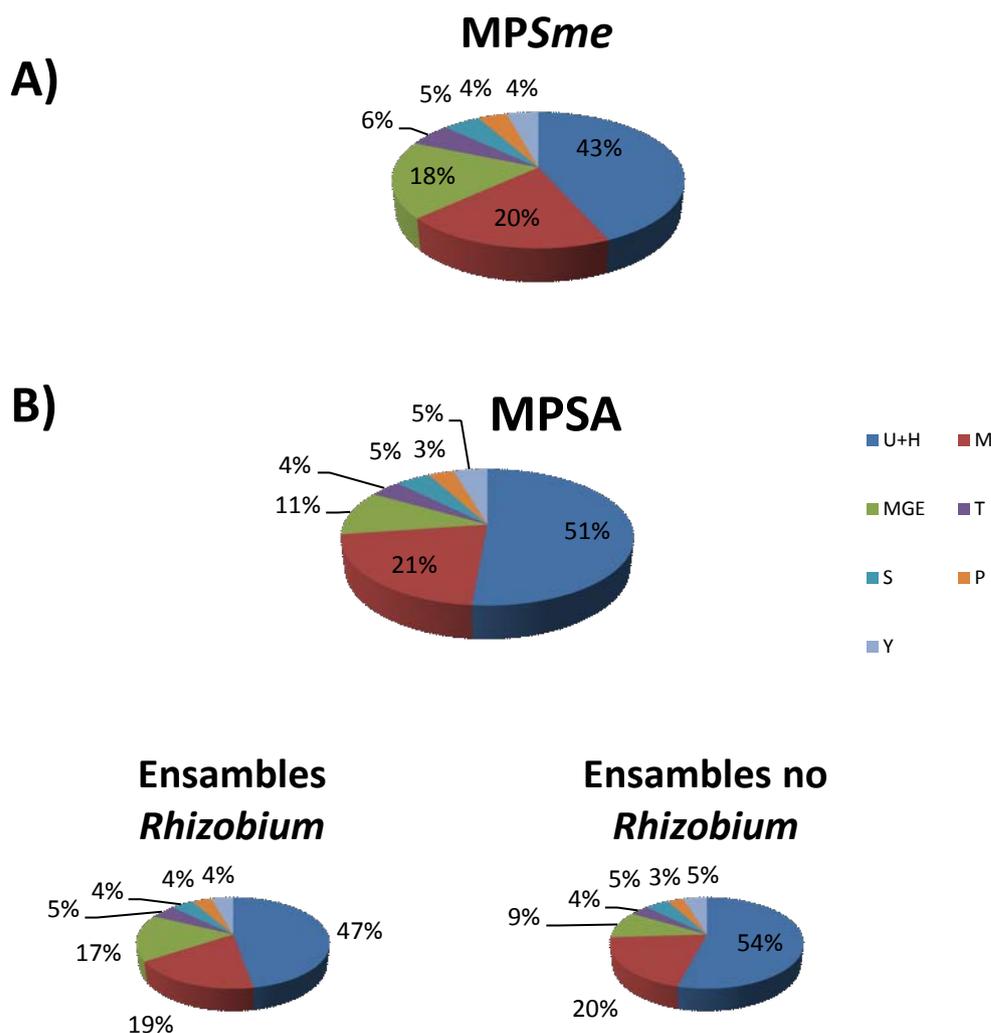


Figura V.5. Comparación de las predicciones funcionales asociadas a los contenidos génicos de los movilomas plasmídicos de *S. meliloti* (MPS_{me})(parte A, capítulo precedente) y de las bacterias asociadas a semillas de alfalfa (MPSA)(parte B, este capítulo). En la parte B se discrimina la composición funcional atendiendo a que se trate de rizobios o de no rizobios. Referencias: U (proteína hipotética), H (proteína hipotética con homólogos en otros organismos, de función desconocida), M (función metabólica general), F (gen relacionado a metabolismo de bacteriófagos), T (función de transporte de metabolitos), R (regulador transcripcional), A (procesamiento del ADN).

Como puede apreciarse para la mayoría de las categorías funcionales el porcentaje es similar entre ambos movilomas. Las diferencias principales se encuentran en la fracción de elementos

genéticos móviles (conjugación, transposición y elementos relacionados a fagos) y en la fracción de genes sin función asignada (U + H), los primeros más abundantes en el MPSme (18% vs 11%) y los segundos en el MPSA (51% vs 43%). Es particularmente interesante que al distinguir dentro del MPSA los plásmidos provenientes de *Rhizobium* de los de otro origen, la composición de los primeros (Ensamblados *Rhizobium*, representando el 35% de los marcos de lectura abiertos del MPSA) se aproxima mejor a la composición funcional del del MPSme (con 17% asignado a elementos genéticos móviles), sugiriendo un vínculo por transferencia lateral entre los plásmidos de los sinorizobios y los rizobios asociados a las semillas.

V.5. Evaluación de relaciones de ortología entre el MPSA y el MPSme

Como ya hemos mencionado en el análisis de MPSme, una de las formas de investigar relaciones filogenéticas entre dos conjuntos génicos (genomas, grupos de genomas) es la evaluación de la proporción de ortologías entre los mismos. En nuestro caso, empleando nuevamente el software EDGAR [268] que busca ortólogos sobre la base de las mejores similitudes de modo bidireccional (expect < 10^{-5}), evaluamos el grado de ortología de diferentes plásmidos/ensamblados del MPSA y también de movilomas plasmídicos del GenBank (consolidados por género de las bacterias huésped de los replicones)(eje "x" en la Figura V.7) respecto del MPSme. El resultado del análisis lo expresamos como la proporción (porcentual) de genes del plásmido/ensamblado/moviloma de GenBank para los que encontramos ortólogos en el MPSme (índice OPCG en el eje "y" de la Figura V.7). Resultó particularmente interesante que al comparar los diferentes ensamblados/plásmidos del MPSA con el MPSme, los pertenecientes a aislamientos del género *Rhizobium* (barra roja del extremo derecho) son los que portaron mayor proporción de ortólogos (mayor OPCG) respecto al MPSme. Uno de cada cuatro genes en los plásmidos de *Rhizobium* provenientes de las semillas de alfalfa mostró tener un ortólogo en el MPSme, hallazgo indicativo de una relación evolutiva cercana entre ambos movilomas. La relación de ortología entre genes de los plásmidos en ambos movilomas (más de 25%) tiene desde luego un vínculo ancestral asociado a la filogenia cercana de los géneros *Rhizobium* y *Sinorhizobium*. Sin embargo, debe tenerse en cuenta que la relación de ortología de otros plásmidos como de rizobios y agrobacterias fueron todos inferiores al 25% observados para los rizobios del MPSA (Figura V.7), sugiriendo que efectivamente los plásmidos de los rizobios presentes en semillas de alfalfa pueden tener un grado mayor de intercambio lateral de información con *S. meliloti*.

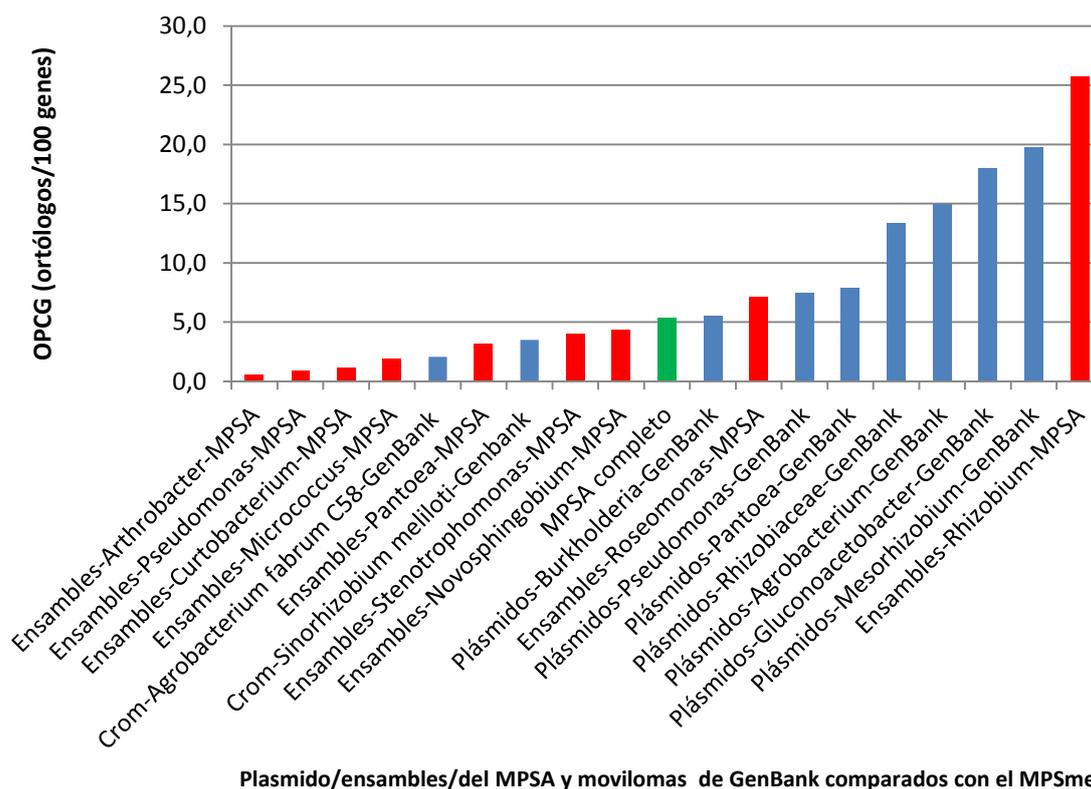


Figura V.7. Porcentaje de ortólogos entre los ensambles/plásmidos del MPSA o diferentes movilomas del GenBank (consolidados por género bacteriano portador de los replicones) y el MPSme. En el eje “x” se listan los plásmidos/ensambles del MPSA (barras rojo), el MPSA completo (verde), y los movilomas reconstruidos a partir de plásmidos depositados en el GenBank (barras azules); que en cada caso fueron comparados en búsqueda de ortólogos contra el MPSme, empleando la herramienta EDGAR. En el eje “y” se muestra el valor del OPCG resultante (número de ortólogos por cada cien genes del grupo analizado del eje “x”).

Distintas evidencias taxonómicas que hemos obtenido para estos rizobios de semillas (fragmentos de secuencias cromosomales, secuencias ARNr 16S y tipificaciones por espectrometría de masas UV-MALDI-TOF) nos han indicado que estos aislamientos tienen cercanía a *Rhizobium radiobacter*, especie sinónima de *Agrobacterium tumefaciens*. Es posible que en su composición plasmídica estos aislamientos estén más adaptados al intercambio de genes con *S. meliloti*, comparados con otros *A. tumefaciens* depositados en las bases de datos. Será importante analizar el tipo de cromosomas que portan los mismos, compararlos con las estructuras ya disponibles para las cepas “clásicas” de *A. tumefaciens*, y evaluar si existen sesgos particulares asociados a unos y a otros, atendiendo a que podría tratarse de ecotipos diferentes.

En contraste con las evidencias anteriores, los ensambles que presentan similitud o que pudieron identificarse por PCR en aislamientos de géneros tales como *Pantoea*, *Pseudomonas*, *Roseomonas*, *Arthrobacter*, *Novosphingobium*, *Stenotrophomonas*, *Micrococcus* y *Curtobacterium*, mostraron no compartir grandes proporciones de ortólogos con el MPSme, consolidando evidencias en apoyo de que la transferencia horizontal de genes de *S. meliloti* con bacterias asociadas a la planta se da especialmente con representantes de la misma familia, y posiblemente -por lo que hemos dicho precedentemente- con variantes particulares

de las mismas dentro de Rhizobiaceae. Una limitación en esta comparación de ortólogos fue la cantidad de genes disponibles para cada género en nuestro MPSA, ya que mientras que en el caso de los ensamblajes de *Rhizobium* se contó con más de 1.500 marcos de lectura abiertos, para los demás géneros estuvieron sólo en el orden del centenar de genes. De ahí la posibilidad de que el número de OPG en el género *Pantoea* sea tan diferente al calculado con todos los plásmidos presentes en GenBank para este género.

V.6. Análisis de secuencias cromosomales presentes como subproducto de la secuenciación masiva de los plásmidos crípticos

Un total de 1.160 cóntigos de baja cobertura de secuenciamiento, que no pudieron incorporarse a ningún ensamble, correspondieron a secuencias de origen cromosomal que contaminaron la preparación de plásmidos. Estos cóntigos, cuyos tamaños oscilaron entre 100 pb y 2.055 pb, nos permitieron acceder al compartimento cromosomal de los aislamientos portadores de los plásmidos. Además, recuperamos entre ellos secuencias de plásmidos que, estando presentes en baja representación en la muestra secuenciada, no pudieron identificarse debido a los criterios de selección manual por cobertura o al ensamblado automático. Estos cóntigos fueron alineados por BLASTn contra la base de datos del GenBank, seleccionando luego el replicón con el cual presentaban mayor E-value. El 83,4 % de estos cóntigos presentaron un alineamiento al genoma blanco con una cobertura mayor que 95%, es decir que en su mayoría se trató de cóntigos de replicones más conservados (cromosomas, megaplásmidos), que presentan de todos modos variaciones de secuencia con respecto a las cepas de referencia depositadas en el GenBank. Entre estos cóntigos **77% resultó tener su mayor similitud con algún replicón cromosomal**, o cromosoma secundario (como cromosomas lineales con similitud al de *Agrobacterium*), mientras que el restante **23% mostró mayor similitud con secuencias plasmídicas**. De entre estos últimos cóntigos con similitud a plásmidos, 33% presentó similitud a plásmidos mayores a 500 kpb -entre los cuales se encuentran diferentes plásmidos de la familia *Rhizobiaceae*, y 89% similitud a plásmidos mayores que 100 kpb. El resultado indicó que, en general, los plásmidos que aparecieron en la muestra con baja representatividad correspondieron a replicones de tamaños grandes (posiblemente con números de copias bajos por célula). La información de los cóntigos con similitud a secuencias cromosomales que comentamos precedentemente nos brindó un mayor número de marcadores, distribuidos al azar a lo largo de sus respectivos replicones, que confirmaron las identificaciones de los huéspedes bacterianos realizadas por secuenciamiento parcial de ARNr 16S y/o por UV-MALDI-TOF. La tabla V.6 muestra la distribución de cóntigos de acuerdo a su similitud con diferentes cromosomas de referencia. La tabla V.7 muestra la distribución cuando el mejor alineamiento de los cóntigos fue con algún plásmido de referencia.

Tabla V.6. Alineamiento de cóntigos de baja cobertura con cromosomas de referencia

Cromosoma de referencia del GenBank (1º alineamiento)	Nº de cóntigos asignados a la referencia	% de identidad (promedio)	% cobertura del alineamiento (promedio)
<i>Rhizobium</i> sp. IRBG74 cromosoma circular	342	98	99
<i>Rhizobium</i> sp. IRBG74 cromosoma lineal	342	97	99
<i>Agrobacterium tumefaciens</i> str. C58 linear chromosome	19	89,1	88
<i>Agrobacterium tumefaciens</i>, cepa Ach5 cromosoma lineal	14	92	95
<i>Agrobacterium tumefaciens</i>, cepa Ach5 cromosoma circular	11	92	96
<i>Agrobacterium</i> sp. H13-3 cromosoma lineal	10	89.2	92,6
<i>Agrobacterium tumefaciens</i>, cepa C58 cromosoma circular	7	91,5	98,7

Como podemos observar la mayor parte de las secuencias cromosomales parecen provenir de aislamientos similares a *Rhizobium* sp. IRBG74, con cóntigos similares al cromosoma circular y otros al lineal. Esta cepa secuenciada en 2013 [395] ha sido obtenida de nódulos de una leguminosa acuática del género *Sesbania* [396] y ha mostrado nodular al menos 8 especies de este género [397]. La misma cepa mostró capacidades de promoción del crecimiento en arroz, y se ha propuesto como modelo para estudiar relaciones simbióticas de bacterias endófitas en cereales [398]. Es particularmente interesante que aislamientos con genotipos parecidos a esta cepa con capacidad de promoción de crecimiento hayan sido encontrados en las semillas de alfalfa. La cepa secuenciada y depositada en el GenBank tiene genes plasmídicos *nifH* y *nodA* y se propone como una nueva especie debido a que, a pesar de su similitud a nivel de secuencia genómica con *Rhizobium radiobacter* (*Agrobacterium tumefaciens*), su porcentaje de hibridación ADN:ADN no supera el 44% y no tiene capacidad de formar tumores ni posee el operón *vir*. Será importante explorar si las cepas con similitud a IRBG74 que están en el microbioma de semillas tienen o no el operón *vir* y si son o no virulentas, así como si poseen o no genes asociados a la fijación de nitrógeno. Es posible que, dada la elevada cantidad de cóntigos cromosomales con similitud a esta cepa, nuestros 4 aislamientos pertenecientes al género *Rhizobium* posean un cromosoma semejante a la misma, difiriendo principalmente entre sí en sus perfiles plasmídicos. Además de investigar si son o no virulentas y poseen genes *vir*, será interesante conocer si estos aislamientos endófitos obtenidos tanto de la raíz como de parte aérea de la plántula de alfalfa presentan los genes de nodulación o fijación de nitrógeno y estudiar qué capacidades de promoción del crecimiento presentan en alfalfa y cuál es su relación con el simbionte *S. meliloti* dado que sus plásmidos (Ensamblados *Rhizobium*) muestran una clara relación de ortología con el MPS_{me}.

Tabla V.7. Alineamiento de cóntigos de baja cobertura con plásmidos de referencia

Plásmido de referencia en GenBank	Tamaño referencia	Nº de cóntigos	% de identidad (promedio)	% de (cobertura promedio)
<i>Rhizobium</i> sp. IRBG74 plásmido IRBL74_p	584.965	40	93	92
<i>Pseudomonas putida</i> , cepa DOT-T1E plásmido pGRT1	133.451	39	95	100
<i>Rhizobium leguminosarum</i> bv. trifolii CB782 plásmido	1.563.772	15	87	96
<i>Pantoea vagans</i> C9-1 plásmido pPag1	167.983	13	85	83
<i>Agrobacterium tumefaciens</i> , cepa C58 plásmido At	542.868	12	92	92
<i>Ochrobactrum anthropic</i> , cepa OAB plásmido 1	155.838	10	91	95
<i>Agrobacterium</i> sp. H13-3 plásmido pAspH13-3a	601.551	8	90	96
<i>Pseudomonas mandelii</i> JR-1 plásmido	410.512	6	97	100
<i>Agrobacterium vitis</i> S4 plásmido pAtS4c	21.162	6	86	62
<i>Micrococcus</i> sp. 28 plásmido pSD10	50.709	5	98	91
<i>Pseudomonas aeruginosa</i> PAO1 plásmido pAMBL1	2.644	5	99	100
<i>Pseudomonas putida</i> ND6 plásmido pND6-2	11.7003	4	99	98

El 23% de los cóntigos que aun teniendo baja cobertura tienen similitud nucleotídica con diferentes plásmidos pueden corresponder a: 1) plásmidos con bajo número de copias y tamaño grande que no fueron purificados de modo eficiente con el procedimiento que hemos utilizado, o 2) secuencias plasmídicas integradas a cromosomas, una posibilidad esta última que deberá ser más explorada con herramientas diferentes (ej. secuenciación genómica de las cepas)

En este capítulo hemos comenzado a abordar el compartimento plasmídico de las bacterias que están asociadas naturalmente a las semillas de alfalfa, y que muestran capacidad de colonizar la planta y posiblemente de tener influencia en el crecimiento o el biocontrol de la planta, tal como se mostró en el Capítulo IV. Pudimos obtener una gran cantidad de secuencias plasmídicas y comprender que la evolución de los plásmidos implica el intercambio a lo largo de su historia con una gran diversidad de especies, en general cercanas taxonómicamente a las especies donde se los aíslan pero con algunos claros ejemplos que nos sugieren que la simpatria favorece la HGT entre las especies que forman el microbioma de alfalfa. Asimismo pudimos establecer que una gran cantidad de los plásmidos que secuenciamos se encuentran contenidos en varios aislamientos del género *Rhizobium*, cuyos movilomas plasmídicos parecen tener una relación especial de coevolución con el moviloma plasmídico de *S. meliloti* analizado en el Capítulo III. En este sentido la semejanza de estos aislamientos a una cepa secuenciada recientemente, no patógena y perteneciente al complejo *Rhizobium/Agrobacterium* nos sugiere la posibilidad de que estos aislamientos formen parte de una adaptación de un miembro de su microbioma y será interesante avanzar en la caracterización de dicho vínculo en relación al rol que desempeñan estas especies en la planta

y el rol de sus plásmidos crípticos. Asimismo, el moviloma plasmídico de varios géneros pertenecientes al microbioma (algunos de cuyos plásmidos hay muy poca información disponible) nos da la oportunidad de avanzar puntualmente en genes que codifiquen funciones de interés que puedan tener relevancia ecológica o práctica en el desarrollo de bioinoculantes.

Conclusiones y perspectivas

El trabajo presentado en esta tesis doctoral se propuso como objetivo central abordar con herramientas genómicas la caracterización de la información contenida en plásmidos crípticos de bacterias que se asocian naturalmente a una planta modelo de relevancia agronómica como lo es alfalfa. *Sinorhizobium meliloti*, rizobio que se asocia con alfalfa para dar lugar a nódulos fijadores de nitrógeno, porta con frecuencia un número variable de plásmidos crípticos además de un cromosoma y dos megaplásmidos (pSymA y pSymB), estos últimos imprescindibles para la simbiosis. Tal constitución genómica proveyó un sistema muy conveniente para el estudio básico de un moviloma plasmídico incluyendo el análisis de su tamaño, diversidad, características de movilización, rol en la evolución genómica de los rizobios, funciones derivadas de su carga génica, y vínculo con movilomas plasmídicos de otras bacterias que cohabitan el ambiente del rizobio entre otros aspectos. En tal sentido, y con este objetivo a mediano plazo, nuestro laboratorio ya había caracterizado en la década precedente la diversidad y características de movilización conjugativa de los plásmidos presentes en una colección local de aislamientos de *S. meliloti* que habíamos constituido en nuestro laboratorio (aislamientos que representaron 22 OTUs/grupos diferentes de diversidad plasmídica)[1]. A través de experimentos clásicos dicho análisis reveló una marcada diversidad de plásmidos en la colección de rizobios, además de una destacada capacidad de transferencia conjugativa en varios de ellos. Tal caracterización sirvió como punto de partida para abordar el estudio genómico molecular de un moviloma plasmídico cuyo tamaño excede la megabase, en una colección tamizada de bacterias de una misma especie y de relevancia agronómica. El secuenciamiento de los plásmidos crípticos que se presentó en el Capítulo 3 de esta tesis nos permitió obtener más de 1,4 Mb de información exclusivamente plasmídica de *S. meliloti*, la cual codifica genes que mostraron un perfil funcional en base a COGs claramente distinto del correspondiente a los genes de los otros replicones presentes en *el rizobio* (pSymA, pSymB y cromosoma), con predominio de funciones asociadas al tráfico y secreción (clase COG "U"), y de funciones -como se espera- relacionadas al mantenimiento y estabilidad de plásmidos (clases COG "L" y "D"). Más allá de estas diferencias, el análisis multivariado que hemos utilizado para comparar el tipo de funciones presentes en cada replicón mostró que, en relación a los otros replicones, el moviloma plasmídico presenta mayor similitud funcional al megaplásmido simbiótico pSymA; en particular en las categorías COG relacionadas a transporte y metabolismo de iones inorgánicos, transporte y metabolismo de carbohidratos, metabolismo de lípidos, biogénesis de membrana celular y pared celular, transporte y metabolismo de nucleótidos y funciones de recambio proteico, modificación post traduccional y chaperonas (clases P, G, I, M, F y O, respectivamente) en comparación a los otros dos replicones (pSymB y cromosoma). Por otra parte, la disponibilidad de la secuencia de siete genomas completos de *S. meliloti* por un lado, y la del moviloma plasmídico que presentamos en esta tesis por el otro; nos permitió investigar -con distintas herramientas- vínculos evolutivos entre los plásmidos crípticos y los diferentes replicones del rizobio. Para dichos análisis realizamos comparaciones basadas en el contenido GC y en el uso de codones de diferentes fracciones génicas, incluyendo los replicones completos, los singletons (genes presentes en un dado replicón en sólo una de las 7 cepas), y en el genoma core (fracción de genes con homólogos presentes en las 7 cepas). Pudimos así poner en evidencia una estrecha relación entre los genes codificados en los plásmidos crípticos y los genes presentes en la fracción singletons de cada uno de los replicones presentes en *S. meliloti*. Según discutimos en el Capítulo 3 de esta tesis dicho vínculo entre genes del moviloma y las fracciones

singletons resultó evidente tanto a nivel de sus composiciones GC% como a través de las frecuencias de uso de codones (parámetros que por otra parte resultaron diferentes a las de las fracciones core de esos mismos replicones). Tales evidencias apoyan fuertemente la hipótesis de que la mayoría de los genes en la fracción singletons (tanto de los pSyms como del cromosoma) podrían haberse adquirido –en un tiempo evolutivo relativamente recientemente- desde los plásmidos crípticos (subyace la idea que los genes singletons están presentes en una cepa y no en otra, como producto de adquisiciones horizontales recientes por parte de un rizobio y no de los otros generando diversidad entre los replicones). De acuerdo a dicho modelo de flujo génico, es esperable que dentro del moviloma puedan encontrarse homólogos (como la fuente originaria) de los singletons de los distintos replicones. Congruente con este modelo identificamos en efecto homólogos al MPSme en las distintas fracciones singletons de los diferentes replicones, mostrando ambas versiones de los genes un cercano uso de codones como resultaría esperable. **Las evidencias provistas fueron así consistentes con la existencia de un flujo génico desde los plásmidos crípticos hacia las fracciones singleton de los megaplásmidos y cromosoma.** Pudimos observar asimismo que los singletons para el caso del pSymB y del cromosoma –probables representantes como hemos dicho de genes transferidos en tiempos recientes desde el MPSme-, predominan funciones relacionadas a fagos y transposasas (elementos móviles) así como genes de función desconocida. En contraposición con esta observación, los homólogos recientemente adquiridos por el pSymA desde los plásmidos crípticos mostraron un perfil funcional más amplio que incluía genes metabólicos como grupo principal, así como genes regulatorios y de transporte además de los genes relacionados a movilidad y de función desconocida. **Esta diferencia funcional característica de los singletons del pSymA homólogos al MPSme fue también observada en la totalidad de los genes singletons en el pSymA, y sumado a la observación de una mayor densidad de genes singletons en el pSymA, su conjunto core relativamente más acotado y su perfil funcional en base a COGs más disperso de las distintas cepas incluídas en el análisis en comparación al pSymB y cromosoma, nos sugirió que los plásmidos crípticos transfieren (o han transferido) genes con mayor frecuencia al pSymA, aportando mayor diversidad y plasticidad funcional a la bacteria.** En relación a esta observación será interesante explorar la posibilidad de que el pSymA pueda ser producto de la incorporación sucesiva de secuencias provenientes de diferentes plásmidos crípticos a lo largo de su formación como megaplásmido simbiótico, siendo el mismo más permisivo a la incorporación de genes desde los plásmidos, quizás debido a un rol menos esencial para la sobrevivencia de la célula o a una mayor presencia de elementos recombinogénicos.

Al analizar con un criterio similar lo que sucedía en la totalidad de los genes core en el pSymA, pSymB y cromosoma pudimos observar que los mismos presentan un uso de codones diferentes entre sí, sugiriendo que la evolución de cada uno de los megaplásmidos se dio en diferentes huéspedes antes de su incorporación en *S. meliloti*. Apoyándonos en la evidencia aceptada que indica que el uso de codones se moldea de acuerdo a factores intrínsecos de cada huésped bacteriano, el análisis sobre los genes homólogos al MPSme en las fracciones core de cada replicón (es decir, genes con un predecesor común por ser homólogos, y que además por ser del core corresponden a los más ancestrales) mostró que **los genes homólogos al MPSme presentes en el pSymA y pSymB debieron transferirse desde (¿o hacia?) los plásmidos crípticos ancestrales antes de la adquisición de los mismos por parte de *S. meliloti*.**

Dicho conjunto debe estar constituido por funciones presentes desde antes de la incorporación de los megaplásmidos a *S. meliloti*, y que han sido preservadas a lo largo de la evolución de la especie seguramente por su importancia para sostener la competencia y adaptabilidad (“*fitness*”) a los nichos que ha transitado el rizobio. Para intentar avanzar en esta dirección será interesante investigar el perfil funcional del conjunto completo de los genes core de los pSyms, y en particular el tipo de funciones del sub-conjunto de genes cuya relación con los genes móviles hemos podido identificar en esta tesis (los homólogos al MPS_{me}).

El análisis de los homólogos al MPS_{me} en las fracciones core también nos permitió observar un corrimiento gradual en el uso de codones de este conjunto de genes con respecto al del conjunto core completo, que en el caso del cromosoma resultó muy evidente. **La observación de que un corrimiento similar se observa en los conjuntos de genes del core cromosomal que son homólogos a cromosomas de especies cada vez más alejadas taxonómicamente, nos permitió inferir que los genes homólogos al MPS_{me} en el cromosoma constituyen posiblemente un subconjunto de genes muy ancestrales dentro del core, adquiridos/compartidos con un moviloma ancestral en los estadios primordiales de la constitución del cromosoma como replicón.** El análisis que hemos realizado nos permitió además comenzar a explorar qué tipo de funciones fueron intercambiadas ancestralmente en la formación del cromosoma, y hoy persisten en el compartimiento plasmídico móvil (el MPS_{me}). De un modo más general, a través del análisis de los homólogos al MPS_{me} en las distintas fracciones core (cromosoma y megaplásmidos) será interesante comparar con detenimiento el tipo de patrón/es funcional/es (actividades particulares/metabolismos/rutas involucradas) asociados a las características más históricas (iniciales) de cada uno de los tres replicones. Será también interesante evaluar cómo fue este intercambio en otras especies de cara a una mayor comprensión del rol que el moviloma plasmídico pueda tener en la evolución de los genomas bacterianos. Los resultados que presentamos en esta parte del trabajo nos han permitido no sólo obtener información de los genes que son parte del moviloma plasmídico accesible a la especie, sino también nuevas evidencias sobre la conformación de los diferentes replicones

El análisis precedente que hemos realizado se enfocó esencialmente en el empleo de la secuencia génica del MPS_{me} para avanzar en el análisis de su vínculo con la evolución de los diferentes replicones de *S. meliloti*. En un análisis subsiguiente, y con el propósito de explorar relaciones del MPS_{me} con genomas de otras bacterias de suelo, investigamos vínculos de ortología con (plásmidos y cromosomas de) bacterias con distinto grado de cercanía a los rizobios. El análisis de genes homólogos entre el MPS_{me} y otros movilomas plasmídicos y cromosomas de especies cercanas a *S. meliloti* mostró que los plásmidos presentes en otras especies noduladoras de plantas eran los que presentaban mayor porcentaje de ortología al moviloma aquí caracterizado. Tal observación era esperable si el MPS_{me} arrastra genes ancestrales comunes a las distintas especies de rizobios (menos presentes en otras), así como genes que se comparten por transferencia horizontal (ingresan y se establecen) con más frecuencia entre rizobios como grupo (aunque heterogéneo) taxonómicamente relacionado. Géneros como *Ochrobactrum* y especies como *S. fredii*, por ejemplo, muestran plásmidos con alta densidad de ortólogos al MPS_{me}. Será interesante investigar el tipo de funciones que codifican los ortólogos identificados en bacterias noduladoras, para analizar sus funciones y estudiar en qué proporción se relacionan a la vida libre o a la simbiótica de este grupo de

bacterias. Con las funciones de mantenimiento concentradas en el cromosoma y las simbióticas especialmente en los dos megaplásmidos (pSyms), podrá usarse la secuencia del MPS_{me} para avanzar en entender a qué funciones y aspectos de la vida de los rizobios se vincula la información del compartimento móvil, y qué proporción de ella es preservada en ortólogos que traspasan la barrera de género y especie en moviloma de otras bacterias noduladoras (y no noduladoras). Resultó particularmente interesante observar que varios genes del MPS_{me} para los que (¿aun?) no encontramos homólogos en plásmidos, sí observamos ortólogos en algunos replicones cromosomales (ej. *M. loti*, *S. medicae*, *M. australicum*, *E. adhaerens*) poniendo nuevamente en evidencia intercambios de genes cromosomas-plásmidos (según ya hemos analizado en el estudio de las fracciones core y singletons del cromosoma de *S. meliloti*) y proveyendo también indicios respecto de las especies por las que el MPS_{me} ha circulado.

Según hemos propuesto en los objetivos de este trabajo de tesis, el aislamiento, secuenciamiento y caracterización del MPS_{me} estaría dirigido no sólo a caracterizar su contenido génico/funcional y participación en la dinámica de la evolución genómica de la especie, sino a indagar en el grado de relación de dicho moviloma con otros rizobios (párrafo precedente) y aun bacterias de otros géneros pero que viven en simpatria con *S. meliloti* vinculadas de algún modo a las plantas de alfalfa. En capítulo IV de esta tesis hemos descrito en detalle la estrategia utilizada para el aislamiento y caracterización de bacterias asociadas a plantas de alfalfa. Más allá del tipo de bacterias que se asocian con la planta en condiciones de suelo, en este trabajo nos hemos centrado en el análisis de aquellas bacterias que están íntimamente asociadas a las semillas y protegidas del entorno (resistentes a la esterilización superficial de las mismas) y que por tanto conllevan con seguridad largos períodos de co-evolución junto a la planta y a los rizobios homólogos. La búsqueda de bacterias en semillas esterilizadas en superficie y en plántulas derivadas de ellas resultó en el aislamiento e identificación de un conjunto de cerca de 40 géneros bacterianos diferentes. Varios de los géneros que hemos observado en la variedad de semillas Super Monarca pudieron ser también aislados de otras variedades, sugiriendo que la diversidad bacteriana asociada de manera muy íntima a las semillas de alfalfa es un hecho más general. La presencia de bacterias dentro de la semilla, la conservación de varios de esos géneros entre variedades, y la expresión de actividades con capacidades de promoción directa del crecimiento de plantas y de control biológico por parte de varias de las bacterias asociadas a semillas fueron todas evidencias que sugirieron la posible presencia de un proceso de coevolución microbiota-huésped orientado a la selección y preservación en el tiempo de aquellas bacterias que mejoran la nutrición y sanidad de la planta, las que por su localización en semillas estarán disponibles en la cercanía de la raíz durante los primeros estadios posteriores a la germinación. En un contexto tal será interesante investigar cuáles de los géneros que caracterizamos en el Capítulo IV tienen capacidad de perpetuarse en sucesivas generaciones de la planta para formar parte del microbioma más estable asociado a la especie, para analizar luego si lo mismos son coincidentes con aquellos cuya presencia ha sido más ubicua en las variedades que hemos analizado. Atendiendo a todas estas consideraciones que presentan al microbioma de semillas de alfalfa como representante cierto de la comunidad que vive en simpatria con *S. meliloti* (y probablemente desde larga data), en el Capítulo V y como parte de los objetivos de este trabajo nos abocamos al aislamiento, secuenciamiento y análisis del moviloma de las bacterias

asociadas a semillas y a la evaluación posterior de sus relaciones con el moviloma de los rizobios que hemos presentado en el Capítulo III. Dicho análisis nos permitió conocer y explorar vínculos entre conjuntos de genes asociados a dos movilomas plasmídicos diferentes: el de *S. meliloti* por un lado, y el aislado de un conjunto de bacterias simpátricas íntimamente asociadas a la misma planta hospedadora. El secuenciamiento de un pool plasmídico proveniente de 18 aislamientos que presentaron diferentes perfiles de plásmidos presentes en aislamientos de al menos 12 géneros diferentes (de los más de 40 que hemos reconocido como colonizadores de semillas), nos brindaron 3,34 Mpb de secuencia no redundante de ADN incluyendo 15 plásmidos cerrados. Al evaluar en primer lugar las relaciones de ortología entre los dos movilomas plasmídicos secuenciados en este trabajo (MPS_{me} y sus bacterias simpátricas en el entorno del hospedador), **podimos poner en evidencia que los ensamblajes pertenecientes a bacterias del género *Rhizobium* incluyen una gran proporción de genes ortólogos al MPS_{me}, hecho que pone en evidencia el intercambio de genes entre *S. meliloti* y otros rizobios simpátricos.** Por otra parte, y en consonancia con la observación precedente, el análisis comparativo de secuencias nos ha permitido poner también en evidencia los conjuntos preferenciales de géneros bacterianos frecuentados por un plásmido en particular, donde uno de ellos se correspondió en varios casos con el género del huésped del que hemos recuperado el plásmido y el resto con géneros que en varios casos coincidieron con los de bacterias presentes dentro del microbioma asociado a alfalfa. Esta observación sugirió fuertemente que **la vida simpátrica en las cercanías de la planta muy posiblemente ha favorecido la transferencia horizontal de genes cruzando en varios casos la barrera de género.**

La estrategia empleada en este trabajo de tesis ha permitido obtener y abordar el análisis genético molecular de más de 4,5 Mpb de secuencia de ADN plasmídico recuperado de diferentes aislamientos de la comunidad de bacterias asociadas a semillas y plantas de alfalfa (el equivalente a más de un genoma de *E. coli*). De dicha información, cerca de 1,5 Mpb correspondieron sólo a plásmidos de *S. meliloti*, rizobios modelo de estudio y simbiontes de alfalfa. La cantidad de información residente en plásmidos que hemos podido coleccionar para dicho rizobio permitió, por primera vez, analizar en un rizobio (y en una bacteria de suelo) a escala de múltiples cepas, la estructura del compartimento plasmídico críptico, su contenido informacional (funcional), y emplearlo como herramienta para analizar aspectos relativos a la evolución y conformación en el tiempo de los otros replicones presentes en la misma especie (intercambio de genes con el moviloma a través de la búsqueda de homólogos al moviloma en fracciones core y singletons, plasticidad de los replicones, y evolución en el uso de codones en distintos conjuntos génicos entre otras cuestiones). La evidencia coleccionada según ya hemos marcado presenta al pSymA como el replicón más vinculado al moviloma, según lo evidencia su mayor contenido de singletons homólogos a genes del MPS_{me} comparado con los otros replicones, que mostraron además una gran proporción de genes metabólicos. Los resultados presentan al moviloma como una fuente importante diversa de nuevas funciones para los rizobios, que más allá de su presencia en los propios plásmidos crípticos, su adquisición de manera más estable puede ser alcanzada por incorporación a un replicón más ubicuo como el pSymA. Atendiendo a que el moviloma críptico transita otras rizobiáceas y no rizobiáceas (análisis de homólogos), el tránsito MPS_{me}-pSymA aparece como una estrategia potencialmente importante para la colección y ensayo de funciones presentes en otras bacterias alcanzadas por el mismo moviloma (valiosa por provenir de organismos simpátricos,

y por tanto portadores de “experiencia” adaptativa al mismo nicho y susceptible de ser compartida).

La estrategia y resultados obtenidos han servido para conocer el contenido informacional de movilomas plasmídicos relacionados (por la simpatría de sus huéspedes); y en base a ello avanzar -con un nuevo enfoque- en la genómica de la evolución de los replicones de un rizobio modelo. Con la información que ya hemos colectado del MPSme y con la ampliación del mismo a través del secuenciamiento completo masivo de todas las cepas de la colección (en curso) esperamos avanzar, con un mayor número de genes, en el análisis de aspectos evolutivos de cada uno de los replicones, obtener la estructura cerrada de diferentes plásmidos para conocer la redundancia (informacional) y estructural de los mismos, y para caracterizar con nuevas herramientas informáticas y luego bioquímicas (húmedas) el tipo de actividades asociadas a diferentes genes metabólicos del genoma plasmídico móvil.

Anexo

Tabla A1

Criterio utilizado para la asignación de cóntigos pertenecientes a plásmidos crípticos

Se tomaron los cóntigos cromosomales listados a continuación y se calculó su cobertura como (nº de lecturas/largo de cóntigo) x 1000. Se calculó el promedio y la desviación estándar y en base a las diferencias de cobertura observadas se fijó como valor de corte 65 lecturas/1000 pb de cara a decidir si un cóntigo era de origen plasmídico.

replicón	cromosoma	pSymA	pSymB
promedio	49	55	49
desviación estándar	7,4	5,5	4,9
2 desviaciones estándar (2DE)	14,8	11,0	9,8
promedio + 2DE	64,0	65,8	59,2
<u>valor de corte</u>	65		

Nº cóntigo	Largo (L, pb)	Número de lecturas (R)	Cobertura (R/Lx1000)	Origen del cóntigo
contig00005	28280	1413	50	cromosoma
contig00009	24671	1256	51	cromosoma
contig00017	20125	1176	58	cromosoma
contig00020	19642	1149	58	cromosoma
contig00021	19397	935	48	cromosoma
contig00022	19220	1110	58	cromosoma
contig00023	19073	1030	54	cromosoma
contig00027	18527	991	53	cromosoma
contig00028	18473	832	45	cromosoma
contig00032	18081	1004	56	cromosoma
contig00034	17911	979	55	cromosoma
contig00035	17807	744	42	cromosoma
contig00039	17054	975	57	cromosoma
contig00049	15975	747	47	cromosoma
contig00060	15398	774	50	cromosoma
contig00066	14406	742	52	cromosoma
contig00074	13662	648	47	cromosoma
contig00081	13282	587	44	cromosoma
contig00094	12413	612	49	cromosoma
contig00208	8476	374	44	cromosoma
contig00212	8406	372	44	cromosoma
contig00213	8387	541	65	cromosoma
contig00215	8337	353	42	cromosoma
contig00218	8276	414	50	cromosoma
contig00223	8179	399	49	cromosoma
contig00222	8174	532	65	cromosoma
contig00224	8147	315	39	cromosoma
contig00228	8077	400	50	cromosoma
contig00229	8039	342	43	cromosoma

contig00232	8021	424	53	cromosoma
contig00234	7976	327	41	cromosoma
contig00235	7949	527	66	cromosoma
contig00236	7946	381	48	cromosoma
contig00237	7939	472	59	cromosoma
contig00238	7930	313	39	cromosoma
contig00239	7925	396	50	cromosoma
contig00240	7918	305	39	cromosoma
contig00242	7901	378	48	cromosoma
contig00243	7886	328	42	cromosoma
contig00246	7804	346	44	cromosoma
contig00248	7772	301	39	cromosoma
contig00249	7768	316	41	cromosoma
contig00250	7716	302	39	cromosoma
contig00252	7665	319	42	cromosoma
contig00259	7601	394	52	cromosoma
contig00258	7595	433	57	cromosoma
contig00018	20046	999	50	pSymA
contig00019	19961	1058	53	pSymA
contig00026	18579	1010	54	pSymA
contig00040	17035	981	58	pSymA
contig00054	15825	832	53	pSymA
contig00063	15100	815	54	pSymA
contig00068	14108	779	55	pSymA
contig00071	13943	737	53	pSymA
contig00073	13717	597	44	pSymA
contig00078	13511	744	55	pSymA
contig00083	13199	718	54	pSymA
contig00087	13028	694	53	pSymA
contig00105	11414	671	59	pSymA
contig00122	10917	455	42	pSymA
contig00130	10650	568	53	pSymA
contig00144	10136	628	62	pSymA
contig00147	10012	624	62	pSymA
contig00153	9827	540	55	pSymA
contig00155	9810	460	47	pSymA
contig00158	9742	564	58	pSymA
contig00164	9544	535	56	pSymA
contig00168	9405	670	71	pSymA
contig00194	8829	543	62	pSymA
contig00199	8699	498	57	pSymA
contig00200	8664	552	64	pSymA
contig00202	8623	406	47	pSymA
contig00201	8613	446	52	pSymA
contig00205	8583	468	55	pSymA
contig00216	8278	439	53	pSymA
contig00217	8253	468	57	pSymA
contig00230	8030	465	58	pSymA
contig00231	8028	462	58	pSymA
contig00244	7856	409	52	pSymA
contig00245	7834	391	50	pSymA
contig00256	7655	407	53	pSymA
contig00265	7500	390	52	pSymA
contig00271	7434	433	58	pSymA

contig00007	26147	1387	53	pSymB
contig00010	24008	1369	57	pSymB
contig00015	20994	898	43	pSymB
contig00029	18221	1000	55	pSymB
contig00033	18008	833	46	pSymB
contig00036	17617	927	53	pSymB
contig00042	16925	768	45	pSymB
contig00046	16431	832	51	pSymB
contig00051	15874	815	51	pSymB
contig00055	15670	729	47	pSymB
contig00069	14069	757	54	pSymB
contig00076	13544	577	43	pSymB
contig00085	13139	661	50	pSymB
contig00091	12484	554	44	pSymB
contig00097	12016	596	50	pSymB
contig00104	11446	506	44	pSymB
contig00106	11307	557	49	pSymB
contig00111	11195	562	50	pSymB
contig00113	11170	511	46	pSymB
contig00117	11067	518	47	pSymB
contig00129	10746	491	46	pSymB
contig00131	10623	517	49	pSymB
contig00136	10405	481	46	pSymB
contig00140	10206	516	51	pSymB
contig00150	9899	467	47	pSymB
contig00163	9544	434	45	pSymB
contig00166	9510	461	48	pSymB
contig00173	9307	437	47	pSymB
contig00190	8910	545	61	pSymB
contig00192	8877	490	55	pSymB
contig00193	8864	398	45	pSymB
contig00196	8812	401	46	pSymB
contig00197	8756	510	58	pSymB
contig00203	8611	398	46	pSymB
contig00204	8583	522	61	pSymB
contig00206	8562	409	48	pSymB
contig00209	8461	419	50	pSymB
contig00210	8435	339	40	pSymB
contig00211	8412	469	56	pSymB
contig00214	8375	382	46	pSymB
contig00233	7995	437	55	pSymB
contig00251	7681	409	53	pSymB
contig00253	7669	377	49	pSymB

Tabla A2

Replicones utilizados para el análisis de variación funcional por COGs

Se listan los replicones depositados en GenBank a partir de los cuales se realizó el análisis de variación composicional de COGs, presentado en la Sección III.3. Los archivos multifasta de los ORFs predichos en cada replicón incorporado al análisis fueron utilizados como datos de entrada para el cálculo de composición de COGs. En el caso del moviloma plasmídico de *Rhizobiaceae*, los ORFs predichos en cada uno de los plásmidos listados se agruparon en un solo archivo multifasta que fue utilizado como dato de entrada para el cálculo de composición de COGs mediante el Software WebMGA (Sección II.14).

Replicón	Número de acceso
Cromosomas en <i>S. meliloti</i>	
Cro-Sme AK83	NC_015590.1
Cro-Sme Rm41	NC_018700.1
Cro-Sme SM11	NC_017325.1
Cro-Sme1021	NC_003047.1
Cro-SmeBL225C	NC_017322.1
Cro-SmeGR4	NC_019845.1
Cromosoma 2011	NC_020528.1
Megaplásmidos pSymA en <i>S. meliloti</i>	
pSymA-Sme2011	NC_020527.1
pSymA-Sme1021	NC_003037.1
pRmeGR4c-Sme GR4	NC_019848.1
pSINMEB01-Sme BL225C	NC_017324.1
pSymA-SmeRm41	NC_018683.1
pSmeSM11c-SmeSM11	NC_017327.1
cromosoma 3-SmeAK83	NC_015591.1
Megaplásmidos pSymB en <i>S. meliloti</i>	
cromosoma 2-SmeAK83	NC_015596.1
pSymB-Sme1021	NC_003078.1
pSymB-Sme2011	NC_020560.1
pSmeSM11d-SmeSM11	NC_017326.1
pRmeGR4d-SmeGR4	NC_019849.1
pSINMEB02-SmeBL225C	NC_017323.1
pSymB-SmeRm41	HE995408
Moviloma plasmídico "<i>Rhizobiaceae</i>"	
pMLa	NC_002679.1
pMLb	NC_002682.1
pRetCFN42a	NC_007762.1
pRetCFN42d	NC_004041.2
pRetCFN42b	NC_007763.1
pRetCFN42c	NC_007764.1
pRetCFN42e	NC_007765.1
pRetCFN42f	NC_007766.1
pSMED03	NC_009622.1

pNGR234a	NC_000914.2
pRi2659	NC_010841.1
pTiC58	NC_003065.3
pRle3841-7	NC_008382.1
pRle3841-8	NC_008383.1
pRle3841-10	NC_008381.1
pRle3841-11	NC_008384.1
pRle3841-9	NC_008379.1
pRLG202	NC_011366.1
pRLG203	NC_011370.1
pRLG204	NC_011371.1
pSfrHH103a	NC_016813.1
pSfrHH103b	NC_016836.1
pR132502	NC_012858.1
pR132503	NC_012853.1
pR132504	NC_012852.1
pR132505	NC_012854.1
pSmeSM11a	NC_013545.1
pSmeSM11b	NC_010865.1
pAtS4a vitis	NC_011986.1
pAtS4b vitis	NC_011991.1
pAtS4c vitis	NC_011984.1
pAtS4e vitis	NC_011981.1
pTiS4 vitis	NC_011982.1
pRmeGR4a	NC_019846.1
pRmeGR4b	NC_019847.1
pRtrCIAT899a	NC_020060.1
pRtrCIAT899b	NC_020061.1
pUSDA257	NC_018197.1
pSINME01	NC_015597.1
pSINME02	NC_015592.1
pMESCI01	NC_014918.1
pAgK84a	NC_011994.1
pAgK84b	NC_011990.1
pAgK84c	NC_011987.1
pA etli CIAT 652	NC_010998.1
pB etli CIAT 652	NC_010996.1
pAspH13-3a	NC_015184.1
plásmido 1 chelativorans BNC1	NC_008242.1
plásmido 2 chelativorans BNC1	NC_008243.1
plásmido 3 chelativorans BNC1	NC_008244.1
pAspH13-3a	NC_015184.1
AtuC58 plásmido At	NC_003064.2

Tabla A3

Listado de cromosomas de cepas de la familia Rhizobiaceae y Bradyrhizobiaceae que conforman los diferentes subconjuntos de genomas core de ortólogos ancestrales, calculados mediante EDGAR

Cada subconjunto varía en la cantidad de genes ortólogos que comparten (número de genes, fila 1). A medida que la distancia filogenética es mayor, menor número de ortólogos conforman el core. Se incorporaron los cromosomas completamente secuenciados de *S. meliloti* (ver materiales y métodos, sección II.15) y cromosomas de las especies indicadas cuyo número de acceso se indica entre paréntesis.

Core Anc 1	Core Anc 2	Core Anc 3	Core Anc 4	Core Anc 5	Core Anc 6	Core Anc 7
1446	977	896	855	685	466	462
<i>S. meliloti</i> cepas 1021,2011, AK83,BL225C, GR4, Rm41 y SM11	<i>S. meliloti</i> cepas 1021,2011, AK83,BL225C, GR4, Rm41 y SM11	<i>S. meliloti</i> cepas 1021,2011, AK83,BL225C, GR4, Rm41 y SM11				
<i>Agrobacterium</i> <i>Radiobacter</i> K84 (NC_011985)	<i>Agrobacterium</i> <i>Radiobacter</i> K84 (NC_011985)	<i>Liberibacter</i> <i>crecens</i> (NC 019907)	<i>Agrobacterium</i> <i>Radiobacter</i> K84 (NC_011985)	<i>Candidatus</i> <i>liberobacter</i> <i>asiaticus</i> (NC 020549)	<i>Agromonas</i> <i>oligotropha</i> S_58 (NC_020453)	<i>Agrobacterium</i> <i>Radiobacter</i> K84 (NC_011985)
<i>Agrobacterium</i> <i>sp. H-13_3</i> (NC_015183)	<i>Agrobacterium</i> <i>sp. H-13_3</i> (NC_015183)		<i>Agrobacterium</i> <i>sp. H-13_3</i> (NC_015183)		<i>Agromonas</i> <i>caulinodans</i> ORS_571 (NC 009937)	<i>Agrobacterium</i> <i>sp. H-13_3</i> (NC_015183)
<i>Agrobacterium</i> <i>vitis S4</i> (NC_011989)	<i>Agrobacterium</i> <i>vitis S4</i> (NC_011989)		<i>Agrobacterium</i> <i>vitis S4</i> (NC_011989)		<i>Bradyrhizobium</i> <i>japonicum</i> USDA 110 (NC 004463)	<i>Agrobacterium</i> <i>vitis S4</i> (NC_011989)
	<i>Agrobacterium</i> <i>fabrum C_58</i> (NC_003062)		<i>Agrobacterium</i> <i>fabrum C_58</i> (NC_003062)		<i>Candidatus</i> <i>liberobacter</i> <i>asiaticus</i> (NC 020549)	<i>Agrobacterium</i> <i>fabrum C_58</i> (NC_003062)
	<i>Agromonas</i> <i>oligotropha</i> S_58 (NC_020453)		<i>Agromonas</i> <i>oligotropha</i> S_58 (NC_020453)		<i>Mesorhizobium</i> <i>australicum</i> (NC 019973)	<i>Agromonas</i> <i>oligotropha</i> S_58 (NC_020453)
			<i>Agromonas</i> <i>caulinodans</i> ORS_571 (NC 009937)		<i>Liberibacter</i> <i>crecens</i> (NC 019907)	<i>Agromonas</i> <i>caulinodans</i> ORS_571 (NC 009937)
			<i>Bradyrhizobium</i> <i>japonicum</i> USDA 110 (NC 004463)			<i>Bradyrhizobium</i> <i>japonicum</i> USDA 110 (NC 004463)
						<i>Candidatus</i> <i>liberobacter</i> <i>asiaticus</i> (NC 020549)
						<i>Mesorhizobium</i> <i>australicum</i> (NC 019973)
						<i>Liberibacter</i> <i>crecens</i> (NC 019907)

Tabla A4

Homólogos al MPSme en el core cromosomal de *S. meliloti* (OC-MC) y en conjuntos core ancestrales.

Se indica la predicción bioinformática de cada ORF realizada en base a BLASTp sobre la base de datos del GenBank (nr) y se clasifica manualmente en los grupos funcionales siguientes: H (proteína hipotética con homólogos en otros organismos, de función desconocida), M (función metabólica general), F (gen relacionado a metabolismo de bacteriófagos), T (función de transporte de metabolitos), R (regulador transcripcional), A (procesamiento del ADN). De las columnas 4 a la 10 se indica si dicho gen presenta un homólogo a los diferentes conjuntos de genes ancestrales (Tabla A3 del Anexo), core ancestral (CA) del 1 al 7.

denominación del OFR	predicción bioinformática de la función	cat goría	CA 1	CA 2	CA 3	CA 4	CA 5	CA 6	CA 7
S. meliloti mobilome final 63	Glutathione import ATP-binding protein gsiA EC 3.6.3.-	T							
S. meliloti mobilome final 64	Dipeptide transport system permease protein dppC	T							
S. meliloti mobilome final 65	Probable peptide ABC transporter permease protein y4tP	T							
S. meliloti mobilome final 66	Glutathione-binding protein gsiB Flags: Precursor	M							
S. meliloti mobilome final 71	K00852 ribokinase	M							
S. meliloti mobilome final 75	hypothetical protein	M							
S. meliloti mobilome final 103	SPBc2 prophage-derived DNA-binding protein HU 2	F							
S. meliloti mobilome final 140	hypothetical protein	H							
S. meliloti mobilome final 147	K01946 biotin carboxylase	M							
S. meliloti mobilome final 178	Rv0623-like transcription factor	R							
S. meliloti mobilome final 179	hypothetical protein	H							
S. meliloti mobilome final 182	60 kDa chaperonin AltName: Full Protein Cpn60 AltName: Full groEL protein	M							
S. meliloti mobilome final 231	K02056 simple sugar transport system ATP-binding protein	T							
S. meliloti mobilome final 235	K03335 inosose dehydratase	M							
S. meliloti mobilome final 241	K10680 N-ethylmaleimide reductase	M							
S. meliloti mobilome final 303	D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding	M							
S. meliloti mobilome final 316	Probable peptide ABC transporter ATP-binding protein y4tR	T							
S. meliloti mobilome final 340	K00135 succinate-semialdehyde dehydrogenase (NADP+)	M							
S. meliloti mobilome final 393	K02030 polar amino acid transport system substrate-binding protein	T							
S. meliloti mobilome final 394	Inner membrane amino-acid ABC transporter permease protein yecS	T							
S. meliloti mobilome final 395	K09972 general L-amino acid transport system ATP-binding protein	T							
S. meliloti mobilome final 402	K00626 acetyl-CoA C-acetyltransferase	M							
S. meliloti mobilome final 403	UPF0262 protein y4uD	H							
S. meliloti mobilome final 456	K02053 putative spermidine/putrescine transport system permease protein	T							
S. meliloti mobilome final 457	K02054 putative spermidine/putrescine transport system permease protein	T							
S. meliloti mobilome final 458	K02055 putative spermidine/putrescine transport system substrate-binding protein	T							
S. meliloti mobilome final 512	DNA-damage-inducible protein F	A							
S. meliloti mobilome final 576	hypothetical protein	H							
S. meliloti mobilome final 581	hypothetical protein	H							
S. meliloti mobilome final 612	K01959 pyruvate carboxylase subunit A	M							

S. meliloti mobilome final 640	NUDIX hydrolase	M								
S. meliloti mobilome final 645	CopG/Arc/MetJ DNA-binding domain-containing transcriptional regulator	R								
S. meliloti mobilome final 666	Arginine-binding periplasmic protein 1 Flags: Precursor	M								
S. meliloti mobilome final 696	K01627 2-dehydro-3-deoxyphosphooctonate aldolase (KDO 8-P synthase)	M								
S. meliloti mobilome final 734	K00108 choline dehydrogenase	M								
S. meliloti mobilome final 746	Spermidine/putrescine import ATP-binding protein potA 2 EC 3.6.3.31	T								
S. meliloti mobilome final 803	K00133 aspartate-semialdehyde dehydrogenase	M								
S. meliloti mobilome final 841	K07010 putative glutamine amidotransferase	M								
S. meliloti mobilome final 845	Glutathione transport system permease protein gsiD	T								
S. meliloti mobilome final 919	K01006 pyruvate orthophosphate dikinase	M								
S. meliloti mobilome final 925	Tetracycline resistance protein class A Short TetA(A)	M								
S. meliloti mobilome final 982	aldo/keto reductase	M								
S. meliloti mobilome final 983	LysR family transcriptional regulator	R								
S. meliloti mobilome final 993	hypothetical protein	H								
S. meliloti mobilome final 996	Transcriptional regulatory protein mucR homolog	R								
S. meliloti mobilome final 1011	hypothetical protein	H								
S. meliloti mobilome final 1052	putative transmembrane protein	H								
S. meliloti mobilome final 1059	K01451 hippurate hydrolase	M								
S. meliloti mobilome final 1064	hypothetical protein	H								
S. meliloti mobilome final 1099	K02337 DNA polymerase III subunit alpha	A								
S. meliloti mobilome final 1115	K00479 Rieske 2Fe-2S family protein	H								
S. meliloti mobilome final 1116	ferredoxin	M								
S. meliloti mobilome final 1120	LuxR family transcriptional regulator	R								
S. meliloti mobilome final 1176	K00812 aspartate aminotransferase	M								
S. meliloti mobilome final 1189	K02023 multiple sugar transport system ATP-binding protein	T								
S. meliloti mobilome final 1241	Leucine-responsive regulatory protein	R								
S. meliloti mobilome final 1258	K01915 glutamine synthetase	M								
S. meliloti mobilome final 1296	acetylglutamate kinase (EC:2.7.2.8)	M								
S. meliloti mobilome final 1336	hypothetical protein	H								
S. meliloti mobilome final 1360	K00979 3-deoxy-manno-octulosonate cytidyllyltransferase (CMP-KDO synthetase)	M								
S. meliloti mobilome final 1366	K00384 thioredoxin reductase (NADPH)	M								
S. meliloti mobilome final 1420	hypothetical protein	H								
S. meliloti mobilome final 1431	K07062	H								
S. meliloti mobilome final 1439	10 kDa chaperonin 1 AltName: Full Protein Cpn10 1 AltName: Full groES protein 1	M								
S. meliloti mobilome final 1528	Probable cold shock protein y4cH	M								
S. meliloti mobilome final 1582	putative cell filamentation protein	M								
S. meliloti mobilome final 1587	hypothetical protein	H								
S. meliloti mobilome final 1594	K00832 aromatic-amino-acid transaminase	M								

Tabla A5

Actividades bioquímicas asociadas a la promoción de crecimiento *in vitro* en aislamientos obtenidos de plántulas y semillas de alfalfa

Se abrevian los ensayos como “Prot”: proteasas; “Celu”: celulasas; “Quití”: quitinasas; “Side”: sideróforos; “AIA”: ácido indol acético; “AIAc/Trip”: ácido indol acético en ensayos suplementados con triptófano. NE: no ensayado.

 AISLAMIENTOS PROVENIENTES DE PLÁNTULAS 								
#AIS	GÉNERO	Prot	Celu	Quití	Solubilización fosfatos	Side	AIA	AIA c/Trip
SP1	<i>Microbacterium</i>	-	+	NE	-	+	-	-
SP2	<i>Enhydrobacter</i>	-	+	NE	-	+	-	-
SP3	<i>Exiguobacterium</i>	+	+	-	+	+	-	-
SP4	<i>Corynebacterium</i>	-	+	NE	-	+	-	NE
SP5	<i>Stenotrophomonas</i>	+	+	+	-	-	-	-
SP6	<i>Pantoea</i>	-	+	NE	+	-	-	+
SP7	<i>Rhizobium</i>	-	-	NE	-	-	-	+
SP8	<i>Rhizobium</i>	-	-	NE	-	-	+	NE
SP9	<i>Rhizobium</i>	-	-	NE	-	-	-	-
SP10	<i>Microbacterium</i>	+	+	-	+	+	-	+
SP11	<i>Pseudomonas</i>	-	+	NE	+	+	+	+
SP12	<i>Rhizobium</i>	-	-	NE	-	+	+	NE
SP13	<i>Pseudomonas</i>	NE	+	NE	+	-	+	+
SP14	<i>Stenotrophomonas</i>	NE	NE	NE	-	NE	NE	NE
SP15	<i>Stenotrophomonas</i>	+	-	NE	-	+	+	NE
SP16	<i>Flavobacterium</i>	NE	NE	NE	-	NE	NE	NE
SP17	<i>Rhizobium</i>	-	-	NE	-	+	+	NE
SP18	<i>Rhizobium</i>	-	-	NE	-	+	+	NE
SP19	<i>Paenibacillus</i>	NE	NE	NE	-	NE	NE	NE
SP20	<i>Novosphingobium</i>	+	-	NE	-	+	+	NE
SP21	<i>Pantoea</i>	-	-	-	+	+	-	-
SP22	<i>Pantoea</i>	-	-	-	+	+	-	+
BP1	<i>Paenibacillus</i>	-	+	-	NE	-	-	+
BP2	NI	-	-	-	-	+	-	+
BP3	NI	-	-	-	-	-	-	+
BP4	<i>Staphylococcus</i>	-	-	-	+	+	-	+
CP1	<i>Staphylococcus</i>	-	-	-	NE	-	-	+
CP2	<i>Staphylococcus</i>	-	-	-	-	-	-	-
CP3	NI	+	+	-	-	-	-	+
CP4	NI	-	-	-	NE	+	-	-
CP5	<i>Pantoea</i>	-	-	-	-	-	-	-
FP1	<i>Arthrobacter</i>	+	-	-	NE	-	-	-
FP2	<i>Arthrobacter</i>	-	-	-	-	-	+	+
FP3	<i>Arthrobacter</i>	-	-	-	-	-	-	-

FP4	NI	-	+	-	NE	-	+	+
FP5	NI	+	-	-	NE	-	-	-
FP6	<i>Arthrobacter</i>	-	-	-	-	-	-	-
DP1	<i>Pantoea</i>	+	-	-	+	-	-	-
DP2	<i>Pantoea</i>	-	-	-	+	+	-	+
DP3	<i>Micrococcus</i>	+	+	-	-	-	-	-
GP1	<i>Pantoea</i>	-	-	-	+	+	-	-
GP2	<i>Pantoea</i>	-	-	-	+	+	-	-

AISLAMIENTOS PROVENIENTES DE SEMILLAS

#AIS	GÉNERO	Prot	Celu	Quit	Solubilización fosfatos	Side	AIA	AIA Trip
SS1	<i>Microbacterium</i>	-	+	NE	-	+	+	+
SS2	<i>Arthrobacter</i>	+	+	NE	-	+	-	-
SS3	<i>Brevundimonas</i>	NE	-	NE	+	+	-	-
SS4	<i>Janibacter</i>	+	+	-	-	+	-	-
SS5	<i>Bacillus</i>	-	+	-	-	-	-	-
SS6	<i>Moraxella</i>	-	-	NE	-	-	-	-
SS7	<i>Bacillus</i>	+	-	-	-	+	-	-
SS8	<i>Paantoea</i>	-	+	NE	+	+	+	+
SS9	<i>Microbacterium</i>	+	+	-	+	+	-	+
SS10	<i>Paenibacillus</i>	+	+	-	-	+	-	-
SS11	<i>Microbacterium</i>	-	-	NE	-	-	+	+
SS12	<i>Enhydrobacter</i>	-	-	NE	-	-	-	-
SS13	<i>Kocuria</i>	-	+	NE	-	+	+	+
SS14	<i>Pantoea</i>	-	+	NE	+	+	-	-
SS15	<i>Terracoccus</i>	NE	NE	NE	-	NE	NE	NE
SS16	<i>Microbacterium</i>	-	+	NE	-	+	-	-
SS17	<i>Staphylococcus</i>	-	-	NE	+	NE	NE	NE
SS18	<i>Rhizobium</i>	+	+	NE	-	+	-	NE
SS19	<i>Sphingomonas</i>	-	-	-	NE	-	-	-
SS20	<i>Paenibacillus</i>	-	-	NE	-	-	-	NE
SS21	<i>Kocuria</i>	-	+	NE	NE	+	-	NE
SS22	NI	NE	NE	NE	NE	NE	NE	NE
DS1	<i>Solibacillus</i>	+	-	NE	-	-	-	NE
DS2	<i>Sporosarcina</i>	-	+	NE	-	NE	+	NE
DS3	<i>Sporosarcina</i>	-	-	NE	-	-	-	NE
DS4	<i>Staphylococcus</i>	-	-	NE	-	-	-	NE
DS5	NI	+	+	NE	-	-	+	NE
DS6	<i>Arthrobacter</i>	-	+	NE	-	+	NE	NE
DS7	<i>Arthrobacter</i>	-	-	NE	-	+	-	NE
DS8	<i>Arthrobacter</i>	-	-	NE	-	-	+	NE
DS9	<i>Pseudomonas</i>	+	+	NE	-	+	+	NE
DS10	<i>Sporosarcina</i>	-	-	NE	-	-	-	NE
DS11	<i>Bacillus</i>	+	-	NE	-	+	-	NE
DS12	<i>Sporosarcina</i>	+	-	NE	-	-	-	NE
DS13	<i>Sanguibacter</i>	-	+	NE	-	NE	NE	NE

DS14	NI	NE						
DS15	<i>Planococcus</i>	NE						
DS16	NI	NE						
DS17	<i>Staphylococcus</i>	-	+	NE	NE	-	-	NE
DS18	<i>Bacillus</i>	+	+	NE	NE	-	NE	NE
DS19	<i>Bacillus</i>	-	-	NE	NE	-	-	NE
DS20	<i>Bacillus</i>	-	+	NE	NE	-	-	NE
DS21	<i>Arthrobacter</i>	-	-	NE	-	+	-	NE
DS22	<i>Arthrobacter</i>	NE						
DS23	<i>Sporosarcina</i>	-	-	NE	NE	-	-	NE
GS1	<i>Pantoea</i>	NE	NE	NE	-	NE	NE	NE
GS2	<i>Curtobacterium</i>	+	+	NE	+	NE	+	NE
GS3	<i>Planococcus</i>	NE	NE	NE	-	NE	NE	NE
GS4	<i>Bacillus</i>	NE	NE	NE	-	NE	NE	NE
GS5	<i>Azorcus</i>	NE	NE	NE	-	NE	NE	NE
GS6	<i>Bacillus</i>	-	-	NE	-	-	+	NE
GS7	<i>Bacillus</i>	+	+	NE	+	-	-	NE
GS8	NI	NE	NE	NE	+	NE	NE	NE
GS9	<i>Curtobacterium</i>	NE	NE	NE	-	NE	NE	NE
GS10	<i>Exiguobacterium</i>	+	+	NE	+	+	-	NE
GS11	<i>Arthrobacter</i>	+	-	NE	-	+	+	NE
GS12	<i>Arthrobacter</i>	+	-	NE	-	+	-	NE
GS13	<i>Roseomonas</i>	+	+	NE	-	+	-	NE
GS14	<i>Agrococcus</i>	-	-	NE	-	-	-	NE
GS15	<i>Shingomonas</i>	NE	NE	NE	-	NE	NE	NE
GS16	<i>Bacillus</i>	+	-	NE	-	+	-	NE
GS17	<i>Paenibacillus</i>	-	-	NE	-	+	-	NE
GS18	<i>Arthrobacter</i>	+	-	NE	-	+	-	NE
GS19	<i>Arthrobacter</i>	-	-	NE	-	+	NE	NE
GS20	<i>Roseomonas</i>	-	+	NE	-	-	-	NE
GS21	<i>Microbacterium</i>	+	+	NE	-	+	+	NE
GS22	<i>Arthrobacter</i>	-	+	NE	-	+	+	NE
GS23	<i>Arthrobacter</i>	-	+	NE	-	-	NE	NE
GS24	<i>Arthrobacter</i>	-	+	NE	-	NE	NE	NE
GS25	<i>Sporosarcina</i>	-	-	NE	-	-	-	NE
GS26	NI	NE						
GS27	<i>Curtobacterium</i>	+	+	NE	-	+	-	NE
GS28	<i>Janibacter</i>	+	+	NE	-	+	-	NE
GS29	<i>Arthrobacter</i>	NE	NE	NE	-	NE	NE	NE
GS30	NI	NE	NE	NE	-	NE	NE	NE
GS31	<i>Kocuria</i>	NE						
GS32	<i>Sphingobium</i>	-	+	NE	NE	NE	NE	NE
GS33	NI	NE	NE	NE	-	NE	NE	NE
GS34	<i>Arthrobacter</i>	-	+	NE	-	+	+	NE
GS35	<i>Arthrobacter</i>	NE	NE	NE	-	NE	NE	NE
GS36	NI	NE						

GS37	<i>Paenibacillus</i>	-	-	NE	NE	-	-	NE
GS38	<i>Bacillus</i>	NE						
FS1	<i>Kocuria</i>	NE	NE	NE	-	NE	NE	NE
FS2	<i>Brevibacillus</i>	NE	NE	NE	-	NE	NE	NE
FS3	NI	NE	NE	NE	-	NE	NE	NE
FS4	<i>Bacillus</i>	NE						
BS1	NI	-	+	NE	+	+	NE	NE
BS2	<i>Kocuria</i>	-	+	NE	-	+	+	NE
BS3	<i>Paenibacillus</i>	-	-	NE	-	+	NE	NE
BS4	<i>Frigobacterium</i>	NE	NE	NE	-	NE	NE	NE
BS5	<i>Arthrobacter</i>	NE	NE	NE	-	NE	NE	NE
BS6	<i>Bacillus</i>	NE	NE	NE	-	NE	NE	NE
BS7	<i>Bacillus</i>	+	+	NE	-	NE	+	NE
BS8	<i>Bacillus</i>	+	+	NE	-	-	+	NE
BS9	<i>Kocuria</i>	NE	NE	NE	-	NE	NE	NE
BS10	<i>Kocuria</i>	+	+	NE	-	+	+	NE
BS11	<i>Arthrobacter</i>	+	-	NE	-	+	+	NE
BS12	<i>Pantoea</i>	-	-	NE	-	-	+	NE
BS13	<i>Arthrobacter</i>	+	+	NE	-	-	+	NE
BS14	<i>Bacillus</i>	+	+	NE	-	+	+	NE
BS15	<i>Microbacterium</i>	+	-	-	-	-	-	+
BS16	<i>Terribacillus</i>	+	-	NE	NE	+	+	NE
BS17	<i>Arthrobacter</i>	NE						
BS18	<i>Microbacterium</i>	NE						
BS19	<i>Arthrobacter</i>	-	+	NE	NE	-	+	NE

Tabla A6.**Ensayo de bacteriocinas entre distintos aislamientos de la colección.**

Se indica en la primer fila los aislamientos que sirvieron de césped y sobre los que se ensayó la acción de bacteriocinas por parte de los aislamientos de la primer columna. Se indica como "SI" en color naranja si se observaron halos de inhibición y en color verde si no se observaron halos de inhibición.

	SS8	SP7	SP20	SP9	SP11	SP18	BP2	SP12	BP4	CP2	CP6	DP1	DP3	SS14
SS8			SI											
SP7														
SP20														
SP9														
SP11			SI								SI			
SP18														
BP2			SI											
SP12														
BP4														
CP2														
CP6														
DP1														
DP3														
SS14														

Tabla A7.**Cobertura promedio de los ensamblados mayores a 15kpb del MPAS**

Denominación del ensamblado	Cobertura promedio de secuenciación (lecturas/kpb)
scaffold00001	120,7
scaffold00002	379,0
scaffold00004	668,2
scaffold00005	514,0
scaffold00007	86,2
scaffold00010	50,3
scaffold00012	116,6
scaffold00013	40,2
scaffold00014	107,0
scaffold00016	104,9
scaffold00017	83,2
scaffold00020	746,9
scaffold00023	2846,8
scaffold00026	39,8
scaffold00027	692,0
scaffold00028	2424,5

scaffold00029	1275,8
scaffold00030	360,4
scaffold00032	188,2
scaffold00033	141,8
scaffold00034	1031,9
scaffold00035	58,1
scaffold00036	4736,6
scaffold00037	40,5
scaffold00038	70,1
scaffold00039	370,2
scaffold00040	42,2
scaffold00042	193,2
scaffold00043	43,3
scaffold00044	82,0
scaffold00045	384,1
scaffold00046	253,6
scaffold00047	58,8
scaffold00049	107,0
scaffold00050	62,8
scaffold00051	38,3
scaffold00052	76,6
scaffold00053	38,8
scaffold00054	1386,5
scaffold00055	66,6

Tabla A8

Detalles agronómicos y fitosanitarios de las semillas utilizadas en este trabajo.

Se indica el nombre de la variedad asociado a la letra con la que se denomina dicha variedad a lo largo de este trabajo. Referencias: MR = moderada resistencia (15-30%); R = resistencia intermedia (> 31%) y AR = alta resistencia (>51%).

	"S": Super Monarca	"C": GAPP 810+	"B": GAPP 909+	"F": F401Y	"D": F301Y	"G": F111X
Grado de reposo invernal	8	7-8	9	6	6	5-6
Porte de planta	semi erecto	erecto	erecto	erecto	erecto	semierecto
Corona	intermedia	amplia	amplia	amplia	amplia	muy amplia
Antracnosis (<i>Colletotrichum trifolii</i>)	MR	AR	R	AR	R	R
Fusariosis (<i>Fusarium oxysporum</i>)	AR	AR	R	AR	AR	AR

Fitóftora <i>(Phytophthora megasperma)</i>	MR	AR	R	AR	AR	R
Pulgón verde <i>(Arcyosiphon pisum)</i>	R	R	R	R	R	R
Pulgón azul <i>(Arcyosiphon kondoi)</i>	R	AR		AR		R
Pulgón montado <i>(Therioaphis trifoli)</i>	AR	AR	R	AR	AR	R

Referencias bibliográficas

1. **Pistorio, M., et al.**, Conjugal properties of the *Sinorhizobium meliloti* plasmid mobilome. *FEMS microbiology ecology*, 2008. **65**(3): p. 372-382.
2. **Speidel, J.J., et al.**, Population policies, programmes and the environment. *Philos Trans R Soc Lond B Biol Sci*, 2009. **364**(1532): p. 3049-65.
3. **Rockstrom, J., et al.**, A safe operating space for humanity. *Nature*, 2009. **461**(7263): p. 472-475.
4. **Godfray, H.C., et al.**, Food security: the challenge of feeding 9 billion people. *Science*, 2010. **327**(5967): p. 812-8.
5. **Galloway, J.N., et al.**, Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science*, 2008. **320**(5878): p. 889-892.
6. **Good, A.G. and P.H. Beatty**, Fertilizing nature: a tragedy of excess in the commons. *PLoS Biol*, 2011. **9**(8): p. e1001124.
7. **Glass, A.D.M.**, Nitrogen Use Efficiency of Crop Plants: Physiological Constraints upon Nitrogen Absorption. *Critical Reviews in Plant Sciences*, 2003. **22**(5): p. 453-470.
8. **Schmer, M.R., et al.**, Energy potential and greenhouse gas emissions from bioenergy cropping systems on marginally productive cropland. *PLoS One*, 2014. **9**(3): p. e89501.
9. **Conley, D.J.**, Ecology: save the Baltic Sea. *Nature*, 2012. **486**(7404): p. 463-4.
10. **Mueller, N.D., et al.**, Closing yield gaps through nutrient and water management. *Nature*, 2012. **490**(7419): p. 254-7.
11. **Ray, D.K., et al.**, Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS One*, 2013. **8**(6): p. e66428.
12. **Robinson, T.P., et al.**, Mapping the global distribution of livestock. *PLoS One*, 2014. **9**(5): p. e96084.
13. **Ellis, E.C., et al.**, Used planet: A global history. *Proceedings of the National Academy of Sciences*, 2013. **110**(20): p. 7978-7985.
14. **Hansen, M.C., et al.**, High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 2013. **342**(6160): p. 850-853.
15. **Lindow, S.E. and M.T. Brandl**, Microbiology of the phyllosphere. *Appl Environ Microbiol*, 2003. **69**(4): p. 1875-83.
16. **Vorholt, J.A.**, Microbial life in the phyllosphere. *Nat Rev Micro*, 2012. **10**(12): p. 828-840.
17. **Rastogi, G., G.L. Coaker, and J.H. Leveau**, New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches. *FEMS Microbiol Lett*, 2013. **348**(1): p. 1-10.
18. **Ryan, R.P., et al.**, Bacterial endophytes: recent developments and applications. *FEMS Microbiol Lett*, 2008. **278**(1): p. 1-9.
19. **Costa, R., et al.**, Effects of site and plant species on rhizosphere community structure as revealed by molecular analysis of microbial guilds. *FEMS Microbiol Ecol*, 2006. **56**(2): p. 236-49.
20. **Pierret, A., et al.**, Root Functional Architecture: A Framework for Modeling the Interplay between Roots and Soil. *Vadose Zone Journal*, 2007. **6**(2): p. 269-281.
21. **Jones, D.L. and P. Hinsinger**, The rhizosphere: complex by design. *Plant and Soil*, 2008. **312**(1-2): p. 1-6.

22. **Hinsinger, P., et al.**, Rhizosphere: biophysics, biogeochemistry and ecological relevance. *Plant and Soil*, 2009. **321**(1-2): p. 117-152.
23. **Raaijmakers, J., et al.**, The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant and Soil*, 2009. **321**(1-2): p. 341-361.
24. **Bonkowski, M., C. Villenave, and B. Griffiths**, Rhizosphere fauna: the functional and structural diversity of intimate interactions of soil fauna with plant roots. *Plant and Soil*, 2009. **321**(1-2): p. 213-233.
25. **Buée, M., et al.**, The rhizosphere zoo: An overview of plant-associated communities of microorganisms, including phages, bacteria, archaea, and fungi, and of some of their structuring factors. *Plant and Soil*, 2009. **321**(1-2): p. 189-212.
26. **Berg, G., L. Eberl, and A. Hartmann**, The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. *Environ Microbiol*, 2005. **7**(11): p. 1673-85.
27. **Herron, P.M., D.J. Gage, and Z.G. Cardon**, Micro-scale water potential gradients visualized in soil around plant root tips using microbiosensors. *Plant Cell Environ*, 2010. **33**(2): p. 199-210.
28. **Weller, D.M., et al.**, Microbial populations responsible for specific soil suppressiveness to plant pathogens. *Annu Rev Phytopathol*, 2002. **40**: p. 309-48.
29. **Dias, A.C.F., et al.**, Potato cultivar type affects the structure of ammonia oxidizer communities in field soil under potato beyond the rhizosphere. *Soil Biology and Biochemistry*, 2012. **50**(0): p. 85-95.
30. **Garbeva, P., J.D. van Elsas, and J.A. van Veen**, Rhizosphere microbial community and its response to plant species and soil history. *Plant and Soil*, 2008. **302**(1-2): p. 19-32.
31. **Smalla, K., et al.**, Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: plant-dependent enrichment and seasonal shifts revealed. *Appl Environ Microbiol*, 2001. **67**(10): p. 4742-51.
32. **Teixeira, L.C., et al.**, Bacterial diversity in rhizosphere soil from Antarctic vascular plants of Admiralty Bay, maritime Antarctica. *Isme j*, 2010. **4**(8): p. 989-1001.
33. **de Ridder-Duine, A.S., et al.**, Rhizosphere bacterial community composition in natural stands of *Carex arenaria* (sand sedge) is determined by bulk soil community composition. *Soil Biology and Biochemistry*, 2005. **37**(2): p. 349-357.
34. **Santos-Gonzalez, J.C., et al.**, Soil, but not cultivar, shapes the structure of arbuscular mycorrhizal fungal assemblages associated with strawberry. *Microb Ecol*, 2011. **62**(1): p. 25-35.
35. **Andrew, D.R., et al.**, Abiotic factors shape microbial diversity in Sonoran Desert soils. *Appl Environ Microbiol*, 2012. **78**(21): p. 7527-37.
36. **Inceoglu, O., J. Falcao Salles, and J.D. van Elsas**, Soil and cultivar type shape the bacterial community in the potato rhizosphere. *Microb Ecol*, 2012. **63**(2): p. 460-70.
37. **Lennon, J.T. and S.E. Jones**, Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Micro*, 2011. **9**(2): p. 119-130.
38. **Fierer, N. and R.B. Jackson**, The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, 2006. **103**(3): p. 626-631.

39. **Kowalchuk, G.A., et al.**, Effects of above-ground plant species composition and diversity on the diversity of soil-borne microorganisms. *Antonie Van Leeuwenhoek*, 2002. **81**(1-4): p. 509-20.
40. **Bergsma-Vlami, M., M.E. Prins, and J.M. Raaijmakers**, Influence of plant species on population dynamics, genotypic diversity and antibiotic production in the rhizosphere by indigenous *Pseudomonas spp.* *FEMS Microbiol Ecol*, 2005. **52**(1): p. 59-69.
41. **Pivato, B., et al.**, Medicago species affect the community composition of arbuscular mycorrhizal fungi associated with roots. *New Phytol*, 2007. **176**(1): p. 197-210.
42. **Haichar, F.Z., et al.**, Plant host habitat and root exudates shape soil bacterial community structure. *Isme j*, 2008. **2**(12): p. 1221-30.
43. **Bressan, M., et al.**, Exogenous glucosinolate produced by *Arabidopsis thaliana* has an impact on microbes in the rhizosphere and plant roots. *Isme j*, 2009. **3**(11): p. 1243-57.
44. **Ladygina, N. and K. Hedlund**, Plant species influence microbial diversity and carbon allocation in the rhizosphere. *Soil Biology and Biochemistry*, 2010. **42**(2): p. 162-168.
45. **Mazzola, M., D.L. Funnell, and J.M. Raaijmakers**, Wheat cultivar-specific selection of 2,4-diacetylphloroglucinol-producing fluorescent *Pseudomonas* species from resident soil populations. *Microb Ecol*, 2004. **48**(3): p. 338-48.
46. **Yao, H. and F. Wu**, Soil microbial community structure in cucumber rhizosphere of different resistance cultivars to fusarium wilt. *FEMS Microbiol Ecol*, 2010. **72**(3): p. 456-63.
47. **Hardoim, P.R., et al.**, Rice root-associated bacteria: insights into community structures across 10 cultivars. *FEMS Microbiology Ecology*, 2011. **77**(1): p. 154-164.
48. **Bouffaud, M.L., et al.**, Is diversification history of maize influencing selection of soil bacteria by roots? *Mol Ecol*, 2012. **21**(1): p. 195-206.
49. **Germida, J. and S. Siciliano**, Taxonomic diversity of bacteria associated with the roots of modern, recent and ancient wheat cultivars. *Biology and Fertility of Soils*, 2001. **33**(5): p. 410-415.
50. **Smith, K.P., J. Handelsman, and R.M. Goodman**, Genetic basis in plants for interactions with disease-suppressive bacteria. *Proceedings of the National Academy of Sciences*, 1999. **96**(9): p. 4786-4790.
51. **Nelson, E.B.**, Microbial dynamics and interactions in the spermosphere. *Annu Rev Phytopathol*, 2004. **42**: p. 271-309.
52. **Hardoim, P.R., et al.**, Dynamics of seed-borne rice endophytes on early plant growth stages. *PLoS One*, 2012. **7**(2): p. e30438.
53. **Normander, B. and J.I. Prosser**, Bacterial Origin and Community Composition in the Barley Phytosphere as a Function of Habitat and Presowing Conditions. *Applied and Environmental Microbiology*, 2000. **66**(10): p. 4372-4377.
54. **Johnston-Monje, D. and M.N. Raizada**, Conservation and diversity of seed associated endophytes in *Zea* across boundaries of evolution, ethnography and ecology. *PLoS One*, 2011. **6**(6): p. e20396.
55. **Woodward, F.I. and M.R. Lomas**, Vegetation dynamics--simulating responses to climatic change. *Biol Rev Camb Philos Soc*, 2004. **79**(3): p. 643-70.

56. **Leveau, J.H. and S.E. Lindow**, Appetite of an epiphyte: quantitative monitoring of bacterial sugar consumption in the phyllosphere. *Proc Natl Acad Sci U S A*, 2001. **98**(6): p. 3446-53.
57. **Miller, W.G., et al.**, Biological sensor for sucrose availability: relative sensitivities of various reporter genes. *Appl Environ Microbiol*, 2001. **67**(3): p. 1308-17.
58. **Remus-Emsermann, M.N., et al.**, Variation in local carrying capacity and the individual fate of bacterial colonizers in the phyllosphere. *Isme j*, 2012. **6**(4): p. 756-65.
59. **Boller, T. and G. Felix**, A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu Rev Plant Biol*, 2009. **60**: p. 379-406.
60. **Knief, C., et al.**, Site and plant species are important determinants of the *Methylobacterium* community composition in the plant phyllosphere. *Isme j*, 2010. **4**(6): p. 719-28.
61. **Beattie, G.A.**, Water relations in the interaction of foliar bacterial pathogens with plants. *Annu Rev Phytopathol*, 2011. **49**: p. 533-55.
62. **Monier, J.M. and S.E. Lindow**, Differential survival of solitary and aggregated bacterial cells promotes aggregate formation on leaf surfaces. *Proc Natl Acad Sci U S A*, 2003. **100**(26): p. 15977-82.
63. **Yu, J., et al.**, Involvement of the exopolysaccharide alginate in the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. *Molecular microbiology*, 1999. **33**(4): p. 712-720.
64. **Bunster, L., N.J. Fokkema, and B. Schippers**, Effect of Surface-Active *Pseudomonas* spp. on Leaf Wettability. *Appl Environ Microbiol*, 1989. **55**(6): p. 1340-5.
65. **Schreiber, L., et al.**, Plant-microbe interactions: identification of epiphytic bacteria and their ability to alter leaf surface permeability. *New Phytol*, 2005. **166**(2): p. 589-94.
66. **Chen, C. and G.A. Beattie**, *Pseudomonas syringae* BetT is a low-affinity choline transporter that is responsible for superior osmoprotection by choline over glycine betaine. *J Bacteriol*, 2008. **190**(8): p. 2717-25.
67. **Freeman, B.C., C. Chen, and G.A. Beattie**, Identification of the trehalose biosynthetic loci of *Pseudomonas syringae* and their contribution to fitness in the phyllosphere. *Environ Microbiol*, 2010. **12**(6): p. 1486-97.
68. **Fry, S.C.**, Cellulases, hemicelluloses and auxin-stimulated growth: a possible relationship. *Physiologia Plantarum*, 1989. **75**(4): p. 532-536.
69. **Brandl, M.T. and S.E. Lindow**, Contribution of indole-3-acetic acid production to the epiphytic fitness of *Erwinia herbicola*. *Appl Environ Microbiol*, 1998. **64**(9): p. 3256-63.
70. **Fall, R. and A.A. Benson**, Leaf methanol—the simplest natural product from plants. *Trends in Plant Science*, 1996. **1**(9): p. 296-301.
71. **Sy, A., et al.**, Methylo trophic metabolism is advantageous for *Methylobacterium extorquens* during colonization of *Medicago truncatula* under competitive conditions. *Applied and environmental microbiology*, 2005. **71**(11): p. 7245-7252.
72. **Kawaguchi, K., et al.**, Yeast methylo trophic and autophagy in a methanol-oscillating environment on growing *Arabidopsis thaliana* leaves. *PLoS One*, 2011. **6**(9): p. e25257-e25257.

73. **Abanda-Nkpwatt, D., et al.**, Molecular interaction between *Methylobacterium extorquens* and seedlings: growth promotion, methanol consumption, and localization of the methanol emission site. *Journal of experimental botany*, 2006. **57**(15): p. 4025-4032.
74. **Saito, T. and Y. Yokouchi**, Stable carbon isotope ratio of methyl chloride emitted from glasshouse-grown tropical plants and its implication for the global methyl chloride budget. *Geophysical Research Letters*, 2008. **35**(8).
75. **Nadalig, T., et al.**, Detection and isolation of chloromethane-degrading bacteria from the *Arabidopsis thaliana* phyllosphere, and characterization of chloromethane utilization genes. *FEMS microbiology ecology*, 2011. **77**(2): p. 438-448.
76. **Blanvillain, S., et al.**, Plant Carbohydrate Scavenging through TonB-Dependent Receptors: A. 2007.
77. **Papen, H., et al.**, Chemolithoautotrophic nitrifiers in the phyllosphere of a spruce ecosystem receiving high atmospheric nitrogen input. *Current microbiology*, 2002. **44**(1): p. 56-60.
78. **Fürnkranz, M., et al.**, Nitrogen fixation by phyllosphere bacteria associated with higher plants and their colonizing epiphytes of a tropical lowland rainforest of Costa Rica. *The ISME Journal*, 2008. **2**(5): p. 561-570.
79. **Marco, M.L., J. Legac, and S.E. Lindow**, *Pseudomonas syringae* genes induced during colonization of leaf surfaces. *Environmental microbiology*, 2005. **7**(9): p. 1379-1391.
80. **Karamanoli, K., et al.**, Polyphenolic compounds on leaves limit iron availability and affect growth of epiphytic bacteria. *Annals of Applied Biology*, 2011. **159**(1): p. 99-108.
81. **Atamna-Ismaeel, N., et al.**, Microbial rhodopsins on leaf surfaces of terrestrial plants. *Environmental microbiology*, 2012. **14**(1): p. 140-146.
82. **Atamna-Ismaeel, N., et al.**, Bacterial anoxygenic photosynthesis on plant leaf surfaces. *Environmental microbiology reports*, 2012. **4**(2): p. 209-216.
83. **Gourion, B., M. Rossignol, and J.A. Vorholt**, A proteomic study of *Methylobacterium extorquens* reveals a response regulator essential for epiphytic growth. *Proc Natl Acad Sci U S A*, 2006. **103**(35): p. 13186-91.
84. **Gourion, B., A. Francez-Charlot, and J.A. Vorholt**, PhyR is involved in the general stress response of *Methylobacterium extorquens* AM1. *Journal of bacteriology*, 2008. **190**(3): p. 1027-1035.
85. **Kaczmarczyk, A., et al.**, Role of *Sphingomonas* sp. strain Fr1 PhyR-NepR- σ EcfG cascade in general stress response and identification of a negative regulator of PhyR. *Journal of bacteriology*, 2011. **193**(23): p. 6629-6638.
86. **Hallmann, J., et al.**, Bacterial endophytes in agricultural crops. *Canadian Journal of Microbiology*, 1997. **43**(10): p. 895-914.
87. **Rosenblueth, M. and E. Martinez-Romero**, Bacterial endophytes and their interactions with hosts. *Mol Plant Microbe Interact*, 2006. **19**(8): p. 827-37.
88. **Strobel, G., et al.**, Natural products from endophytic microorganisms. *Journal of Natural products*, 2004. **67**(2): p. 257-268.
89. **Sturz, A.V. and B.G. Matheson**, Populations of endophytic bacteria which influence host-resistance to *Erwinia*-induced bacterial soft rot in potato tubers. *Plant and Soil*, 1996. **184**(2): p. 265-271.

90. **Duijff, B.J., V. Gianinazzi-Pearson, and P. Lemanceau**, Involvement of the outer membrane lipopolysaccharides in the endophytic colonization of tomato roots by biocontrol *Pseudomonas fluorescens* strain WCS417r. *New Phytologist*, 1997. **135**(2): p. 325-334.
91. **Chanway, C.P.**, Inoculation of Tree Roots with Plant Growth Promoting Soil Bacteria: An Emerging Technology for Reforestation. *Forest Science*, 1997. **43**(1): p. 99-112.
92. **Bent, E. and C.P. Chanway**, The growth-promoting effects of a bacterial endophyte on lodgepole pine are partially inhibited by the presence of other rhizobacteria. *Canadian Journal of Microbiology*, 1998. **44**(10): p. 980-988.
93. **Siciliano, S.D., et al.**, Selection of Specific Endophytic Bacterial Genotypes by Plants in Response to Soil Contamination. *Applied and Environmental Microbiology*, 2001. **67**(6): p. 2469-2475.
94. **Barac, T., et al.**, Engineered endophytic bacteria improve phytoremediation of water-soluble, volatile, organic pollutants. *Nat Biotechnol*, 2004. **22**(5): p. 583-8.
95. **Germaine, K., et al.**, Colonisation of poplar trees by gfp expressing bacterial endophytes. *FEMS Microbiol Ecol*, 2004. **48**(1): p. 109-18.
96. **Germaine, K.J., et al.**, Bacterial endophyte-enhanced phytoremediation of the organochlorine herbicide 2,4-dichlorophenoxyacetic acid. *FEMS Microbiol Ecol*, 2006. **57**(2): p. 302-10.
97. **Moore, F.P., et al.**, Endophytic bacterial diversity in poplar trees growing on a BTEX-contaminated site: the characterisation of isolates with potential to enhance phytoremediation. *Syst Appl Microbiol*, 2006. **29**(7): p. 539-56.
98. **Ryan, R.P., et al.**, An acquired efflux system is responsible for copper resistance in *Xanthomonas* strain IG-8 isolated from China. *FEMS Microbiol Lett*, 2007. **268**(1): p. 40-6.
99. **Posada, F. and F.E. Vega**, Establishment of the fungal entomopathogen *Beauveria bassiana* (Ascomycota: Hypocreales) as an endophyte in cocoa seedlings (*Theobroma cacao*). *Mycologia*, 2005. **97**(6): p. 1195-200.
100. **Miche, L. and J. Balandreau**, Effects of rice seed surface sterilization with hypochlorite on inoculated *Burkholderia vietnamiensis*. *Appl Environ Microbiol*, 2001. **67**(7): p. 3046-52.
101. **Elvira-Recuenco, M. and J.W. van Vuurde**, Natural incidence of endophytic bacteria in pea cultivars under field conditions. *Can J Microbiol*, 2000. **46**(11): p. 1036-41.
102. **Zinniel, D.K., et al.**, Isolation and characterization of endophytic colonizing bacteria from agronomic crops and prairie plants. *Appl Environ Microbiol*, 2002. **68**(5): p. 2198-208.
103. **Verma, S.C., J.K. Ladha, and A.K. Tripathi**, Evaluation of plant growth promoting and colonization ability of endophytic diazotrophs from deep water rice. *J Biotechnol*, 2001. **91**(2-3): p. 127-41.
104. **Wakelin, S., et al.**, Phosphate solubilization by *Penicillium spp.* closely associated with wheat roots. *Biology and Fertility of Soils*, 2004. **40**(1): p. 36-43.
105. **Lee, S., et al.**, Indole-3-Acetic Acid Biosynthesis Is Deficient in *Gluconacetobacter diazotrophicus* Strains with Mutations in Cytochrome c Biogenesis Genes. *Journal of Bacteriology*, 2004. **186**(16): p. 5384-5391.

106. **Costa, J.M. and J.E. Loper**, Characterization of siderophore production by the biological control agent *Enterobacter cloacae*. MPMI-Molecular Plant Microbe Interactions, 1994. **7**(4): p. 440-448.
107. **Pirttilä, A.M., et al.**, Bud endophytes of Scots pine produce adenine derivatives and other compounds that affect morphology and mitigate browning of callus cultures. *Physiologia Plantarum*, 2004. **121**(2): p. 305-312.
108. **Compant, S., et al.**, Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. Applied and environmental microbiology, 2005. **71**(9): p. 4951-4959.
109. **Compant, S., et al.**, Endophytic colonization of *Vitis vinifera* L. by plant growth-promoting bacterium *Burkholderia* sp. strain PsJN. Applied and Environmental Microbiology, 2005. **71**(4): p. 1685-1693.
110. **Gray, E. and D. Smith**, Intracellular and extracellular PGPR: commonalities and distinctions in the plant-bacterium signaling processes. Soil Biology and Biochemistry, 2005. **37**(3): p. 395-412.
111. **Kerry, B.R.**, Rhizosphere interactions and the exploitation of microbial agents for the biological control of plant-parasitic nematodes. Annual review of phytopathology, 2000. **38**(1): p. 423-441.
112. **Sturz, A., B. Christie, and J. Nowak**, Bacterial endophytes: potential role in developing sustainable systems of crop production. Critical Reviews in Plant Sciences, 2000. **19**(1): p. 1-30.
113. **Ping, L. and W. Boland**, Signals from the underground: bacterial volatiles promote growth in *Arabidopsis*. Trends in plant science, 2004. **9**(6): p. 263-266.
114. **Berg, G. and J. Hallmann**, Control of plant pathogenic fungi with bacterial endophytes, in Microbial root endophytes. 2006, Springer. p. 53-69.
115. **Van Loon, L., P. Bakker, and C. Pieterse**, Systemic resistance induced by rhizosphere bacteria. Annual review of phytopathology, 1998. **36**(1): p. 453-483.
116. **Lodewyckx, C., et al.**, Endophytic bacteria and their potential applications. Critical Reviews in Plant Sciences, 2002. **21**(6): p. 583-606.
117. **Stierle, A., G. Strobel, and D. Stierle**, Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of Pacific yew. Science, 1993. **260**(5105): p. 214-216.
118. **Miller, C., et al.**, Ecomycins, unique antimycotics from *Pseudomonas viridiflava*. Journal of applied microbiology, 1998. **84**(6): p. 937-944.
119. **Guan, S.-h., et al.**, p-Aminoacetophenonic Acids Produced by a Mangrove Endophyte: *Streptomyces griseus* subsp. Journal of natural products, 2005. **68**(8): p. 1198-1200.
120. **Castillo, U.F., et al.**, Munumbicins, wide-spectrum antibiotics produced by *Streptomyces* NRRL 30562, endophytic on *Kennedia nigriscans*. Microbiology, 2002. **148**(9): p. 2675-2685.
121. **Castillo, U., et al.**, Kakadumycins, novel antibiotics from *Streptomyces* sp. NRRL 30566, an endophyte of *Grevillea pteridifolia*. FEMS Microbiology Letters, 2003. **224**(2): p. 183-190.
122. **Beck, H.C., A.M. Hansen, and F.R. Lauritsen**, Novel pyrazine metabolites found in polymyxin biosynthesis by *Paenibacillus polymyxa*. FEMS microbiology letters, 2003. **220**(1): p. 67-73.

123. **Guo, B., et al.**, Cytonic acids A and B: novel tridepside inhibitors of hCMV protease from the endophytic fungus *Cytonaema* species. *Journal of Natural Products*, 2000. **63**(5): p. 602-604.
124. **Ezra, D., et al.**, Coronamycins, peptide antibiotics produced by a verticillate *Streptomyces* sp.(MSU-2110) endophytic on *Monstera* sp. *Microbiology*, 2004. **150**(4): p. 785-793.
125. **Mcinroy, J.A. and J.W. Kloepper**, Survey of indigenous bacterial endophytes from cotton and sweet corn. *Plant and soil*, 1995. **173**(2): p. 337-342.
126. **Mickler, C.J., K.L. Bowen, and J.W. Kloepper**, Evaluation of selected geocarposphere bacteria for biological control of *Aspergillus flavus* in peanut. *Plant and soil*, 1995. **175**(2): p. 291-299.
127. **Misaghi, I. and C. Donndelinger**, Endophytic bacteria in symptom-free cotton plants. *Phytopathology*, 1990. **80**(9): p. 808-811.
128. **Mundt, J.O. and N.F. Hinkle**, Bacteria within ovules and seeds. *Applied and Environmental Microbiology*, 1976. **32**(5): p. 694-698.
129. **Pleban, S., F. Ingel, and I. Chet**, Control of *Rhizoctonia solani* and *Sclerotium rolfsii* in the greenhouse using endophytic *Bacillus* spp. *European Journal of Plant Pathology*, 1995. **101**(6): p. 665-672.
130. **Rijavec, T., et al.**, Isolation of bacterial endophytes from germinated maize kernels. *Canadian journal of microbiology*, 2007. **53**(6): p. 802-808.
131. **Beattie, G.A. and S.E. Lindow**, The secret life of foliar bacterial pathogens on leaves. *Annual review of phytopathology*, 1995. **33**(1): p. 145-172.
132. **Compant, S., et al.**, Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches of colonization. *Microbial ecology*, 2011. **62**(1): p. 188-197.
133. **Johnston-Monje, D., et al.**, Impact of swapping soils on the endophytic bacterial communities of pre-domesticated, ancient and modern maize. *BMC plant biology*, 2014. **14**(1): p. 233.
134. **Mousa, W.K., et al.**, Bacterial endophytes from wild maize suppress *Fusarium graminearum* in modern maize and inhibit mycotoxin accumulation. *Front Plant Sci*, 2015. **6**: p. 805.
135. **Perret, X., C. Staehelin, and W.J. Broughton**, Molecular basis of symbiotic promiscuity. *Microbiology and Molecular Biology Reviews*, 2000. **64**(1): p. 180-201.
136. **Peck, M.C., R.F. Fisher, and S.R. Long**, Diverse flavonoids stimulate NodD1 binding to nod gene promoters in *Sinorhizobium meliloti*. *Journal of bacteriology*, 2006. **188**(15): p. 5417-5427.
137. **Oldroyd, G.E. and J.A. Downie**, Calcium, kinases and nodulation signalling in legumes. *Nature Reviews Molecular Cell Biology*, 2004. **5**(7): p. 566-576.
138. **Timmers, A., M.-C. Auriac, and G. Truchet**, Refined analysis of early symbiotic steps of the *Rhizobium-Medicago* interaction in relationship with microtubular cytoskeleton rearrangements. *Development*, 1999. **126**(16): p. 3617-3628.
139. **Sieberer, B.J., A.C. Timmers, and A.M.C. Emons**, Nod factors alter the microtubule cytoskeleton in *Medicago truncatula* root hairs to allow root hair reorientation. *Molecular plant-microbe interactions*, 2005. **18**(11): p. 1195-1204.

140. **Cárdenas, L., et al.**, The role of nod factor substituents in actin cytoskeleton rearrangements in *Phaseolus vulgaris*. *Molecular plant-microbe interactions*, 2003. **16**(4): p. 326-334.
141. **de Ruijter, N.C., T. Bisseling, and A.M.C. Emons**, *Rhizobium* Nod factors induce an increase in sub-apical fine bundles of actin filaments in *Vicia sativa* root hairs within minutes. *Molecular plant-microbe interactions*, 1999. **12**(9): p. 829-832.
142. **Esseling, J.J., F.G. Lhuissier, and A.M.C. Emons**, Nod factor-induced root hair curling: continuous polar growth towards the point of nod factor application. *Plant Physiology*, 2003. **132**(4): p. 1982-1988.
143. **Gage, D.J.**, Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiology and Molecular Biology Reviews*, 2004. **68**(2): p. 280-300.
144. **Foucher, F. and E. Kondorosi**, Cell cycle regulation in the course of nodule organogenesis in Medicago. *Plant molecular biology*, 2000. **43**(5-6): p. 773-786.
145. **Wasson, A.P., F.I. Pellerone, and U. Mathesius**, Silencing the flavonoid pathway in *Medicago truncatula* inhibits root nodule formation and prevents auxin transport regulation by rhizobia. *The Plant Cell*, 2006. **18**(7): p. 1617-1629.
146. **Dickstein, R., et al.**, Expression of nodule-specific genes in alfalfa root nodules blocked at an early stage of development. *Genes & development*, 1988. **2**(6): p. 677-687.
147. **Dylan, T., et al.**, Symbiotic pseudorevertants of *Rhizobium meliloti* ndv mutants. *Journal of bacteriology*, 1990. **172**(3): p. 1409-1417.
148. **Yao, S.-Y., et al.**, *Sinorhizobium meliloti* ExoR and ExoS proteins regulate both succinoglycan and flagellum production. *Journal of bacteriology*, 2004. **186**(18): p. 6042-6049.
149. **Klein, S., et al.**, Interaction of nod and exo *Rhizobium meliloti* in alfalfa nodulation. *Mol. Plant-Microbe Interact*, 1988. **1**: p. 94-100.
150. **Gage, D.J.**, Analysis of infection thread development using Gfp-and DsRed-expressing *Sinorhizobium meliloti*. *Journal of Bacteriology*, 2002. **184**(24): p. 7042-7046.
151. **Pellock, B.J., H.-P. Cheng, and G.C. Walker**, Alfalfa Root Nodule Invasion Efficiency Is Dependent on *Sinorhizobium meliloti* polysaccharides. *Journal of bacteriology*, 2000. **182**(15): p. 4310-4318.
152. **Glazebrook, J. and G.C. Walker**, A novel exopolysaccharide can function in place of the calcofluor-binding exopolysaccharide in nodulation of alfalfa by *Rhizobium meliloti*. *Cell*, 1989. **56**(4): p. 661-672.
153. **Bhagwat, A.A., et al.**, Further Studies of the Role of Cyclic β -Glucans in symbiosis. An ndvC mutant of *Bradyrhizobium japonicum* synthesizes Cyclodecakis-(1 \rightarrow 3)- β -Glucosyl. *Plant physiology*, 1999. **119**(3): p. 1057-1064.
154. **Arrighi, J.-F., et al.**, The *Medicago truncatula* lysine motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant physiology*, 2006. **142**(1): p. 265-279.
155. **Limpens, E., et al.**, LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science*, 2003. **302**(5645): p. 630-633.
156. **Ardourel, M., et al.**, *Rhizobium meliloti* lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and

- induction of plant symbiotic developmental responses. *The Plant Cell*, 1994. **6**(10): p. 1357-1374.
157. **Mergaert, P., et al.**, Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*–legume symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, 2006. **103**(13): p. 5230-5235.
158. **Cebolla, A., et al.**, The mitotic inhibitor *ccs52* is required for endoreduplication and ploidy-dependent cell enlargement in plants. *The EMBO Journal*, 1999. **18**(16): p. 4476-4484.
159. **Galitski, T., et al.**, Ploidy regulation of gene expression. *Science*, 1999. **285**(5425): p. 251-254.
160. **Brewin, N.J.**, Plant cell wall remodelling in the *Rhizobium*–legume symbiosis. *Critical Reviews in Plant Sciences*, 2004. **23**(4): p. 293-316.
161. **Robertson, J.G. and P. Lyttleton**, Division of peribacteroid membranes in root nodules of white clover. *Journal of cell science*, 1984. **69**(1): p. 147-157.
162. **Campbell, G.R., B.L. Reuhs, and G.C. Walker**, Chronic intracellular infection of alfalfa nodules by *Sinorhizobium meliloti* requires correct lipopolysaccharide core. *Proceedings of the National Academy of Sciences*, 2002. **99**(6): p. 3938-3943.
163. **Davies, B.W. and G.C. Walker**, Disruption of *sitA* compromises *Sinorhizobium meliloti* for manganese uptake required for protection against oxidative stress. *Journal of bacteriology*, 2007. **189**(5): p. 2101-2109.
164. **Putnoky, P., et al.**, The *pha* gene cluster of *Rhizobium meliloti* involved in pH adaptation and symbiosis encodes a novel type of K⁺ efflux system. *Molecular microbiology*, 1998. **28**(6): p. 1091-1101.
165. **Mitsui, H., et al.**, *Sinorhizobium meliloti* RpoH1 is required for effective nitrogen-fixing symbiosis with alfalfa. *Molecular Genetics and Genomics*, 2004. **271**(4): p. 416-425.
166. **Wells, D.H. and S.R. Long**, The *Sinorhizobium meliloti* stringent response affects multiple aspects of symbiosis. *Molecular microbiology*, 2002. **43**(5): p. 1115-1127.
167. **Bardin, S., et al.**, A phosphate transport system is required for symbiotic nitrogen fixation by *Rhizobium meliloti*. *Journal of bacteriology*, 1996. **178**(15): p. 4540-4547.
168. **Yuan, Z.-C., R. Zaheer, and T.M. Finan**, Regulation and properties of PstSCAB, a high-affinity, high-velocity phosphate transport system of *Sinorhizobium meliloti*. *Journal of bacteriology*, 2006. **188**(3): p. 1089-1102.
169. **Fischer, H.-M.**, Genetic regulation of nitrogen fixation in rhizobia. *Microbiological reviews*, 1994. **58**(3): p. 352.
170. **Barnett, M.J. and R.F. Fisher**, Global gene expression in the rhizobial-legume symbiosis. *Symbiosis (Rehovot)*, 2006. **42**(1): p. 1-24.
171. **Poole, P. and D. Allaway**, Carbon and nitrogen metabolism in *Rhizobium*. *Advances in microbial physiology*, 2000. **43**: p. 117-163.
172. **Prell, J. and P. Poole**, Metabolic changes of rhizobia in legume nodules. *Trends in microbiology*, 2006. **14**(4): p. 161-168.
173. **Day, D., et al.**, Ammonia and amino acid transport across symbiotic membranes in nitrogen-fixing legume nodules. *Cellular and Molecular Life Sciences CMLS*, 2001. **58**(1): p. 61-71.

174. **Barsch, A., et al.**, GC–MS based metabolite profiling implies three interdependent ways of ammonium assimilation in *Medicago truncatula* root nodules. *Journal of biotechnology*, 2006. **127**(1): p. 79-83.
175. **Harrison, J., et al.**, Does lowering glutamine synthetase activity in nodules modify nitrogen metabolism and growth of *Lotus japonicus*? *Plant physiology*, 2003. **133**(1): p. 253-262.
176. **Willis, L.B. and G.C. Walker**, The phbC (poly- β -hydroxybutyrate synthase) gene of *Rhizobium (Sinorhizobium) meliloti* and characterization of phbC mutants. *Canadian journal of microbiology*, 1998. **44**(6): p. 554-564.
177. **Aneja, P., A. Zachertowska, and T. Charles**, Comparison of the symbiotic and competition phenotypes of *Sinorhizobium meliloti* PHB synthesis and degradation pathway mutants. *Canadian journal of microbiology*, 2005. **51**(7): p. 599-604.
178. **Ronson, C.W., P. Lyttleton, and J.G. Robertson**, C4-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. *Proceedings of the National Academy of Sciences*, 1981. **78**(7): p. 4284-4288.
179. **Driscoll, B.T. and T.M. Finan**, NAD⁺-dependent malic enzyme of *Rhizobium meliloti* is required for symbiotic nitrogen fixation. *Molecular microbiology*, 1993. **7**(6): p. 865-873.
180. **Ott, T., et al.**, Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. *Current biology*, 2005. **15**(6): p. 531-535.
181. **Gordon, A.J., et al.**, Sucrose synthase in legume nodules is essential for nitrogen fixation. *Plant Physiology*, 1999. **120**(3): p. 867-878.
182. **Krusell, L., et al.**, The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. *The Plant Cell*, 2005. **17**(5): p. 1625-1636.
183. **Wienkoop, S. and G. Saalbach**, Proteome analysis. Novel proteins identified at the peribacteroid membrane from *Lotus japonicus* root nodules. *Plant Physiology*, 2003. **131**(3): p. 1080-1090.
184. **Felix, G., et al.**, Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal*, 1999. **18**(3): p. 265-276.
185. **Boller, T.**, Peptide signalling in plant development and self/non-self perception. *Current opinion in cell biology*, 2005. **17**(2): p. 116-122.
186. **van Brussel, A.A., et al.**, Autoregulation of root nodule formation: signals of both symbiotic partners studied in a split-root system of *Vicia sativa* subsp. *nigra*. *Molecular plant-microbe interactions*, 2002. **15**(4): p. 341-349.
187. **Mithöfer, A.**, Suppression of plant defence in rhizobia–legume symbiosis. *Trends in plant science*, 2002. **7**(10): p. 440-444.
188. **Scheidle, H., A. Groß, and K. Niehaus**, The lipid A substructure of the *Sinorhizobium meliloti* lipopolysaccharides is sufficient to suppress the oxidative burst in host plants. *New phytologist*, 2005. **165**(2): p. 559-566.
189. **Tellström, V., et al.**, The lipopolysaccharide of *Sinorhizobium meliloti* suppresses defense-associated gene expression in cell cultures of the host plant *Medicago truncatula*. *Plant physiology*, 2007. **143**(2): p. 825-837.

190. **Handelsman, J. and W.J. Brill**, *Erwinia herbicola* isolates from alfalfa plants may play a role in nodulation of alfalfa by *Rhizobium meliloti*. Applied and environmental microbiology, 1985. **49**(4): p. 818-821.
191. **Stajković, O., et al.**, Isolation and characterization of endophytic non-rhizobial bacteria from root nodules of alfalfa (*Medicago sativa* L.). Botanica serbica, 2009. **33**(1): p. 107-114.
192. **Pini, F., et al.**, Exploring the plant-associated bacterial communities in *Medicago sativa* L. BMC microbiology, 2012. **12**(1): p. 78.
193. **Bentley, S.D. and J. Parkhill**, Comparative genomic structure of prokaryotes. Annual Review of Genetics, 2004. **38**: p. 771-792.
194. **Guerrero, R. and M. Berlanga**, Life's unity and flexibility: The ecological link. International Microbiology, 2006. **9**(3): p. 225-235.
195. **Sallstrom, B. and S.G.E. Andersson**, Genome reduction in the α -proteobacteria. Current Opinion in Microbiology, 2005. **8**(5): p. 579-585.
196. **Binnewies, T.T., et al.**, Ten years of bacterial genome sequencing: Comparative-genomics-based discoveries. Functional and Integrative Genomics, 2006. **6**(3): p. 165-185.
197. **Ilyina, T.S.**, Bacterial superintegrons, a source of new genes with adaptive functions. Russian Journal of Genetics, 2006. **42**(11): p. 1294-1302.
198. **Sundin, G.W.**, Genomic insights into the contribution of phytopathogenic bacterial plasmids to the evolutionary history of their hosts. Annual review of phytopathology, 2007. **45**: p. 129-151.
199. **Sorensen, S.J., et al.**, Studying plasmid horizontal transfer in situ: a critical review. Nat Rev Microbiol, 2005. **3**(9): p. 700-10.
200. **Lawrence, J.G. and H. Ochman**, Amelioration of bacterial genomes: Rates of change and exchange. Journal of Molecular Evolution, 1997. **44**(4): p. 383-397.
201. **Syvanen, M.**, Horizontal gene transfer: Evidence and possible consequences. Annual Review of Genetics, 1994. **28**: p. 237-261.
202. **Lawrence, J.G. and H. Ochman**, Molecular archaeology of the *Escherichia coli* genome. Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(16): p. 9413-9417.
203. **Janssen, D.B., et al.**, Bacterial degradation of xenobiotic compounds: Evolution and distribution of novel enzyme activities. Environmental Microbiology, 2005. **7**(12): p. 1868-1882.
204. **Arber, W.**, Genetic variation: molecular mechanisms and impact on microbial evolution. FEMS Microbiol Rev, 2000. **24**(1): p. 1-7.
205. **Davison, J.**, Genetic exchange between bacteria in the environment. Plasmid, 1999. **42**(2): p. 73-91.
206. **Øvreås, L.**, Population and community level approaches for analysing microbial diversity in natural environments. Ecology Letters, 2000. **3**(3): p. 236-251.
207. **Torsvik, V., et al.**, Novel techniques for analysing microbial diversity in natural and perturbed environments. Journal of Biotechnology, 1998. **64**(1): p. 53-62.

208. **Courtois, S., et al.**, Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl Environ Microbiol*, 2003. **69**(1): p. 49-55.
209. **Daniel, R.**, Construction of environmental libraries for functional screening of enzyme activity., in *Directed Molecular Evolution of Proteins.*, S. Brakmann, Johnsson, K., Weinheim,, Editor. 2002, Wiley-VCH Verlag GmbH.
210. **Osburne, M.S., et al.**, Tapping into microbial diversity for natural products drug discovery. *ASM News*, 2000. **66**: p. 411-417.
211. **Sentchilo, V., et al.**, Community-wide plasmid gene mobilization and selection. *The ISME journal*, 2013. **7**(6): p. 1173-1186.
212. **Dealtry, S., et al.**, Cultivation-independent screening revealed hot spots of IncP-1, IncP-7 and IncP-9 plasmid occurrence in different environmental habitats. *PloS one*, 2014. **9**(2): p. e89922.
213. **Palenik, B., et al.**, Coastal *Synechococcus* metagenome reveals major roles for horizontal gene transfer and plasmids in population diversity. *Environmental microbiology*, 2009. **11**(2): p. 349-359.
214. **Jones, B.V. and J.R. Marchesi**, Transposon-aided capture (TRACA) of plasmids resident in the human gut mobile metagenome. *Nature Methods*, 2007. **4**(1): p. 55-61.
215. **Kav, A.B., I. Benhar, and I. Mizrahi**, A method for purifying high quality and high yield plasmid DNA for metagenomic and deep sequencing approaches. *Journal of microbiological methods*, 2013. **95**(2): p. 272-279.
216. **Garcia-de Los Santos, A., S. Brom, and D. Romero**, *Rhizobium plasmids* in bacteria-legume interactions. *World Journal of Microbiology and Biotechnology*, 1996. **12**(2): p. 119-125.
217. **Mercado-Blanco, J. and N. Toro**, Plasmids in rhizobia: the role of nonsymbiotic plasmids. *MPMI-Molecular Plant Microbe Interactions*, 1996. **9**(7): p. 535-545.
218. **Banfalvi, Z., E. Kondorosi, and A. Kondorosi**, *Rhizobium meliloti* carries two megaplasmids. *Plasmid*, 1985. **13**(2): p. 129-138.
219. **Hynes, M.F., et al.**, The two megaplasmids of *Rhizobium meliloti* are involved in the effective nodulation of alfalfa. *Molecular and General Genetics MGG*, 1986. **202**(3): p. 356-362.
220. **Banfalvi, Z., et al.**, Location of nodulation and nitrogen fixation genes on a high molecular weight plasmid of *R. meliloti*. *Molecular and General Genetics MGG*, 1981. **184**(2): p. 318-325.
221. **Smets, B.F. and T. Barkay**, Horizontal gene transfer: perspectives at a crossroads of scientific disciplines. *Nature Reviews Microbiology*, 2005. **3**(9): p. 675-678.
222. **Pistorio, M., et al.**, Identification of a transmissible plasmid from an Argentine *Sinorhizobium meliloti* strain which can be mobilised by conjugative helper functions of the European strain *S. meliloti* GR4. *FEMS microbiology letters*, 2003. **225**(1): p. 15-21.
223. **Roumiantseva, M.L., et al.**, Diversity of *Sinorhizobium meliloti* from the Central Asian alfalfa gene center. *Applied and environmental microbiology*, 2002. **68**(9): p. 4694-4697.
224. **Stiens, M., et al.**, Sequence analysis of the 144-kilobase accessory plasmid pSmeSM11a, isolated from a dominant *Sinorhizobium meliloti* strain identified during a

- long-term field release experiment. Applied and environmental microbiology, 2006. **72**(5): p. 3662-3672.
225. **Kosier, B., A. Pühler, and R. Simon**, Monitoring the diversity of *Rhizobium meliloti* field and microcosm isolates with a novel rapid genotyping method using insertion elements. Molecular Ecology, 1993. **2**(1): p. 35-46.
226. **Mercado-Blanco, J. and J. Olivares**, Stability and transmissibility of the cryptic plasmids of *Rhizobium meliloti* GR4. Archives of microbiology, 1993. **160**(6): p. 477-485.
227. **Barnett, M.J., et al.**, Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. Proceedings of the National Academy of Sciences, 2001. **98**(17): p. 9883-9888.
228. **Finan, T.M., et al.**, The complete sequence of the 1,683-kb pSymB megaplasmid from the N₂-fixing endosymbiont *Sinorhizobium meliloti*. Proceedings of the National Academy of Sciences, 2001. **98**(17): p. 9889-9894.
229. **Giuntini, E., et al.**, Large-scale genetic variation of the symbiosis-required megaplasmid pSymA revealed by comparative genomic analysis of *Sinorhizobium meliloti* natural strains. BMC genomics, 2005. **6**(1): p. 158.
230. **Oresnik, I.J., et al.**, Megaplasmid pRme2011a of *Sinorhizobium meliloti* is not required for viability. Journal of bacteriology, 2000. **182**(12): p. 3582-3586.
231. **Galardini, M., et al.**, Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*. Genome biology and evolution, 2013. **5**(3): p. 542-558.
232. **Galardini, M., et al.**, Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. BMC genomics, 2011. **12**(1): p. 235.
233. **Schneiker-Bekel, S., et al.**, The complete genome sequence of the dominant *Sinorhizobium meliloti* field isolate SM11 extends the *S. meliloti* pan-genome. Journal of biotechnology, 2011. **155**(1): p. 20-33.
234. **Hahn, M. and H. Hennecke**, Mapping of a *Bradyrhizobium japonicum* DNA region carrying genes for symbiosis and an asymmetric accumulation of reiterated sequences. Applied and environmental microbiology, 1987. **53**(9): p. 2247-2252.
235. **Kaluza, K., M. Hahn, and H. Hennecke**, Repeated sequences similar to insertion elements clustered around the nif region of the *Rhizobium japonicum* genome. Journal of bacteriology, 1985. **162**(2): p. 535-542.
236. **Brom, S., et al.**, High-frequency rearrangements in *Rhizobium leguminosarum* bv. phaseoli plasmids. Journal of bacteriology, 1991. **173**(3): p. 1344-1346.
237. **Lagares, A., J. Sanjuán, and M. Pistorio**, The Plasmid Mobilome of the Model Plant-Symbiont *Sinorhizobium meliloti*: Coming up with New Questions and Answers. Microbiology spectrum, 2014. **2**(5).
238. **Kuhn, S., et al.**, Prevalence of pSmeSM11a-like plasmids in indigenous *Sinorhizobium meliloti* strains isolated in the course of a field release experiment with genetically modified *S. meliloti* strains. FEMS microbiology ecology, 2008. **63**(1): p. 118-131.
239. **Bromfield, E., D. Lewis, and L. Barran**, Cryptic plasmid and rifampin resistance in *Rhizobium meliloti* influencing nodulation competitiveness. Journal of bacteriology, 1985. **164**(1): p. 410-413.

240. **García-Rodríguez, F.M. and N. Toro**, *Sinorhizobium meliloti* nfe (nodulation formation efficiency) genes exhibit temporal and spatial expression patterns similar to those of genes involved in symbiotic nitrogen fixation. *Molecular plant-microbe interactions*, 2000. **13**(6): p. 583-591.
241. **Velazquez, E., et al.**, Attenuation of symbiotic effectiveness by *Rhizobium meliloti* SAF22 related to the presence of a cryptic plasmid. *Applied and environmental microbiology*, 1995. **61**(5): p. 2033-2036.
242. **Eckhardt, T.**, A rapid method for the identification of plasmid desoxyribonucleic acid in bacteria. *Plasmid*, 1978. **1**(4): p. 584-588.
243. **Coutinho, H.L., et al.**, Evaluation of the diversity of rhizobia in Brazilian agricultural soils cultivated with soybeans. *Applied Soil Ecology*, 1999. **13**(2): p. 159-167.
244. **Cevallos, M.A., R. Cervantes-Rivera, and R.M. Gutiérrez-Ríos**, The repABC plasmid family. *Plasmid*, 2008. **60**(1): p. 19-37.
245. **Pinto, U.M., K.M. Pappas, and S.C. Winans**, The ABCs of plasmid replication and segregation. *Nature Reviews Microbiology*, 2012. **10**(11): p. 755-765.
246. **Bertani, G.**, STUDIES ON LYSOGENESIS I.: The Mode of Phage Liberation by Lysogenic *Escherichia coli*1. *Journal of bacteriology*, 1951. **62**(3): p. 293.
247. **Beringer, J.**, R factor transfer in *Rhizobium leguminosarum*. *Journal of general microbiology*, 1974. **84**(1): p. 188-198.
248. **Froni, L.**, Ciclo biológico del fósforo. *Procesos microbianos*, 1999: p. 273.
249. **Takiguchi, Y. and K. Shimahara**, N,N-Diacetylchitobiose production from chitin by *Vibrio anguillarum* strain E-383a.
250. **Schwyn, B. and J.B. Neilands.**
251. **Hankin, L. and S.L. Anagnostakis**, Solid media containing carboxymethylcellulose to detect CX cellulose activity of micro-organisms. *J Gen Microbiol*, 1977. **98**(1): p. 109-15.
252. **Patten, C.L. and B.R. Glick**, Role of *Pseudomonas putida* indoleacetic acid in development of the host plant root system. *Appl Environ Microbiol*, 2002. **68**(8): p. 3795-801.
253. **Walsh, G.A., et al.**, Technical note: detection and quantification of supplemental fungal beta-glucanase activity in animal feed. *J Anim Sci*, 1995. **73**(4): p. 1074-6.
254. **Cordero, P., et al.**, Inhibition of the phytopathogenic fungus *Fusarium proliferatum* by volatile compounds produced by *Pseudomonas*. *Archives of microbiology*, 2014. **196**(11): p. 803-809.
255. **Parret, A.H.A., et al.**, Plant Lectin-Like Bacteriocin from a Rhizosphere-Colonizing *Pseudomonas* Isolate. *Journal of Bacteriology*, 2003. **185**(3): p. 897-908.
256. **Weisburg, W.G., et al.**, 16S ribosomal DNA amplification for phylogenetic study. *Journal of bacteriology*, 1991. **173**(2): p. 697-703.
257. **Young, J., H. Downer, and B. Eardly**, Phylogeny of the phototrophic rhizobium strain BTAi1 by polymerase chain reaction-based sequencing of a 16S rRNA gene segment. *Journal of Bacteriology*, 1991. **173**(7): p. 2271-2277.
258. **Maier, T., et al.**, Fast and reliable maldi-tof ms–based microorganism identification. *Nature Methods| Application Notes*, 2006.

259. **Ferreira, L., et al.**, MALDI-TOF mass spectrometry is a fast and reliable platform for identification and ecological studies of species from family Rhizobiaceae. *PLoS One*, 2011. **6**(5): p. e20223.
260. **Wheatcroft, R., D. McRae, and R. Miller**, Changes in the *Rhizobium meliloti* genome and the ability to detect supercoiled plasmids during bacteroid development. *Mol. Plant-Microbe Interact*, 1990. **3**(1): p. 9-17.
261. **Jouanin, L., et al.**, DNA sequence homology in *Rhizobium meliloti* plasmids. *Molecular and General Genetics MGG*, 1981. **182**(2): p. 189-195.
262. **Rothberg, J.M., et al.**, An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. **475**(7356): p. 348-52.
263. **Gordon, D., C. Abajian, and P. Green**, Consed: a graphical tool for sequence finishing. *Genome research*, 1998. **8**(3): p. 195-202.
264. **Bartels, D., et al.**, BACCardI—a tool for the validation of genomic assemblies, assisting genome finishing and intergenome comparison. *Bioinformatics*, 2005. **21**(7): p. 853-859.
265. **Meyer, F., et al.**, GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic acids research*, 2003. **31**(8): p. 2187-2195.
266. **Wu, S., et al.**, WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC genomics*, 2011. **12**(1): p. 444.
267. **Xia, J., et al.**, MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic acids research*, 2015. **43**(W1): p. W251-W257.
268. **Blom, J., et al.**, EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *Bmc Bioinformatics*, 2009. **10**(1): p. 154.
269. **Fitch, W.M.**, Distinguishing homologous from analogous proteins. *Systematic Biology*, 1970. **19**(2): p. 99-113.
270. **Altenhoff, A.M. and C. Dessimoz**, Phylogenetic and functional assessment of orthologs inference projects and methods. *PLoS Comput Biol*, 2009. **5**(1): p. e1000262.
271. **Hulsen, T., et al.**, Benchmarking ortholog identification methods using functional genomics data. *Genome biology*, 2006. **7**(4): p. R31.
272. **Lerat, E., V. Daubin, and N.A. Moran**, From gene trees to organismal phylogeny in prokaryotes: the case of the γ -Proteobacteria. *PLoS Biol*, 2003. **1**(1): p. e19.
273. **Sharp, P.M. and W.-H. Li**, An evolutionary perspective on synonymous codon usage in unicellular organisms. *Journal of molecular evolution*, 1986. **24**(1-2): p. 28-38.
274. **McInerney, J.O.**, GCUA: general codon usage analysis. *Bioinformatics*, 1998. **14**(4): p. 372-373.
275. **Harrison, P.W., et al.**, Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol*, 2010. **18**(4): p. 141-8.
276. **Galibert, F., et al.**, The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science*, 2001. **293**(5530): p. 668-672.
277. **Tatusov, R.L., E.V. Koonin, and D.J. Lipman**, A genomic perspective on protein families. *Science*, 1997. **278**(5338): p. 631-637.

278. **Zheng, J., et al.**, Evolution and dynamics of megaplasmids with genome sizes larger than 100 kb in the *Bacillus cereus* group. *BMC evolutionary biology*, 2013. **13**(1): p. 262.
279. **Plotkin, J.B. and G. Kudla**, Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics*, 2011. **12**(1): p. 32-42.
280. **Shabalina, S.A., N.A. Spiridonov, and A. Kashina**, Sounds of silence: synonymous nucleotides as a key to biological regulation and complexity. *Nucleic acids research*, 2013. **41**(4): p. 2073-2094.
281. **Mehnaz, S., D.N. Baig, and G. Lazarovits**, Genetic and phenotypic diversity of plant growth promoting rhizobacteria isolated from sugarcane plants growing in Pakistan. *J Microbiol Biotechnol*, 2010. **20**(12): p. 1614-1623.
282. **Syvanen, M.**, Horizontal gene transfer: evidence and possible consequences. *Annu Rev Genet*, 1994. **28**: p. 237-61.
283. **Lawrence, J.G. and H. Ochman**, Amelioration of bacterial genomes: rates of change and exchange. *J Mol Evol*, 1997. **44**(4): p. 383-97.
284. **Lawrence, J.G. and H. Ochman**, Molecular archaeology of the *Escherichia coli* genome. *Proc Natl Acad Sci U S A*, 1998. **95**(16): p. 9413-7.
285. **Arber, W.**, Genetic variation: molecular mechanisms and impact on microbial evolution.
286. **Francia, M.V., et al.**, A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiol Rev*, 2004. **28**(1): p. 79-100.
287. **Ding, H. and M.F. Hynes**, Plasmid transfer systems in the rhizobia. *Can J Microbiol*, 2009. **55**(8): p. 917-27.
288. **Giusti Mde, L., et al.**, Genetic and functional characterization of a yet-unclassified rhizobial Dtr (DNA-transfer-and-replication) region from a ubiquitous plasmid conjugal system present in *Sinorhizobium meliloti*, in *Sinorhizobium medicae*, and in other nonrhizobial Gram-negative bacteria. *Plasmid*, 2012. **67**(3): p. 199-210.
290. **Seaver, L.C. and J.A. Imlay**, Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *J Bacteriol*, 2001. **183**(24): p. 7173-81.
291. **Holmgren, A.**, Thioredoxin and glutaredoxin systems. *J Biol Chem*, 1989. **264**(24): p. 13963-6.
292. **Rossjohn, J., et al.**, A mixed disulfide bond in bacterial glutathione transferase: functional and evolutionary implications. *Structure*, 1998. **6**(6): p. 721-34.
293. **Ezraty, B., L. Aussel, and F. Barras**, Methionine sulfoxide reductases in prokaryotes. *Biochim Biophys Acta*, 2005. **1703**(2): p. 221-9.
294. **Tsilibaris, V., G. Maenhaut-Michel, and L. Van Melderen**, Biological roles of the Lon ATP-dependent protease. *Research in microbiology*, 2006. **157**(8): p. 701-713.
295. **Brader, G., et al.**, Metabolic potential of endophytic bacteria. *Current opinion in biotechnology*, 2014. **27**: p. 30-37.
296. **Jones, K.M., et al.**, How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. *Nature Reviews Microbiology*, 2007. **5**(8): p. 619-633.
297. **Liu, Y., et al.**, Study on diversity of endophytic bacterial communities in seeds of hybrid maize and their parental lines. *Archives of microbiology*, 2012. **194**(12): p. 1001-1012.

298. **Kaga, H., et al.**, Rice seeds as sources of endophytic bacteria. *Microbes and environments*, 2009. **24**(2): p. 154-162.
299. **Cottyn, B., et al.**, Bacterial populations associated with rice seed in the tropical environment. *Phytopathology*, 2001. **91**(3): p. 282-292.
300. **Kremer, R.J.**, Identity and properties of bacteria inhabiting seeds of selected broadleaf weed species. *Microbial ecology*, 1987. **14**(1): p. 29-37.
301. **Ramaprasad, E. and S. Ch.**, *Roseomonas oryzae* sp. nov., isolated from paddy rhizosphere soil. *International Journal of Systematic and Evolutionary Microbiology*, 2015.
302. **Guiñazú, L.B., et al.**, Evaluation of rhizobacterial isolates from Argentina, Uruguay and Chile for plant growth-promoting characteristics and antagonistic activity towards *Rhizoctonia* sp. and *Macrophomina* sp. in vitro. *European Journal of Soil Biology*, 2013. **54**: p. 69-77.
303. **Kadyan, S., et al.**, Assessment of functional and genetic diversity of aerobic endospore forming *Bacilli* from rhizospheric soil of *Phyllanthus amarus* L. *World Journal of Microbiology and Biotechnology*, 2013. **29**(9): p. 1597-1610.
304. **Preveena, J. and S.J. Bhore**, Identification of bacterial endophytes associated with traditional medicinal plant *Tridax procumbens* Linn. *Ancient science of life*, 2013. **32**(3): p. 173.
305. **An, S.-Y., et al.**, *Terribacillus saccharophilus* gen. nov., sp. nov. and *Terribacillus halophilus* sp. nov., spore-forming bacteria isolated from field soil in Japan. *International journal of systematic and evolutionary microbiology*, 2007. **57**(1): p. 51-55.
306. **Liu, W., et al.**, *Terribacillus ailingensis* sp. nov., a moderately halophilic bacterium. *International journal of systematic and evolutionary microbiology*, 2010. **60**(12): p. 2940-2945.
307. **Enya, J., et al.**, Culturable leaf-associated bacteria on tomato plants and their potential as biological control agents. *Microbial ecology*, 2007. **53**(4): p. 524-536.
308. **Gao, S., et al.**, *Novosphingobium fluoreni* sp. nov., isolated from rice seeds. *Int J Syst Evol Microbiol*, 2015. **65**(Pt 5): p. 1409-14.
309. **Adesemoye, A., H. Torbert, and J. Kloepper**, Plant growth-promoting rhizobacteria allow reduced application rates of chemical fertilizers. *Microbial Ecology*, 2009. **58**(4): p. 921-929.
310. **Martínez-Viveros, O., et al.**, Mechanisms and practical considerations involved in plant growth promotion by rhizobacteria. *Journal of soil science and plant nutrition*, 2010. **10**(3): p. 293-319.
311. **Majeed, A., et al.**, Isolation and characterization of plant growth-promoting rhizobacteria from wheat rhizosphere and their effect on plant growth promotion. *Frontiers in microbiology*, 2015. **6**.
312. **Crowley, D.E.**, Microbial siderophores in the plant rhizosphere, in *Iron nutrition in plants and rhizospheric microorganisms*. 2006, Springer. p. 169-198.
313. **Siddiqui, Z.A.**, PGPR: prospective biocontrol agents of plant pathogens, in *PGPR: biocontrol and biofertilization*. 2006, Springer. p. 111-142.

314. **Hider, R.C. and X. Kong**, Chemistry and biology of siderophores. Natural product reports, 2010. **27**(5): p. 637-657.
315. **Wilson, M.K., et al.**, Siderophores of *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*. Biochemical and biophysical research communications, 2006. **348**(1): p. 320-325.
316. **Budzikiewicz, H., et al.**, Corynebactin, a cyclic catecholate siderophore from. Zeitschrift für Naturforschung C, 1997. **52**(7-8): p. 551-554.
317. **Cornelis, P.**, Iron uptake and metabolism in pseudomonads. Applied microbiology and biotechnology, 2010. **86**(6): p. 1637-1645.
318. **Smith, M.J., et al.**, Rhizobactin, a structurally novel siderophore from *Rhizobium meliloti*. Journal of the American Chemical Society, 1985. **107**(6): p. 1739-1743.
319. **Charlang, G., et al.**, Extracellular siderophores of rapidly growing *Aspergillus nidulans* and *Penicillium chrysogenum*. Journal of bacteriology, 1982. **150**(2): p. 785-787.
320. **Noordman, W., et al.**, Growth stimulation of *Brevibacterium* sp. by siderophores. Journal of applied microbiology, 2006. **101**(3): p. 637-646.
321. **Fenwick, B., et al.**, Iron repressible outer membrane proteins of *Moraxella bovis* and demonstration of siderophore-like activity. Veterinary microbiology, 1996. **48**(3): p. 315-324.
322. **Rajendran, G., M.H. Patel, and S.J. Joshi**, Isolation and characterization of nodule-associated *Exiguobacterium* sp. from the root nodules of Fenugreek (*Trigonella foenum-graecum*) and their possible role in plant growth promotion. International journal of microbiology, 2012. **2012**.
323. **Matzanke, B.F., et al.**, Transport and utilization of ferrioxamine-E-bound iron in *Erwinia herbicola* (*Pantoea agglomerans*). Biology of metals, 1991. **4**(3): p. 181-185.
324. **Haag, H., et al.**, Isolation and biological characterization of staphyloferrin B, a compound with siderophore activity from staphylococci. FEMS microbiology letters, 1994. **115**(2-3): p. 125-130.
325. **Chaiharn, M., S. Chunhaleuchanon, and S. Lumyong**, Screening siderophore producing bacteria as potential biological control agent for fungal rice pathogens in Thailand. World Journal of Microbiology and Biotechnology, 2009. **25**(11): p. 1919-1928.
326. **Berg, G., P. Marten, and G. Ballin**, *Stenotrophomonas maltophilia* in the rhizosphere of oilseed rape—occurrence, characterization and interaction with phytopathogenic fungi. Microbiological Research, 1996. **151**(1): p. 19-27.
327. **Farina, R., et al.**, Diversity of plant growth-promoting rhizobacteria communities associated with the stages of canola growth. Applied soil ecology, 2012. **55**: p. 44-52.
328. **Son, H.-J., et al.**, Solubilization of insoluble inorganic phosphates by a novel salt- and pH-tolerant *Pantoea agglomerans* R-42 isolated from soybean rhizosphere. Bioresource technology, 2006. **97**(2): p. 204-210.
329. **Oteino, N., et al.**, Plant growth promotion induced by phosphate solubilizing endophytic *Pseudomonas* isolates. Frontiers in microbiology, 2015. **6**.
330. **Kang, S.-M., et al.**, Phosphate solubilizing *Bacillus megaterium* mj1212 regulates endogenous plant carbohydrates and amino acids contents to promote mustard plant growth. Indian journal of microbiology, 2014. **54**(4): p. 427-433.

331. **Selvakumar, G., et al.**, Growth promotion of wheat seedlings by *Exiguobacterium acetylicum* 1P (MTCC 8707) a cold tolerant bacterial strain from the Uttarakhand Himalayas. *Indian journal of microbiology*, 2010. **50**(1): p. 50-56.
332. **Teale, W.D., I.A. Paponov, and K. Palme**, Auxin in action: signalling, transport and the control of plant growth and development. *Nature Reviews Molecular Cell Biology*, 2006. **7**(11): p. 847-859.
333. **Magie, A.R., E. Wilson, and T. Kosuge**, Indoleacetamide as an intermediate in the synthesis of indoleacetic acid in *Pseudomonas savastanoi*. *Science*, 1963. **141**(3587): p. 1281-1282.
334. **Zambryski, P., J. Tempe, and J. Schell**, Transfer and function of T-DNA genes from *Agrobacterium* Ti and Ri plasmids in plants. *Cell*, 1989. **56**(2): p. 193-201.
335. **Tien, T., M. Gaskins, and D. Hubbell**, Plant growth substances produced by *Azospirillum brasilense* and their effect on the growth of pearl millet (*Pennisetum americanum* L.). *Applied and Environmental Microbiology*, 1979. **37**(5): p. 1016-1024.
336. **Fett, W.F., S.F. Osman, and M.F. Dunn**, Auxin production by plant-pathogenic pseudomonads and xanthomonads. *Applied and environmental microbiology*, 1987. **53**(8): p. 1839-1845.
337. **Ernstsen, A., et al.**, Endogenous indoles and the biosynthesis and metabolism of indole-3-acetic acid in cultures of *Rhizobium phaseoli*. *Planta*, 1987. **171**(3): p. 422-428.
338. **Sekine, M., et al.**, Detection of the IAA biosynthetic pathway from tryptophan via indole-3-acetamide in *Bradyrhizobium* spp. *Plant and cell physiology*, 1988. **29**(5): p. 867-874.
339. **Spaepen, S., J. Vanderleyden, and R. Remans**, Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS microbiology reviews*, 2007. **31**(4): p. 425-448.
340. **Costacurta, A. and J. Vanderleyden**, Synthesis of phytohormones by plant-associated bacteria. *Critical reviews in microbiology*, 1995. **21**(1): p. 1-18.
341. **Pollmann, S., D. Neu, and E.W. Weiler**, Molecular cloning and characterization of an amidase from *Arabidopsis thaliana* capable of converting indole-3-acetamide into the plant growth hormone, indole-3-acetic acid. *Phytochemistry*, 2003. **62**(3): p. 293-300.
342. **Costacurta, A., P. Mazzafera, and Y.B. Rosato**, Indole-3-acetic acid biosynthesis by *Xanthomonas axonopodis* pv. *citri* is increased in the presence of plant leaf extracts. *FEMS microbiology letters*, 1998. **159**(2): p. 215-220.
343. **Vandeputte, O., et al.**, Biosynthesis of auxin by the gram-positive phytopathogen *Rhodococcus fascians* is controlled by compounds specific to infected plant tissues. *Applied and environmental microbiology*, 2005. **71**(3): p. 1169-1177.
344. **Brandl, M.T. and S.E. Lindow**, Environmental signals modulate the expression of an indole-3-acetic acid biosynthetic gene in *Erwinia herbicola*. *Molecular plant-microbe interactions*, 1997. **10**(4): p. 499-505.
345. **Brandl, M., B. Quinones, and S. Lindow**, Heterogeneous transcription of an indoleacetic acid biosynthetic gene in *Erwinia herbicola* on plant surfaces. *Proceedings of the National Academy of Sciences*, 2001. **98**(6): p. 3454-3459.
346. **Hartmann, A. and W. Zimmer**, Physiology of *Azospirillum*. *Azospirillum/plant associations*, 1994: p. 15-39.

347. **Kulkarni, G.B., et al.**, Indole-3-acetic acid biosynthetic pathway and aromatic amino acid aminotransferase activities in *Pantoea dispersa* strain GPK. *Lett Appl Microbiol*, 2013. **56**(5): p. 340-7.
348. **Pereira, S. and P. Castro**, Diversity and characterization of culturable bacterial endophytes from *Zea mays* and their potential as plant growth-promoting agents in metal-degraded soils. *Environmental Science and Pollution Research*, 2014. **21**(24): p. 14110-14123.
349. **Bal, H.B., et al.**, ACC deaminase and IAA producing growth promoting bacteria from the rhizosphere soil of tropical rice plants. *Journal of Basic Microbiology*, 2013. **53**(12): p. 972-984.
350. **Sziderics, A.H., et al.**, Bacterial endophytes contribute to abiotic stress adaptation in pepper plants *Capsicum annuum*. *Canadian Journal of Microbiology*, 2007. **53**(11): p. 1195-1202.
351. **Goswami, D., et al.**, Delineating *Kocuria turfanensis* 2M4 as a credible PGPR: a novel IAA-producing bacteria isolated from saline desert. *Journal of Plant Interactions*, 2014. **9**(1): p. 566-576.
352. **Pugsley, A. and B. Oudega**, Methods for studying colicins and their plasmids. . 1987: p. 105-161.
353. **Michel-Briand, Y. and C. Baysse**, The pyocins of *Pseudomonas aeruginosa*. *Biochimie*, 2002. **84**(5): p. 499-510.
354. **Klaenhammer, T.R.**, Bacteriocins of lactic acid bacteria. *Biochimie*, 1988. **70**(3): p. 337-349.
355. **Grinter, R., J. Milner, and D. Walker**, Bacteriocins active against plant pathogenic bacteria. *Biochemical Society Transactions*, 2012. **40**(6): p. 1498.
356. **Boland, G. and R. Hall**, Index of plant hosts of *Sclerotinia sclerotiorum*. *Canadian Journal of Plant Pathology*, 1994. **16**(2): p. 93-108.
357. **Lim, H.-S., Y.-S. Kim, and S.-D. Kim**, *Pseudomonas stutzeri* YPL-1 genetic transformation and antifungal mechanism against *Fusarium solani*, an agent of plant root rot. *Applied and Environmental Microbiology*, 1991. **57**(2): p. 510-516.
358. **Wrather, J., W. Stienstra, and S. Koenning**, Soybean disease loss estimates for the United States from 1996 to 1998. *Canadian Journal of Plant Pathology*, 2001. **23**(2): p. 122-131.
359. **Goswami, R.S. and H.C. Kistler**, Heading for disaster: *Fusarium graminearum* on cereal crops. *Molecular plant pathology*, 2004. **5**(6): p. 515-525.
360. **Leonard, K.J. and W.R. Bushnell**, *Fusarium* head blight of wheat and barley. 2003: American Phytopathological Society (APS Press).
361. **Gevers, D., et al.**, Re-evaluating prokaryotic species. *Nature Reviews Microbiology*, 2005. **3**(9): p. 733-739.
362. **Bredow, C., et al.**, In silico analysis of the 16S rRNA gene of endophytic bacteria, isolated from the aerial parts and seeds of important agricultural crops. *Genetics and molecular research: GMR*, 2015. **14**(3): p. 9703.
363. **Nadarasah, G. and J. Stavrinides**, Quantitative evaluation of the host-colonizing capabilities of the enteric bacterium *Pantoea* using plant and insect hosts. *Microbiology*, 2014. **160**(Pt 3): p. 602-615.

364. **Stavrínides, J., A. No, and H. Ochman**, A single genetic locus in the phytopathogen *Pantoea stewartii* enables gut colonization and pathogenicity in an insect host. *Environmental microbiology*, 2010. **12**(1): p. 147-155.
365. **Wolf, A., et al.**, *Stenotrophomonas rhizophila* sp. nov., a novel plant-associated bacterium with antifungal properties. *International journal of systematic and evolutionary microbiology*, 2002. **52**(6): p. 1937-1944.
366. **Ryan, R.P., et al.**, The versatility and adaptation of bacteria from the genus *Stenotrophomonas*. *Nature Reviews Microbiology*, 2009. **7**(7): p. 514-525.
367. **McSpadden Gardener, B.B.**, Ecology of *Bacillus* and *Paenibacillus* spp. in agricultural systems. *Phytopathology*, 2004. **94**(11): p. 1252-1258.
368. **Lebuhn, M., T. Heulin, and A. Hartmann**, Production of auxin and other indolic and phenolic compounds by *Paenibacillus polymyxa* strains isolated from different proximity to plant roots. *FEMS Microbiology Ecology*, 1997. **22**(4): p. 325-334.
369. **Lal, S. and S. Tabacchioni**, Ecology and biotechnological potential of *Paenibacillus polymyxa*: a minireview. *Indian Journal of Microbiology*, 2009. **49**(1): p. 2-10.
370. **Ryu, C.-M., et al.**, Bacterial volatiles promote growth in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, 2003. **100**(8): p. 4927-4932.
371. **König, H.**, *Bacillus* species in the intestine of termites and other soil invertebrates. *Journal of Applied Microbiology*, 2006. **101**(3): p. 620-627.
372. **Harrison, H., R. Patel, and A.A. Yousten**, *Paenibacillus* associated with milky disease in Central and South American scarabs. *Journal of invertebrate pathology*, 2000. **76**(3): p. 169-175.
373. **Puri, A., K.P. Padda, and C.P. Chanway**, Can a diazotrophic endophyte originally isolated from lodgepole pine colonize an agricultural crop (corn) and promote its growth? *Soil Biology and Biochemistry*, 2015. **89**: p. 210-216.
374. **D'Costa, V.M., et al.**, Sampling the antibiotic resistome. *Science*, 2006. **311**(5759): p. 374-377.
375. **Ouyang, J., et al.**, *Paenibacillus thiaminolyticus*: a new cause of human infection, inducing bacteremia in a patient on hemodialysis. *Annals of Clinical & Laboratory Science*, 2008. **38**(4): p. 393-400.
376. **Duval, D., et al.**, A Novel Bacterial Pathogen of *Biomphalaria glabrata*: A Potential Weapon for Schistosomiasis Control? *PLoS neglected tropical diseases*, 2015. **9**(2): p. e0003489-e0003489.
377. **Gholami, M., Z. Etemadifar, and M. Bouzari**, Isolation a new strain of *Kocuria rosea* capable of tolerating extreme conditions. *Journal of environmental radioactivity*, 2015. **144**: p. 113-119.
378. **Romanovskaia, V., P. Rokitko, and I. Malashenko**, [Unique properties of highly radioresistant bacteria]. *Mikrobiolohichniy zhurnal (Kiev, Ukraine: 1993)*, 1999. **62**(1): p. 40-63.
379. **Wang, K., et al.**, *Kocuria dechangensis* sp. nov., a novel actinobacterium isolated from saline and alkaline soils. *International journal of systematic and evolutionary microbiology*, 2015.
380. **Jiang, Z., et al.**, *Kocuria subflava* sp. nov., isolated from marine sediment from the Indian Ocean. *Antonie van Leeuwenhoek*, 2015: p. 1-7.

381. **Dotis, J., et al.**, *Kocuria* species peritonitis: although rare, we have to care. *Peritoneal Dialysis International*, 2015. **35**(1): p. 26-30.
382. **Savini, V., et al.**, Drug sensitivity and clinical impact of members of the genus *Kocuria*. *Journal of medical microbiology*, 2010. **59**(12): p. 1395-1402.
383. **Haiyambo, D., P. Chimwamurombe, and B. Reinhold-Hurek**, Isolation and Screening of Rhizosphere Bacteria from Grasses in East Kavango Region of Namibia for Plant Growth Promoting Characteristics. *Current microbiology*, 2015. **71**(5): p. 566-571.
384. **Ormeño-Orrillo, E., et al.**, Taxonomy of Rhizobia and agrobacteria from the Rhizobiaceae family in light of genomics. *Systematic and applied microbiology*, 2015. **38**(4): p. 287-291.
385. **Peeters, K. and A. Willems**, The *gyrB* gene is a useful phylogenetic marker for exploring the diversity of *Flavobacterium* strains isolated from terrestrial and aquatic habitats in Antarctica. *FEMS microbiology letters*, 2011. **321**(2): p. 130-140.
386. **Nedashkovskaya, O.I., et al.**, *Flavobacterium ahnfeltiae* sp. nov., a new marine polysaccharide-degrading bacterium isolated from a Pacific red alga. *Archives of microbiology*, 2014. **196**(10): p. 745-752.
387. **Michaud, L., et al.**, Predominance of *Flavobacterium*, *Pseudomonas*, and *Polaromonas* within the prokaryotic community of freshwater shallow lakes in the northern Victoria Land, East Antarctica. *FEMS microbiology ecology*, 2012. **82**(2): p. 391-404.
388. **Sang, M. and K. Kim**, The volatile-producing *Flavobacterium johnsoniae* strain GSE09 shows biocontrol activity against *Phytophthora capsici* in pepper. *Journal of applied microbiology*, 2012. **113**(2): p. 383-398.
389. **Declercq, A.M., et al.**, Columnaris disease in fish: a review with emphasis on bacterium-host interactions. *Vet Res*, 2013. **44**(1): p. 27-44.
390. **Kolton, M., et al.**, Comparative genomic analysis indicates that niche adaptation of terrestrial *Flavobacteria* is strongly linked to plant glycan metabolism. 2013.
391. **Aylward, F.O., et al.**, Comparison of 26 sphingomonad genomes reveals diverse environmental adaptations and biodegradative capabilities. *Applied and environmental microbiology*, 2013. **79**(12): p. 3724-3733.
392. **Liu, Z.-P., et al.**, *Novosphingobium taihuense* sp. nov., a novel aromatic-compound-degrading bacterium isolated from Taihu Lake, China. *International journal of systematic and evolutionary microbiology*, 2005. **55**(3): p. 1229-1232.
393. **Rogers, E.E. and E.A. Backus**, Anterior foregut microbiota of the glassy-winged sharpshooter explored using deep 16S rRNA gene sequencing from individual insects. 2014.
394. **Wetzel, J., C. Kingsford, and M. Pop**, Assessing the benefits of using mate-pairs to resolve repeats in de novo short-read prokaryotic assemblies. *BMC bioinformatics*, 2011. **12**(1): p. 95.
395. **Crook, M.B., et al.**, Complete genome sequence of the *Sesbania* symbiont and rice growth-promoting endophyte *Rhizobium* sp. strain IRBG74. *Genome announcements*, 2013. **1**(6): p. e00934-13.
396. **Biswas, J., J. Ladha, and F. Dazzo**, Rhizobia inoculation improves nutrient uptake and growth of lowland rice. 2000.

397. **Cummings, S.P., et al.**, Nodulation of *Sesbania* species by *Rhizobium* (*Agrobacterium*) strain IRBG74 and other rhizobia. *Environmental microbiology*, 2009. **11**(10): p. 2510-2525.
398. **Biswas, J.C., et al.**, Rhizobial inoculation influences seedling vigor and yield of rice. *Agronomy Journal*, 2000. **92**(5): p. 880-886.