
**Universidad Nacional Autónoma de México**

**Doctorado en Ciencias Biomédicas**

**Centro de Ciencias Genómicas**

# Inferencias taxonómicas y funcionales a partir del estudio del genoma de *Rhizobium phaseoli* Ch24-10 y de su transcriptoma en el rizoplano de maíz y de frijol

**Tesis que para optar por el grado de Doctora en Ciencias**

**Presenta:**

**Biol. Martha Guadalupe López Guerrero**

**Directora de tesis**

**Dra. Esperanza Martínez Romero**

**Centro de Ciencias Genómicas, UNAM**

**Comité Tutor**

**Dr. Juan Miranda Ríos**

**Instituto de Investigaciones Biomédicas, UNAM**

**Dr. Víctor González**

**Centro de Ciencias Genómicas, UNAM**

**Cuernavaca, Morelos, marzo de 2013**

“*Knowing is not enough; we must apply.*

*Willing is not enough; we must do*”

"Saber no es suficiente,  debemos aplicar

Estar dispuesto no es suficiente, hay que actuar"

Johann Wolfgang von Goethe

## Agradecimientos.

A la Universidad Nacional Autónoma de México (UNAM) que me ha privilegiado con educación en todos sentidos desde mi ingreso a la Escuela Nacional Preparatoria.

A la Dra. Esperanza Martínez, por su dirección, tutoría y apoyo. Por estimular mi interés por la ciencia y motivarme a seguir mi vocación.

Al comité tutoral inicial, parte importante del desarrollo de este proyecto, Dr. Juan Miranda Ríos y Dr. Miguel Ángel Ramírez.

A los sinodales que evaluaron este trabajo por su tiempo y comentarios, mismos que permitieron, sin duda, mejorar este trabajo: Dr. Enrique Morett, Dr. Lorenzo Segovia, Dr. Jesús Silva y Dra. Carmen Wacher.

A todas y cada una de las personas que laboran en el Centro de Ciencias Genómicas (CCG), cuyo trabajo es fundamental para que sean exitosas todas las actividades que se llevan a cabo en él. En particular a varias personas que durante todo el tiempo de mi estancia en el CCG me han brindado apoyo y amistad, a los técnicos José Espíritu, Jorge Muñoz, Martín García, Araceli Sánchez, y Antonia Jaimes y a las asistentes Gladys Avilés y Lucy Lulo. De manera especial, agradezco a la dirección del CCG a cargo del Dr. David Romero, por su apoyo invaluable al final de este proceso. A la Unidad de Administración de Tecnologías de la Información, UATI, particularmente al Ing. Víctor del Moral, M. en C. Romualdo Zayas y L.I. Waldo Díaz.

De manera especial agradezco al Dr. Ernesto Ormeño, Dra. Karla Martínez, Dr. José L. Acosta, M. en C. Marco A. Rogel, Dra. Mónica Rosenblueth, Biól. Rafael Díaz y M. en C. Alfredo Mendoza, su apoyo fue fundamental en la realización de este proyecto.

Al CONACYT institución que me otorgó una beca de manutención durante mis estudios de posgrado.

Al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la UNAM por el apoyo IN205412 otorgado a la Dra. Esperanza Martínez Romero.

## Dedicatoria.

Es difícil escribir y describir en unas líneas todo el cariño y agradecimiento que tengo hacia las personas a quienes dedico este trabajo, me faltaran palabras para decirles lo importante que son y que han sido en mi vida. Dedico este trabajo a la familia, la que me conoce desde siempre y la que he tenido la fortuna de conocer con los años, grandes amigos, la familia en conjunto que me ha dado su apoyo y cariño en todo momento. Es una fortuna ser parte de una familia unida, solidaria y en la que todos y cada uno de sus miembros son un ejemplo de superación y fortaleza. Gracias por permitirme aprender de todos y cada uno de ustedes. Familia, amigos, los amo.

## Resumen.

Se denominan rizobios a un conjunto de α-proteobacterias y β-proteobacterias capaces de establecer una simbiosis mutualista con plantas leguminosas que culmina en la fijación biológica de nitrógeno. La cepa Ch24-10 es un α-proteobacterias capaz de formar nódulos fijadores de nitrógeno con el frijol (*Phaseolus vulgaris*), es un endófito natural del maíz (*Zea mays*) y promueve el crecimiento vegetal de éste.

Para entender cómo se lleva a cabo la interacción de *Rhizobium* con el maíz y con el frijol se obtuvo la secuencia de su genoma y su perfil de expresión en el rizoplano de ambas plantas.

Mediante genómica comparativa fue posible conocer algunos atributos genómicos de la cepa Ch24-10, mismos que permitieron proponer la reclasificación taxonómica de la cepa originalmente clasificada como *Rhizobium etli* a *Rhizobium phaseoli*.

Utilizando la metodología *RNA-seq* se obtuvo la secuencia de su transcriptoma en asociación con frijol y maíz. La mayoría de los genes altamente expresados son de función desconocida. Se encontró un conjunto de genes que se expresan tanto en el rizoplano del frijol como en el de maíz, lo que sugiere que existe un grupo común necesarios para la colonización de la rizósfera.

Ya que en *Rhizobium* los genes necesarios para establecer simbiosis con plantas leguminosas residen en los plásmidos simbióticos, se supuso que los genes importantes en la colonización de las raíces pudieran estar codificados en estos replicones (crómidos o plásmidos) por lo que se puso especial atención al análisis de estos genes. En este grupo se encuentran genes que codifican para transportadores ABC de péptidos y glucósidos; responsables de la síntesis de aminoácidos: glutamina, glutamato, aspartato, asparagina; degradación de arginina y ornitina; descomposición de urea y degradación de lisina y de prolina.

## Abstract.

Several α-proteobacteria and β-proteobacteria called rizobia have the capacity of establishing symbiosis with leguminous plants to fix atmospheric nitrogen. Strain Ch24-10 it is an α-proteobacteria that belongs to this group, capable to form nitrogen-fixing nodules with common bean (*Phaseolus vulgaris*) and it is also a natural endophyte of maize (*Zea mays*) promoting their vegetal growth.

It is important to understand how the interaction with maize and bean takes place. To elucidate the genetic elements in strain Ch24-10 that allow it to promote vegetal growth we got their genome and transcriptome sequences in rhizoplane of maize and beans.

Through comparative genomics we understand some genomic attributes of strain Ch24-10 besides we get knowledge about its evolutionary relationships. This allows us to propose the taxonomical reclassification of the strain formerly classified as *Rhizobium etli* as *Rhizobium phaseoli*.

Using *RNA-seq* we obtained the transcriptome of *R. phaseoli* Ch24-10 in association with bean and corn allowing us to have an overview of the expression profile in the rhizoplane of both plants.

We found common genes expressed in the rhizoplane of beans and maize, suggesting that there is a common set of genes important in the rhizosphere colonization. In *Rhizobium,* most of the essential genes to establish symbiosis with legumes reside in symbiotic plasmids. Similarly, we assumed that the genes important in the rhizosphere colonization could also be encoded in these replicons. Therefore, description of the highly expressed genes is focused on extrachromosomal elements. Among the highly expressed genes are those with unknown function and several encoding for specific and ABC transporters suggesting that some of them may participate in the catabolism of root exudates.

# Índice.

## Índice de Tablas.



## Índice de Figuras.

# 1. Introducción.

## 1.1 La rizósfera y el rizoplano.

La rizósfera se define como el volumen de suelo alrededor de la raíz de las plantas, un cilindro milimétrico que es influenciado por la actividad biológica de las raíces [1]. Está compuesta por tres elementos, el suelo, la raíz y el rizoplano (superficie de la raíz y las partículas de suelo que están adheridas a ésta) [2] y es colonizada por bacterias, hongos, protistas y nemátodos [3].

Las raíces secretan compuestos a la rizósfera (conocidos en general como *exudados*), mediante un proceso denominado rizodeposición [4]. El carbono transferido a la rizósfera puede oscilar entre el 5-20% del total generado por fotosíntesis [5,6]. En condiciones de estrés nutricional, la planta puede secretar más del 40% [4]. Los exudados contienen un amplio rango de compuestos orgánicos: azúcares, polisacáridos, aminoácidos, ácidos orgánicos, péptidos, proteínas (incluyendo enzimas), lisados provenientes de las células muertas de las plantas, mucílago, compuestos fenólicos y metabolitos secundarios [6,7].

La rizodeposición tiene diversas funciones en el suelo, mantenimiento de la zona de contacto suelo-raíz, lubricación de las puntas de los pelos radiculares, protección contra la desecación (de la raíz y de los organismos asociados), estabilización de los microagregados del suelo y absorción, así como almacenamiento selectivo de iones [7].

La interacción entre microorganismos con la materia orgánica y con las plantas vivas genera un recambio de la materia orgánica y reciclamiento de nutrientes, procesos agrícola y ecológicamente importantes [2]. Los exudados contienen moléculas de señalización que dan inicio a interacciones entre las raíces y bacterias como la fijación biológica de nitrógeno [6]. Además, la rizodeposición regula las comunidades bacterianas en la rizósfera [4].

## 1.2 Rizobacterias Promotoras del Crecimiento Vegetal.

Entre las bacterias que se asocian a la rizósfera de las plantas existe un grupo que es capaz de promover su crecimiento y se han denominado rizobacterias promotoras del crecimiento vegetal, *PGPR* (*Plant Growth Promoting Rhizobacteria*) [8]. Las *PGPR* resultan de interés agronómico ya que permiten el aumento en el rendimiento de los cultivos y la disminución del uso de agroquímicos, que tiene como resultado la disminución del impacto ambiental y económico que generan [9,10]. El incremento en el rendimiento del cultivo es medido en parámetros agronómicos, como el aumento en el porcentaje de germinación, altura de la planta, área foliar, contenido de clorofila, biomasa, longitud de la raíz, contenido de nitrógeno, así como la cantidad de grano producido por hectárea [8,11].

Las *PGPR* promueven el crecimiento vegetal mediante distintos mecanismos: producción de fitohormonas y vitaminas; aumento de la toma de nutrientes como fósforo y potasio, ya sea porque inducen el aumento del tamaño y número de pelos radiculares o porque son capaces de transformarlos a una forma asimilable para las plantas; supresión de patógenos e inducción de la resistencia sistémica; producción de sideróforos, involucrados en captar hierro y por la fijación biológica de nitrógeno asociada a las raíces [8,12,13].

Algunas *PGPR* son capaces de llegar al interior de las raíces y establecerse como organismos *endófitos*[a] sin causar daño a la planta. Algunas de ellas pueden pasar la barrera endodérmica y atravesar el *córtex* radicular hacia el sistema vascular y colonizar tallo, hojas y frutos. La capacidad de las bacterias de colonizar diferentes tejidos refleja su habilidad para ocupar diferentes nichos lo que les permite asociarse de distintas maneras con su hospedero [9].

En general, las bacterias rizosféricas y las endofíticas utilizan mecanismos similares para promover el crecimiento de las plantas y para suprimir patógenos: inducción de la resistencia sistémica, competencia en la colonización y/o la producción de sustancias antimicrobianas [9,12,14]. Gracias a su uso potencial como control biológico son una alternativa al uso de pesticidas [15].

[a] Endófito: Microorganismo capaces de colonizar el interior de los tejidos de las plantas sin causarles daño aparente [14].

La capacidad competitiva de las *PGPR* debe tomarse en consideración si se desea utilizarlas como inoculantes en campo. Las bacterias deben sobrevivir y proliferar en la rizósfera, utilizar los nutrientes exudados por la planta, ser capaces de colonizar el sistema radicular y competir efectivamente con los microorganismos nativos [16].

## 1.3 Los rizobios.

Los rizobios son bacterias conocidas por su capacidad de asociarse con plantas leguminosas en una simbiosis mutualista[b] en la que se induce la formación de nódulos en las raíces, estructuras especializadas en las que se lleva a cabo la fijación biológica del nitrógeno atmosférico ($N_2$) [17]. Son organismos ubicuos en el suelo, donde sobreviven y se reproducen [18]. En este ambiente tienen la capacidad de moverse, a diferencia de lo que ocurre en los nódulos. Hasta el momento no existe evidencia de que tengan la capacidad de fijar nitrógeno atmosférico en esas condiciones [19].

Taxonómicamente, los rizobios se agrupan en los géneros *Rhizobium, Sinorhizobium, Bradyrhizobium, Azorhizobium* y *Mesorhizobium* [20] de las α-proteobacterias y en los géneros *Burkholderia, Cupriavidus (Ralstonia)* [19] y *Herbaspirillum* [21] de las β-proteobacterias.

Este grupo de bacterias puede colonizar la rizósfera de plantas leguminosas y no leguminosas, así como el interior de los tejidos vegetales como endófitos [14]. *Rhizobium etli* puede colonizar las raíces de maíz, alcanzando un densidad celular de $10^9$ células por gramo de raíz (peso fresco) [22]. *Rhizobium leguminosarum* coloniza la rizósfera del arroz, la canola y la lechuga, promoviendo el crecimiento vegetal [23,24]. *Rhizobium endophyticum* puede colonizar semillas de inmaduras de frijol como endófito [25].

[b] Simbiosis: Asociación entre dos o más organismos de distinta especie en algún momento de su ciclo de vida. Se definen varios tipos, dependiendo del aumento en la adecuación, beneficio /costo que tiene cada integrante en la interacción. Mutualismo: ambos aumentan su adecuación y se benefician. Comensalismo: un integrante se beneficia aumentando su adecuación sin afectar al otro. Parasitismo: un integrante aumenta su adecuación mientras que en el otro disminuye [116].

La capacidad de los rizobios de asociarse simbióticamente con las leguminosas está determinada por plásmidos o islas simbióticas en las que se encuentran los genes necesarios para establecer esta interacción. Los plásmidos son replicones extracromosomales que pueden perderse y transferirse entre bacterias. Se encuentran plásmidos simbióticos en *Rhizobium, Sinorhizobium,* en algunas cepas de *Mesorhizobium*, en *Cupriavidus taiwanensis* y en *Burkholderia* CCGE1002; mientras que en *Bradyrhizobium, Azorhizobium caulinodans* y la mayoría de las cepas *de Mesorhizobium* la información necesaria para realizar la fijación biológica de nitrógeno está codificada en islas simbióticas en el cromosoma [26].

Estos elementos, plásmidos e islas simbióticas, determinan tanto la especificidad por el hospedero como la efectividad de la simbiosis (tasa de fijación de nitrógeno). Dependiendo de esta especificidad, se han reconocido grupos simbióticos denominados *simbiovares*. Por ejemplo, las cepas de *R. leguminosarum* se agrupan en tres: *phaseoli*, *trifolii* y *viciae*, por su capacidad de nodular frijol; chícharo o haba, respectivamente. Las cepas de *R. leguminosarum* de diferentes *simbiovares* tienen el mismo fondo cromosomal, pero diferente plásmido simbiótico [27].

Por otro lado, diferentes especies pueden pertenecer a un mismo *simbiovar*; por ejemplo, *Rhizobium gallicum, R. giardinii, R. leguminosarum* [20] *R. vallis* [28], *R. lusitanum* [29] y *Rhizobium etli* [20] tienen la capacidad de nodular al frijol y pertenecen al *simbiovar phaseoli* [27].

Además de los plásmidos simbióticos, los rizobios pueden tener otros replicones extracromosomales: plásmidos no simbióticos de diferentes tamaños y estos pueden constituir aproximadamente entre el 30-45% del total del contenido genómico [30].

Un grupo particular de estos plásmidos, normalmente de mayor tamaño que el resto de los plásmidos presentes en la célula (aproximadamente de 1Mb), han recibido el nombre de *crómido.* Estos replicones son estables en la célula y tienen una composición nucleotídica similar a la del cromosoma, contienen genes conservados *esenciales* [31].

A algunos de estos plásmidos se les ha dado el nombre de *cromosoma secundario* [32]. Sin embargo, ya que su sistema de mantenimiento y de replicación es tipo plasmídico se propone que el termino *crómido* es más adecuado [31].

Desde el punto de vista taxonómico, las cepas pertenecientes al género *Rhizobium* han sufrido cambios a lo largo del tiempo. Se ha discutido el tipo de caracteres que son utilizados para su clasificación, en particular, la relevancia que se le ha dado a caracteres importantes en la simbiosis con plantas pues éstos no permiten inferir la evolución de un grupo, sino sólo de un caracter adaptativo [33].

Un ejemplo de clasificación taxonómica artificial como resultado del uso de caracteres simbióticos, es la asignación de las todas las cepas capaces de nodular frijol a la especie *Rhizobium phaseoli.* Este mismo criterio se utilizaba para clasificar las cepas capaces de nodular chícharo y trébol en las especies *R. leguminosarum* y *R. trifolii*, respectivamente. En 1984, estas tres especies fueron reclasificadas como *R. leguminosarum,* y divididas en tres *biovares* dependiendo de su planta hospedera: chícharo, trébol y frijol como *viciae*, *trifolii* y *phaseoli*, respectivamente [34].

Cuando fueron reclasificadas las cepas de *R. leguminosarun* en tres *biovares* [34] no fueron rechazadas las especies *R. trifolii* y *R. phaseoli*. La cepa tipo de *R. phaseoli* ATCC 14482 fue recientemente caracterizada y se encontró que la secuencia del gene *rRNA16s* presentaba una identidad menor al 97% (punto de corte) respecto a otras secuencias reportadas. Los valores de *reasociación DNA-DNA* con las especies cercanas, *R. leguminosarum* y *R. etli* estuvieron por debajo del 70%. Con estas evidencias se reconsideró a *R. phaseoli* como una especie válida [35].

Otros ejemplos de cambios en la clasificación se pueden ver en *Rhizobium leguminosarum,* las cepas pertenecientes al grupo II se renombraron como *R. tropici* [36] y las cepas del tipo I fueron designadas como *Rhizobium etli*, un grupo con una diversidad genética alta y en la que no se encontró variación en las secuencias del gen ribosomal [37]. Recientemente, las cepas del grupo A de *R. tropici* han sido clasificadas en una nueva especie, *R. leucaenae* [38].

## 1.4 Frijol y maíz. Origen e importancia.

El frijol (*Phaseolus vulgaris* L.) es una leguminosa, taxonómicamente ubicada en la Familia *Fabaceae* de la Clase Magnoliopsida (dicotiledóneas) [39]. Junto con la yuca y el maíz ocupa la mayor superficie sembrada en América y es la leguminosa más consumida en el mundo, como grano maduro, semilla inmadura y como vegetal (hojas y tallos). Constituye la fuente más importante de proteína en la dieta de los países latinoamericanos y del este de África [40].

Se ha presentado evidencia reciente que indica que México es el único centro de origen del frijol [41] Se propone que la domesticación del frijol ocurrió hace aproximadamente 4-10 mil años y que fue un proceso múltiple e independiente en los dos centros de origen mencionados (Mesoamérica y los Andes) [39]. Mediante estudios morfológicos, de tipificación genética utilizando *RAPD (Random Amplification of Polymorphic DNA)*, *RFLP (Restriction Fragment Lengh Polymorphism),* isoenzimas (faseolina en particular); se ha propuesto que existen dos centros de diversificación del frijol: Mesoamérica (México, Centroamérica y Colombia) y la región de los Andes en Sudamérica (Ecuador, Perú y Argentina) [42,43]. Recientemente, mediante el análisis de *microsatélites*[c] y de la estructura genética del frijol, se ha planteado que Brasil es un centro secundario de diversificación [44].

El maíz (*Zea mays spp. mays*) se localiza taxonómicamente en la familia *Poaceae* (=*Gramineae*) de la Clase *Liliopsida* (monocotiledoneas) [45]. En conjunto con el trigo y el arroz son los cereales con mayor superficie sembrada en el mundo. El maíz es principal cultivo en América Latina [46] y es utilizado como grano, forraje y como materia prima básica de la industria de transformación (producción de almidón, aceite, proteínas, bebidas alcohólicas, edulcorantes alimenticios y combustible) [47].

[c] *Microsatélite*: también conocidos como SSR (*Simple Sequence Repeat*) o STR (*Short Tandem Repeat*) por sus siglas en inglés, son secuencias de *DNA* de 2-6 pb repetidas de manera consecutiva, generalmente en regiones no codificantes del genoma [117].

Una de las evidencias fósiles más antiguas del uso del maíz como alimento por el hombre se encontró en México, en el valle de Tehuacán, Puebla, y tiene una antigüedad de más de 6 mil años [48]. Otros trabajos proponen que el maíz comenzó a domesticarse en México hace aproximadamente 9 mil años [49].

Existen prácticas agrícolas en las que se siembra un cereal y una leguminosa, ya sea en intercultivo o en rotación, mismas que permiten el mejoramiento del suelo. La simbiosis leguminosa-rizobio aporta nitrógeno, un elemento limitante y fundamental para que el cereal tenga un buen desarrollo y rendimiento [10]. Con estos sistemas se minimizan los riesgos agronómicos pues se aprovecha mejor la superficie cultivable, hay retención de la humedad y se reduce la dispersión de patógenos [39].

Un ejemplo de lo anterior lo encontramos en Egipto, donde se siembra el trébol alejandrino en rotación con arroz. Es interesante que Yanni y colaboradores [23] encontraron que *R. leguminosarum,* la bacteria que forma nódulos con el trébol, es capaz de colonizar endofíticamente las raíces y tallos del arroz.

En México se ha sembrado maíz y frijol en asociación desde tiempos prehispánicos, en un sistema denominado *milpa*, comúnmente acompañados por calabaza, chile, otras hortalizas y árboles frutales [50]. Se ha propuesto, dada esta asociación entre frijol y maíz, que ambos cultivos pudieron haberse co-domesticado [20].

Es importante hacer notar que al igual que en el ejemplo anterior, en este sistema agrícola se ha encontrado a cepas clasificadas como *R. etli*, bacteria que forma nódulos con el frijol, colonizando la rizósfera del maíz y el interior de raíces y tallos como endófito [22]. Además, algunas cepas aisladas de maíz fueron capaces de promover su crecimiento [22].

# 2. Antecedentes.

## 2.1 La cepa de *Rhizobium* Ch24-10.

Como se mencionó anteriormente diferentes especies del género *Rhizobium* son capaces de formar nódulos fijadores de nitrógeno en el frijol y se han descrito cepas clasificadas como *R. etli* mediante *electroforesis de enzimas multilocus*, *MLEE* (*MultiLocus Enzyme Electrophoresis*), organización del gen *nifH* y patrones de digestión del gen *rRNA16s*[d], capaces de colonizar la rizósfera del maíz y el interior de raíces y tallos. En detalle, cuando se buscó a *R. etli en* plantas de maíz cultivadas en un sistema agrícola en *milpa* en México se encontró a esta bacteria en todas ellas, mientras que sólo estuvo presente en el 20% de las plantas analizadas en monocultivo. Se propone que este sistema agrícola en intercultivo con el frijol favorece que el maíz sea colonizado por *Rhizobium* [22].

Posteriormente, para conocer la diversidad de *Rhizobium* en la *milpa* en comparación con monocultivos de maíz y frijol, se aislaron bacterias de nódulos de frijol, de rizósfera, raíces y tallos de maíz. En general, los perfiles plasmídicos de las cepas aisladas fueron muy diversos, sin embargo, en el interior de tallos de maíz, en uno de los dos sitios de muestreo (cultivo en milpa en Cholula, Puebla, México), se encontró un único perfil de plásmidos aparentemente idéntico que también fue el más abundante en raíz y en rizósfera de maíz. Se seleccionó a la cepa Ch24-10 como un representante de este grupo y se asignó a la especie *R. etli* [51].

La cepa Ch24-10 fue competitiva en ensayos de colonización de rizósfera de maíz respecto a otros aislados de nódulo y rizósfera de esta planta, así como respecto a la cepa CFN 42 [51]. Además, es capaz de colonizar la rizósfera del arroz y en competencia con la cepa CFN 42 ocupa la mayoría de los nódulos en frijol. La tasa de FBN es mayor respecto a la cepa CFN 42 (Esperanza Martínez-Romero, datos no publicados).

[d] *rRNA16s*: subunidad 16s del *RNA* ribosomal.

Se observaron diferencias fenotípicas marcadas entre las cepas Ch24-10 y CFN 42 en la interacción con frijol y maíz y como una aproximación para entender los determinantes génicos de ésta se propuso conocer el genoma presente en la cepa Ch24-10 y ausente en la cepa CFN 42 mediante la construcción de una biblioteca con la metodología denominada *hibridación sustractiva*, *SSH* (*Suppression Subtrative Hybridation*) [52]. Entre las secuencias obtenidas que no estuvieron presentes en el genoma de la cepa CFN 42 [53] se encontraron algunas con identidad a proteínas anotadas con función putativa o conocida como reguladores transcripcionales de las familias *AraC* y *LysR*, transportadores ABC y proteínas con motivos *GGDEF/EAL* [52], mismos que por su función putativa podrían tener un papel en el fenotipo de la cepa Ch24-10.

Un conjunto de las secuencias de la cepa Ch24-10 generadas por hibridación *sustractiva* que sí presentaron identidad al genoma de la cepa CFN 42 se utilizaron para calcular los índices *ANI (Average Nucleotide Identity) y AAI (Average Amino acid Identity)* (ver más adelante) respecto a las secuencias ortólogas de las cepas CFN 42 y CIAT 62. Los valores de *ANI* fueron de 93% respecto a la cepa CFN 42 y de 99% respecto a la cepa CIAT 652, mientras que los de *AAI* fueron de 92% y 99% respecto a la cepa CFN 42 y CIAT 652, respectivamente. Estos datos mostraron, de manera preliminar que la cercanía filogenética entre las cepas Ch24-10 y CIAT 652 es mayor que entre las cepas Ch24-10 y CFN 42, pues los valores de *ANI* y *AAI* están por debajo del corte propuesto para que una cepa pertenezca a la misma especie (96%) [54].

Posteriormente, Richter & Rosselló-Mora en el año 2009 [55] obtuvieron el *ANI* entre los genomas secuenciados de las cepas CFN42 [53] y CIAT 652 [56]. El valor que se obtuvo fue menor al 96%, por lo que se propuso revisar la clasificación taxonómica de la cepa CIAT 652. Posteriormente, Robledo y colaboradores [57] basándose en los resultados de la comparación de las secuencias parciales de los genes *recA*, *atpD* y *celC* determinaron que la cepa CIAT 652 pertenece a la especie *R. phaseoli* y no a *Rhizobium etli*.

Tomando en cuenta estos resultados, así como los valores de *ANI* y *AAI* entre las cepas Ch24-10 y CFN 42 y CIAT 652, la clasificación taxonómica de la cepa Ch24-10 debe ser reconsiderada también.

## 2.2 Interacción de la cepa Ch24-10 con el maíz.

En general se sabe poco de la interacción entre rizobios y plantas no leguminosas,aunque existen algunos trabajos que han abordado este tema en *Rhizobium leguminosarum* [58–60]. Resulta de interés conocer los genes que intervieren en la interacción de la cepa Ch24-10 con el maíz, tanto por el valor de la interacción rizobio no leguminosa, como por la capacidad de promoción de crecimiento que le confiere potencial agronómico.

Anteriormente se mencionó que es posible que el frijol y el maíz se co-domesticaran en el sistema agrícola de milpa [20], por lo que también resulta de interés saber si existen genes comunes involucrados en la colonización de ámbas plantas.

En el grupo de trabajo dónde se desarrolló este proyecto se ha utilizado una estrategia de análisis de mutantes construidas al azar con un transposón unido a un gen reportero (Tn5*gusA*) para conocer genes que se expresan en presencia de exudados de maíz de la cepa Ch24-10. Después de analizar más de mil colonias bacterianas, se encontró la expresión de un gen que codifica para una ATPasa que se encuentra junto a un gen que codifica para una poligalacturonasa, misma que al mutarse no presentó un fenotipo en la colonización de maíz diferente a la cepa silvestre [61].

Utilizando esta misma estrategia de mutagénesis, se seleccionaron cepas incapaces de crecer utilizando prolina como fuente de carbono y nitrógeno. El uso de la prolina resulta de interés pues es uno de los aminoácidos más abundantes en los exudados de maíz [62] y tiene un papel importante en el establecimiento de la simbiósis con frijol [63]. Una de las mutantes analizadas corresponde a una una permeasa putativa que al mutarse de manera dirigida no presentó fenotipo comparada con la cepa silvestre en colonización del maíz ni del frijol [64].

Posteriormente, se construyó una mutante del gen *putA* por inserción de un *cassette* con un gen que confiere resitencia a kanamicina y usando como gen reportero el gen *lacZ*. El gen *putA* codifica para una prolina deshidrogenasa. Se encontró la inducción de este gen en presencia de prolina y de exudados de maíz y de frijol. La mutante no es capaz de crecer en medio mínimo cuando se utiliza este aminoácido como única fuente de carbono y no presentó diferencias en el fenotipo de colonización de la rizósfera del maíz respecto a la cepa silvestre, Ch24-10, sin embargo, sí estuvo afectada en la competencia en nodulación respecto a la cepa de *R. etli* CFN42 [64].

Se han desarrollado metodologías de secuenciación masiva de *DNA* adaptadas para conocer el transcriptoma de un organismo en una condición dada, se ha denominado *RNA-seq*. Esta metodología tiene varias ventajas como la detección de transcritos cuya expresión es baja y del *RNA* no codificante y en comparación con los microarreglos, no está limitada a aquellos genes blanco conocidos previamente [65]. Se decidió utilizar esta estrategia para conocer los genes que expresa la cepa Ch24-10 en el rizoplano tanto del maíz como del frijol.

# 3. Objetivos.

**Objetivos generales.**

a) Conocer las características genómicas de la cepa Ch24-10.
b) Revisar la clasificación taxonómica de la cepa Ch24-10.
c) Determinar los genes de la cepa Ch24-10 que participan en la interacción con el rizoplano de maíz y de frijol.

**Objetivos particulares.**

a) Obtener la secuencia genómica de la cepa Ch24-10.
b) Realizar experimentos de genómica comparativa de la cepa Ch24-10 utilizando como referencia a otros rizobios (i.e. *R. etli* CFN 42, *R. etli* CIAT 652, *R. etli* CFPAF512).
c) Determinar el perfil de expresión génica de la cepa Ch24-10 en el rizoplano del maíz.
d) Obtener el transcriptoma de la cepa Ch24-10 en la rizósfera del frijol
e) Conocer los genes comunes expresados en la rizósfera de maíz y de frijol.

# 4. Materiales y Método.

## 4.1 Cepas y semillas.

La lista y descripción de las cepas y semillas que se utilizaron en este trabajo se encuentran en la Tabla 1.

**Tabla 1. Cepas y semillas utilizadas.**

| Cepas | Hospedero / Origen / Características | Referencia |
|---|---|---|
| *R. etli* CFN42 | *P. vulgaris* / Guanajuato, México / Cepa tipo de *Rhizobium etli* | 66 |
| *R. phaseoli* Ch24-10 | *Zea mays* / Puebla, México / Aislado de tallo de maíz | 51 |
| *R. phaseoli* CIAT 652 | *Phaseolus vulgaris* / Buitrera, Colombia | 67 |
| *R. phaseoli* CNPAF512 | *P. vulgaris* / Brasil | 68,69 |
| *R. phaseoli* ATCC 14482 | *P. vulgaris* / Betlsville, Maryland | 70 |
| **Semillas** | **Características** | |
| Maíz, *Zea mays* *Criollo* Amatlán | Variedad criolla, proveniente de Amatlán, Morelos. Al ser inoculada con la cepa Ch24-10 presenta un fenotipo en promoción de crecimiento superior a otras variedades analizadas | 51 |
| Frijol, (*P. vulgaris* BAT 477 | Variedad mejorada, adaptada a los trópicos. Resistente a sequía, acidez y adaptado a la baja disponibilidad de fósforo disponible | 71 |

## 4.2 Obtención de DNA.

Las cepas de *Rhizobium* fueron cultivadas en medio líquido PY (peptona de caseína, 5 g; extracto de levadura, 3 g; $CaCl_2 \cdot 2H_2O$, 1 g/L), se mantuvieron en agitación a 150 *RPM* a una temperatura de 28°C por 12 horas. El *DNA* se extrajo con el *DNA Isolation Kit for Cells and Tissues* (Roche, USA). Se cuantificó utilizando el equipo *NanoDrop2000* (*Thermo Scientific*) y se verificó su integridad por electroforesis en un gel de agarosa 1% en Tris Acetato (TA) (Trizma base 48.4 g; 11. 42 ml, 0.5 M EDTA, pH8 / L), a 90 volts por 60 minutos. El *DNA* se marcó con bromuro de etidio (0.5 µg/ml) y se expuso a luz ultravioleta en el equipo *Eagle Eye II* (*Stratagene*).

## 4.3 Secuenciación y ensamble del genoma de la cepa Ch24-10.

La secuencia del genoma de la cepa Ch24-10 se obtuvo utilizando dos plataformas. En primer lugar, con la tecnología *Roche-454 Life Science* con el protocolo *Long-Tag Paired End* utilizando un servicio comercial. Se obtuvieron dos bibliotecas, protocolos para 3 y 8 Kb, de una longitud de 350 pares de bases y se secuenciaron 150 nucleótidos de cada extremo.

En segundo lugar, se utilizó el sistema *Genome Analyzer GAIIx* de *Illumina* con el protocolo *Pair-end sequencing assay protocol* en la Unidad Universitaria de Secuenciación Masiva de *DNA* (UUSMD), de la Universidad Nacional Autónoma de México (UNAM). Se construyó una biblioteca de 200 pares de bases y se secuenciaron 36 nucleótidos de cada extremo.

Las lecturas que se obtuvieron con la tecnología *Roche-454 Life Science* fueron ensambladas *de novo* con el programa *GSassembler* (*Newbler*) versión 2.53 (*Roche Diagnostics Corporation*), utilizando los parámetros por defecto.

Las lecturas *pair-end* que se obtuvieron por *Illumina*, se ensamblaron con los programas *VELVET* [72], k=27; *SOAPdenovo* [73] con los parámetros: –k 25 -R -o graph_prefix y con el programa *AMOS-Cmp*-shortReads [74], utilizando como referencia el genoma de la cepa CIAT 652 y parámetros por defecto.

Las lecturas *forward* generadas por *Illumina* fueron ensambladas con *TAIPAN* (*k =19)* y *SSAKE (k = 17)*. Los *contigs*[e] generados con estos dos programas se unieron para generar un ensamble híbrido utilizando la aplicación *Minimus2* del programa *AMOS* [75], con los siguientes parámetros: -d conserr = 0.001, -d minid = 99.

Posteriormente, utilizando las lecturas y los ensambles generados por *Illumina* (secuenciación del genoma y de *RNA-seq*), así como la información generada por 454, se generaron ensambles híbridos utilizando los programas *Newbler* (parámetros por defecto) *y* Minimus2 (-D CONSERR=0.001 -D MINID=99) para generar el ensamble final.

[e] *Contig*: De *contiguous*, es un grupo de fragmentos de *DNA* que sobrelapan y que representan el consenso de una región de *DNA* sin huecos (*gaps*) [118].

## 4.4 Construcción de pseudomoléculas y anotación del genoma.

Los *contigs* contenidos en los *andamios* del ensamble final se alinearon, ordenaron y orientaron con el programa *ABACAS* [76] utilizando los genomas de las cepas CFN 42 y CIAT 652 como referencia, lo que permitió construir *pseudoreplicones* del genoma de la cepa Ch24-10, y asignar los contigs al *pseudocromosoma* o a los *pseudoplásmidos*. Para asegurar que en la secuencias pudieran leerse los 6 marcos de lectura, para hacer la anotación de cada *pseudomolécula*, se agregó en el extremo 5´ de cada *contig* la secuencia *NNNNCATTCCATTCATTAATTAATTAATGAATGAATGNNNNN*. La predicción y anotación de genes de se realizó con el programa *CG-Pipeline* [77].

## 4.5 Electroforesis de campos pulsados.

Las cepas Ch24-10, CFN 42 y CIAT 652 se inocularon en 3 mL de PY líquido y se mantuvieron en agitación a 150 RPM a 28°C hasta alcanzar una densidad óptica (DO) de 1 medida a 620 nm en el *Spectrophotometer DU650* de *Beckman*. Se centrifrugó 1.5 mL de cultivo a 6,000 RPM y el botón celular se resuspendió en 400 µL de Solución de Resuspención (10 mM Tris pH 7.2, 20 mM NaCl, 50mM EDTA) y 400 µL de agarosa 1.8% en TA a 55°C. Utilizando el *CHEF Mapper plug mold* de *Bio-Rad* se generaron *plugs* con el cultivo celular embebido en agarosa. Los *plugs* se colocaron en tubos de polipropileno de 50 mL y se incubaron en 20 mL de Solución de Lisis (50 mM Tris-HCl, 50 mM EDTA, N-lauroylsarcosine 1%) más 100 µL de proteinasa K (20 mg/mL) durante 6 horas a temperatura ambiente. Pasado este tiempo fueron lavadas con agua destilada filtrada por ósmosis inversa en el equipo *Mili-Q* de *Merk Millipore* (agua destilada grado *milli-Q*). El proceso se repitió dos veces incubando por 12 horas. Finalmente, los *plugs* se lavaron con agua destilada grado *Mili-Q* y se incubaron 3 veces por 20 minutos en 20 mL de Solución de Lavado (20 mM Tris pH8, 50 mM EDTA). Finalmente, se agregaron 10 mL de EDTA 0.5 M y se mantuvieron a 4°C hasta su uso posterior.

Para visualizar los plásmidos, los *plugs* se embebieron en un gel de agarosa 1% en Tris Boratos (TB) al 0.5 X (TB 1X: Trizma base, 10.8 g; EDTA, 0.93 g; ácido bórico, 5.5 g / L). Utilizando el equipo *CHEF Mapper XA Pulsed Field Electrophoresis System* (Byo-Rad, EEUU) se realizó una electroforesis de campos pulsados con las siguientes condiciones: tiempo total 23 horas, *initial switch* 24 segundos, *final switch* 2 minutos 30 segundos, V/cm 6.0, 200 V, ángulo 120°.

## 4.6 Geles tipo *Eckhardt.*

Para visualizar los plásmidos de *Rhizobium* se utiliza el protocolo de *Eckhardt* [78], modificado por Hynes & O'Conell [79]. Las cepas CFN42 y Ch24-10 se inocularon en $PYNaI_{30}$ y se incubaron a 28°C en agitación a 150 RPM, toda la noche. Posteriormente, se tomaron 200 µl de cultivo al que se le añadieron 0.5 mL de *Sarcosyl* 0.3% en TB y se centrifugaron a 13,000 RPM por 5 segundos. El botón celular se mantuvo en hielo, se agregaron 20 µl de solución E1 (sacarosa 10% en TB 1X; *RNAse*, 10mg /ml en TB; más lisozima que se añade en el momento) para cargar las células en un gel de agarosa 0.7% en TB 1X y dodecil sulfato de sodio 1%. La electroforesis se realizó durante 5 minutos a 15 volts y posteriormente a 80 volts por 8 horas a 4°C.

## 4.7 Regiones compartidas en el genoma.

Para obtener las regiones genómicas compartidas de la cepa Ch24-10 contra el genoma de las cepas CFN 42 y CIAT 652, estos se alinearon utilizando el programa MAUVE [80], con los siguientes parámetros= *--max-backbone-gap=50, --weight=90, --backbone-size=100, --max-backbone-gap=5, --island-size=10*.

## 4.8 Obtención de regiones codificantes y familias de proteínas.

Se realizó una comparación de las regiones codificantes (secuencias de aminoácidos) de las cepas Ch24-10, CFN 42 y CIAT 652 utilizando el programa *BLASTp* [81] con los siguientes parámetros: $E > 1e^{-7}$, no filtrar regiones de baja complejidad. Las matrices de similitud resultantes se

utilizaron como archivos de entrada para agrupar a las proteínas ortólogas (y parálogos recientes) con el programa *SCPS* [82] con los parámetros: épsilon 1.05 y k-max 250. Las familias fueron graficadas en un diagrama de *Venn* con el programa *VennDiagram* [83].

## 4.9 Análisis filogenéticos a partir de las secuencias de los genes *rRNA16s atpD, recA y rpoB*.

Las reacciones de PCR se hicieron en un volumen final de 100 µl como sigue: 200 mM de cada *dNTP*, 1 U de *Taq polymerase* (Invitrogen), 20 pmol de cada oligonucleótido y 1.5mM de $MgCl_2$. Para amplificar los genes *rRNA16s, atpD, recA y rpoB* se utilizaron los oligonucleótidos-temperatura de *annealing* como sigue (en orden respectivo): *fd1*/*rd1-55°C* [84], atpD255F/atpD754R- [85], *recA6F/ recA555R-50°C, Br3200F/Br39050R-57°C* [86].

En todos los casos se utilizó el *termociclador GeneAmp ® PCRSystem* 2700 (*Applied Biosystems,* ahora *Life Technologies*). Los productos de *PCR* se observaron en un gel de agarosa 1% en TA y se marcaron con bromuro de etidio para su visualización en gel de agarosa 1%. Los oligonucleótidos y *dNTPs* se eliminaron con el *High Pure PCR Product Purification Kit* (Roche). La secuencia se obtuvo utilizando un servicio comercial.

Las secuencias se alinearon y editaron con el programa *BioEdit 7* [87]. El porcentaje de identidad entre las secuencias se obtuvo después de remover todas las columnas con gaps del alineamiento. Se seleccionó el mejor modelo de evolución con el programa *JModelTest 0.1.1* usando el criterio Akaike. Se construyeron las filogenias por *máxima verosimilitud* (*maximum likelihood*) y *neighbor-joining* con el programa MEGA5 [88]. El soporte de los nodos se evaluó con réplicas de *bootstrap* de 100 y 1000 para *máxima verosimilitud* y *neighbor-joining* respectivamente.

## 4.10 Construcción de árboles filogenéticos a partir de secuencias genómicas.

Las regiones compartidas entre todos los genomas de *Rhizobium* CIAT 652, CFN 42*, CNPAF 512, Rhizobium leguminosarum bv. viciae 3841 y Sinorhizobium meliloti 1021* se identificaron

utilizando el programa *Mugsy* [89]*,* optimizado para alinear genomas de cepas/especies con cercanía filogenética. Se utilizaron los siguientes parámetros: *--minlength=30 --distance=100 -duplications 1 -fullsearch –refine.* Cada región ortóloga se extrajo utilizando códigos de Perl escritos para tal propósito (JL Acosta, no publicados) y filtrados para eliminar regiones con alineamientos de baja consistencia utilizando el programa trimAL [90] con el parámetro *-automated1.*

Todas las regiones que alinearon se concatenaron para construir dos árboles filogenéticos. En primer lugar, utilizando el programa *RaxML* [91] se construyó un árbol mediante *máxima verosimilitud* se utilizó el modelo de sustitución *GTR y GAMMA+P-Invar* para estimar la tasa de heterogeneidad *(rate heterogeneity),* se realizaron 1000 réplicas con los parámetros*: -p 12345, -e 0.0000001 -c 8, -j STRICT -k 1000.* El segundo árbol se construyó con la metodología *neighbor-net network* con el programa *Splits tree4* [92]*.*

## 4.11 Hibridación *in silico* y obtención de los índices *ANI*.

La hibridación *in silico* se realizó alineando pares de genomas utilizando *MUMmer* [93], utilizando los parámetros por defecto. Se escribieron códigos en *Perl* para analizar el archivo de salida (Ernesto Ormeño, datos no publicados). Se obtuvo el valor de hibridación *in silico* entre cada par de genomas expresado como el porcentaje del número de nucleótidos idénticos en las regiones alineadas en relación a la longitud del genoma utilizado como referencia.

Los índices *ANI* entre los genomas de las cepas Ch24-10, CNPAF512, CFN 42 y CIAT 652 se obtuvieron con el programa *Jspecies* [55].

## 4.12 Cultivo de maíz y frijol.

Para esterilizar las semillas de maíz se lavaron con agua destilada estéril y se sumergieron en etanol 75% durante 5 minutos y posteriormente en hipoclorito de sodio 1% durante una hora en agitación, se lavaron 10 veces con agua destilada estéril, y se colocaron en cajas *Petri* con agar 0.7% por 48 horas en obscuridad para su germinación.

Las semillas de frijol se lavaron con agua destilada estéril y se mantuvieron en agitación en etanol al 75% por 15 minutos y posteriormente en hipoclorito de sodio 0.6% durante 20 minutos. Se lavaron 10 veces con agua destilada estéril y se colocaron en cajas *Petri* con agar 0.7% por 48 horas en obscuridad para su germinación.

Transcurridas las 48 horas posteriores a la esterilización, cada semilla se colocó en un sistema hidropónico en tubos de vidrio de 50 mL sobre un soporte de acero inoxidable (Figura 1) y se inocularon con una concentración bacteriana final de $10^8$ bacterias por planta (0.005 DO, 620 nm), resuspendidas en medio *Fahraeus* [94]. Se cubrió la parte inferior del tubo donde se encuentra la semilla para impedir el paso de luz. Las plántulas de maíz se mantuvieron por 5 días a 28°C con ciclos de luz-obscuridad de 12 horas. Las plántulas de frijol se mantuvieron en las mismas condiciones por 3 días (Figura 1).

## 4.13 Obtención de bacterias adheridas al rizoplano de maíz.

Transcurridos 3 y 5 días para maíz o frijol respectivamente, las raíces de 20 plantas sin síntomas de estrés ni contaminación por otros microorganismos, se cortaron y se sumergieron en una solución de *RNAlater* (*Ambion*) 10% v/v; rifampicina, concentración final 200 µg; MgSO4, 20 g/L en un volumen final de 130 mL. Para desprender a las bacterias adheridas a la raíces, éstas se sometieron a sonicación indirecta durante 20 minutos y se agitaron en vórtex por 30 segundos, como se describe Ormeño y colaboradores [95] este proceso no afecta la viabilidad celular.

La suspensión resultante se filtró con 4 membranas superpuestas estériles de *Miracloth* (*Merck Millipore*) para eliminar el tejido vegetal que se hubiera desprendido en el procedimiento, se centrifugó a 14°C, a 6,000 RPM durante 5 minutos para recuperar las células desprendidas del rizoplano de cada planta.

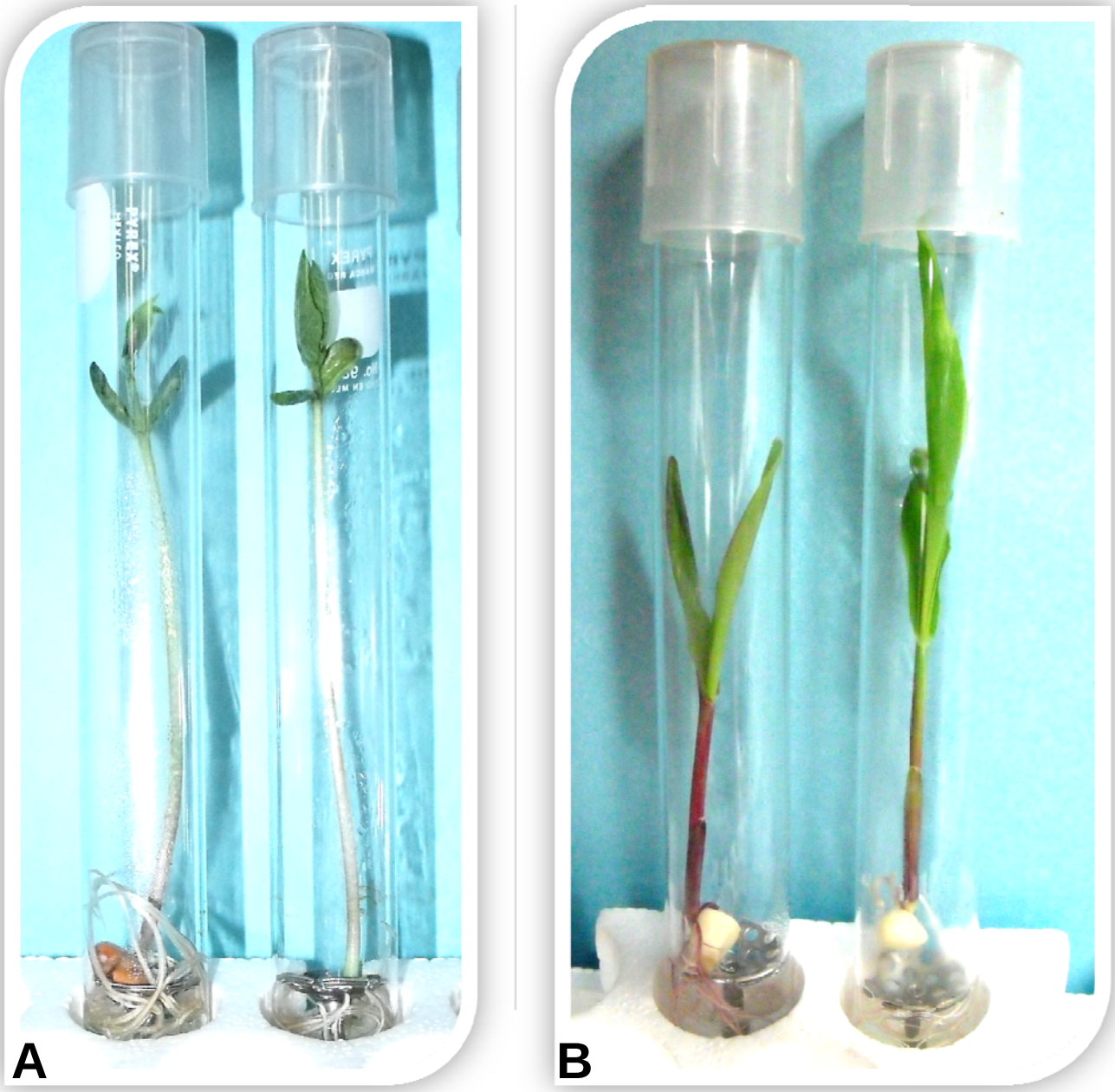



**Figura 1. Sistema hidropónico de cultivo de maíz.**

Se muestran el sistema hidropónico de cultivo de maíz y de frijol. **A**. Plántulas de frijol 3 días posteriores a la inoculación. **B**. Plántulas de maíz 5 días posteriores a la inoculación Las raíces de las plantas se encuentran sumergidas en medio Fahraeus [94] inoculado con la cepa Ch24-10. La semilla está sostenida en un soporte de acero inoxidable con la finalidad de evitar el contacto de ésta con el medio.

## 4.14 Extracción y purificación de *RNA*. Secuenciación del transcriptoma mediante *RNA-seq*.

A partir del botón celular que se obtuvo en el paso anterior, se extrajo el *RNA* total utilizando *TRIzol Reagent* (*Invitrogen*), se cuantificó y verificó la calidad con el *NanoDrop2000* (*Thermo Scientific*). Sólo aquellas muestras con una relación 260/280 mayor a 1.2 se utilizaron para los pasos siguientes.

Un microgramo de *RNA* total se trató con la enzima *DNAse RNAse free* de *Invitrogen* para eliminar el *DNA* presente en la muestra Posteriormente, se eliminó el *RNA ribosomal* con el *kit RiboMinus Transcriptome Isolation Kits* (Yeast and Bacteria) de *Invitrogen*.

La secuencia del transcriptoma de la cepa Ch24-10 fue obtenida en la UUSMD, UNAM, con el sistema *Genome Analyzer GAIIx* (*Illumina*). La calidad y cantidad del *RNA* se verificó con el *Bioanalyzer-2010 RNA600 Kit* (*Agilent Technologies*). Para el caso del primer experimento de rizoplano de maíz, MR1-RNA, se obtuvo *DNA* complementario mediante *PCR* a partir del *RNA* utilizando oligonucleótidos al azar con 22 y 7 ciclos posteriores de amplificación. Se siguió el protocolo *Single-read sequencing assay*. Se obtuvo una biblioteca de 200 pares de bases y se secuenciaron 36 nucleótidos. En este experimento se utilizó el *RNA* que se extrajo del botón celular que se obtuvo de 9 experimentos independientes de 20 plántulas cada uno.

Se realizó una repetición en condiciones de rizoplano de maíz, MR2-RNA y se obtuvieron dos muestras de rizoplano de frijol independientes FR1-RNA y FR2-RNA. En estos tres casos el *RNA* se extrajo a partir del botón bacteriano que se obtuvo de 40 plántulas cada uno. El *DNA* complementario se obtuvo utilizando oligonucleótidos al azar con 22 ciclos de amplificación. Estas muestras se trataron con la enzima *Terminator™ 5´-Phosphate-Dependent Exonuclease TER51020* (*Epicentre)* para eliminar el *RNA* degradado presente en las muestras. Se siguió el protocolo *Single-read sequencing assay*, se obtuvo una biblioteca de 200 pares de bases y se secuenciaron 36 nucleótidos.

## 4.15 Ensamble y anotación del transcriptoma.

El ensamble de las lecturas generadas por *RNA-seq* se realizó con el programa *OASES* [96], el cual permite ensamblar en *andamios* y posteriormente dividirlos en *locus* que se consideran unidades transcripcionales [96]. *OASES* recibe como entrada un alineamiento generado en *VELVET* [72], se utilizaron los siguiente parámetros para cada muestra:

FR1-RNA: k=21, -exp_cov 2, -cov_cutoff 0.3777216 -amos_file yes -read_trkg yes.

FR2-RNA: k=25 -exp_cov 5 -cov_cutoff 2.832724128 -amos_file yes -read_trkg yes.

MR2-RNA: k=23 -exp_cov 11 -cov_cutoff 7.8190213571712 -amos_file yes -read_trkg yes.

MR2-RNA: k=25, exp_cov 4 -cov_cutoff 1.0439650887168 -amos_file yes -read_trkg yes.

Posteriormente, a partir de los resultados de *VELVET* se ensambló cada una de las muestras con *OASES* utilizando los parámetros por defecto. Los *loci* obtenidos se anotaron en el servidor del Sistema para una rápida anotación basado en subsistemas*, RAST (Rapid Annotation using Subsystems Technlogy)* [97].

## 4.16 Determinación de los genes con mayor expresión.

Las lecturas que se obtuvieron de cada experimento de RNA-seq, FR1-RNA, FR2-RNA,MR2-RNA y MR2-RNA se mapearon contra la secuencia de los *pseudoreplicones* del genoma de la cepa Ch24-10, utilizando el programa *PALMapper* [98]. Utilizando códigos en *Python* (JL Acosta, no publicados) se obtuvo el número de lecturas alineadas en los genes predichos de cada pseudoreplicón del genoma de la cepa Ch24-10. Se tomaron en cuenta sólo aquellas lecturas que alinearon en un solo gen.

Se obtuvo la frecuencia de mapeo de las lecturas por gen, misma que se ajustó a una distribución estadística *gamma*. Se obtuvieron los genes cuya distribución estuvo a la derecha de la curva después del punto de inflexión de ésta y se consideró como sobre-expresado si se encontraba presente en las dos repeticiones de cada experimento, en rizoplano de maíz por un lado y rizoplano de frijol por otro.

# 5. Resultados. Genómica y taxonomía.

## 5.1 Ensamble del genoma de la cepa Ch24-10.

La secuencia del genoma de la cepa Ch24-10 se obtuvo con dos tecnologías: *Roche-454 Life Science*, a partir de dos bibliotecas con los protocolos 8K y 3K *Long-Tag Paired End* e *Illumina Genome Analyzer* (*Paired end Protocol*).

En la Tabla 2 se muestran las estadísticas generales de las secuencias obtenidas con la tecnología *Roche-454 Life Science*. Estas lecturas se ensamblaron con el programa *gsAssembler* (*Newbler)* (Tabla 3).

**Tabla 2. Estadísticas generales de las bibliotecas obtenidas por la tecnología *Roche-454 Life Science.***

| Biblioteca | Lecturas obtenidas sin par | | Lecturas pareadas | Longitud de las secuencias (nucleótidos) | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Forward | Reverse | | Forward | | | Reverse | | |
| | | | | media | mínima | máxima | media | mínima | máxima |
| 8 Kb | 148,316 | 235,989 | 87,866 | 224 | 20 | 703 | 209 | 20 | 501 |
| 3 Kb | 67,924,526 | 63,007,839 | 156,195 | 150 | 20 | 457 | 146 | 20 | 459 |

Mediante la tecnología de *Illumina* se obtuvieron 17, 836, 006 lecturas y se eliminaron todas aquellas que tuvieran en su secuencia una o más *N* quedando un total de 17, 754,122. En una primera aproximación se ensamblaron las lecturas pareadas utilizando los programas *VELVET* [72], *SOAPdenovo* [73] y *AMOS*-*Cmp*-shortReads [75]. Se observó que los tres programas dejaron la mayoría de las lecturas pareadas fuera del ensamble: *VELVET*, 10, 154,055; *SOAPdenovo*, 15, 246,25 y *AMOS*, 8, 663,232.

En la Figura 2 se puede observar la calidad de las lecturas generadas por *Illumina*. En el caso de las lecturas *reverse* ésta es menor a 25 a partir de la segunda posición, mientras que las lecturas *forward* tienen una calidad mayor a 30 que decae a partir de la posición 20.

Para evaluar si existían diferencias en los ensambles generados utilizando por separado las lecturas *forward* y *reverse* se generaron ensambles *de novo* con *VELVET* (por la rapidez con que se generan los resultados), en el primer caso de obtuvieron 6,791 *contigs*[f] con una *N50*[g] de 1,881 quedando fuera del ensamble 1, 267,398 lecturas. El ensamble con las lecturas *reverse* generó 13,746 *contigs,* con una *N50* de 852 y quedaron fuera del ensamble 8, 198,363 lecturas. El número de lecturas *forward* que son utilizadas para construir el ensamble es considerablemente mayor al de las lecturas *reverse*, el número de *contigs* generados es menor y de mayor longitud respecto a las lecturas *reverse.*

Debido a estos resultados se decidió ensamblar las lecturas *forward* Se evaluaron los programas *TAIPAN* y SSAKE. Se obtuvieron los mejores resultados utilizando un *k* de 19 y de 17, respectivamente y se construyó un ensamble híbrido con *Minimus2* a partir de los *contigs* que estuvieran presentes en los resultados de los dos programas para excluir errores de ensamblado. Como resultado de este ensamble híbrido se obtuvieron 386 *contigs* unidos en 14 *andamios* con un total de 6, 776,943 bases, una N50 de 1, 738,464 bases y el tamaño del *andamio* mayor de 1, 789,161 bases.

[f] *Contig*: De *contiguous*, es un grupo de fragmentos de *DNA* que sobrelapan y que representan un consenso de una región de *DNA* sin huecos (*gaps*) [118].

[g] N50: Es la longitud del *contig* más pequeño del grupo que contiene a los *contigs* de mayor tamaño que en conjunto representan la longitud de al menos el 50% del ensamble [119].

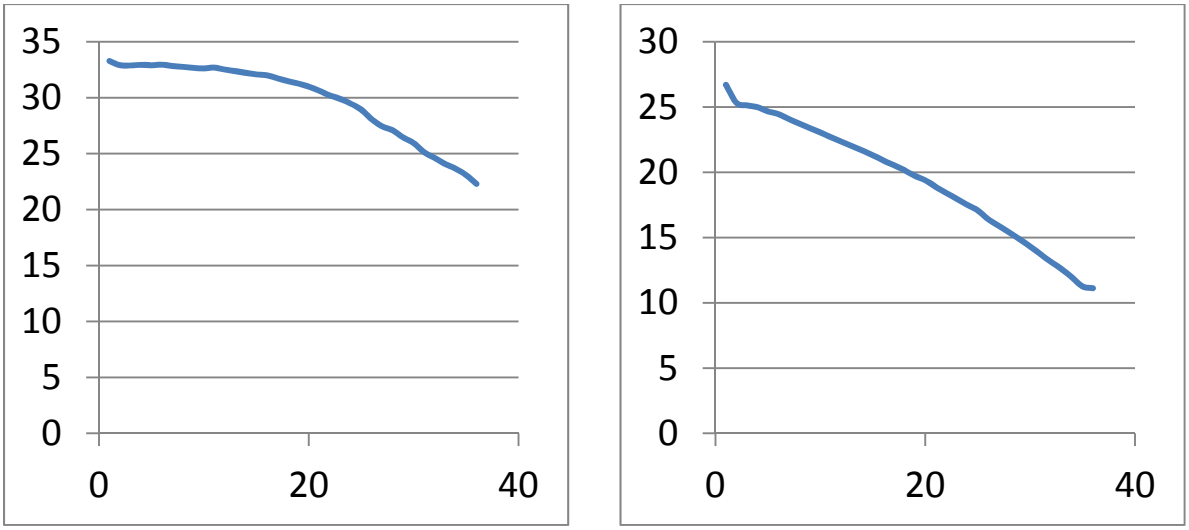



**Figura 2. Calidad de las lecturas generadas por *Illumina.***

Se muestra el promedio de la calidad por posición de las 17836006 lecturas obtenidas por *Illumina.* En cada gráfica se muestra en el eje de las abscisas la posición de la secuencia y en el eje de las ordenadas el promedio de la calidad por posición. En el panel A se ven los resultados para las lecturas *forward* y en el panel B los resultados para las lecturas *reverse*. Se puede ver que la calidad promedio en las lecturas *reverse* desde el inicio de la lectura es menor a 30 y menor a la calidad de las lecturas *forward* también.

Para obtener un ensamble final del genoma de la cepa Ch24-10 se realizaron diferentes experimentos bioinformáticos. Se generaron ensambles híbridos utilizando como referencia el ensamble *E1Ref454* (generado partir de las secuencias de *Roche-454 Life Science*) en combinación con el ensamble de las lecturas *forward-Minimus2*, las 440 secuencias generadas por el método *Sanger* del proyecto precedente[52] y las lecturas pareadas generadas por *Illumina*. En la Tabla 3 se muestra en detalle la información utilizada y en la Tabla 4 se muestran algunas estadísticas de los ensambles.

Los ensambles *E2RCIF* y *E4RCIFS* se construyeron utilizando como referencia el ensamble *E1Ref454 y* los *contigs* generados a partir de las secuencias *forward* de *Illumina.* Al añadir las secuencias *Sanger* en el ensamble *E4RCIFS* se observa que ensamble se fragmenta, pasando de

14 a 15 *andamios*[h], además disminuye la longitud promedio de éstos, así como la N50. Sin embargo, se obtiene un *andamio* de longitud mayor (Tabla 4).

En el ensamble *E6RCIFLFSCseq* se utilizó como información adicional el ensamble de las secuencias que se obtuvieron por *RNA-seq*. Si éste se compara con el *E5RCIFLFS* que utiliza la misma información pero sin en ensamble de las lecturas de *RNA-seq*, vemos que el ensamble se fragmenta, como ocurre en el caso anterior, pasando de 17 *andamios* a 21 de menor tamaño pero se obtiene un *andamio* de mayor tamaño (Tabla 4).

Adicionalmente, para determinar si la información pareada de las secuencias de *Illumina* (la distancia entre lecturas es de 200 nucleótidos aproximadamente) permitiría mejorar en ensamble híbrido, se realizó un mapeo de las lecturas *forward* utilizando BWA [99]contra el ensamble *E1Ref454* y posteriormente se recuperó su par *reverse* utilizando un código en Perl (José Luis Acosta, datos no publicados). Se recuperaron 15, 093,202 lecturas de un total de 17, 836,006. Esta información junto con las lecturas *forward* se utilizó para generar el ensamble *E7LIFPE* usando como referencia el ensamble *E1Ref454*. Al utilizar estas lecturas el resultado no fue mejor respecto al ensamble que utiliza la misma información sin las lecturas pareadas, *E2RCIF, e*l ensamble se fragmentó pasando de 14 a 17 *andamios*, disminuyó el tamaño promedio de éstos y la N50, pero generó un *andamio* de mayor longitud respecto al ensamble *E2RCIF.*

Finalmente, para construir el ensamble final del genoma se decidió utilizar los ensambles *E4RCIFS, E5RCIFLFS, E6RCIFLFSCseq para* generar en ensamble final *E8F456* pues éstos contienen toda la información disponible y se obtuvieron los *andamios* de mayor tamaño. Se utilizó en ensamble *E4RCIFS* como referencia por contener el menor número de *andamios*, el tamaño promedio y N50 mayores. En la Tabla 5 se muestran las estadísticas del ensamble final.

[h] *Andamio* o *Scaffold*: Una serie de *contigs* ordenados y orientados correctamente, forman una secuencia de *DNA* que puede estar separada por huecos de longitud conocida [120].

## Tabla 3. Ensambles híbridos del genoma de la cepa Ch24-10.

| Ensamble | Datos utilizados |
|---|---|
| *E1Ref454* | Secuencias 8K y 3K Long-Tag Paired End |
| *E2RCIF* | *E1Ref454* + *Contigs Illumina forward (Minimus2)* |
| *E3RCIFLF* | *E1Ref454* + *Contigs Illumina forward (Minimus2)* + Lecturas *forward Ilumina* |
| *E4RCIFS* | *E1Ref454* + *Contigs Illumina forward (Minimus2)* + Lecturas *Sanger* |
| *E5RCIFLFS* | *E1Ref454* + *Contigs Illumina forward (Minimus2)* + Lecturas *forward Ilumina* + *Lecturas Sanger* |
| *E6RCIFLFSCseq* | *E1Ref454* + *Contigs Illumina forward (Minimus2)* + Lecturas *forward Ilumina* + Lecturas Sanger + *Contigs RNA-seq* |
| *E7LIFPE* | *E1Ref454* + *Lecturas forward Illumina* + *illumina paired-end* |
| *E8F456* | *E4RCIFS* + *E5RCIFLFS* + *E6RCIFLFSCseq* |

Se muestran las combinaciones que se utilizaron para evaluar distintos ensambles utilizando diferente información de entrada. Todos los ensambles híbridos fueron generados con *Newbler*, excepto el *E8F456* que se generó con *Minimus2.* A partir del ensamble *E2RCIF* se utilizó como referencia el *ensamble E1Ref454* como referencia, excepto en el ensamble *E8E456* dónde se utilizó el ensamble *E4RCIFS*.

## Tabla 4. Estadísticas de los ensambles del genoma de la cepa Ch24-10.

| Parámetro | *E1Ref454* | *E2RCIF* | *E3RCIFLF* | *E4RCIFS* | *E5RCIFLFS* | *E6RCIFLFSCseq* | *E7LIFPE* | *E8F456* |
|---|---|---|---|---|---|---|---|---|
| Número | 16 | 14 | 16 | 15 | 17 | 21 | 17 | 14 |
| Longitud promedio | 424,120 | 498,067 | 423,243 | 451,389 | 398,027 | 323,245 | 398,499 | 473,557 |
| N50 | 1,736,075 | 1,738,464 | 1,735,285 | 1,736,012 | 1,735,191 | 1,740,719 | 1,106,433 | 1,757,397 |
| Longitud del mayor | 1,796,264 | 1,789,161 | 1,787,229 | 1,791,714 | 1,789,339 | 1,797,758 | 1,792,656 | 2,219,033 |

Se muestran los valores para los *andamios* generados en cada alineamiento, utilizando diferente información de entrada. En ensamble *E8F456* se logró obtener el menor número de *andamios*, así como el de mayor tamaño. Además es el que tiene menor número de huecos (*gaps*) en cada *andamio*. Este se considera en ensamble final del genoma de la cepa Ch24-10.

## Tabla 5. Estadísticas del ensamble final del genoma de la cepa Ch24-10.

| Parámetros | *Ensamble Final E8F456* |
|---|---|
| Tamaño total de los *andamios* | 6,629,798 |
| *Andamio* de mayor longitud | 2,219,033 |
| *Andamio* de menor longitud | 2,116 |
| Número de *andamios* > 500 | 14 (100%) |
| Número de *andamios* > 1Kb | 14 (100%) |
| Número de *andamios* > 10Kb | 10 (71.4%) |
| Número de *andamios*> 100Kb | 7 (50%) |
| Número de *andamios*> 1Mb | 3 (21.4%) |
| Tamaño promedio de los *andamios* | 473,557 |
| Mediana del tamaño de los *andamios* | 295,425 |
| N50 de la longitud de los *andamios* | 1,757,397 |

Se muestran las estadísticas del ensamble final *E8F456 del genoma de la cepa Ch24-10 generado por Minimus2*[75]

## 5.2 El genoma de la cepa Ch24-10.

Cuando se aisló la cepa Ch24-10 se observó el perfil plasmídico utilizando la metodología desarrollada por *Eckhardt* [51] y modificada por Hynes y McGregor [79]. En éste se observaron dos bandas menores a 500 Kb y una de aproximadamente 1 Mb, se supuso que la cepa tiene 3 plásmidos (Figura 1.A). Sin embargo, al analizar el ensamble del genoma generado en el apartado anterior se encontró la presencia de 5 operones completos *repABC*, uno de ellos repetido en el *andamio* mayor a 1 Mb.

Para verificar la presencia de 4 plásmidos en la cepa se realizó una electroforesis de campos pulsados, que permitió visualizar los dos plásmidos de tamaño similar (entre 400-415 Kb) y corroborar la presencia de 4 plásmidos en la cepa (Figura 1.B.).

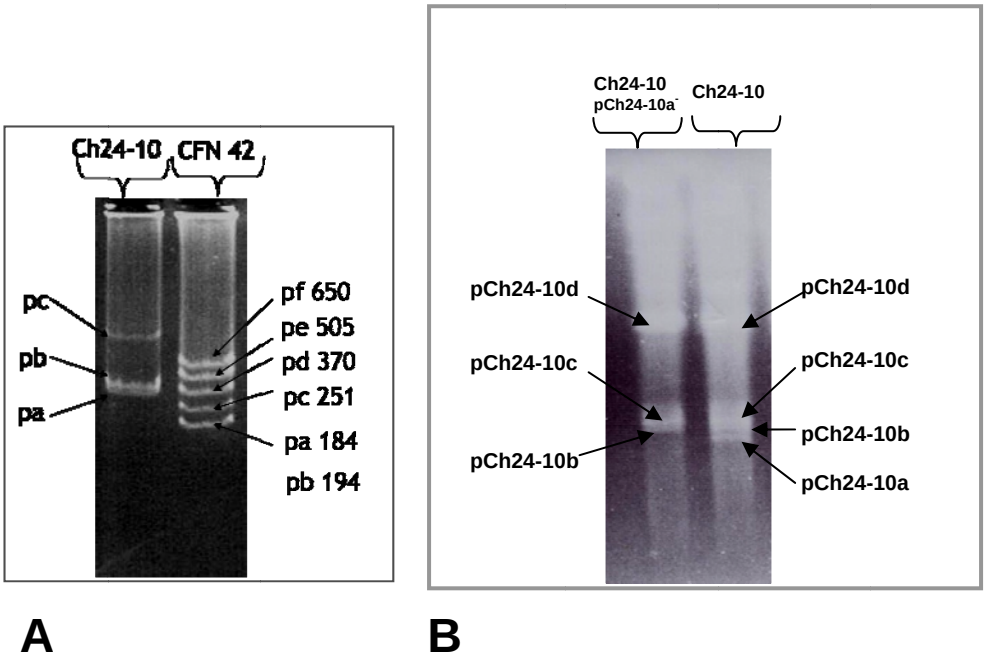



**Figura 3. Perfil plasmídico de la cepa Ch24-10**

A. Perfil plasmídico obtenido mediante geles tipo *Eckhardt* [79]. Se muestran los tres plásmidos que se planteó tenía la cepa Ch24-10 cuando se aisló (pa, pb y pc). Se muestra también el perfil plasmídico de la cepa *R. etli* CFN 42, así como el tamaño en kilobases. B. Perfil plasmídico de la cepa Ch24-10 mediante electroforesis de campos pulsados. En el carril uno se muestra a la cepa Ch24-10 curada del plásmido más pequeño pCh24-10a y la cepa silvestre con los 4 plásmidos.

A partir del ensamble final *E8F456* se generaron 5 *pseudoreplicones*[i], utilizando como referencia la secuencia de los genomas completos de las cepas CFN 42 y CIAT 652, para ordenar y orientar los *andamios* y para asignarlos a cada uno de los replicones. En la Tabla 6 se muestra la longitud predicha para cada uno, el número y longitud de los huecos en la secuencia, así como el número de sitios en los que no fue posible obtener un consenso.

Por localización del los genes *nif y fix*, así como por un alineamiento global contra los plásmidos simbióticos de las cepas CFN 42 y CIAT 652, se deduce que el plásmido simbiótico es el pCh24-10c (Tabla 6).

Los resultados del ensamble del genoma de la cepa Ch24-10 se encuentran depositados en *DDBJ/EMBL/GenBank* con el número de acceso AHJU00000000.

## Tabla 6. Longitud de los pseudoreplicones de la cepa Ch24-10.

| Replicón | Tamaño (b) | Huecos (b) (menor, mayor ) / Total huecos b / Sitios sin resolver b |
|---|---|---|
| Cromosoma | 4,425,198 | 24 (15, 2,946) / 14,506 / 36 |
| pCH241d | 1,100,215 | 2 (433, 763) / 1,196 / 30 |
| pCH241c pSym | 436,817 | 8 (161,2,178) / 8,434 / 11 |
| pCH241b | 419,653 | 2 (161,348) / 509 / 0 |
| pCH241a | 365,574 | 1 (720) / 720 / 1 |

Se muestra la longitud predicha de los *pseudoreplicones* construidos a partir del ensamble final. Los *andamios* del ensamble *E8F456* se alinearon, ordenaron y orientaron con el programa *ABACAS* [76] utilizando como referencia los genomas completos de las cepas CFN 42 y CIAT 652. Se muestra el tamaño calculado de cada replicón, considerando el número y tamaño de los huecos mayor y menor y los sitios en los que no se obtuvo un consenso. b=bases.

[i] Replicón: Segmento de DNA que puede autoreplicarse de manera autónoma [31].

## 5.3 Taxonomía y genómica comparativa de la cepa Ch24-10.

Cuando la cepa Ch24-10 fue aislada se clasificó como *Rhizobium etli* de acuerdo a características fenotípicas como la morfología de crecimiento en medio PY: colonias aperladas, gomosas y redondas; la incapacidad de crecer en medio *LB* y por la propiedad de formar nódulos fijadores de nitrógeno con el frijol [51] .

Otra tipificación fenotípica utilizada fue la *electroforesis de enzimas multilocus* (*MLEE*). Esta metodología permite distinguir entre diferentes alelos presentes en distintas cepas, por movilidad electroforética de enzimas, lo que permite agrupar a las cepas utilizadas en tipos electroforéticos (TE). Se obtiene la distancia genética entre cada par de TE, se construye una matriz de distancia y se realiza el agrupamiento por el método de *unión promedio (average-linkage)*. En este caso se evaluó la actividad de ocho enzimas codificadas en cromosoma: xantino deshidrogenasa, malato deshidrogenasa, indofenol oxidasa, glucosa-6-fosfato deshidrogenasa, isocitrato deshidrogenasa, fosfoglucomutasa, hexoquinasa y alanina deshidrogenasa. Los resultados colocaron a la cepa Ch24-10 en el mismo TE que la cepa tipo de *Rhizobium etli* CFN 42 [51].

La identidad de la secuencia génica del gen *rRNA16S* de la cepa Ch24-10 es del 99% (e = 0.0) respecto al gen de las cepas *R. etli* CFN 42 y CIAT 652. Como se puede ver en la Figura 4, cuando se genera un árbol filogenético utilizando una estrategia basada en distancia (identidad entre secuencias), *neighbor-joining*, se puede ver que las cepas CIAT652, CNPAF512, *R. phaseoli* ATCC 14482, *R. fabae* CCBAU3202 y *R. pisi* DSM30132 forman un clado, adyacente al que forman las cepas de *R. etli* CFN42 y MIM1 y en medio de éstos se encuentra a la cepa Ch24-10 y a *R. leguminosarum* USDA2370.

Se puede ver que la secuencia del *rRNA16s* no tiene carácter resolutivo al nivel taxonómico de especie [100] . Se ha propuesto que se utilicen otros marcadores moleculares. En este caso se seleccionaron los genes *atpD*, *recA* y *rpoB* (considerados genes constitutivos o *housekeeping* y que han mostrado ser adecuados para resolver las relaciones filogenéticas de los rizobios [85,86]) para

hacer una reconstrucción filogenética mediante *máxima verosimilitud.* Al utilizar las secuencias concatenadas de éstos para generar un árbol filogenético se observa que las cepas Ch24-10, CIAT652, CNPAF512 y *R. phaseoli* ATCC 14482 se agrupan en un clado distinto al que forman la cepas de *R. etli* CFN 42 y MIM1 (Figura 5 ).

El criterio estándar para asignar una cepa a una especie bacteriana es la r*easociación DNA-DNA* o la *ΔTm*[j][101]. Por ello se determinaron estos valores entre las cepas Ch24-10, las cepas tipo de *R. etli* CFN 42 y *R. phaseoli* ATCC 14482 y la cepa CNPAF512 (Tabla 7).

En la Tabla 7 se pueden ver los resultados de la *reasociación DNA-DNA*. Cuando se compara a la cepa Ch24-10 como referencia contra la cepa de *R. etli* CFN 42 el valor está por arriba del punto de corte, 75% ± 6, cuando se compara contra la cepa ATCC 14482 este valor es de 88% ± 6 y respecto a la cepa CNPAF 512 es de 78%. Al utilizar a la cepa *R. phaseoli* ATCC 14482 como referencia contra la cepa *R. etli* CFN 42 el valor de reasociación *DNA-DNA* es de 69% ± 4 y de 78% ± 5 respecto a la cepa CNPAF512.

Es claro que los valores que se obtienen al comparar las cepas Ch24-10, CIAT 652 y CNPAF 512 son mayores respecto a la cepa tipo de *R. phaseoli* ATCC 14482 que cuando se comparan contra la cepa tipo de *R. etli* CFN 42, sin embargo, estos últimos son iguales o cercanos al punto de corte de 70% propuesto.

Si se revisa el texto de Wayne y colaboradores [101] donde se propone la *reasociación DNA-DNA* para realizar inferencias taxonómicas dice textualmente: "…*species generally would include strains with approximately 70% or greater DNA-DNA relatedness and with 5°C or less ΔTm"* (Una especie generalmente incluirá cepas con valores de reasociación de DNA-DNA de aproximadamente 70% o mayor y una diferencia en la ΔTm de 5°C o menor). Se puede ver, que si bien se establecen puntos de corte en ambos valores, no es un criterio cerrado. En el texto se hace hincapié en que un

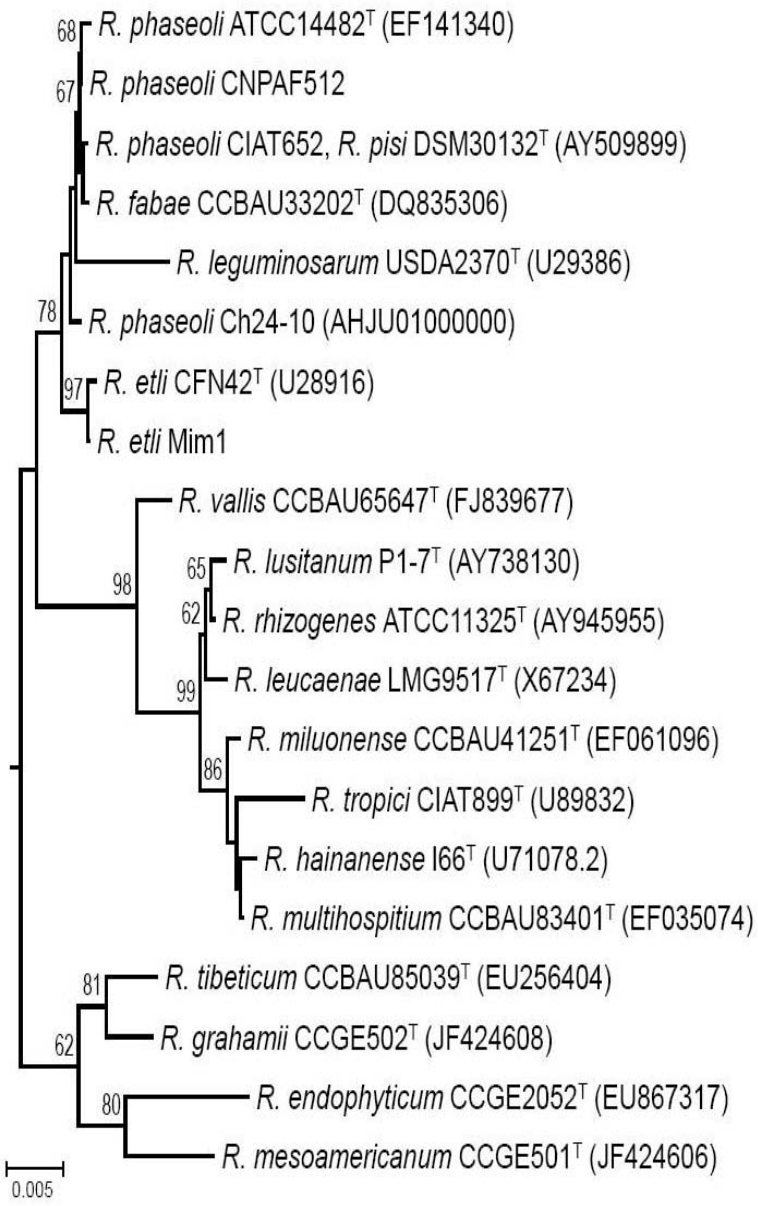
[j] *ΔTm*: Diferencia en la temperatura de asociación entre *DNA* dúplex y heteroduplex 102</sup>.

experto, considerando las particularidades de un grupo bacteriano, puede decidir considerar un valor más relajado o más estricto y colocar a una cepa dentro de una especie.

Considerando el carácter laxo del punto de corte recomendado para la *reasociación DNA-DNA* y que las cepas Ch24-10, CIAT 652 y CNPAF512 tienen porcentajes mayores entre si y respecto a la cepa tipo de *R. phaseoli* ATCC 14482 que a la cepa tipo de *R. etli* CFN 42 (aunque en un rango cercano al punto de corte) se propone que pertenecen a la especie *Rhizobium phaseoli* y no a *R. etli*.

Anteriormente se ha propuesto que las cepas CIAT 652 y CFN 42 pertenecen a linajes independientes dentro de *R. etli* y que entre ellos existe una recombinación baja de DNA [102] . Sin embargo los datos de *reasociación DNA-DNA* apuntan a que pertenecen a especies distintas.

Para respaldar los resultados anteriores, debido a la variabilidad y error estándar asociada a los experimentos de hibridación *DNA-DNA*, se utilizaron los genomas completos de las cepas CFN 42, CIAT 652, Ch24-10 y CNPAF512 para obtener datos de hibridación *in sillico*. Los resultados se muestran en la Tabla 7 y confirman los obtenidos en la hibridación *DNA-DNA*.




## Figura 4. Árbol filogenético construido a partir de las secuencias de los genes *rRNA16s* de diferentes cepas de *Rhizobium.*

Árbol filogenético construido por *neighbor-joining* con las secuencias parciales de los genes *rRNA16s* de cepas y especies de *Rhizobium* con el programa *MEGA5* [88]. Se muestran los valores de *bootstrap* mayores a 50%. La barra indica 5 sustituciones por cada 1000 nucleótidos. Entre paréntesis se muestra el número de acceso a las secuencias utilizadas. T = cepa tipo

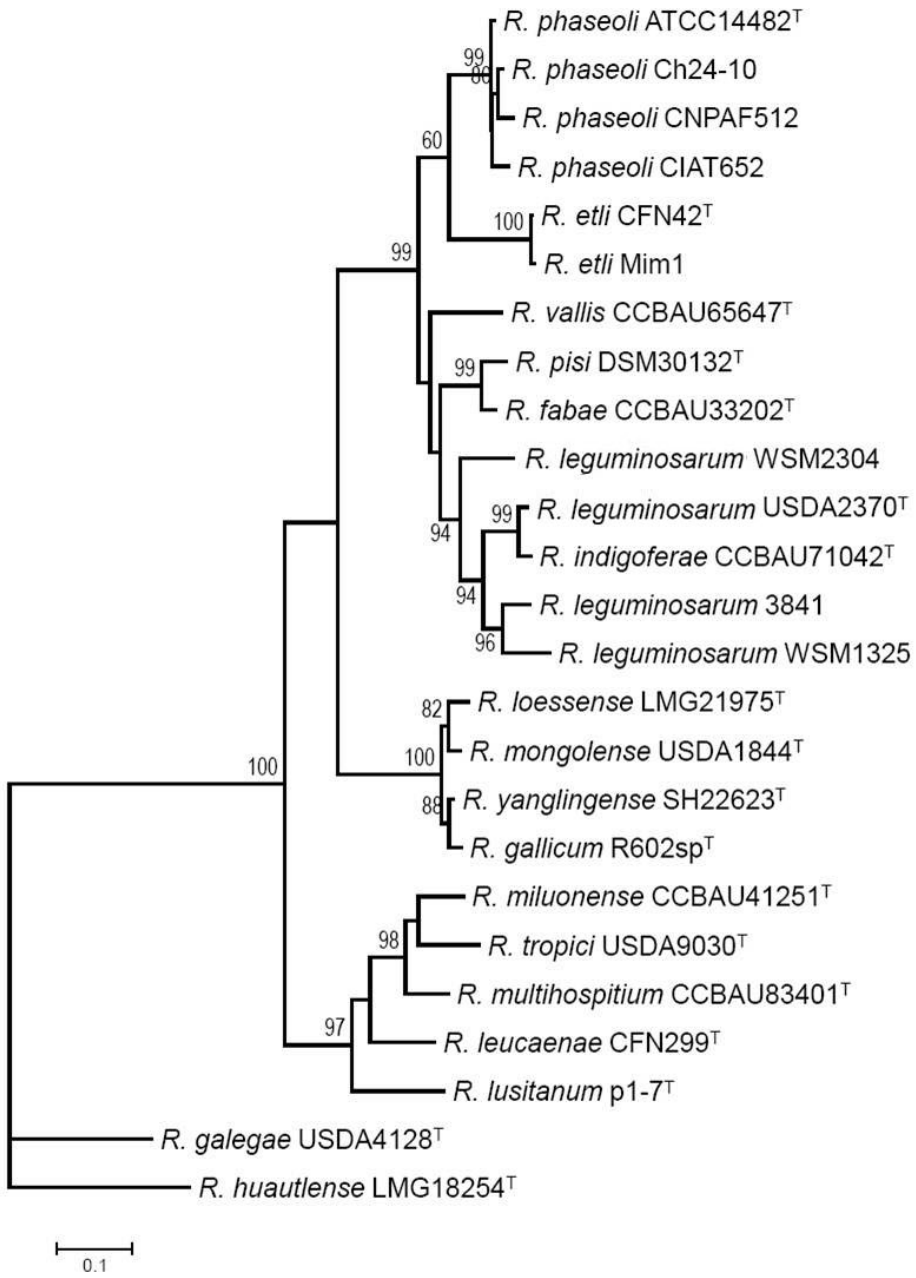



**Figura 5. Árbol filogenético construido partir de secuencias parciales de los genes *atpD, recA* y *rpoB* de diferentes cepas de *Rhizobium.***

Árbol filogenético construido por *maximum likelihood* con el programa *MEGA5* [88]. Se muestran los valores de *bootstrap* mayores a 50%. La barra indica una sustitución por cada 10 nucleótidos. T=cepa tipo

## Tabla 7. Porcentajes de hibridación DNA-DNA y *Average Nucleotide Identity* entre distintas cepas de *Rhizobium*.

| | Porcentaje de *reasociación DNA-DNA* | | | Porcentaje de hibridación i*n silico* | | | | *Average Nucleotide Identity* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Referencia | CFN 42 | ATCC 14482 | CNPAF 512 | CFN 42 | CIAT 652 | Ch24-10 | CNPAF512 | CFN 42 | CIAT 652 | Ch24-10 | CNPAF512 |
| CFN 42 | 100 | **70 ±2** | 67 ± 3 | 100 | **69** | **72** | 71 | 100 | | | |
| CIAT 652 | 68 ± 3 | **84 ± 3** | 76 ± 4 | **70** | 100 | **88.7** | 88 | **90.5** | 100 | | |
| Ch24-10 | **75 ± 6** | **88 ± 6** | 78 | **71** | **86** | 100 | 88 | **90.5** | **98.1** | 100 | |
| CNPAF512 | 67 ± 3 | 84 ±2 | 100 | 69 | 85 | 87 | 100 | 90.5 | 97.3 | 97.1 | 100 |
| ATCC 14482 | **69 ± 4** | 100 | 78 ± 5 | - | - | - | - | - | - | - | - |

Todos los datos están expresados en porcentaje, ± = error estándar. Se muestran en las columnas 2-4 los valores de *reasociación DNA-DNA* y en las columnas 5-8 los porcentajes de hibridación *in silico*. En las columnas 9-12 se muestran los resultados del índice *ANI*. Ya que el genoma de la cepa ATCC 1482 no ha sido secuenciado, no se incluyó en la comparación *in silico* ni para la obtención del *ANI*. Se muestran en negritas los valores obtenidos de la comparación entre las cepas Ch24-10, CFN 42 ATCC 14481 y CIAT 652.

Cuando se utiliza el genoma de la cepa CFN 42 como referencia los porcentajes de hibridación *in silico* son 71.8%, 69% y 70.6% respecto a las cepas Ch24-10, CIAT 652 y CNPAF512, respectivamente. Cuando la cepa Ch24-10 es utilizada como referencia los valores son de 71.5% respecto a la cepa CFN 42 (vs. 75% ±6 *de reasociación*), 86.4% respecto a CIAT 652 y 87.7 % respecto a CNPAF512.

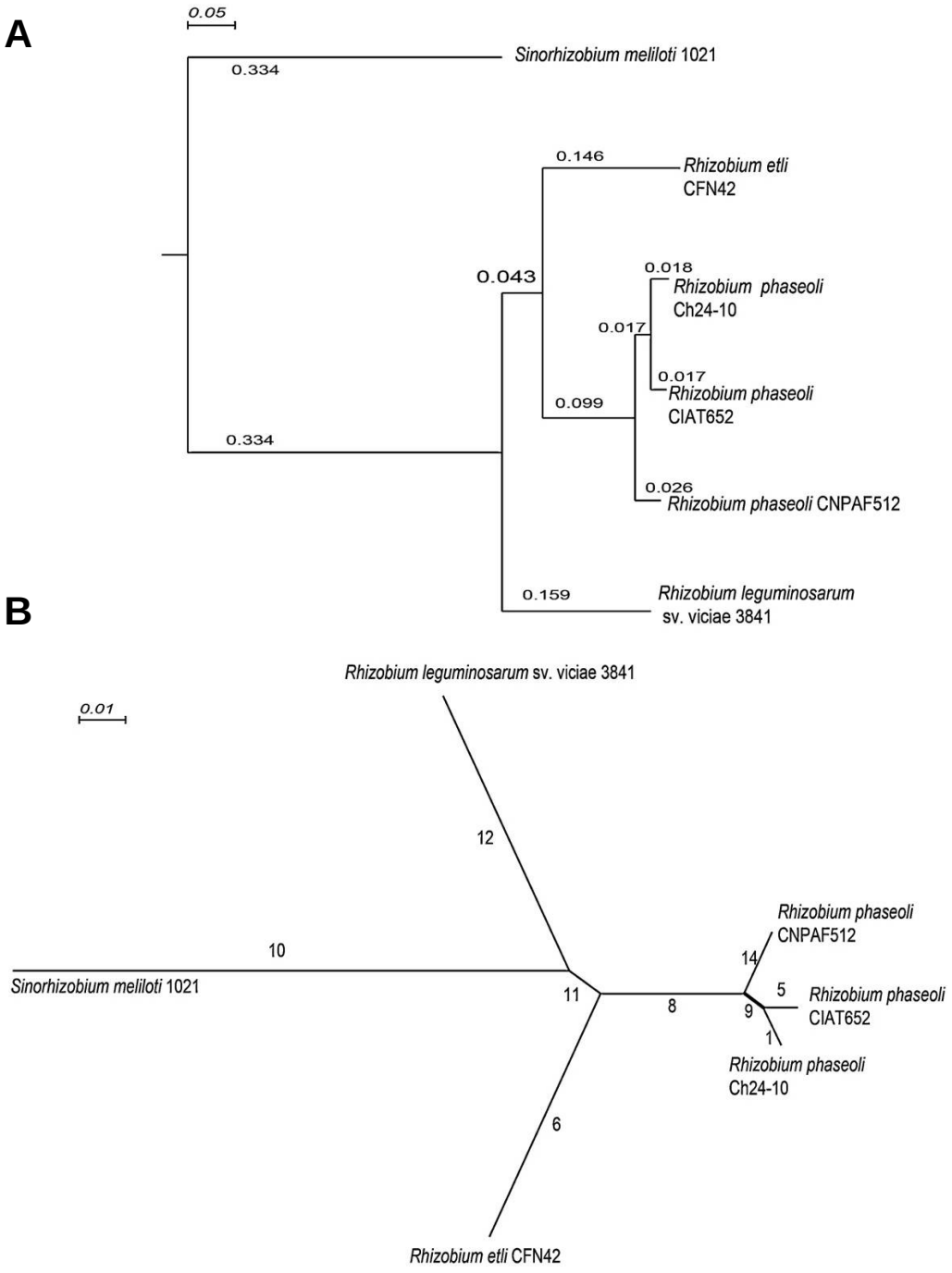
Otro criterio taxonómico que utiliza información de genomas completos o parciales es el *ANI* [55] (*Average Nucleotide Identity*) que indica el porcentaje promedio de identidad en la secuencia de nucleótidos compartida entre dos genomas. Por correlación, un valor de *ANI* del 96% corresponde a un valor del 70% de *reasociación DNA-DNA*, por tanto se utiliza este valor como punto de corte para discriminar entre especies [100].

Se obtuvo este índice para los genomas de las cepas CFN 42, Ch24-10, CIAT652 y CNPAF512. Los resultados se muestran en la Tabla 7. Se puede ver que cuando se utiliza como referencia el genoma de la cepa CFN 42, los valores de *ANI* son todos menores a 91%. Ahora bien, cuando se utiliza el genoma de la cepa CIAT 652 como referencia, los valores que se obtienen respecto a la cepa Ch24-10, y CNPAF512 son mayores al 96% (=punto de corte).

A partir de todas las regiones compartidas entre los genomas de las cepas Ch24-10, CIAT652, CNPAF512 y CFN 42 se obtuvieron dos reconstrucciones filogenéticas utilizando estrategias diferentes. Por un lado, mediante *máxima verosimilitud* se obtuvo un árbol enraizado con *S. meliloti*. Por otro lado, mediante *network joining* se construyó una red filogenética sin enraizar (Figura 6). Ambas reconstrucciones reflejan la historia evolutiva de la cepa Ch24-10.

Estos datos confirman los resultados de *reasociación DNA-DNA*, la hibridación *in silico* y la filogenia que se obtiene con tres marcadores moleculares (*atpD, recA y rpoB*)*: la* cepa Ch24-10, lo mismo que CIAT 652 y CNPAF 512, pertenece a la especie *R. phaseoli*. Se refuerza la propuesta de Richter y Rosselló-Mora de que la cepa CIAT 652 no pertenece a la especie *R. etli* [55] si no a *R. phaseoli* como propusieron adelante Robledo y colaboradores (utilizando los marcadores moleculares *recA, atpD* y *celC*)[57].

Si bien, la información genómica permite realizar inferencias taxonómicas con un grado de confianza alto, cuando no se cuenta con ésta, el uso de marcadores moleculares, en este caso permitió obtener la misma topología en los árboles filogenéticos y obtener el mismo resultado.




## Figura 6. Árboles filogenéticos a partir de las regiones conservadas entre los genomas de diferentes especies y cepas de rizobios.

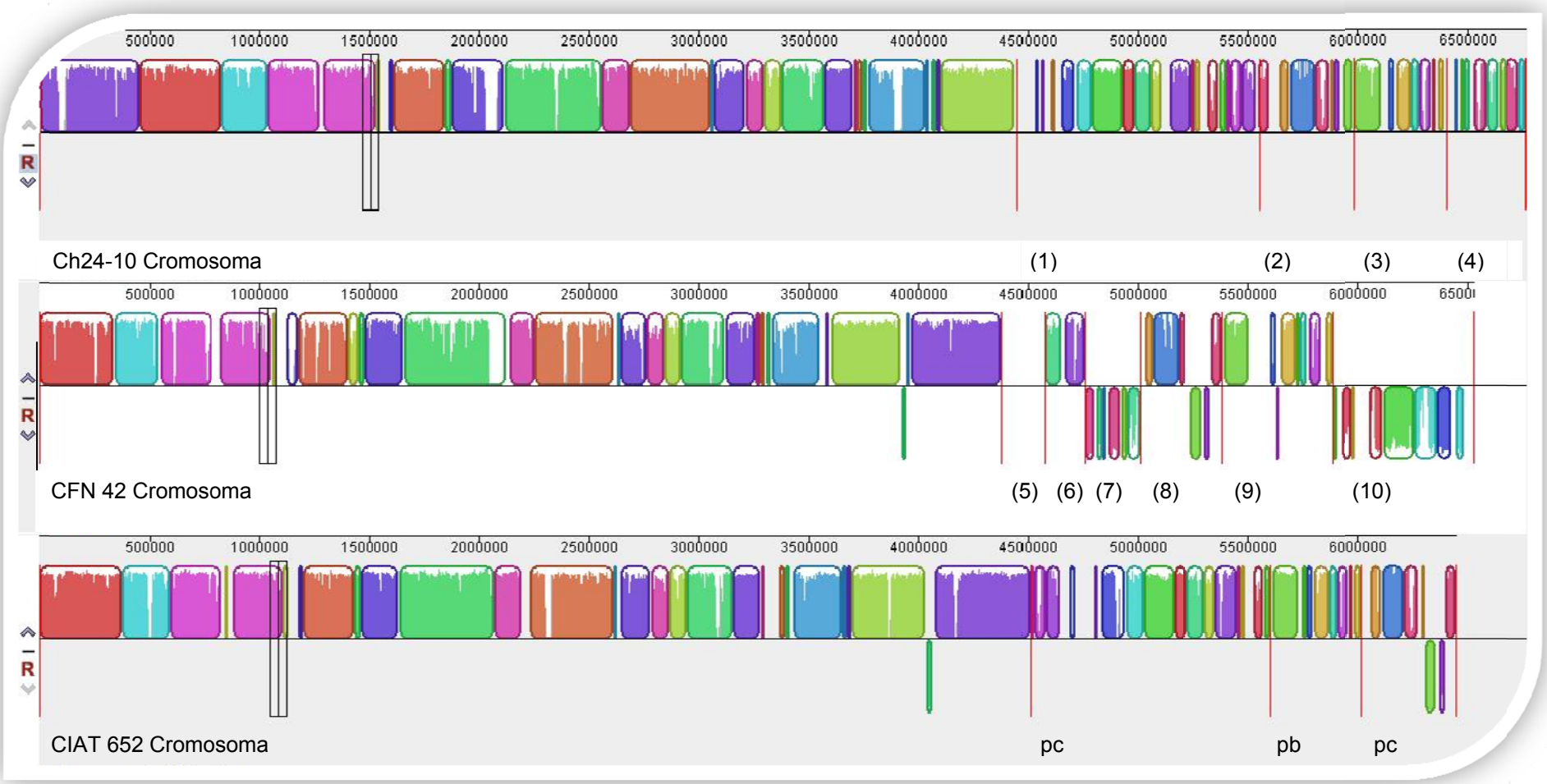
Reconstrucción filogenética a partir de las regiones compartidas concatenadas de los genomas completos de diferentes cepas y especies de rizobios, obtenidas con el programa *Mugsy* [89] a partir de 1373 fragmentos concatenados (codificantes y no codificantes) con un promedio de 2.5 Kb de longitud (3,441,383 bases). **A**. Árbol filogenético enraizado con el genoma de *S. meliloti* 1021 construido mediante *máxima verosimilitud* con el programa *RaxML* [91]. La barra indica 5 sustituciones por cada 100 bases. Los números sobre las ramas indican las sustituciones esperadas por sitio. **B** Reconstrucción filogenética sin enraizar por *network joining* con el programa *Splits tree 4* [92]. La barra indica 1 sustitución por cada 100 bases. Los números sobre las ramas indican las biparticiones (incongruencias) respecto al consenso.

Se obtuvieron alineamientos globales de los genomas de la cepa Ch24-10 en comparación con las cepas CFN 42 y CIAT 652 (Figuras 7, 8 y 9). Se obtuvo un alineamiento global entre los genomas de las tres cepas (Figura 7) y entre las cepas Ch24-10-CFN42 (Figura 8.A) y entre las cepas Ch24-10 CIAT652 (Figura 8.B). En la Figura 9 se puede ver un alineamiento entre los replicones extracromosomales.

Cuando se comparan los genomas de las tres cepas se ve que existen bloques de sintenia con identidad alta a nivel de nucleótidos (que se ve reflejado en los índices *ANI* que se muestran previamente) (Figura 6). Cuando se comparan sólo las cepas Ch24-10 y CIAT 652 la longitud de las regiones conservadas es mayor, respecto a lo que se observa cuando se comparan sólo las cepas Ch24-10 y CFN42 (Figura 8). Este resultado es un reflejo de la cercanía filogenética entre las cepas Ch24-10 y CIAT 652 y de su divergencia respecto a la cepa CFN 42.

Al obtener las regiones sinténicas de los elementos extracromosomales únicamente (Figura 9), se ve que los plásmidos de la cepa Ch24-10 son prácticamente iguales a los plásmidos de la cepa CIAT 652, excepto por el plásmido más pequeño, pCh24-10a, que no tiene secuencias compartidas en ningún replicón de ésta (Figura 9.A), sin embargo, sí encontramos secuencias homólogas en la cepa CFN 42, en su mayoría en el plásmido p42C y también los plásmidos p42E y p42F (Figura 9.B). Este resultado explica podría explicar los valores de hibridación *DNA-DNA* respecto a la cepa CFN 42.

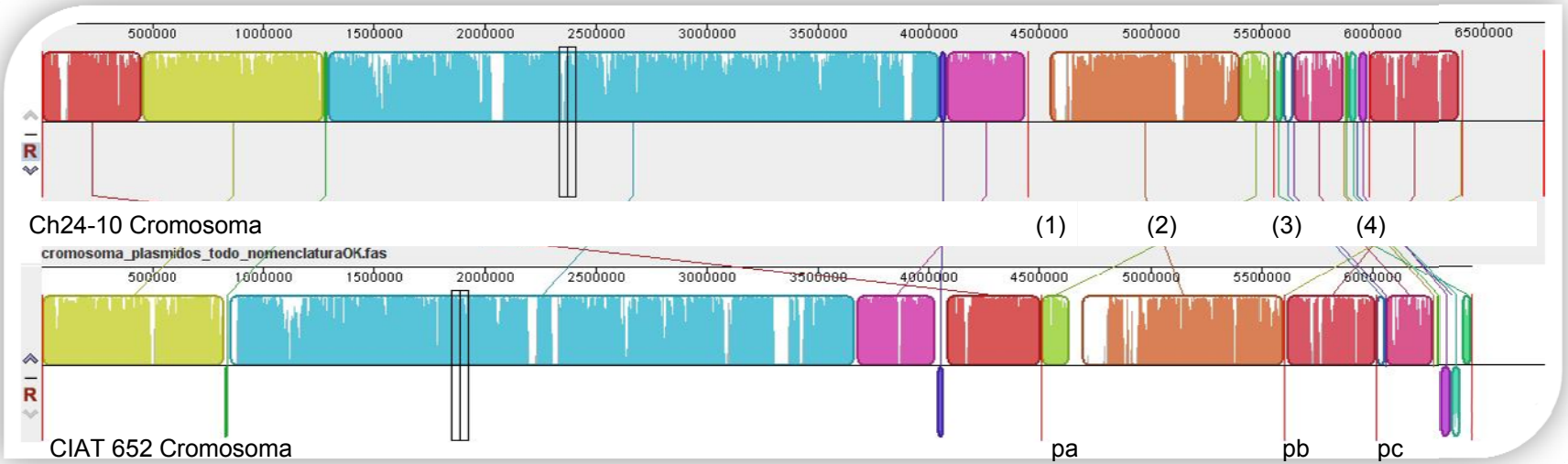
Adicionalmente, se ve que los plásmidos simbióticos, pCh24-10c, p42d y pb de las cepas Ch24-10, CFN 42 y CIAT 652 son muy conservados, aunque son más parecidos entre las cepas Ch24-10 y CIAT652. Desde el punto de vista de la asociación simbiótica esto tiene sentido pues las tres cepas son capaces de asociarse con el frijol y pertenecen al *biovar phaseoli*.




## Figura 7. Alineamiento global entre los genomas de las cepas Ch24-10, CIAT 652 y CFN 42.

Alineamientos globales entre los genomas concatenados de las cepas Ch24-10, CFN 42 y CIAT 652 generados con el programa MAUVE [80]. Los cuadros en color indican bloques de sintenia entre los genomas. Dentro de cada cuadro está graficado el porcentaje de identidad de las secuencias. Los huecos indican zonas donde no existen regiones compartidas. Las líneas rojas marcan el inicio y final de cada replicón. (1) pCh24-10d, (2) pCh24-10c, (3) pCh24-10b, (4) pCh24-10a, 5) p42a, (6) p42b, (7) p42c, (8) p42d, (9) p42e, (10) p42f.

A




B

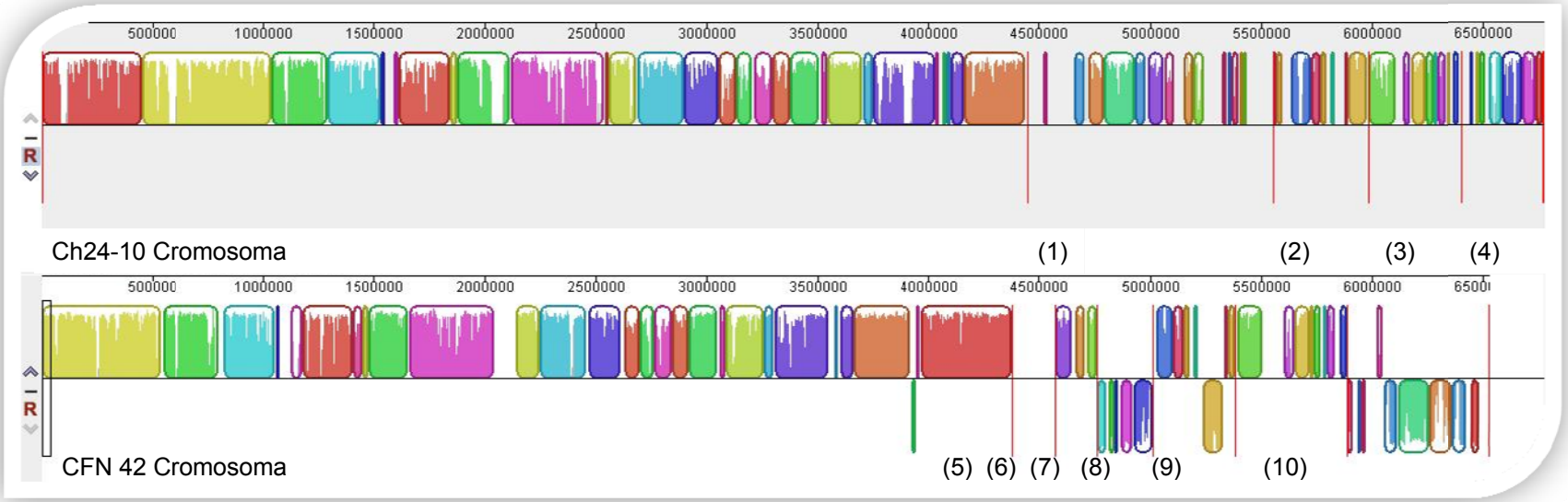



**Figura 8. Alineamientos globales entre los genomas de las cepas Ch24-10 / CIAT 652 y Ch24-10 / CFN 42**

Alineamientos globales entre los genomas concatenados de las cepas Ch24-10, CFN 42 y CIAT 652 generados con el programa MAUVE [80]. Los cuadros en color indican bloques de sintenia entre los genomas Dentro de cada cuadro está graficado el porcentaje de identidad de las secuencias. Los huecos indican zonas donde no existen regiones compartidas. Las líneas rojas marcan el inicio y final de cada replicón. **A.** Alineamiento entre los genomas de las cepas Ch24-10 y CIAT 652. **B.** Alineamiento entre los genomas de las cepas Ch24-10 y CFN 42. (1) pCh24-10d, (2) pCh24-10c, (3) pCh24-10b, (4) pCh24-10a, (5) p42a, (6) p42b, (7) p42c, (8) p42d, (9) p42e, (10) p42f.

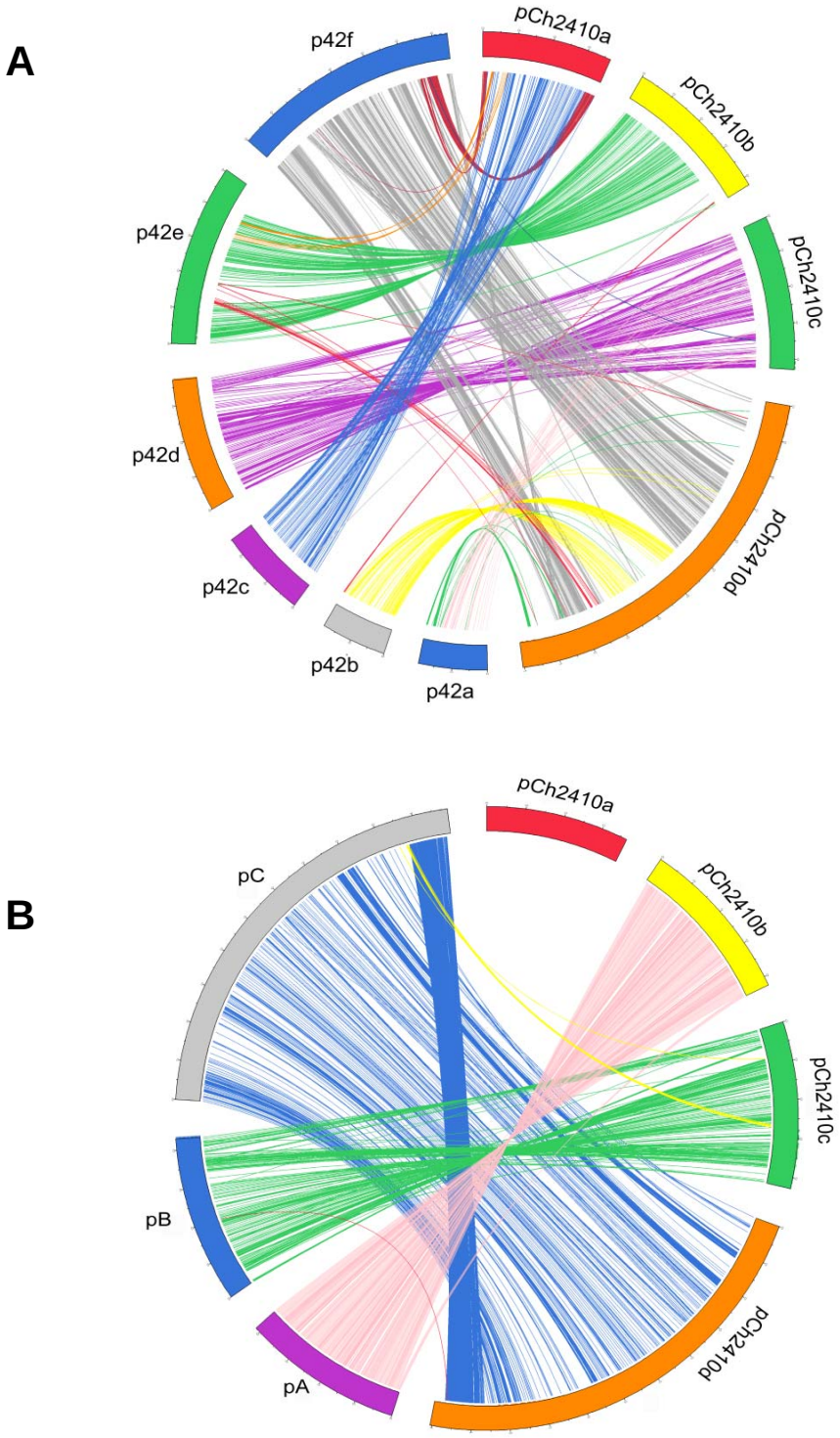


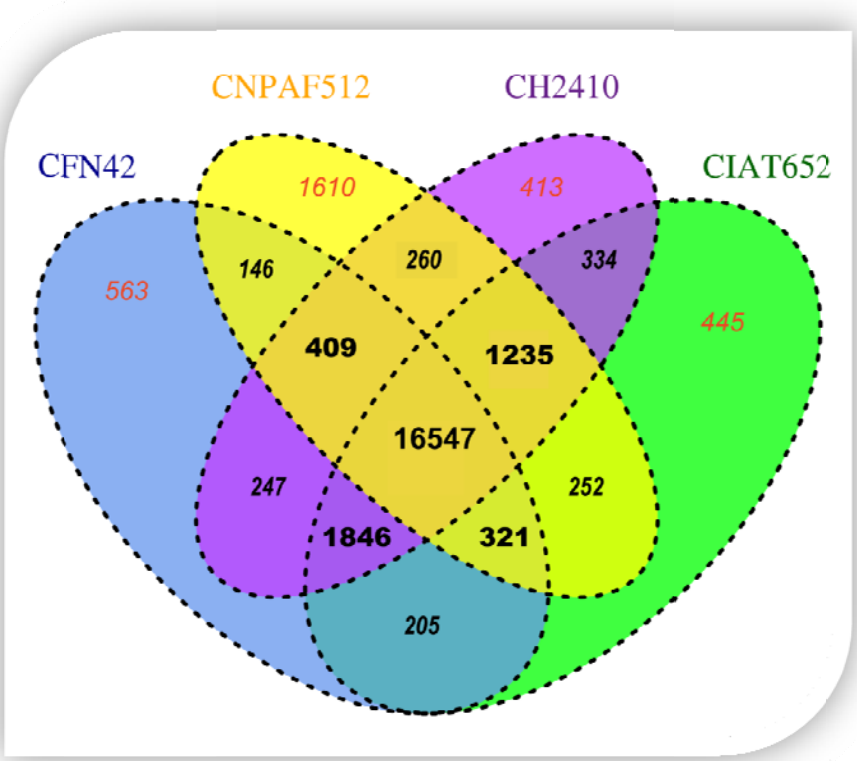

**Figura 9. Comparación de las regiones homólogas entre los plásmidos de las cepas Ch24-10, CFN 42 y CIAT 652.**

Las regiones homólogas se obtuvieron con *Satsuma* [103] A. Comparación de los replicones extracromosomales de las cepas Ch24-10 y CFN 42. B. Comparación de los replicones extracromosomales de las cepas Ch24-10 y CIAT 652. Las líneas indican bloques sinténicos compartidos entre replicones.

En los alineamientos globales entre los genomas de las cepas Ch24-10, CFN 42 y CIAT 652 se consideran todas las regiones, codificantes y no codificantes, por ello se obtuvieron las regiones codificantes compartidas, considerando también a la cepa CNPAF512. En la Figura 10 se puede ver el número de proteínas compartidas entre las cuatro cepas. El número de proteínas que comparte la cepa Ch24-10 con la cepa CIAT 652 (334) es mayor que respecto la cepas CNPAF512 (260) y CFN42 (247). A su vez el número de proteínas ortólogas que comparte la cepa CIAT 652 con la cepa CFN42 es menor (205) que respecto a la cepa CNPAF512 (252). El número menor de proteínas compartidas lo encontramos entre las cepas CNPAF512 y CFN 42 (146).

En general, la cepa Ch24-10 comparte un mayor número de genes ortólogos con las cepas CIAT652 y CNPAF 512 que con a la cepa CFN42 y entre las tres primeras cepas comparten un mayor número de genes codificantes respecto a ésta última.

A




B

| Cepa | Número de genes ortólogos compartidos entre cepas | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Ch24-10 | CIAT 652 | CNPAF512 | CFN 42 | Ch24-10+ CIAT652 | CIAT 652 + CNPAF512 | CNPAF512 + Ch24-10 | Ch24-10 + CIAT6 652 + CNPAF512 |
| Ch24-10 | 413 | | | | | | | |
| CIAT 652 | 334 | 445 | | | | | | |
| CNPAF512 | 260 | 252 | 1610 | | 1235 | | | |
| CFN 42 | 247 | 205 | 146 | 563 | 1846 | 321 | 409 | 16547 |

## Figura 10. Proteínas ortólogas compartidas entre las cepas Ch24-10, CIAT652, CNPAF512 y CFN42.

Las regiones codificantes compartidas entre cada par de genomas se obtuvo con el programa *BLASTp* [81] y se agruparon con el programa *SCPS* [82] **A**. Diagrama de Venn construido con el programa [83] que muestra los números de proteínas ortólogas compartidas entre cada grupo de genomas. **B**. Tabla que muestra el número de regiones codificantes compartidas entre grupos de genomas.

## 5.4 Discusión y perspectivas

### Ensamble del genoma.

En ensamble del genoma de la cepa Ch24-10 se generó mediante una estrategia que combinó los resultados obtenidos por dos plataformas de secuenciación y varios programas que utilizan diferentes algoritmos y que tienen eficiencia distinta para manejar distintos datos de entrada. Los ensambles híbridos permitieron conservar sólo aquellos resultados que fueran repetibles y reproducibles. La estrategia que se sigue para ensamblar un genoma es diferente en cada caso, por ello se evaluaron programas y parámetros distintos (se muestran sólo los resultados de los mejores ensambles).

Se decidió utilizar como referencia el ensamble generado por *Newbler* partir de los datos *Roche-454 Life Science* como referencia, pues resultó ser el de menor número de *contigs* contenidos en un menor número de *andamios*, es decir, un mayor número de *contigs* ordenados y separados por distancias conocidas, lo que permite tener certeza de que las lecturas se ensamblen correctamente sin generar errores o quimeras y permite resolver regiones con secuencias repetidas.

Las *pseudomoléculas* que se obtuvieron aún tienen huecos y es necesario unir los extremos de estas, por lo que se implementará una estrategia de amplificación mediante *PCR*, con oligonucleótidos específicos para obtener las secuencias faltantes.

### Taxonomía y genómica comparativa.

La taxonomía bacteriana es un campo de estudio en Biología que con el tiempo ha ido adquiriendo más herramientas para hacer inferencias con mayor grado de certeza y que reflejen la historia evolutiva de los organismos estudiados.

La *reasociación DNA-DNA* es una metodología con un error estándar asociado alto, presenta limitaciones metodológicas que la hacen difícil de aplicar y los resultados pueden diferir dependiendo del grupo de trabajo que realice los experimentos [100]. Por ejemplo, el valor de *reasociación DNA-*

*DNA* que se obtuvo en este trabajo entre las cepas *R. phaseoli* ATCC 14482 y *R. etli* CFN 42 es de 69% ± 4 usando a la primera como referencia y de 70% ± 2 utilizando a la segunda. Ramírez-Bahena y colaboradores [35] al hacer la revisión taxonómica *R. phaseoli* reportan un valor de 52%. La variación es grande y las inferencias que se pueden hacer al respecto distintas. En el caso de los resultados de Ramírez-Bahena y colaboradores, la conclusión sería que las cepas no pertenecen a la misma especie.

Los valores de *reasociación DNA-DNA* en este trabajo son cercanos o iguales al punto de corte, y se validaron con los datos de hibridación *in silico*, así como los índices *ANI*.

En el caso de no contar con información genómica es importante considerar la elección de marcadores moleculares adecuados para hacer reconstrucciones filogenéticas. Como se ha discutido las filogenias generadas utilizando la secuencia del gen *rRNA16S*, tienen resolución a nivel de género pero no de especie, sin embargo, esta información permite conocer el grupo taxonómico del organismo lo que a su vez permitirá elegir la estrategia adecuada para resolver su filogenia. En el caso de la cepa Ch24-10 la filogenia generada con las secuencias parciales concatenadas de los genes *atpd, recA y ropB* es congruente con la que se obtuvo al analizar todas las regiones compartidas entre genomas de cepas y especies cercanas.

Si bien se recomienda utilizar caracteres fenotípicos en la clasificación taxonómica, aún queda por explorar este aspecto. Se sabe que las cepas Ch24-10 y CIAT 652 son capaces de colonizar al maíz y se ha encontrado evidencia, que necesita corroborarse, de que ésta última es capaz también de promover el crecimiento del maíz (no se muestran los datos). También es sabido que las cepas Ch24-10, CIAT 652 y CNPAF512 tienen mayores tasas de fijación de nitrógeno que la cepa CFN 42.

# 6. Resultados. Interacción de *Rhizobium phaseoli* Ch24-10 con el maíz y el frijol.

## *6.1* Transcriptoma de la cepa Ch24-10 sobre el rizoplano de maíz y de frijol obtenido mediante la tecnología de *RNA-seq.*

Para conocer el perfil de expresión de la cepa Ch24-10 en interacción con el maíz y con el frijol, se realizaron experimentos de *RNA-seq*. Se obtuvo la secuencia del transcriptoma de las bacterias creciendo en el rizoplano de ambas plantas por separado con la tecnología *Illumina*. Se generaron 4 bibliotecas, 2 experimentos independientes por cada planta, MR1-RNA y MR2-RNA para maíz; FR1-RNA y FR2-RNA para frijol.

Las lecturas de cada experimento se mapearon contra los genes anotados y predichos de los pseudoreplicones de la cepa Ch24-10 (Figura 11). En la Tabla 7 se muestra el número de lecturas que se utilizaron para el mapeo, así como el total de lecturas alineadas.

Se obtuvo la frecuencia de mapeo de lecturas por gen y se ajustó a una distribución estadística *Gamma* de acuerdo a lo que se ha observado en transcriptomas generados por *RNA-seq* en *Escherichia coli* en un sistema -*una sola célula*- [104]. Se consideró que los genes altamente expresados corresponden a aquellos cuya distribución se encontró a la derecha de la curva después del punto de inflexión de ésta y que cumplieran esta condición en las dos repeticiones de cada experimento. Los genes altamente expresados corresponden a aquellos en los que alinean las lecturas más representadas en las muestras, considerando la distribución de todas las lecturas en todos los genes en los que se mapearon lecturas.

Se graficó para cada gen de cada *pseudoreplicón* generado en el ensamble del genoma de la cepa Ch24-10 el número total de lecturas mapeadas, estos resultados se ven en la Figura 11. Se puede observar que el número de lecturas que alinearon con el pCh24-10a fue cero para la mayoría de los genes predichos. Revisando a detalle las lecturas que sí alinearon en este replicón se vio que

corresponden a genes sin anotación (no presentaron identidad significativa cuando se realizó la anotación del genoma), proteínas hipotéticas y transportadores ABC, elementos abundantes en el genoma de la cepa Ch24-10. Estos resultados permiten deducir que la cepa con la que se realizaron los experimentos de *RNA-seq* es una cepa curada del plásmido pequeño, pCh24-10a, lo que fue verificado mediante electroforesis de campos pulsados (Figura 3).

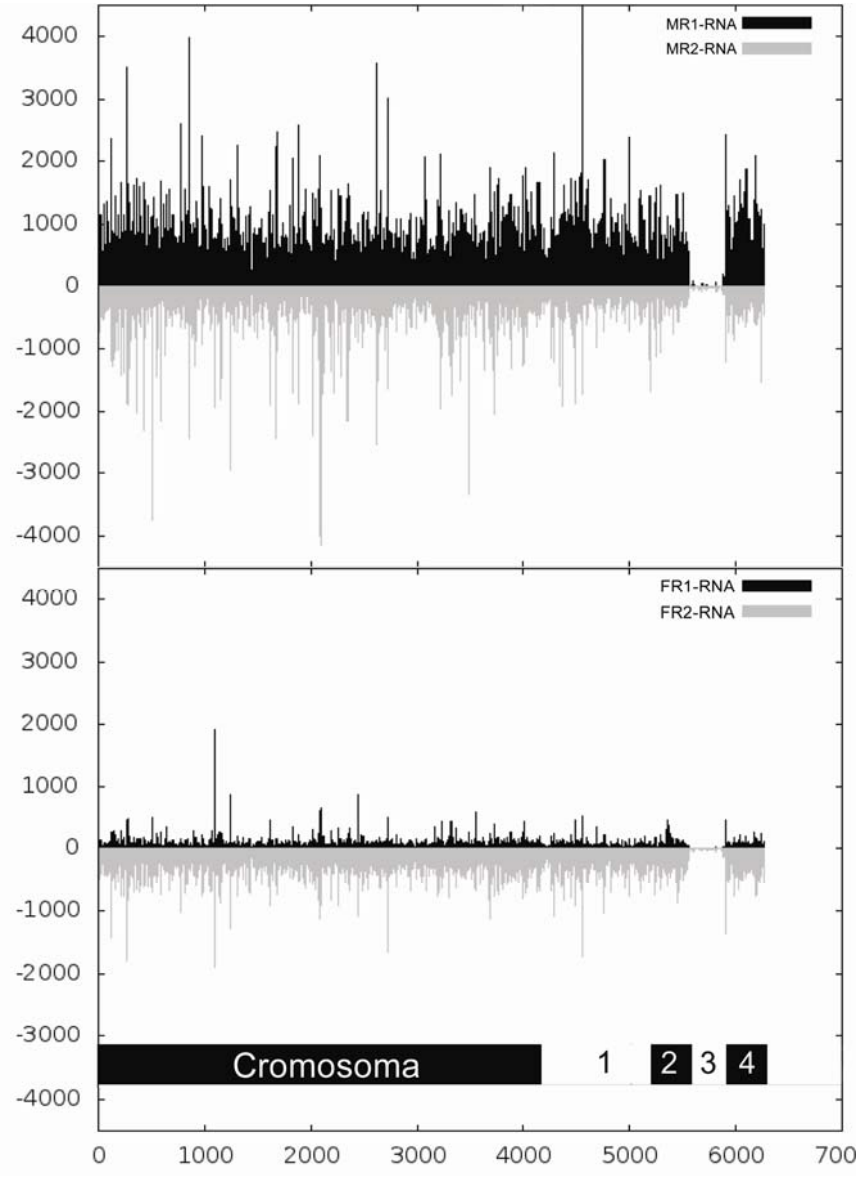



**Figura 11. Mapeo de las lecturas obtenidas por *RNA-seq* contra los replicones de la cepa Ch24-10.**

Las lecturas del transcriptoma obtenidas por *Illumina* se mapearon contra los pseudoreplicones de la cepa Ch24-10 con el programa PALMapper [98]. Se muestra de en el eje de las abscisas 1 = pCh24-10d, 2 = pCh24-10c, 3 = pCh24-10b, 3 = pCh24-10a, 4 = pCh24-10b. En el eje de las ordenadas se muestra el número de lecturas mapeadas para cada experimento, MR1-RNA y MR2-RNA para maíz y FR1-RNA y FR2-RNA para frijol.

**Tabla 8. Lecturas (*RNA-seq*) utilizadas y mapeadas en el genoma de la cepa Ch24-10.**

| Experimento | Número de Lecturas | Lecturas no mapeadas / (%) |
|---|---|---|
| MR1-RNA | 23,825,247 | 15, 169,936 / (63.67%) |
| MR2-RNA | 32,161,264 | 18, 189,677 / (56.55%) |
| FR1-RNA | 25,814,778 | 4, 564,454 / (17.68%) |
| FR2-RNA | 22,178,199 | 5, 632,905 / (25.39%) |

A partir de los resultados de la distribución, los datos se organizaron en tres grupos:

1) Genes altamente expresados en rizoplano de maíz y de frijol (Tabla 9.Anexo I).

2) Genes altamente expresados sólo en la rizósfera de maíz (Tabla 10, Anexo I).

3) Genes altamente expresados sólo en el rizoplano de frijol (Tabla 11, Anexo I)

El grupo 1 está constituido por 598 genes: 441 cromosomales, 82 en plásmido pCh24-10d, 29 en el plásmido pCh24-10c y 46 en el plásmido pCh24-10b. El 11% de estos genes corresponde a 67 proteínas hipotéticas, 40 en el cromosoma, 13 en pCh24-10d, 11 en el pCh24-10c y 3 en el pCh24-10b.

En el grupo 2, genes altamente expresados en maíz y no en frijol, se encuentran 49 genes, 37 de ellos son cromosomales, 6 se ubican en el pCh24-10d, 3 en el pCh24-10c y 3 en el pCh24-10b. En el grupo 3, genes altamente expresados en frijol y no en maíz, se encuentran 52 genes, 33 cromosomales, 10 localizados en el pCh24-10d, 5 en el pCh24-10c y 4 en el pCh24-10b.

El número de genes altamente expresados en el rizoplano de maíz y de frijol es notablemente mayor al de los genes mayormente expresados en una sola planta (598 contra 37 y 33 en maíz y frijol, respectivamente). Cabe resaltar que analizar los datos detalladamente, había genes que se consideraron como no altamente expresados y presentan valores cercanos al punto de corte o bien,

entraban en la categoría de altamente expresados en una repetición del experimento y no en la otra, por lo que no se tomaron en cuenta.

Para conocer las categorías funcionales en las que se agruparon estos genes, se obtuvieron las secuencias generadas a partir del ensamble del transcriptoma que se obtuvo con el programa *OASES*. Fueron anotadas y se generaron mapas metabólicos utilizando el sistema Anotación rápida utilizando subsistemas (*Rapid Annotation using subsystem technology*, *RAST*) [97] (Figura 12).

Debido a que la anotación del *pseudoreplicones* de la cepa Ch24-10 se hizo de manera automática, la anotación de los genes que se identificaron como altamente expresados se corrigió de manualmente con los resultados obtenidos del *RAST* en caso de que fuera diferente. La anotación generada en el *RAST* es altamente confiable debido a que se hace la comparación contra bases de datos curadas manualmente y se identifican dominios conservados, además de la identidad de las secuencias.

Entre los genes altamente expresados se encuentran algunos involucrados en quimiotáxis y en síntesis de flagelo, lo que permite suponer que las células se encuentran en movimiento sobre la superficie de la raíz, o bien, de la raíz al medio circundante y viceversa.

La presencia de genes involucrados en el metabolismo de *DNA*, síntesis de membrana celular (cardiolipina), así como del procesamiento de *RNA* y transcripción activa, la síntesis de proteínas (y degradación), sugiere que las células se encuentran metabólicamente activas y en crecimiento.

En este grupo se encontró una secuencia cromosomal con identidad del 99% (e = 0.0) a la proteína *RelA/SpoT* de *R. etli* CFN 42. Esta enzima es responsable de la síntesis y degradación de guanosina pentafosfato y guanosina tetrafosfato, llamadas en conjunto alarmonas, *(p)ppGpp*, moléculas involucradas en la *respuesta astringente*, mecanismo que permite a las bacterias adaptarse a la limitación de aminoácidos y de fuentes de carbono en el medio [105].

En *Escherichia coli*, la limitación de aminoácidos ocasiona el aumento de *tRNA*'s sin cargar, lo que activa la síntesis de *(p)ppGpp*. En *R. etli*, la enzima *RelA* está involucrada en la utilización de nitrato y aminoácidos como fuente de nitrógeno[106]. Una mutante en el gen que codifica para esta enzima en la cepa CFN 42 es incapaz de utilizar glutamina como fuente de nitrógeno o de nitrógeno y carbono [107]. En el caso de la interacción de *R. etli* con el frijol, también se mostró que *relA* es importante para la síntesis del factor *Nod* [107], y para que se realice la nodulación y fijación de nitrógeno de manera eficiente [106]. En este caso, puede estar jugando un papel en etapas tempranas para el establecimiento de nódulos fijadores de nitrógeno.

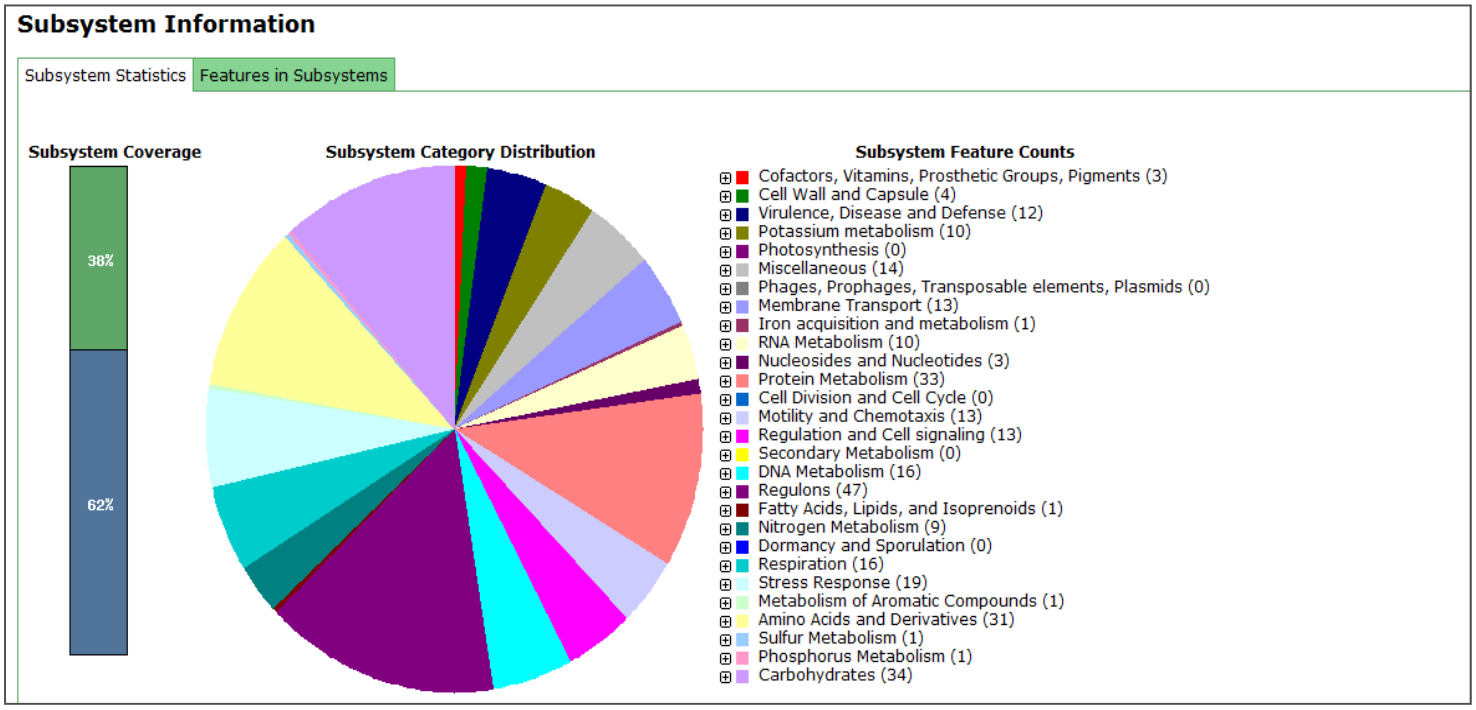



**Figura 12. Genes sobre-expresados en el rizoplano de maíz y de frijol organizados en sistemas de acuerdo al *RAST.***

Se muestran los resultados de la anotación y organización de todos los genes sobre-expresados tanto en rizoplano de maíz como de frijol mediante el *RAST* [97]. El porcentaje de genes que fueron organizados en alguno de los subsistemas corresponde al 38%. El resto de los genes fueron anotados pero no clasificados en algún sistema. La lista completa de los subsistemas y número de genes anotados en cada uno se encuentra en la Tabla 9 del Anexo 1. Los números en paréntesis indican el número de genes anotados en cada subsistema.

Es importante aclarar que el medio hidropónico en el que se realizaron los experimentos no fue suplementado con ninguna fuente de nitrógeno, suponemos que ésta proviene de los recursos de los que dispone la planta en la semilla, así como del metabolismo de la plántula después de germinar. Es posible que al momento de extraer las bacterias del rizoplano de ambas plantas, el medio ya tenía deficiencias en nutrientes. Otra posibilidad es que las bacterias estén fijando nitrógeno atmosférico, aunque en el estudio transcriptómico no se encontró evidencia de esto. Sin embargo, anteriormente se observado que la FBN puede ocurrir a tasas muy bajas en interacción con maíz (Esperanza Martínez-Romero, datos no publicados), desafortunadamente no existen más estudios al respecto, por lo que es un tema que puede ser explorado en detalle.

Los exudados constituyen una fuente de carbohidratos y otros compuestos como aminoácidos. En el caso de maíz, bajo condiciones anóxicas y en bajo potasio se ha visto que el mayor porcentaje de los compuestos en los exudados corresponden a azúcares (glucosa) y ácidos orgánicos (fumárico, oxaloacético, glutámico, aspártico), mientras que los aminoácidos se encuentran en un porcentaje menor al 2%. Cuando el medio se adiciona con nitrato como fuente de nitrógeno, a los 5 días los azúcares se encuentran en baja proporción, mientras que los ácidos orgánicos se encuentran en la misma proporción que en las condiciones de bajo potasio [108].

En un trabajo anterior del grupo de investigación dónde se realizó este proyecto, ya se había mostrado que el gen que codifica para la enzima prolina deshidrogenasa de la cepa Ch24-10 es expresado en exudados de maíz, la mutante por interrupción de este gen no presenta un fenotipo contrastante en la colonización del cereal [64]. La prolina es uno de los aminoácidos exudados en mayor concentración por el maíz y la exudación de esta aumenta en presencia de microorganismos [62]. Es posible que tanto el frijol como el maíz estén exudando aminoácidos al medio y que la cepa Ch24-10 los utilice como fuente de nitrógeno y/o carbono.

Otro indicio del metabolismo bacteriano sobre el rizoplano del maíz lo encontramos al ver expresados genes que son parte del ciclo de los ácidos tricarboxílicos, metabolismo del piruvato (PEP), biosíntesis de trehalosa, utilización de lactosa y xilosa.

Se encontraron genes sobre-expresados, que se agrupan en el subsistema del *RAST* de estrés osmótico. Por un lado, una proteína que participa en la síntesis de glucanos periplásmicos osmoregulados (descrita en *Escherichia coli*), cuya concentración aumenta al incrementar la osmolaridad en el medio. Por otro lado, se encontró sobre-expresado un gen que participa en la síntesis o transporte de glicina-betaína que ha sido descrito como un osmolito en plantas y bacterias que se acumula en el citoplasma en respuesta a estrés osmótico. En *Sinorhizobium meliloti* (*Ensifer meliloti*) a diferencia de lo que ocurre en *E. coli y Bacillus subtilis,* también puede ser utilizada como fuente de carbono y nitrógeno. La glicina-betaína puede ser transportada de manera específica a la célula o sintetizada a partir de colina [109,110]. Estos resultados indican que existe estrés osmótico en el medio de cultivo hidropónico. Se desconoce el papel que tiene la glicina-betaína en *Rhizobium,* pero podría ser similar al que tiene en *S. meliloti*.

También se encontraron genes altamente expresados relacionados con el estrés oxidativo, enzimas que intervienen en la protección sobre especies reactivas de oxígeno (catalasa, peroxidasa). Las plantas secretan estos tipos de especies reactivas en condiciones de estrés biótico o abiótico como parte de la señalización de su ciclo de vida, aunque también se ha observado que son importantes en la interacción *Rhizobium*-leguminosa [111]. O bien, las células se encuentran bajo estrés oxidativo o estas moléculas participan en la interacción planta bacteria, éstas se han descrito en la interacción con leguminosas ¿podrían ser moléculas de una vía común de señalización entre rizobios y leguminosas y no leguminosas?

El trabajo publicado dónde se reportan los datos obtenidos por RNA-seq, se centra en los genes extracromosomales, principalmente en aquellos altamente expresados en ambas condiciones, se estableció un punto de corte más laxo respecto al que se presenta aquí. Entre los genes

altamente expresados en maíz y frijol se encuentran algunos involucrados en el metabolismo de la prolina, toma de hierro, biosíntesis de tiamina y giberelina, un sistema de secreción tipo VI, transportadores para oligopéptidos y azúcares y bombas de extrusión, por mencionar algunos [30]. Además de analizar los genes sobre-expresados en los plásmidos se hace una comparación respecto a los genes no cromosomales que se encontraron sobre-expresados en *R. leguminosarum* bv. *viciae* 3841 cuando crece en la rizósfera de su hospedero, chícharo (*Pisum sativum),* remolacha (no leguminosa), y alfalfa (leguminosa, no hospedero) [60] .

La evidencia muestra que existen genes comunes que se expresan en la rizósfera de varias plantas y otros específicos que sólo se expresan en plantas en particular. Existe un grupo de genes conservados extracromosomales de la cepa Ch24-10 expresados en el rizoplano del maíz y de frijol, y que también expresa *R. leguminosarum* en la rizósfera de trébol y alfalfa y remolacha [30,60]. Llama la atención que tanto en como en *R. phaseoli*, la mayoría de genes altamente expresados son de función desconocida.

## 6.2 Discusión y perspectivas.

Entender cómo funcionan las bacterias en su ambiente y cómo interactúan con él es una de las metas de la microbiología. Las nuevas tecnologías genómicas permiten obtener aproximaciones globales pues se puede estudiar la expresión de genes y presencia de proteínas en condiciones diversas.

En particular, la tecnología denominada *RNA-seq* permite conocer los genes expresados en una condición dada, incluso aquellos genes cuya expresión es baja y sin dependencia del conocimiento previo. En este trabajo sólo se analizaron aquellos genes que se sobre-expresaron, con énfasis en los genes extracromosomales. Como perspectiva queda aún analizar a detalle los genes cromosomales sobre expresados. También es importante conocer de manera global la expresión génica de la cepa Ch24-10 en la interacción con el maíz y el frijol, y no centrarse en los genes con abundancia relativa respecto al resto y aplicar otras estrategias para cuantificar la expresión de los genes.

Como perspectivas, en este sentido se repetirán los experimentos de *RNA-seq* con la cepa que no ha perdido el plásmido pCh24-10a. Cabe aclarar que aunque la cepa perdió el plásmido pCh24-10a no perdió la capacidad de colonizar el rizoplano del maíz. El número de bacterias que se recupera a los 5 es de $1x10^9$ unidades formadoras de colonia por raíz después de ser inoculado; número similar al que se recupera con la cepa que no carece del plásmido pCh24-10a ($1x10^9$ / g de raíz, [51]). Sin embargo, datos preliminares obtenidos en el laboratorio muestran que sí pierde competitividad en la nodulación de frijol (M.A. Rogel, datos no publicados).

Una vez que se tenga el genoma completo y cerrado de la cepa Ch24-10 será posible obtener otro tipo de índices para cuantificar la expresión relativa de los genes expresados, como es el *RPKM* (*Reads Per Kilobase of gene length per Million reads*, lecturas por kilobase de gen por millón de lecturas mapeadas [112]) , un índice normalizado respecto al tamaño del genoma y al número de lecturas totales.

Otra perspectiva es obtener información más específica sobre de los genes expresados en diferentes zonas del rizoplano, se podría explorar el uso de genes reporteros que permitan evaluar su actividad *in situ* en la raíz de las plantas hospederas. Otra alternativa es la construcción de mutantes dirigidas y evaluar su fenotipo en la colonización y promoción del crecimiento del maíz.

Es importante considerar que la expresión de genes en las condiciones de laboratorio y aún en suelo puede ser afectada y limitada por las condiciones en las que se realicen los experimentos. El suelo es un sistema heterogéneo y se ha propuesto que un centímetro de suelo no es igual al centímetro aledaño [113]. En este ambiente hay micronichos y microhábitats y a su vez las raíces son heterogéneas en estructura y en fisiología, así que la rizósfera se considera un ambiente hipervariable. Por ello, la expresión génica bacteriana no será homogénea en la superficie de una raíz. Hay que considerar también la variación individual de planta a planta dentro de una especie vegetal. Sin embargo, el modelo experimental que se utilizó, dado que es un modelo hidropónico, permitiría la difusión de los exudados y esto podría disminuir las diferencias en el medio circundante, haciendo que la expresión de genes por parte de *Rhizobium* sea más parecida en toda la superficie de la raíz.

Es de interés en el grupo de trabajo dónde se realizó esta proyecto analizar la interacción entre rizobios y sus plantas hospederas. *Rhizobium* es un buen modelo de estudio, pues es un organismo competitivo en la colonización de la rizósfera, es capaz de desplazar a otras bacterias en ésta y puede alcanzar una densidad celular alta en maíz de hasta $10^9$ células/g de raíz cuando se inocula [22]. Como un resultado adicional a este proyecto se escribió una capítulo en libro *Microbial Ecology of The Rhizosphere* en el que se hace una breve revisión sobre genes de *Rhizobium* que son expresados en la rizósfera de plantas leguminosas y no leguminosas (Anexo II) [114]

Se añade interés al modelo, considerando que el maíz y el frijol son plantas importantes cultural, agrícola y biológicamente y son plantas hospederas de *R. phaseoli* Ch24-10. La cepa Ch24-10 promueve el crecimiento del maíz, cultivo que requiere ser suplementado con una fuente de

nitrógeno para obtener buenos rendimientos. En diversos estudios a nivel nacional y en Latinoamérica se ha visto que la inoculación de este cereal con bacterias promotoras de crecimiento, como *Azospirillum,* permite la disminución de las dosis de nitrógeno aplicado [11,115].

Analizar el conjunto de genes que se expresan tanto en el rizoplano de maíz como de frijol, (número que es mayor, por mucho, a los genes sobre-expresados de manera individual en cada planta), tendría sentido pensar que el genoma de *R. phaseoli* Ch24-10, al ser huésped de ambas plantas ha sufrido la presión de selección de éstas y que por tanto los genes expresados pueden dar indicios de la adaptación a la colonización en ambientes dónde crecen juntas como la milpa.

# 7. Consideraciones finales.

La información que nos proporciona la secuenciación de los genomas, particularmente el de la cepa Ch24-10 ha permitido proponer una modificación taxonómica de ésta, cambiándola de de *R. etli* a *R. phaseoli*. Los resultados obtenidos de la secuenciación del transcriptoma ha dado un panorama general de cómo se lleva a cabo la interacción de esta bacteria con el frijol y con el maíz en etapas tempranas y permitió identificar genes que son comunes en la vía de comunicación con ambas. Sin embargo, aún queda por explorar en detalle los resultados para entender mejor la información que proporcionan.

El genoma de la cepa C24-10 es un reflejo de su historia evolutiva, de los cambios que ha sufrido a través del tiempo y de la presión de selección a la que ha sido sometida. El genoma refleja cómo han cambiado dos poblaciones de *Rhizobium* a través del tiempo, lo que las ha separado filogenéticamente. El transcriptoma muestra un conjunto de genes que han sido seleccionados bajo la presión de selección en la interacción con las plantas y que le han permitido adaptarse.

Es importante considerar que la investigación realizada puede tener una aplicación práctica, con la perspectiva de que los resultados generados en la UNAM den un beneficio directo e inmediato a quienes la sustentan. La propuesta es trabajar en la elaboración de un inoculante de la cepa Ch24-10, un producto que sea disponible de manera comercial para ser aplicado en los campos agrícolas de maíz. Tenemos evidencia de que al inocular el maíz con esta cepa se incrementa significativamente el rendimiento comparando con tratamientos sin inoculante bacteriano (Esperanza Martínez-Romero, datos no publicados y [51].

De pensarse en la aplicación en agricultura de los resultados de este trabajo, se podría generar un inoculante bacteriano de la cepa Ch24-10. La información generada permitiría generar, por ejemplo, un vehículo que le permita sobrevivir, reproducirse y competir mejor en la rizósfera del maíz, para que los efectos en campo se vean reflejados en el aumento de la productividad del cultivo.

# 8. Referencias bibliográficas.


1. Hinsinger, P., Gobran, G. R., Gregory, P. J. & Wenzel, W. W. Rhizosphere geometry and heterogeneity arising from root-mediated physical and chemical processes. *The New Phytologist* **168**, 293–303 (2005).

2. Barea, J. M., Pozo, M. J., Azcón, R. & Azcón-Aguilar, C. Microbial co-operation in the rhizosphere. *Journal of Experimental Botany* **56**, 1761–1778 (2005).

3. Brencic, A. & Winans, S. C. Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiology and Molecular Biology Reviews* **69**, 155–194 (2005).

4. Singh, B. K., Millard, P., Whiteley, A. S. & Murrell, J. C. Unravelling rhizosphere-microbial interactions: opportunities and limitations. *Trends in Microbiology* **12**, 386–393 (2004).

5. Personeni, E., Nguyen, C., Marchal, P. & Pagès, L. Experimental evaluation of an efflux-influx model of C exudation by individual apical root segments. *Journal of Experimental Botany* **58**, 2091–2099 (2007).

6. Walker, T. S., Bais, H. B., Grotewold, E. & Vivanco, J. M. Root exudation and rhizosphere biology. *Plant Physiology* **132**, 44–51 (2003).

7. Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S. & Vivanco, J. M. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annual Review of Plant Biology* **57**, 233–266 (2006).

8. Saharan, B. S. & Nehra, V. Plant Growth Promoting Rhizobacteria: A critical review. *Life Sciences and Medicine Research* **2011: LSMR**, 1–30 (2011).

9. Compant, S., Duffy, B., Nowak, J. & Cle, C. Use of Plant Growth-Promoting Bacteria for biocontrol of plant diseases: Principles, mechanisms of action, and future prospects. *Applied and Environmental Microbiology* **71**, 4951–4959 (2005).

10. Crews, T. E. & Peoples, M. B. Legume versus fertilizer sources of nitrogen: ecological tradeoffs and human needs. *Agriculture, Ecosystems & Environment* **102**, 279–297 (2004).

11. Aguado-Santacruz G.A. *Introducción al uso y manejo de los biofertilizantes en Agricultura*. (INIFAP/SAGARPA: 2012).

12. Kennedy, I. R., Choudhury, A. T. M. A. & Kecskés, M. L. Non-symbiotic bacterial diazotrophs in crop-farming systems: can their potential for plant growth promotion be better exploited? *Soil Biology and Biochemistry* **36**, 1229–1244 (2004).

13. Kiely, P. D. *et al.* Exploiting new systems-based strategies to elucidate plant-bacterial interactions in the rhizosphere. *Microbial Ecology* **51**, 257–266 (2006).

14. Rosenblueth, M. & Martínez-Romero, E. Bacterial endophytes and their interactions with hosts. *Molecular Plant-Microbe Interactions* **19**, 827–837 (2006).

15. Hufbauer, R. A. & Roderick, G. K. Microevolution in biological control: Mechanisms, patterns, and processes. *Biological Control* **35**, 227–239 (2005).

16. Bloemberg, G. V. & Lugtenberg, B. J. Molecular basis of plant growth promotion and biocontrol by rhizobacteria. *Current opinion in Plant Biology* **4**, 343–50 (2001).

17. Lodwig, E. & Poole, P. Metabolism of *Rhizobium* Bacteroids. *Critical Reviews in Plant Sciences* **22**, 37–78 (2003).

18. Denison, R. F. & Toby Kiers, E. Why are most rhizobia beneficial to their plant hosts, rather than parasitic? *Microbes and Infection* **6**, 1235–1239 (2004).

19. Chen, G., Zhu, H. & Zhang, Y. Soil microbial activities and carbon and nitrogen fixation. *Research in Microbiology* **154**, 393–398 (2003).

20. Martínez-Romero, E. Diversity of *Rhizobium-Phaseolus vulgaris* symbiosis: overview and perspectives. *Plant and Soil* **252**, 11–23 (2003).

21. Valverde, A. *et al. Herbaspirillum lusitanum* sp. nov., a novel nitrogen-fixing bacterium associated with root nodules of *Phaseolus vulgaris*. *International Journal of Systematic and Evolutionary Microbiology* **53**, 1979–1983 (2003).

22. Gutiérrez-Zamora, M. L. & Martínez-Romero, E. Natural endophytic association between *Rhizobium etl*i and maize (*Zea mays* L.). *Journal of Biotechnology* **91**, 117–126 (2001).

23. Yanni, Y. G. *et al.* Natural endophytic association between *Rhizobium leguminosarum bv . trifolii* and rice roots and assessment of its potential to promote rice growth. *Plant and Soil* 99–114 (1997).

24. Noel, T. C., Sheng, C., Yost, C. K., Pharis, R. P. & Hynes, M. F. *Rhizobium leguminosarum* as a plant growth-promoting rhizobacterium: direct growth promotion of canola and lettuce. *Canadian Journal of Microbiology* **42**, 279–283 (1996).

25. López-López, A., Rogel, M. a, Ormeño-Orrillo, E., Martínez-Romero, J. & Martínez-Romero, E. *Phaseolus vulgaris* seed-borne endophytic community with novel bacterial species such as *Rhizobium endophyticum* sp. nov. *Systematic and Applied Microbiology* **33**, 322–327 (2010).

26. Ormeño-Orrillo, E., Hungria, M. & Martinez-Romero, E. Dinitrogen-fixing prokaryotes. *The Prokaryotes Vol. 1: Symbiotic Associations, Biotechnology, Applied Microbiology 4th ed.* (En prensa)

27. Rogel, M. A., Ormeño-Orrillo, E. & Martinez Romero, E. Symbiovars in rhizobia reflect bacterial adaptation to legumes. *Systematic and Applied Microbiology* **34**, 96–104 (2011).

28. Wang, F. *et al. Rhizobium vallis* sp. nov., isolated from nodules of three leguminous species. *International Journal of Systematic and Evolutionary Microbiology* **61**, 2582–2588 (2011).

29. Valverde, A., Igual, J. M., Peix, A., Cervantes, E. & Velázquez, E. *Rhizobium lusitanum* sp. nov. a bacterium that nodulates *Phaseolus vulgaris*. *International Journal of Systematic and Evolutionary Microbiology* **56**, 2631–2637 (2006).

30. López-Guerrero, M. G. *et al.* Rhizobial extrachromosomal replicon variability, stability and expression in natural niches. *Plasmid* **68**, 149–158 (2012).

31. Harrison, P. W., Lower, R. P. J., Kim, N. K. D. & Young, J. P. W. Introducing the bacterial "chromid": not a chromosome, not a plasmid. *Trends in Microbiology* **18**, 141–14–8 (2010).

32. Slater, S. C. *et al.* Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. *Journal of Bacteriology* **191**, 2501–2511 (2009).

33. Young, J. M., Kuykendall, L. D., Martínez-Romero, E., Kerr, A. & Sawada, H. A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R vitis*. *International Journal of Systematic and Evolutionary Microbiology* **51**, 89–103 (2001).

34. Jordan, D. C. & Allen, O. N. Family III. *Rhizobiaceae* Conn 1938. *Bergey's Manual of Determinative Bacteriology* 261–267 (1974).

35. Ramírez-Bahena, M. H. *et al.* Revision of the taxonomic status of the species *Rhizobium leguminosarum* (Frank 1879) Frank 1889, *Rhizobium phaseoli* Dangeard 1926 and *Rhizobium trifolii* Dangeard 1926. *R. trifolii* is a later synonym of *R. leguminosarum*. Reclassification of the strain *R. leguminosarum* DSM 30132 (5NCIMB 11478) as *Rhizobium pisi* sp. nov.. l. *International journal of systematic and Evolutionary Microbiology* **58**, 2484–2490 (2008).

36. Martínez-Romero, E. *et al. Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. trees. *International Journal of Systematic Bacteriology* **41**, 417–426 (1991).

37. Segovia, L., Young, P. W. & Martínez-Romero, E. Reclassification of American *Rhizobium leguminosarum biovar phaseoli* type I strains as *Rhizobium etli* sp. nov. *International journal of systematic bacteriology* **43**, 374–7 (1993).

38. Ribeiro, R. A. *et al.* Reclassification of *Rhizobium tropici* type A strains as *Rhizobium leucaenae* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **62**, 1179–1184 (2012).

39. Graham, P. Common bean (*Phaseolus vulgaris* L.). *Field Crops Research* **53**, 131–146 (1997).

40. Broughton, W. J. *et al.* Beans ( *Phaseolus* spp .) – model food legumes. *Plant and Soil* 55–128 (2003).

41. Bitocchi, E. *et al.* Mesoamerican origin of the common bean ( *Phaseolus vulgaris* L .) is revealed by sequence data. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 788–796 (2011).

42. Beebe, S. S. *et al.* Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. *Crop Science* **40**, 264–273 (2000).

43. Aguilar, O. M., Riva, O. & Peltzer, E. Analysis of *Rhizobium etli* and of its symbiosis with wild *Phaseolus vulgaris* supports coevolution in centers of host diversification. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 13548–53 (2004).

44. Burle, M. L., Fonseca, J. R., Kami, J. a & Gepts, P. Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. *Theoretical and Applied Genetics* **121**, 801–813 (2010).

45. Iltis, H. H. From teosinte to maize: the catastrophic sexual transmutation. *Science* **222**, 886–894 (1983).

46. CIMMYT *La economía del maíz en América Latina*. 4–6 (1998). at <http://apps.cimmyt.org/research/Economics/map/impact_studies/ImpactsMaize66_97/ImpactosLA/pdfs/ImpactosLA_economia.pdf>

47. Vega Valdivia, D. D. & Ramírez Moreno, P. *Situación y perspectivas del maíz en México*. 1–56 (2004).at <http://200.77.231.100/pics/p/p1763/Maiz__270304.pdf>

48. Piperno, D. R. & Flannery, K. V The earliest archaeological maize (*Zea mays* L.) from highland Mexico: new accelerator mass spectrometry dates and their implications. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2101–2103 (2001).

49. Matsuoka, Y. *et al.* A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6080–6084 (2002).

50. Kaplan, L. & Lynch, T. F. *Phaseolus* (*Fabaceae*) in Archaeology: AMPS radioacarbon dates and their significance for pre-colombian agriculture. *Economic Botany* **53**, 261–272 (1999).

51. Rosenblueth, M. & Martínez-Romero, E. *Rhizobium etli* maize populations and their competitiveness for root colonization. *Archives of microbiology* **181**, 337–44 (2004).

52. López-Guerrero, M. G. & Martinez-Romero, E. Análisis de una Genoteca construida a partir de Hibridación Sustractiva entre las cepas Ch24-10 y CFN 42 de *Rhizobium etli*. (2006).Facultad de Ciencias, Universidad Nacional Autónoma de México.

53. González, V. *et al.* The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3834–9 (2006).

54. Diatchenko, L. *et al.* Suppresion Subtractive Hybridization. A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6025–6030 (1996).

55. Richter, M. & Rosselló-Mora, R. Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 19126–19131 (2009).

56. González, V. *et al.* Conserved symbiotic plasmid DNA sequences in the multireplicon pangenomic structure of *Rhizobium etli*. *Applied and environmental microbiology* **76**, 1604–14 (2010).

57. Robledo, M. *et al.* The celC gene, a new phylogenetic marker useful for taxonomic studies in *Rhizobium. Systematic and Applied Microbiology* **34**, 393–399 (2011).

58. Barr, M., East, A. K., Leonard, M., Mauchline, T. H. & Poole, P. S. In vivo expression technology (IVET) selection of genes of *Rhizobium leguminosarum biovar viciae* A34 expressed in the rhizosphere. *FEMS Microbiology Letters* **282**, 219–227 (2008).

59. Karunakaran, R. *et al.* Transcriptomic analysis of Rhizobium leguminosarum biovar viciae in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. *Journal of Bacteriology* **191**, 4002–4014 (2009).

60. Ramachandran, V. K., East, A. K., Karunakaran, R., Downie, J. A. & Poole, P. S. Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. *Genome biology* **12**, R106 (2011).

61. Vilchis, A. & Rosenblueth, M. Mutantes de *Rhizobium etli* que expresan β-glucoronidasa en presencia de exudados de maíz. (2004).Facultad de Ciencias Biológicas, Universidad Autónoma del Estado de Morelos.

62. Phillips, D. A., Fox, T. C., King, M. D., Bhuvaneswari, T. & Teuber, L. R. Microbial products trigger amino acid exudation from plant roots. *Plant Physiology* **136**, 2887–2894 (2004).

63. Jiménez-Zurdo, J. I., García-Rodríguez, F. M. & Toro, N. The *Rhizobium meliloti* putA gene: its role in the establishment of the symbiotic interaction with alfalfa. *Molecular Microbiology* **23**, 85–93 (1997).

64. Escalante, V. & Rosenblueth, M. Obtención de mutantes que *Rhizobium etli* (Ch24-10) deficientes para utilizar la prolina y análisis de su fenotipo para colonizar plantas de maíz. (2006). Facultad de Ciencias Biológicas, Universidad Autónoma del Estado de Morelos.

65. Ozsolak, F. & Milos, P. M. RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics* **12**, 87–98 (2010).

66. Quinto, C. *et al.* Reiteration of nitrogen fixation gene sequences in *Rhizobium phaseoli*. *Nature* **299**, 724–726 (1982).

67. Uribe, L. Evaluación de medios para la selección de cepas de *Rhizobium leguminosarum bv. phaseoli* tolerantes a baja concentración de fosfato en medio de cultivo. *Agronomia Costarricense* **17**, 103–109 (1993).

68. Fauvart, M., Sánchez-Rodríguez, A., Beullens, S., Marchal, K. & Michiels, J. Genome sequence of *Rhizobium etli* CNPAF512, a nitrogen-fixing symbiont isolated from bean root nodules in Brazil. *Journal of Bacteriology* **193**, 3158–3159 (2011).

69. Michiels, J., Pelemans, H., Vlassak, K., Verreth, C. & Vanderleyden, J. Identification and characterization of a *Rhizobium leguminosarum bv. phaseoli* gene that is important for nodulation competitiveness and shows structural homology to a *Rhizobium fredii* host-inducible gene. *MPMI-Molecular Plant Microbe Interactions* **8**, 468–472 (1995).

70. American Type Culture Collection. at <www.atcc.org>

71. Frahm, M. a. *et al.* Breeding beans for resistance to terminal drought in the Lowland tropics. *Euphytica* **136**, 223–232 (2004).

72. Zerbino, D. R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* **18**, 821–829 (2008).

73. Li, R. *et al.* De novo assembly of human genomes with massively parallel short read sequencing. *Genome Research* **20**, 265–272 (2010).

74. Pop, M., Phillippy, A., Delcher, A. L. & Salzberg, S. L. Comparative genome assembly. *Briefings in Bioinformatics* **5**, 237–248 (2004).

75. Treangen, T. J., Sommer, D. D., Angly, F. E., Koren, S. & Pop, M. Next generation sequence assembly with AMOS. *Current Protocols in Bioinformatics* **Chapter 11**, Unit 11.8 (2011).

76. Assefa, S., Keane, T. M., Otto, T. D., Newbold, C. & Berriman, M. ABACAS: algorithm-based automatic contiguation of assembled sequences. *Bioinformatics* **25**, 1968–1969 (2009).

77. Kislyuk, A. O. *et al.* A computational genomics pipeline for prokaryotic sequencing projects. *Bioinformatics* **26**, 1819–1826 (2010).

78. Eckhardt, T. A rapid method for the identification of plasmid desoxyribonucleic acid in bacteria. *Plasmid* **1**, 584–588 (1978).

79. Hynes, M. F. & McGregor, N. F. Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. *Molecular Microbiology* **4**, 567–574 (1990).

80. Darling, A. E., Mau, B. & Perna, N. T. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PloS one* **5**, e11147 (2010).

81. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402 (1997).

82. Nepusz, T., Sasidharan, R. & Paccanaro, A. SCPS: a fast implementation of a spectral method for detecting protein families on a genome-wide scale. *BMC bioinformatics* **11**, 1–13 (2010).

83. Chen, H. & Boutros, P. C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC bioinformatics* **12**, 1–7 (2011).

84. Weisburg, W. G., Barns, S. M., Pelletier, D. a & Lane, D. J. 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology* **173**, 697–703 (1991).

85. Vinuesa, P. *et al.* Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from *rrs, atpD, recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetlands. *Systematic and Applied Microbiology* **28**, 702–716 (2005).

86. Lloret, L. *et al. Ensifer mexicanus* sp. nov. a new species nodulating *Acacia angustissima* (Mill.) Kuntze in Mexico. *Systematic and Applied Microbiology* **30**, 280–290 (2007).

87. Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 95–98 (1999).

88. Tamura, K. *et al.* MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**, 2731–2739 (2011).

89. Angiuoli, S. V & Salzberg, S. L. Mugsy: fast multiple alignment of closely related whole genomes. *Bioinformatics* **27**, 334–342 (2011).

90. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).

91. Stamatakis, a, Ludwig, T. & Meier, H. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21**, 456–463 (2005).

92. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**, 254–267 (2006).

93. Kurtz, S. *et al.* Versatile and open software for comparing large genomes. *Genome Biology* **5**, R12 (2004).

94. Fahraeus, G. The infection of clover root hair by nodule bacteria studied by a single glass technique. *Journal of General Microbiology* **16**, 374–381 (1957).

95. Ormeño-Orrillo, E., Rosenblueth, M., Luyten, E., Vanderleyden, J. & Martínez-Romero, E. Mutations in lipopolysaccharide biosynthetic genes impair maize rhizosphere and root colonization of *Rhizobium tropici* CIAT899. *Environmental Microbiology* **10**, 1271–1284 (2008).

96. Schulz, M. H., Zerbino, D. R., Vingron, M. & Birney, E. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* **28**, 1086–1892 (2012).

97. Aziz, R. K. *et al.* The RAST Server: rapid annotations using subsystems technology. *BMC genomics* **9**, 75 (2008).

98. Jean, G. *et al.* PALMapper: Fast and Accurate Spliced Alignments of Sequence Reads. (2011).at <http://www.raetschlab.org/suppl/palmapper>

99. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

100. Konstantinidis, K. T. & Tiedje, J. M. Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2567–72 (2005).

101. Wayne, L. G. *et al.* Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *International Journal of Systematic Bacteriology* **37**, 463–464 (1987).

102. Acosta, J. L. *et al.* Genomic lineages of *Rhizobium etli* revealed by the extent of nucleotide polymorphisms and low recombination. *BMC Evolutionary biology* **11**, 1–13 (2011).

103. Grabherr, M. G. *et al.* Genome-wide synteny through highly sensitive sequence alignment: Satsuma. *Bioinformatics* **26**, 1145–1151 (2010).

104. Taniguchi, Y. *et al.* Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533–538 (2010).

105. Chatterji, D. & Kumar Ojha, A. Revisiting the stringent response, ppGpp and starvation signaling. *Current Opinion in Microbiology* **4**, 160–165 (2001).

106. Moris, M. *et al.* Effective symbiosis between *Rhizobium etli* and *Phaseolus vulgaris* requires the alarmone ppGpp. *Journal of Bacteriology* **187**, 5460–5469 (2005).

107. Calderón-Flores, A. *et al.* The stringent response is required for amino acid and nitrate utilization, nod factor regulation, nodulation, and nitrogen fixation in *Rhizobium etli*. *Journal of Bacteriology* **187**, 5075–5083 (2005).

108. Kraffczyk, I., Trolldenier, G. & Beringer, H. Soluble root exudates of maize: Influence of potassium supply and rhizosphere microorganisms. *Soil Biology and Biochemistry* **16**, 315–322 (1984).

109. Boncompagni, E. *et al.* Characterization of a *Sinorhizobium meliloti* ATP-Binding Cassette Histidine Transporter Also Involved in Betaine and Proline Uptake. *Journal of Bacteriology* **182**, 3717–3725 (2000).

110. Mandon, K. *et al.* The *Sinorhizobium meliloti* glycine betaine biosynthetic genes (betlCBA) are induced by choline and highly expressed in bacteroids. *Molecular Plant-Microbe Interactions* **16**, 709–719 (2003).

111. Pauly, N. *et al.* Reactive oxygen and nitrogen species and glutathione: key players in the legume-*Rhizobium* symbiosis. *Journal of Experimental Botany* **57**, 1769–1776 (2006).

112. Zheng, W., Chung, L. M. & Zhao, H. Bias detection and correction in RNA-Sequencing data. *BMC bioinformatics* **12**, 1–14 (2011).

113. Whitfield, J. Underground networking. *Nature* **449**, 136–138 (2007).

114. López-Guerrero, M. G., Ramírez-Romero, M. & Martínez-Romero, E. Rhizobial genetic repertoire to inhabit legume and non-legume rhizospheres. *Microbial Ecology of The Rhizosphere*

115. Izaguirre-Mayoral, M. & Labandera, C. Biofertilizantes en Iberoamérica: Visión técnica, científica y empresarial. (2007).

116. Gil, R., Latorre, A. & Moya, A. Bacterial endosymbionts of insects: insights from comparative genomics. *Environmental Microbiology* **6**, 1109–1122 (2004).

117. Quellerm, D. Q., Strassmann, J. E. & Hughes, C. R. Microsatellites and kinship. *Trends in Ecology & Evolution* **8**, 285–288 (1993).

118. Gregory, S. G. Contig Assembly. *Encyclopedia of Life Sciences* 1–4 (2005).doi:10.1038/npg.els.0005365

119. Miller, J. R., Koren, S. & Sutton, G. Assembly algorithms for next-generation sequencing data. *Genomics* **95**, 315–327 (2010).

120. NCBI Genome Reference Consortium. (2012).at <http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/info/definitions.shtml >

# 9. Anexo I. Resultados adicionales.

## Tabla 9. Genes altamente expresados en maíz y en frijol.

| Gen | Función putativa |
|---|---|
| CromosomaCh24-10_draft_0001 | *Membrane proteins related to metalloendopeptidases* |
| CromosomaCh24-10_draft_0002 | *ClpB protein* |
| CromosomaCh24-10_draft_0006 | *Phosphocarrier protein kinase/phosphorylase, nitrogen regulation associated* |
| CromosomaCh24-10_draft_0020 | *Protein export cytoplasm protein SecA ATPase RNA helicase (TC 3.A.5.1.1)* |
| CromosomaCh24-10_draft_0023 | *Acetyl-coenzyme A synthetase (EC 6.2.1.1)* |
| CromosomaCh24-10_draft_0030 | *cAMP-dependent Kef-type K transport system* |
| CromosomaCh24-10_draft_0048 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_0071 | *Aerobic cobaltochelatase CobT subunit (EC 6.6.1.2)* |
| CromosomaCh24-10_draft_0086 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0089 | *Helicase (Snf2/Rad54 family)* |
| CromosomaCh24-10_draft_0091 | *Type III restriction-modification enzyme helicase subunit* |
| CromosomaCh24-10_draft_0109 | *Putative glycosyl hydrolase of unknown function (DUF1680)* |
| CromosomaCh24-10_draft_0115 | *Outer membrane protein* |
| CromosomaCh24-10_draft_0116 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0122 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_0126 | *NAD-dependent formate dehydrogenase alpha subunit* |
| CromosomaCh24-10_draft_0139 | *ATP synthase beta chain (EC 3.6.3.14)* |
| CromosomaCh24-10_draft_0144 | *Helicase PriA essential for oriC/DnaA-independent DNA replication* |
| CromosomaCh24-10_draft_0151 | *Dihydrolipoamide dehydrogenase of 2-oxoglutarate dehydrogenase (EC 1.8.1.4)* |
| CromosomaCh24-10_draft_0157 | *2-oxoglutarate dehydrogenase E1 component (EC 1.2.4.2)* |
| CromosomaCh24-10_draft_0168 | *Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)* |
| CromosomaCh24-10_draft_0188 | *GTP-binding protein TypA/BipA* |
| CromosomaCh24-10_draft_0189 | *Dipeptidyl carboxypeptidase Dcp (EC 3.4.15.5)* |
| CromosomaCh24-10_draft_0210 | *ABC transporter, permease protein, putative* |
| CromosomaCh24-10_draft_0215 | *Aconitate hydratase (EC 4.2.1.3)* |
| CromosomaCh24-10_draft_0227 | *Signal recognition particle, subunit Ffh SRP54 (TC 3.A.5.1.1)* |
| CromosomaCh24-10_draft_0235 | *3-isopropylmalate dehydratase large subunit (EC 4.2.1.33)* |
| CromosomaCh24-10_draft_0244 | *Cell division protein FtsK* |
| CromosomaCh24-10_draft_0245 | *Ammonium transporter* |
| CromosomaCh24-10_draft_0262 | *Ribose ABC transport system, ATP-binding protein RbsA (TC 3.A.1.2.1)* |
| CromosomaCh24-10_draft_0265 | *Dipeptide-binding ABC transporter, periplasmic substrate-binding component (TC 3.A.1.5.2)* |
| CromosomaCh24-10_draft_0272 | *Cyclic beta-1,2-glucan synthase (EC 2.4.1.-)* |
| CromosomaCh24-10_draft_0275 | *Beta-(1-->2)glucan export ATP-binding/permease protein NdvA (EC 3.6.3.42)* |

| | |
|---|---|
| CromosomaCh24-10_draft_0277 | *Pyruvate carboxyl transferase (EC 6.4.1.1)* |
| CromosomaCh24-10_draft_0284 | *1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (EC 1.17.7.1)* |
| CromosomaCh24-10_draft_0297 | *ATP-dependent DNA helicas* |
| CromosomaCh24-10_draft_0306 | *ABC transporter, ATP-binding protein* |
| CromosomaCh24-10_draft_0307 | *ABC transporter permease / ATP-binding protein* |
| CromosomaCh24-10_draft_0326 | *Beta-mannosidase (EC 3.2.1.25)* |
| CromosomaCh24-10_draft_0328 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0334 | *Aldehyde dehydrogenase B (EC 1.2.1.22)* |
| CromosomaCh24-10_draft_0338 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0359 | *C-terminal processing peptidase protein* |
| CromosomaCh24-10_draft_0363 | *DNA polymerase III alpha subunit (EC 2.7.7.7)* |
| CromosomaCh24-10_draft_0376 | *Major facilitator family transporter* |
| CromosomaCh24-10_draft_0384 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0387 | *NAD-glutamate dehydrogenase* |
| CromosomaCh24-10_draft_0388 | *NAD-specific glutamate dehydrogenase (EC 1.4.1.2), large form* |
| CromosomaCh24-10_draft_0389 | *Major facilitator superfamily MFS_1* |
| CromosomaCh24-10_draft_0390 | *IMP cyclohydrolase (EC 3.5.4.10) / Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3)* |
| CromosomaCh24-10_draft_0395 | *Acetyl-coenzyme A synthetase (EC 6.2.1.1)* |
| CromosomaCh24-10_draft_0398 | *Acetyl-coenzyme A synthetase (EC 6.2.1.1)* |
| CromosomaCh24-10_draft_0400 | *Leucyl-tRNA synthetase (EC 6.1.1.4)* |
| CromosomaCh24-10_draft_0406 | *tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA* |
| CromosomaCh24-10_draft_0422 | *DNA gyrase subunit B (EC 5.99.1.3)* |
| CromosomaCh24-10_draft_0434 | *Dihydrofolate synthase (EC 6.3.2.12) / Folylpolyglutamate synthase (EC 6.3.2.17)* |
| CromosomaCh24-10_draft_0436 | *ATP-dependent nuclease subunit A* |
| CromosomaCh24-10_draft_0437 | *ATP-dependent nuclease subunit B* |
| CromosomaCh24-10_draft_0440 | *Two-component sensor histidine kinase* |
| CromosomaCh24-10_draft_0441 | *Adenosylhomocysteinase (EC 3.3.1.1)* |
| CromosomaCh24-10_draft_0445 | *Sensor histidine kinase ChvG (EC 2.7.3.-)* |
| CromosomaCh24-10_draft_0463 | *Malate synthase G (EC 2.3.3.9)* |
| CromosomaCh24-10_draft_0468 | *ATP-dependent helicase protein* |
| CromosomaCh24-10_draft_0470 | *Superfamily I DNA/RNA helicase protein* |
| CromosomaCh24-10_draft_0475 | *Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12)* |
| CromosomaCh24-10_draft_0494 | *Sarcosine dehydrogenase (EC 1.5.99.1)* |
| CromosomaCh24-10_draft_0503 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0504 | *SSU ribosomal protein S1p* |
| CromosomaCh24-10_draft_0508 | *Adenylate cyclase (EC 4.6.1.1)* |
| CromosomaCh24-10_draft_0524 | *Ornithine decarboxylase (EC 4.1.1.17)* |
| CromosomaCh24-10_draft_0525 | *Arginine decarboxylase (EC 4.1.1.19); Ornithine decarboxylase (EC 4.1.1.17); Lysine decarboxylase (EC 4.1.1.18)* |

| | |
|---|---|
| CromosomaCh24-10_draft_0538 | *Potassium efflux system KefA protein / Small-conductance mechanosensitive channel* |
| CromosomaCh24-10_draft_0547 | *3-oxoacyl-[acyl-carrier-protein] synthase, KASI (EC 2.3.1.41)* |
| CromosomaCh24-10_draft_0550 | *Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8)* |
| CromosomaCh24-10_draft_0555 | *Translation initiation factor 2* |
| CromosomaCh24-10_draft_0557 | *Transcription termination protein NusA* |
| CromosomaCh24-10_draft_0565 | *DNA polymerase III subunits gamma and tau (EC 2.7.7.7)* |
| CromosomaCh24-10_draft_0584 | *Chaperone protein DnaK* |
| CromosomaCh24-10_draft_0585 | *Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)* |
| CromosomaCh24-10_draft_0589 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0591 | *DNA polymerase I (EC 2.7.7.7)* |
| CromosomaCh24-10_draft_0592 | *Cell division protein FtsK* |
| CromosomaCh24-10_draft_0623 | *Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)* |
| CromosomaCh24-10_draft_0632 | *Predicted nucleoside ABC transporter, substrate-binding component* |
| CromosomaCh24-10_draft_0633 | *Predicted nucleoside ABC transporter, ATP-binding component* |
| CromosomaCh24-10_draft_0647 | *Flp pilus assembly protein, pilin Flp* |
| CromosomaCh24-10_draft_0653 | *Type II/IV secretion system ATP hydrolase TadA/VirB11/CpaF, TadA subfamily* |
| CromosomaCh24-10_draft_0668 | *Aminotransferase, class III* |
| CromosomaCh24-10_draft_0669 | *Formyltetrahydrofolate deformylase (EC 3.5.1.10)* |
| CromosomaCh24-10_draft_0684 | *Translation elongation factor LepA* |
| CromosomaCh24-10_draft_0701 | *Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)* |
| CromosomaCh24-10_draft_0712 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0717 | *Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)* |
| CromosomaCh24-10_draft_0734 | *GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2)* |
| CromosomaCh24-10_draft_0760 | *MG(2+) chelatase family protein / ComM-related protein* |
| CromosomaCh24-10_draft_0770 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0776 | *Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36) / Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)* |
| CromosomaCh24-10_draft_0786 | *Aromatic amino acid ammonia-lyase family* |
| CromosomaCh24-10_draft_0787 | *Ubiquinone biosynthesis monooxygenase UbiB* |
| CromosomaCh24-10_draft_0792 | *Chromosomal replication initiator protein DnaA* |
| CromosomaCh24-10_draft_0820 | *[Protein-PII] uridylyltransferase (EC 2.7.7.59)* |
| CromosomaCh24-10_draft_0822 | *DNA mismatch repair protein MutS* |
| CromosomaCh24-10_draft_0823 | *NADP-dependent malic enzyme (EC 1.1.1.40)* |
| CromosomaCh24-10_draft_0826 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0843 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_0870 | *Inner membrane protein translocase component YidC, long form* |
| CromosomaCh24-10_draft_0873 | *Sensor histidine kinase in cluster with mercury reductase* |
| CromosomaCh24-10_draft_0913 | *Glucose-6-phosphate isomerase (EC 5.3.1.9)* |
| CromosomaCh24-10_draft_0914 | *Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)* |

| | |
|---|---|
| CromosomaCh24-10_draft_0915 | *Hypothetical protein* |
| CromosomaCh24-10_draft_0916 | *Beta-mannosidase (EC 3.2.1.25)* |
| CromosomaCh24-10_draft_0926 | *Ribose transport ATP-binding protein RbsA* |
| CromosomaCh24-10_draft_0933 | *Putative ABC transporter (fused ATP-binding and permease components)* |
| CromosomaCh24-10_draft_0975 | *Putative ATP-dependent helicase* |
| CromosomaCh24-10_draft_0981 | *Mll5622 protein* |
| CromosomaCh24-10_draft_0984 | *Acyl-CoA dehydrogenase (EC 1.3.99.3)* |
| CromosomaCh24-10_draft_0986 | *Enoyl-CoA hydratase [isoleucine degradation] (EC 4.2.1.17) / 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35) / 3-hydroxybutyryl-CoA epimerase (EC 5.1.2.3)* |
| CromosomaCh24-10_draft_0993 | *Sensor histidine kinase* |
| CromosomaCh24-10_draft_1000 | *ATP-dependent DNA helicase UvrD/PcrA, proteobacterial paralog* |
| CromosomaCh24-10_draft_1045 | *ATP-dependent DNA ligase (EC 6.5.1.1) clustered with Ku protein, LigD* |
| CromosomaCh24-10_draft_1055 | *Phytochrome, two-component sensor histidine kinase (EC 2.7.3.-)* |
| CromosomaCh24-10_draft_1060 | *Protein-export membrane protein SecD (TC 3.A.5.1.1) / Protein-export membrane protein SecF (TC 3.A.5.1.1)* |
| CromosomaCh24-10_draft_1062 | *Binding-protein-dependent transport systems inner membrane component:ATP/GTP-binding site motif A (P-loop) :TrkA-N:Potassium* |
| CromosomaCh24-10_draft_1069 | *Signal transduction histidine kinase CheA (EC 2.7.3.-)* |
| CromosomaCh24-10_draft_1076 | *Flagellar M-ring protein FliF* |
| CromosomaCh24-10_draft_1098 | *Hypothetical protein* |
| CromosomaCh24-10_draft_1099 | *Flagellin protein FlaA* |
| CromosomaCh24-10_draft_1109 | *Flagellar hook protein FlgE* |
| CromosomaCh24-10_draft_1116 | *Flagellar biosynthesis protein FlhA* |
| CromosomaCh24-10_draft_1126 | *Alpha-glucosides-binding periplasmic protein AglE precursor* |
| CromosomaCh24-10_draft_1129 | *Alpha-glucosidase AglA* |
| CromosomaCh24-10_draft_1130 | *Alpha-glucoside transport ATP-binding protein AglK* |
| CromosomaCh24-10_draft_1132 | *Phosphogluconate dehydratase (EC 4.2.1.12)* |
| CromosomaCh24-10_draft_1139 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_1141 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_1151 | *Acetoacetyl-CoA synthetase (EC 6.2.1.16)* |
| CromosomaCh24-10_draft_1152 | *Sensory histidine kinase AtoS* |
| CromosomaCh24-10_draft_1154 | *Hemolysin-type calcium binding protein* |
| CromosomaCh24-10_draft_1155 | *Hemolysin-type calcium binding protein* |
| CromosomaCh24-10_draft_1162 | *Dipeptide-binding ABC transporter, periplasmic substrate-binding component (TC 3.A.1.5.2)* |
| CromosomaCh24-10_draft_1216 | *Glycolate dehydrogenase (EC 1.1.99.14), subunit GlcD* |
| CromosomaCh24-10_draft_1231 | *Histidyl-tRNA synthetase (EC 6.1.1.21)* |
| CromosomaCh24-10_draft_1237 | *Heat shock protein 60 family chaperone GroEL* |
| CromosomaCh24-10_draft_1243 | *Isoleucyl-tRNA synthetase (EC 6.1.1.5)* |
| CromosomaCh24-10_draft_1246 | *Ribosomal large subunit pseudouridine synthase B (EC 4.2.1.70)* |

| | |
|---|---|
| CromosomaCh24-10_draft_1251 | *Putative Glutathione-regulated potassium-efflux system protein KefB* |
| CromosomaCh24-10_draft_1255 | *3-deoxy-D-manno-octulosonic-acid transferase (EC 2.-.-.-)* |
| CromosomaCh24-10_draft_1260 | *DNA mismatch repair protein MutL* |
| CromosomaCh24-10_draft_1274 | *Kup system potassium uptake protein* |
| CromosomaCh24-10_draft_1287 | *TPR domain protein* |
| CromosomaCh24-10_draft_1299 | *5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)* |
| CromosomaCh24-10_draft_1304 | *Glycyl-tRNA synthetase beta chain (EC 6.1.1.14)* |
| CromosomaCh24-10_draft_1305 | *Transcription accessory protein (S1 RNA-binding domain)* |
| CromosomaCh24-10_draft_1311 | *Flagellar basal-body rod protein FlgF* |
| CromosomaCh24-10_draft_1315 | *Benzoate 1,2-dioxygenase (EC 1.14.12.10)* |
| CromosomaCh24-10_draft_1319 | *Acyl-CoA dehydrogenase (EC 1.3.99.3)* |
| CromosomaCh24-10_draft_1327 | *1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7)* |
| CromosomaCh24-10_draft_1328 | *Putative uncharacterized protein* |
| CromosomaCh24-10_draft_1365 | *TldD protein, part of proposed TldE/TldD proteolytic complex (PMID 12029038)* |
| CromosomaCh24-10_draft_1368 | *Cytochrome c oxidase polypeptide I (EC 1.9.3.1)* |
| CromosomaCh24-10_draft_1390 | *Sensory box/GGDEF domain/EAL domain protein* |
| CromosomaCh24-10_draft_1403 | *Modification methylase* |
| CromosomaCh24-10_draft_1416 | *Chromosome partition protein Smc* |
| CromosomaCh24-10_draft_1419 | *Pyruvate,phosphate dikinase (EC 2.7.9.1)* |
| CromosomaCh24-10_draft_1427 | *Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)* |
| CromosomaCh24-10_draft_1433 | *Long-chain-fatty-acid-CoA ligase (EC 6.2.1.3)* |
| CromosomaCh24-10_draft_1435 | *D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)* |
| CromosomaCh24-10_draft_1486 | *Hypothetical protein* |
| CromosomaCh24-10_draft_1490 | *HtrA protease/chaperone protein* |
| CromosomaCh24-10_draft_1495 | *Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)* |
| CromosomaCh24-10_draft_1511 | *Choline dehydrogenase (EC 1.1.99.1)* |
| CromosomaCh24-10_draft_1536 | *Hypothetical protein* |
| CromosomaCh24-10_draft_1538 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_1542 | *ATP-dependent DNA ligase (EC 6.5.1.1) LigC* |
| CromosomaCh24-10_draft_1597 | *Glucose dehydrogenase, PQQ-dependent (EC 1.1.5.2)* |
| CromosomaCh24-10_draft_1598 | *Selenoprotein O and cysteine-containing homologs* |
| CromosomaCh24-10_draft_1601 | *Protease II (EC 3.4.21.83)* |
| CromosomaCh24-10_draft_1614 | *Heat shock protein 60 family chaperone GroEL* |
| CromosomaCh24-10_draft_1624 | *Na(+) H(+) antiporter subunit A* |
| CromosomaCh24-10_draft_1636 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_1642 | *Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)* |
| CromosomaCh24-10_draft_1669 | *Macromolecule metabolism; Macromolecule degradation; degradation of proteins, peptides, glycopeptides* |
| CromosomaCh24-10_draft_1672 | *Two-component sensor histidine kinase PleC* |
| CromosomaCh24-10_draft_1673 | *Glutamate-ammonia-ligase adenylyltransferase (EC 2.7.7.42)* |

| | |
|---|---|
| CromosomaCh24-10_draft_1674 | *Membrane alanine aminopeptidase N (EC 3.4.11.2)* |
| CromosomaCh24-10_draft_1681 | *Acriflavin resistance protein* |
| CromosomaCh24-10_draft_1682 | *Probable Co/Zn/Cd efflux system membrane fusion protein* |
| CromosomaCh24-10_draft_1685 | *Electron transfer flavoprotein-ubiquinone oxidoreductase (EC 1.5.5.1)* |
| CromosomaCh24-10_draft_1703 | *AMP nucleosidase (EC 3.2.2.4)* |
| CromosomaCh24-10_draft_1721 | *Epi-inositol hydrolase (EC 3.7.1.-)* |
| CromosomaCh24-10_draft_1725 | *Ferredoxin reductase* |
| CromosomaCh24-10_draft_1727 | *GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase II / Guanosine-;-bis(diphosphate);-pyrophosphohydrolase (EC 3.1.7.2)* |
| CromosomaCh24-10_draft_1732 | *LysM domain protein* |
| CromosomaCh24-10_draft_1762 | *Cysteinyl-tRNA synthetase (EC 6.1.1.16)* |
| CromosomaCh24-10_draft_1774 | *Amidophosphoribosyltransferase (EC 2.4.2.14)* |
| CromosomaCh24-10_draft_1777 | *Outer membrane protein Imp, required for envelope biogenesis / Organic solvent tolerance protein precursor* |
| CromosomaCh24-10_draft_1798 | *Prolyl oligopeptidase family protein* |
| CromosomaCh24-10_draft_1805 | *Prolyl oligopeptidase family protein* |
| CromosomaCh24-10_draft_1827 | *Surface antigen* |
| CromosomaCh24-10_draft_1830 | *Anaerobic dehydrogenases, typically selenocysteine-containing / Nitrate reductase* |
| CromosomaCh24-10_draft_1837 | *Polyphosphate kinase (EC 2.7.4.1)* |
| CromosomaCh24-10_draft_1863 | *Hypothetical protein* |
| CromosomaCh24-10_draft_1877 | *Putative beta-D-1,6 glucosyltransferase protein* |
| CromosomaCh24-10_draft_1884 | *Pyrophosphate-energized proton pump (EC 3.6.1.1)* |
| CromosomaCh24-10_draft_1894 | *Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12)* |
| CromosomaCh24-10_draft_1895 | *Cyclic di-GMP binding protein precursor* |
| CromosomaCh24-10_draft_1897 | *FOG: TPR repeat* |
| CromosomaCh24-10_draft_1920 | *Hypothetical protein* |
| CromosomaCh24-10_draft_1933 | *Probable bacteriophage protein STY1048* |
| CromosomaCh24-10_draft_1934 | *Hypothetical protein* |
| CromosomaCh24-10_draft_1950 | *Hypothetical protein* |
| CromosomaCh24-10_draft_1954 | *Unknown protein* |
| CromosomaCh24-10_draft_1955 | *Hypothetical protein* |
| CromosomaCh24-10_draft_1978 | *ABC transport system, sugar-binding protein* |
| CromosomaCh24-10_draft_1985 | *Putative transmembrane succinoglycan biosynthesis transport protein* |
| CromosomaCh24-10_draft_2013 | *ATP-dependent Clp protease ATP-binding subunit ClpX* |
| CromosomaCh24-10_draft_2014 | *ATP-dependent protease Type I (EC 3.4.21.53)* |
| CromosomaCh24-10_draft_2028 | *NADH-ubiquinone oxidoreductase chain E (EC 1.6.5.3)* |
| CromosomaCh24-10_draft_2029 | *NADH-ubiquinone oxidoreductase chain F (EC 1.6.5.3)* |
| CromosomaCh24-10_draft_2031 | *NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)* |
| CromosomaCh24-10_draft_2036 | *NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)* |
| CromosomaCh24-10_draft_2037 | *NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)* |

| | |
|---|---|
| CromosomaCh24-10_draft_2040 | *Metallo-beta-lactamase family protein, RNA-specific* |
| CromosomaCh24-10_draft_2047 | *DNA polymerase III alpha subunit (EC 2.7.7.7)* |
| CromosomaCh24-10_draft_2058 | *Exoribonuclease RNase R* |
| CromosomaCh24-10_draft_2059 | *DNA topoisomerase I (EC 5.99.1.2)* |
| CromosomaCh24-10_draft_2065 | *Cytoplasmic axial filament protein CafA and Ribonuclease G (EC 3.1.4.-)* |
| CromosomaCh24-10_draft_2067 | *Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)* |
| CromosomaCh24-10_draft_2088 | *DNA-directed RNA polymerase beta subunit (EC 2.7.7.6)* |
| CromosomaCh24-10_draft_2089 | *DNA-directed RNA polymerase beta subunit (EC 2.7.7.6)* |
| CromosomaCh24-10_draft_2093 | *Translation elongation factor G* |
| CromosomaCh24-10_draft_2116 | *Preprotein translocase secY subunit (TC 3.A.5.1.1)* |
| CromosomaCh24-10_draft_2120 | *DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)* |
| CromosomaCh24-10_draft_2126 | *Dihydroxy-acid dehydratase (EC 4.2.1.9)* |
| CromosomaCh24-10_draft_2141 | *Putative two component sensor histidine kinase transcriptional regulatory protein* |
| CromosomaCh24-10_draft_2152 | *Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)* |
| CromosomaCh24-10_draft_2187 | *Ribonucleotide reductase of class II (coenzyme B12-dependent) (EC 1.17.4.1)* |
| CromosomaCh24-10_draft_2200 | *Cyclic-nucleotide;-phosphodiesterase (EC 3.1.4.16)* |
| CromosomaCh24-10_draft_2218 | *Assimilatory nitrate reductase large subunit (EC:1.7.99.4)* |
| CromosomaCh24-10_draft_2220 | *Nitrite reductase [NAD(P)H] large subunit (EC 1.7.1.4)* |
| CromosomaCh24-10_draft_2253 | *Valyl-tRNA synthetase (EC 6.1.1.9)* |
| CromosomaCh24-10_draft_2259 | *Arginyl-tRNA synthetase (EC 6.1.1.19)* |
| CromosomaCh24-10_draft_2260 | *Soluble lytic murein transglycosylase (EC 3.2.1.-)* |
| CromosomaCh24-10_draft_2270 | *Lipoprotein NlpD* |
| CromosomaCh24-10_draft_2273 | *Protein-export membrane protein SecD (TC 3.A.5.1.1) / Protein-export membrane protein SecF (TC 3.A.5.1.1)* |
| CromosomaCh24-10_draft_2282 | *Cell division trigger factor (EC 5.2.1.8)* |
| CromosomaCh24-10_draft_2284 | *Possible P-methylase (EC 2.1.1.-)* |
| CromosomaCh24-10_draft_2301 | *Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)* |
| CromosomaCh24-10_draft_2314 | *Polyhydroxyalkanoic acid synthase* |
| CromosomaCh24-10_draft_2328 | *Hypothetical protein* |
| CromosomaCh24-10_draft_2346 | *ATP-dependent Clp protease ATP-binding subunit ClpA* |
| CromosomaCh24-10_draft_2363 | *Outer membrane protein assembly factor YaeT precursor* |
| CromosomaCh24-10_draft_2369 | *Citrate synthase (EC 2.3.3.1)* |
| CromosomaCh24-10_draft_2371 | *DNA internalization-related competence protein ComEC/Rec2* |
| CromosomaCh24-10_draft_2408 | *Nitrogen regulation protein NtrY (EC 2.7.3.-)* |
| CromosomaCh24-10_draft_2416 | *Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)* |
| CromosomaCh24-10_draft_2428 | *Transport ATP-binding protein CydCD* |
| CromosomaCh24-10_draft_2445 | *Mannonate dehydratase (EC 4.2.1.8)* |
| CromosomaCh24-10_draft_2447 | *Dipeptide-binding ABC transporter, periplasmic substrate-binding component (TC 3.A.1.5.2)* |

| | |
|---|---|
| CromosomaCh24-10_draft_2454 | *Hypothetical protein* |
| CromosomaCh24-10_draft_2479 | *Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)* |
| CromosomaCh24-10_draft_2481 | *ATP-dependent DNA helicase RecG (EC 3.6.1.-)* |
| CromosomaCh24-10_draft_2483 | *Transcription-repair coupling factor* |
| CromosomaCh24-10_draft_2488 | *ABC transporter, periplasmic substrate-binding protein* |
| CromosomaCh24-10_draft_2491 | *Glutamine synthetase type I (EC 6.3.1.2)* |
| CromosomaCh24-10_draft_2495 | *Excinuclease ABC subunit A* |
| CromosomaCh24-10_draft_2498 | *DNA gyrase subunit A (EC 5.99.1.3)* |
| CromosomaCh24-10_draft_2514 | *Ferric iron ABC transporter, permease protein* |
| CromosomaCh24-10_draft_2519 | *Hypothetical protein* |
| CromosomaCh24-10_draft_2529 | *Hypothetical protein* |
| CromosomaCh24-10_draft_2533 | *Adenylate cyclase (EC 4.6.1.1) / Guanylate cyclase (EC 4.6.1.2)* |
| CromosomaCh24-10_draft_2543 | *Putative sensory box/GGDEF family protein* |
| CromosomaCh24-10_draft_2547 | *Potassium efflux system KefA protein / Small-conductance mechanosensitive channel* |
| CromosomaCh24-10_draft_2557 | *Methionyl-tRNA synthetase (EC 6.1.1.10)* |
| CromosomaCh24-10_draft_2576 | *Peptidyl-prolyl cis-trans isomerase ppiD (EC 5.2.1.8)* |
| CromosomaCh24-10_draft_2582 | *Macromolecule metabolism; macromolecule synthesis, modification; RNA synthesis, modification , DNA transcription* |
| CromosomaCh24-10_draft_2591 | *Topoisomerase IV subunit B (EC 5.99.1.-)* |
| CromosomaCh24-10_draft_2593 | *Membrane protein, putative* |
| CromosomaCh24-10_draft_2600 | *Threonyl-tRNA synthetase (EC 6.1.1.3)* |
| CromosomaCh24-10_draft_2614 | *D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)* |
| CromosomaCh24-10_draft_2629 | *Biotin carboxylase (EC 6.3.4.14) / Biotin carboxyl carrier protein* |
| CromosomaCh24-10_draft_2646 | *Glycine dehydrogenase [decarboxylating] (glycine cleavage system P protein) (EC 1.4.4.2)* |
| CromosomaCh24-10_draft_2652 | *Iron-sulfur cluster assembly protein SufD* |
| CromosomaCh24-10_draft_2663 | *Large exoproteins involved in heme utilization or adhesion* |
| CromosomaCh24-10_draft_2679 | *Sensory box/GGDEF family protein* |
| CromosomaCh24-10_draft_2689 | *Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)* |
| CromosomaCh24-10_draft_2702 | *Hypothetical / partial homology* |
| CromosomaCh24-10_draft_2707 | *Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)* |
| CromosomaCh24-10_draft_2712 | *Alanyl-tRNA synthetase (EC 6.1.1.7)* |
| CromosomaCh24-10_draft_2713 | *RecA protein* |
| CromosomaCh24-10_draft_2716 | *Sensory box histidine kinase/response regulator* |
| CromosomaCh24-10_draft_2721 | *Amino acid regulated cytosolic protein* |
| CromosomaCh24-10_draft_2730 | *Kinesin-like protein* |
| CromosomaCh24-10_draft_2744 | *NAD-dependent malic enzyme (EC 1.1.1.38)* |
| CromosomaCh24-10_draft_2750 | *NAD synthetase (EC 6.3.1.5) / Glutamine amidotransferase chain of NAD synthetase* |

| | |
|---|---|
| CromosomaCh24-10_draft_2754 | *2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase II (EC 2.5.1.54)* |
| CromosomaCh24-10_draft_2762 | *Hypothetical protein* |
| CromosomaCh24-10_draft_2766 | *Fumarate hydratase class I, aerobic (EC 4.2.1.2)* |
| CromosomaCh24-10_draft_2803 | *Ferric siderophore transport system, periplasmic binding protein TonB* |
| CromosomaCh24-10_draft_2825 | *Sensor histidine kinase (EC 2.7.3.-)* |
| CromosomaCh24-10_draft_2843 | *Beta-mannosidase (EC 3.2.1.25)* |
| CromosomaCh24-10_draft_2878 | *Chloride channel protein* |
| CromosomaCh24-10_draft_2906 | *Putative mechanosensitive channel protein* |
| CromosomaCh24-10_draft_2925 | *Two-component sensor histidine kinase* |
| CromosomaCh24-10_draft_2928 | *Hypothetical conserved protein* |
| CromosomaCh24-10_draft_2942 | *Excinuclease ABC subunit B* |
| CromosomaCh24-10_draft_2990 | *Na+/H+ antiporter* |
| CromosomaCh24-10_draft_2991 | *Kup system potassium uptake protein* |
| CromosomaCh24-10_draft_2994 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3008 | *Alternative cytochrome c oxidase polypeptide CoxN (EC 1.9.3.1)* |
| CromosomaCh24-10_draft_3016 | *Putative autoaggregation protein* |
| CromosomaCh24-10_draft_3038 | *Serine phosphatase RsbU, regulator of sigma subunit* |
| CromosomaCh24-10_draft_3045 | *Diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)* |
| CromosomaCh24-10_draft_3051 | *Adenylate cyclase (EC 4.6.1.1)* |
| CromosomaCh24-10_draft_3086 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3087 | *Adenylate cyclase (EC 4.6.1.1)* |
| CromosomaCh24-10_draft_3152 | *Ribosome protection-type tetracycline resistance related proteins* |
| CromosomaCh24-10_draft_3174 | *Cobalt-containing nitrile hydratase subunit beta (EC 4.2.1.84)* |
| CromosomaCh24-10_draft_3180 | *Acetolactate synthase large subunit (EC 2.2.1.6)* |
| CromosomaCh24-10_draft_3182 | *Bipolar DNA helicase* |
| CromosomaCh24-10_draft_3187 | *HtrA protease/chaperone protein* |
| CromosomaCh24-10_draft_3195 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3196 | *D-Lactate dehydrogenase, cytochrome c-dependent (EC 1.1.2.4)* |
| CromosomaCh24-10_draft_3208 | *RND multidrug efflux transporter; Acriflavin resistance protein* |
| CromosomaCh24-10_draft_3209 | *Probable Co/Zn/Cd efflux system membrane fusion protein* |
| CromosomaCh24-10_draft_3214 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3215 | *Serine protein kinase (prkA protein), P-loop containing* |
| CromosomaCh24-10_draft_3217 | *ATP-dependent DNA helicase UvrD/PcrA* |
| CromosomaCh24-10_draft_3225 | *Xaa-Pro aminopeptidase (EC 3.4.11.9)* |
| CromosomaCh24-10_draft_3229 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_3234 | *DNA ligase (EC 6.5.1.2)* |
| CromosomaCh24-10_draft_3236 | *DNA repair protein RecN* |
| CromosomaCh24-10_draft_3239 | *Cell division protein FtsZ (EC 3.4.24.-)* |
| CromosomaCh24-10_draft_3245 | *UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8)* |

| | |
|---|---|
| CromosomaCh24-10_draft_3252 | *Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)* |
| CromosomaCh24-10_draft_3272 | *ABC transporter, ATP-binding protein* |
| CromosomaCh24-10_draft_3277 | *Thiamine pyrophosphate-requiring enzymes* |
| CromosomaCh24-10_draft_3281 | *Protein ErfK/SrfK* |
| CromosomaCh24-10_draft_3286 | *Putative amino acid permease protein* |
| CromosomaCh24-10_draft_3302 | *5-methyltetrahydrofolate-homocysteine methyltransferase (EC 2.1.1.13)* |
| CromosomaCh24-10_draft_3344 | *RNA polymerase sigma factor RpoD* |
| CromosomaCh24-10_draft_3348 | *ATP-dependent DNA helicase RecQ* |
| CromosomaCh24-10_draft_3362 | *Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)* |
| CromosomaCh24-10_draft_3368 | *Aerobic C4-dicarboxylate transporter for fumarate, L-malate, D-malate, succinate* |
| CromosomaCh24-10_draft_3375 | *Oxidoreductase, FMN-binding* |
| CromosomaCh24-10_draft_3386 | *Inter-alpha-trypsin inhibitor domain protein* |
| CromosomaCh24-10_draft_3408 | *NaCromosomalecule metabolism; maCromosomalecule degradation; degradation of proteins, peptides, glycopeptides* |
| CromosomaCh24-10_draft_3413 | *Diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)* |
| CromosomaCh24-10_draft_3424 | *Ubiquinol--cytochrome c reductase, cytochrome B subunit (EC 1.10.2.2)* |
| CromosomaCh24-10_draft_3426 | *ATP-binding protein of ABC transporter* |
| CromosomaCh24-10_draft_3428 | *Lipid A export ATP-binding/permease protein MsbA* |
| CromosomaCh24-10_draft_3430 | *Sodium-dependent phosphate transporter* |
| CromosomaCh24-10_draft_3440 | *Putative membrane protein* |
| CromosomaCh24-10_draft_3441 | *Threonine dehydrogenase and related Zn-dependent dehydrogenases* |
| CromosomaCh24-10_draft_3460 | *Anthranilate synthase, aminase component (EC 4.1.3.27) / Anthranilate synthase, amidotransferase component (EC 4.1.3.27)/ Para-aminobenzoate synthase,amidotransferase component (EC 2.6.1.85)* |
| CromosomaCh24-10_draft_3463 | *Formate-tetrahydrofolate ligase (EC 6.3.4.3)* |
| CromosomaCh24-10_draft_3475 | *Adenosylmethionine-8-amino-7-oxononanoate aminotransferase (EC 2.6.1.62)* |
| CromosomaCh24-10_draft_3478 | *Adenine deaminase (EC 3.5.4.2)* |
| CromosomaCh24-10_draft_3481 | *Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)* |
| CromosomaCh24-10_draft_3489 | *Glutamine synthetase type II, eukaryotic (EC 6.3.1.2)* |
| CromosomaCh24-10_draft_3493 | *GTP-binding protein EngA* |
| CromosomaCh24-10_draft_3511 | *Xanthine dehydrogenase, molybdenum binding subunit (EC 1.17.1.4)* |
| CromosomaCh24-10_draft_3526 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3536 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3542 | *ATP-dependent RNA helicase RhlE* |
| CromosomaCh24-10_draft_3553 | *Sugar ABC transporter, ATP-binding protein* |
| CromosomaCh24-10_draft_3589 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3600 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3622 | *Hypothetical protein* |

| | |
|---|---|
| CromosomaCh24-10_draft_3625 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3680 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3681 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3682 | *Type III restriction-modification enzyme, helicase subunit* |
| CromosomaCh24-10_draft_3685 | *Adenine specific DNA methylase (Mod-related)* |
| CromosomaCh24-10_draft_3687 | *Putative DNA helicase* |
| CromosomaCh24-10_draft_3693 | *Phytochrome, two-component sensor histidine kinase (EC 2.7.3.-)* |
| CromosomaCh24-10_draft_3696 | *Dihydropyrimidinase (EC 3.5.2.2)* |
| CromosomaCh24-10_draft_3703 | *RND multidrug efflux transporter; Acriflavin resistance protein* |
| CromosomaCh24-10_draft_3709 | *Urease alpha subunit (EC 3.5.1.5)* |
| CromosomaCh24-10_draft_3720 | *Urea ABC transporter, permease protein UrtB* |
| CromosomaCh24-10_draft_3722 | *Putative phosphatase protein* |
| CromosomaCh24-10_draft_3723 | *Two-component hybrid sensor and regulator* |
| CromosomaCh24-10_draft_3725 | *High-affinity leucine-specific transport system, periplasmic binding protein LivK (TC 3.A.1.4.1)* |
| CromosomaCh24-10_draft_3749 | *Hypothetical transcription regulator* |
| CromosomaCh24-10_draft_3750 | *Adenylosuccinate synthetase (EC 6.3.4.4)* |
| CromosomaCh24-10_draft_3753 | *Diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)* |
| CromosomaCh24-10_draft_3754 | *Xylonate dehydratase (EC 4.2.1.82)* |
| CromosomaCh24-10_draft_3755 | *RND multidrug efflux transporter; Acriflavin resistance protein* |
| CromosomaCh24-10_draft_3756 | *RND efflux system, membrane fusion protein CmeA* |
| CromosomaCh24-10_draft_3768 | *RND efflux system, inner membrane transporter CmeB* |
| CromosomaCh24-10_draft_3791 | *NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)* |
| CromosomaCh24-10_draft_3800 | *NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)* |
| CromosomaCh24-10_draft_3810 | *RND multidrug efflux transporter; Acriflavin resistance protein* |
| CromosomaCh24-10_draft_3818 | *Aldehyde dehydrogenase (EC 1.2.1.3)* |
| CromosomaCh24-10_draft_3823 | *Lead, cadmium, zinc and mercury transporting ATPase (EC 3.6.3.3) (EC 3.6.3.5); Copper-translocating P-type ATPase (EC 3.6.3.4)* |
| CromosomaCh24-10_draft_3826 | *Ribonucleotide reductase of class Ib (aerobic), alpha subunit (EC 1.17.4.1)* |
| CromosomaCh24-10_draft_3841 | *RND efflux system, inner membrane transporter CmeB* |
| CromosomaCh24-10_draft_3855 | *Not classified regulator* |
| CromosomaCh24-10_draft_3864 | *Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9); Fructose-6-phosphate phosphoketolase (EC 4.1.2.22)* |
| CromosomaCh24-10_draft_3877 | *D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)* |
| CromosomaCh24-10_draft_3880 | *Phosphoglucosamine mutase (EC 5.4.2.10)* |
| CromosomaCh24-10_draft_3881 | *Cell division protein FtsH (EC 3.4.24.-)* |
| CromosomaCh24-10_draft_3900 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_3909 | *Putative membrane protein* |
| CromosomaCh24-10_draft_3911 | *Potassium efflux system KefA protein / Small-conductance mechanosensitive* |

| | |
|---|---|
| | *channel* |
| CromosomaCh24-10_draft_3919 | *Transketolase (EC 2.2.1.1)* |
| CromosomaCh24-10_draft_3922 | *Na+/H+ antiporter* |
| CromosomaCh24-10_draft_3931 | *ABC-type multidrug transport system, ATPase and permease components* |
| CromosomaCh24-10_draft_3949 | *Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)* |
| CromosomaCh24-10_draft_3950 | *Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)* |
| CromosomaCh24-10_draft_3952 | *Signal transduction histidine kinase CheA (EC 2.7.3.-)* |
| CromosomaCh24-10_draft_3954 | *Methyl-accepting chemotaxis receptor/sensory transducer* |
| CromosomaCh24-10_draft_3973 | *Pyruvate kinase (EC 2.7.1.40)* |
| CromosomaCh24-10_draft_3978 | *Putative exported protein* |
| CromosomaCh24-10_draft_3979 | *Response regulatory protein* |
| CromosomaCh24-10_draft_3980 | *Oligoendopeptidase F* |
| CromosomaCh24-10_draft_3984 | *Homospermidine synthase (EC 2.5.1.44)* |
| CromosomaCh24-10_draft_3994 | *Membrane-bound lytic murein transglycosylase B precursor (EC 3.2.1.-)* |
| CromosomaCh24-10_draft_3997 | *Glutamate synthase [NADPH] small chain (EC 1.4.1.13)* |
| CromosomaCh24-10_draft_3998 | *Glutamate synthase [NADPH] large chain (EC 1.4.1.13)* |
| CromosomaCh24-10_draft_4025 | *Large protein containing transglutaminase-like domain* |
| CromosomaCh24-10_draft_4026 | *Protein containing domains DUF404, DUF407, DUF403* |
| CromosomaCh24-10_draft_4029 | *Glycogen phosphorylase (EC 2.4.1.1)* |
| CromosomaCh24-10_draft_4033 | *Phosphoglucomutase (EC 5.4.2.2)* |
| CromosomaCh24-10_draft_4038 | *Sarcosine oxidase alpha subunit (EC 1.5.3.1)* |
| CromosomaCh24-10_draft_4040 | *Sarcosine oxidase beta subunit (EC 1.5.3.1)* |
| CromosomaCh24-10_draft_4042 | *Adenylate cyclase (EC 4.6.1.1)* |
| CromosomaCh24-10_draft_4046 | *Sensor histidine kinase, putative* |
| CromosomaCh24-10_draft_4053 | *Diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)* |
| CromosomaCh24-10_draft_4062 | *Putative metal chaperone, involved in Zn homeostasis, GTPase of COG0523 family* |
| CromosomaCh24-10_draft_4079 | *Hypothetical protein* |
| CromosomaCh24-10_draft_4086 | *D-xylose transport ATP-binding protein XylG* |
| CromosomaCh24-10_draft_4089 | *Xylulose kinase (EC 2.7.1.17)* |
| CromosomaCh24-10_draft_4096 | *Glutamyl-tRNA synthetase (EC 6.1.1.17) / Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)* |
| CromosomaCh24-10_draft_4097 | *Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)* |
| CromosomaCh24-10_draft_4127 | *Beta-hexosaminidase (EC 3.2.1.52)* |
| CromosomaCh24-10_draft_4129 | *Component of multidrug efflux system* |
| CromosomaCh24-10_draft_4147 | *Beta-galactosidase (EC 3.2.1.23)* |
| CromosomaCh24-10_draft_4148 | *Insertion sequence transposase* |
| CromosomaCh24-10_draft_4151 | *Cytochrome c biosynthesis protein* |
| CromosomaCh24-10_draft_4152 | *ABC-type sugar transport system, periplasmic component* |
| CromosomaCh24-10_draft_4153 | *Inositol transport system permease protein* |

| | |
|---|---|
| pCh-2410d_draft_0007 | *Diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)* |
| pCh24-10d_draft_0008 | *Diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)* |
| pCh-2410d_draft_0040 | *Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)* |
| pCh-2410d_draft_0103 | *DNA polymerase III alpha subunit (EC 2.7.7.7)* |
| pCh-2410d_draft_0119 | *Putative methyl-accepting chemotaxis protein* |
| pCh-2410d_draft_0127 | *Signal transduction histidine kinase CheA (EC 2.7.3.-)* |
| pCh-2410d_draft_0147 | *Putative glycosyltransferase protein* |
| pCh-2410d_draft_0167 | *Biotin sulfoxide reductase (EC 1.-.-.-)* |
| pCh-2410d_draft_0194 | *Multi-sensor hybrid histidine kinase* |
| pCh-2410d_draft_0202 | *Alpha-mannosidase (EC 3.2.1.24)* |
| pCh-2410d_draft_0213 | *Beta-glucosidase (EC 3.2.1.21)* |
| pCh-2410d_draft_0219 | *Ferrichrome-iron receptor* |
| pCh-2410d_draft_0230 | *Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)* |
| pCh-2410d_draft_0236 | *Methyl-accepting chemotaxis protein* |
| pCh-2410d_draft_0263 | *L-sorbosone dehydrogenase* |
| pCh-2410d_draft_0269 | *Hypothetical protein* |
| pCh-2410d_draft_0270 | *Putative formate dehydrogenase oxidoreductase protein* |
| pCh-2410d_draft_0277 | *L-proline glycine betaine binding ABC transporter protein ProX (TC 3.A.1.12.1)* |
| pCh-2410d_draft_0297 | *Proline dehydrogenase (EC 1.5.99.8) (Proline oxidase) / Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)* |
| pCh-2410d_draft_0298 | *Hypothetical protein* |
| pCh-2410d_draft_0302 | *ATPase components of various ABC-type transport systems, contain duplicated ATPase* |
| pCh-2410d_draft_0305 | *ABC transporter, binding protein* |
| pCh-2410d_draft_0317 | *Succinate-semialdehyde dehydrogenase [NADP+] (EC 1.2.1.16)* |
| pCh-2410d_draft_0329 | *Fe-S protein, homolog of lactate dehydrogenase SO1521* |
| pCh-2410d_draft_0331 | *Heat shock protein 60 family chaperone GroEL* |
| pCh-2410d_draft_0338 | *Putative two component sensor histidine kinase transcriptional regulatory protein* |
| pCh-2410d_draft_0339 | *Hypothetical/Partial homology* |
| pCh-2410d_draft_0346 | *Hypothetical protein* |
| pCh-2410d_draft_0362 | *Methyl-accepting chemotaxis protein* |
| pCh-2410d_draft_0365 | *ABC transporter, substrate binding protein* |
| pCh-2410d_draft_0366 | *Hypothetical protein* |
| pCh-2410d_draft_0378 | *RND efflux transporter* |
| pCh-2410d_draft_0381 | *Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)* |
| pCh-2410d_draft_0389 | *Alpha-amylase (EC 3.2.1.1)* |
| pCh-2410d_draft_0390 | *Trehalose synthase (EC 5.4.99.16)* |
| pCh-2410d_draft_0391 | *1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)* |

| | |
|---|---|
| pCh-2410d_draft_0401 | *Two-component hybrid sensor and regulator* |
| pCh-2410d_draft_0402 | *Hypothetical protein* |
| pCh-2410d_draft_0431 | *Dipeptide-binding ABC transporter, periplasmic substrate-binding component (TC 3.A.1.5.2)* |
| pCh-2410d_draft_0436 | *ABC-type dipeptide transport system, periplasmic component* |
| pCh-2410d_draft_0449 | *Hypothetical protein* |
| pCh-2410d_draft_0511 | *dTDP-glucose 4,6-dehydratase protein* |
| pCh-2410d_draft_0518 | *Multidrug resistance protein A* |
| pCh-2410d_draft_0522 | *ATP-binding protein of ABC transporter* |
| pCh-2410d_draft_0529 | *Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)* |
| pCh-2410d_draft_0535 | *Glycerol-3-phosphate ABC transporter, permease protein UgpA (TC 3.A.1.1.3)* |
| pCh-2410d_draft_0567 | *Nucleoside-diphosphate-sugar epimerases* |
| pCh-2410d_draft_0581 | *Putative formate dehydrogenase oxidoreductase protein* |
| pCh-2410d_draft_0589 | *Hypothetical protein* |
| pCh-2410d_draft_0605 | *Flagellar hook-length control protein FliK* |
| pCh-2410d_draft_0607 | *Hypothetical protein* |
| pCh-2410d_draft_0610 | *Probable secretion ATP-binding protein* |
| pCh-2410d_draft_0666 | *Chromosome partition protein smc* |
| pCh-2410d_draft_0675 | *Cell processes; transport of small molecules; amino acids, amines, peptides* |
| pCh-2410d_draft_0687 | *Probable ABC transporter, ATP-binding protein* |
| pCh-2410d_draft_0690 | *Cytosine deaminase (EC 3.5.4.1)* |
| pCh-2410d_draft_0702 | *Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)* |
| pCh-2410d_draft_0715 | *Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE* |
| pCh-2410d_draft_0726 | *Putative oxidoreductase* |
| pCh-2410d_draft_0736 | *Methylmalonyl-CoA mutase (EC 5.4.99.2)* |
| pCh-2410d_draft_0773 | *Putative transmembrane GGDEF/EAL domain protein* |
| pCh-2410d_draft_0795 | *Urea carboxylase (EC 6.3.4.6)* |
| pCh-2410d_draft_0796 | *Allophanate hydrolase (EC 3.5.1.54)* |
| pCh-2410d_draft_0817 | *Exoenzymes regulatory protein AepA precursor* |
| pCh-2410d_draft_0839 | *Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)* |
| pCh-2410d_draft_0900 | *Catalase (EC 1.11.1.6) / Peroxidase (EC 1.11.1.7)* |
| pCh-2410d_draft_0916 | *Hypothetical protein* |
| pCh-2410d_draft_0935 | *Unknown protein* |
| pCh-2410d_draft_0938 | *Hypothetical protein* |
| pCh-2410d_draft_0939 | *VgrG protein* |
| pCh-2410d_draft_0946 | *IcmF-related protein* |
| pCh-2410d_draft_0958 | *Serine/Threonine protein kinase* |
| pCh-2410d_draft_0959 | *ClpB protein* |
| pCh-2410d_draft_0963 | *IcmF-related protein* |

| | |
|---|---|
| pCh-2410d_draft_0975 | *Long-chain-fatty-acid-CoA ligase (EC 6.2.1.3)* |
| pCh-2410d_draft_0981 | *Putative plasmid stabilization protein* |
| pCh-2410d_draft_0985 | *Hypothetical protein* |
| pCh-2410d_draft_0986 | *Bores hole in peptidoglycan layer allowing type IV secretion complex assembly to occur (VirB1)* |
| pCh-2410d_draft_0989 | *ATPase provides energy for both assembly of type IV secretion complex and secretion of T-DNA complex (VirB4)* |
| pCh-2410d_draft_1027 | *ATPase involved in DNA repair* |
| pCh-2410d_draft_1028 | *Cyclopropane-fatty-acyl-phospholipid synthase (EC 2.1.1.79)* |
| pCh-2410d_draft_1035 | *Aminotransferase, class III* |
| pCh24-10c_draft_0043 | *Hypothetical protein* |
| pCh24-10c_draft_0044 | *Hypothetical protein* |
| pCh24-10c_draft_0045 | *Hypothetical protein* |
| pCh24-10c_draft_0051 | *Catalase (EC 1.11.1.6)* |
| pCh24-10c_draft_0052 | *Hypothetical protein* |
| pCh24-10c_draft_0054 | *Hypothetical protein* |
| pCh24-10c_draft_0058 | *Peptidase M4, thermolysin:Propeptide, peptidase M4 and M36( EC:3.4.24.28 )* |
| pCh24-10c_draft_0060 | *Tyrosinase* |
| pCh24-10c_draft_0083 | *Blue copper oxidase CueO precursor* |
| pCh24-10c_draft_0090 | *Beta-glucosidase (EC 3.2.1.21)* |
| pCh24-10c_draft_0111 | *Hypothetical protein* |
| pCh24-10c_draft_0112 | *Hypothetical protein* |
| pCh24-10c_draft_0138 | *Tyrosinase* |
| pCh24-10c_draft_0156 | *Methyl-accepting chemotaxis protein* |
| pCh24-10c_draft_0162 | *Diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)* |
| pCh24-10c_draft_0188 | *Diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)* |
| pCh24-10c_draft_0221 | *Hypothetical protein* |
| pCh24-10c_draft_0222 | *Hypothetical protein* |
| pCh24-10c_draft_0237 | *VgrG protein* |
| pCh24-10c_draft_0250 | *IcmF-related protein* |
| pCh24-10c_draft_0259 | *Hypothetical protein RetlK5_32521* |
| pCh24-10c_draft_0274 | *Nitrogenase (molybdenum-iron) beta chain (EC 1.18.6.1)* |
| pCh24-10c_draft_0275 | *Nitrogenase (molybdenum-iron) alpha chain (EC 1.18.6.1)* |
| pCh24-10c_draft_0287 | *FUPA27 P-type ATPase* |
| pCh24-10c_draft_0308 | *Nodulation protein nolO (EC 2.1.3.-)* |
| pCh24-10c_draft_0324 | *Putative cytochrome P450 hydroxylase* |
| pCh24-10c_draft_0327 | *Hypothetical protein* |
| pCh24-10c_draft_0336 | *Mobile element protein* |

| | |
|---|---|
| pCh24-10b_draft_0006 | *PAN domain protein* |
| pCh24-10b_draft_0015 | *ABC-transport protein, ATP-binding component* |
| pCh24-10b_draft_0016 | *ABC-transport protein, ATP-binding component* |
| pCh24-10b_draft_0019 | *Probable two-component sensor histidine kinase/response regulator hybrid protein* |
| pCh24-10b_draft_0024 | *Cytochrome O ubiquinol oxidase subunit I (EC 1.10.3.-)* |
| pCh24-10b_draft_0026 | *MFS permease* |
| pCh24-10b_draft_0031 | *Hypothetical protein* |
| pCh24-10b_draft_0037 | *Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein* |
| pCh24-10b_draft_0038 | *ATP-binding component of a ABC transport system (oligopeptide)* |
| pCh24-10b_draft_0045 | *Adenylate cyclase (EC 4.6.1.1)* |
| pCh24-10b_draft_0047 | *Adenylate cyclase (EC 4.6.1.1)* |
| pCh24-10b_draft_0073 | *Urocanate hydratase (EC 4.2.1.49)* |
| pCh24-10b_draft_0087 | *Quinone-dependent oxidoreductase domain / Putative aminomethyl transferase domain* |
| pCh24-10b_draft_0093 | *Periplasmic alpha-galactoside-binding protein precursor* |
| pCh24-10b_draft_0096 | *Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)* |
| pCh24-10b_draft_0103 | *Sugar binding protein of sugar ABC transporter* |
| pCh24-10b_draft_0109 | *Translation elongation factor G-related protein* |
| pCh24-10b_draft_0112 | *Putative peptidase protein* |
| pCh24-10b_draft_0121 | *Adenylate cyclase (EC 4.6.1.1) / Guanylate cyclase (EC 4.6.1.2)* |
| pCh24-10b_draft_0122 | *Long-chain-fatty-acid-CoA ligase (EC 6.2.1.3)* |
| pCh24-10b_draft_0124 | *Adenylate cyclase (EC 4.6.1.1)* |
| pCh24-10b_draft_0129 | *N-Acetyl-D-glucosamine ABC transport system, sugar-binding protein* |
| pCh24-10b_draft_0130 | *Coenzyme F420-dependent N5,N10-methylene tetrahydromethanopterin reductase and related flavin-dependent oxidoreductases* |
| pCh24-10b_draft_0134 | *Adenylate cyclase (EC 4.6.1.1)* |
| pCh24-10b_draft_0141 | *Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)* |
| pCh24-10b_draft_0142 | *Alpha-N-arabinofuranosidase (EC 3.2.1.55)* |
| pCh24-10b_draft_0152 | *Glycogen debranching enzyme (EC 3.2.1.-)* |
| pCh24-10b_draft_0158 | *Multi-sensor hybrid histidine kinase* |
| pCh24-10b_draft_0163 | *Cardiolipin synthetase (EC 2.7.8.-)* |
| pCh24-10b_draft_0183 | *ATP-dependent DNA ligase (EC 6.5.1.1) clustered with Ku protein, LigD* |
| pCh24-10b_draft_0194 | *Osmosensitive K+ channel histidine kinase KdpD (EC 2.7.3.-)* |
| pCh24-10b_draft_0196 | *Potassium-transporting ATPase B chain (EC 3.6.3.12) (TC 3.A.3.7.1)* |
| pCh24-10b_draft_0212 | *Choline dehydrogenase (EC 1.1.99.1)* |
| pCh24-10b_draft_0223 | *Predicted rhamnulose-1-phosphate aldolase (EC 4.1.2.19) / Predicted lactaldehyde dehydrogenase (EC 1.2.1.22)* |
| pCh24-10b_draft_0258 | *Putative glutathione transporter, ATP-binding component* |
| pCh24-10b_draft_0271 | *Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12)* |

| | |
|---|---|
| pCh24-10b_draft_0285 | *RecD-like DNA helicase YrrC* |
| pCh24-10b_draft_0292 | *Beta-hexosaminidase (EC 3.2.1.52)* |
| pCh24-10b_draft_0295 | *Hypothetical conserved protein* |
| pCh24-10b_draft_0310 | *Formyltetrahydrofolate deformylase (EC 3.5.1.10)* |
| pCh24-10b_draft_0319 | *Sarcosine oxidase alpha subunit (EC 1.5.3.1)* |
| pCh24-10b_draft_0337 | *Hypothetical protein* |
| pCh24-10b_draft_0339 | *Putative trifolitoxin related protein* |
| | *Lead, cadmium, zinc and mercury transporting ATPase (EC 3.6.3.3) (EC 3.6.3.5);* |
| pCh24-10b_draft_0367 | *Copper-translocating P-type ATPase (EC 3.6.3.4)* |
| pCh24-10b_draft_0368 | *Malto-oligosyltrehalose synthase (EC 5.4.99.15)* |

**Lista de genes cromosomales y plasmídicos de *Rhizobium phaseoli* Ch24-10 altamente expresados en la rizósfera de maíz y en frijol agrupados en sistemas de clasificación de acuerdo al *RAST*.**

Se muestran la lista desglosada de los genes agrupados en sistemas y subsistemas, clasificación que se obtuvo al anotar los genes en el servidor utilizando el RAST (Rpaid Anotation using Subsistem Technology) [97]. El número entre paréntesis indica cuántos genes se localizaron en el subsistema.

a) *Cofactors, Vitamins, Prosthetic Groups, Pigments (3)*
   a. *Thiamin biosynthesis (1)*
   b. *NAD and NADP cofactor biosynthesis global (2)*

b) *Cell Wall and Capsule (4)*
   a. *Murein Hydrolases (2)*
   b. *UDP-N-acetylmuramate from Fructose-6-phosphate Biosynthesis (2)*

c) *Virulence, Disease and Defense (12)*
   a. *Resistance to antibiotics and toxic compounds (10)*
      i. *Copper homeostasis (3)*
      ii. *Tetracycline resistance, ribosome protection type (2)*
      iii. *Resistance to fluoroquinolones (3)*
      iv. *Tetracycline resistance, ribosome protection type (2)*
   b. ⊟ *Invasion and intracellular resistance (2)*
      i. *Mycobacterium virulence operon involved in DNA transcription (2)*

d) *Potassium metabolism (10)*
   a. *Potassium homeostasis (10)*

e) *Membrane Transport (13)*
   a. *ABC transporters (12)*
      i. *Periplasmic-Binding-Protein-Dependent Transport System for α-Glucosides (3)*
      ii. *ABC transporter oligopeptide (TC 3.A.1.5.1) (5)*
      iii. *ABC transporter dipeptide (TC 3.A.1.5.2) (4)*

f) *Uni- Sym- and Antiporters (1)*
   a. *NhaA, NhaD and Sodium-dependent phosphate transporters (1)*
   b. *Iron acquisition and metabolism (1)*
      i. *Hemin transport system (1)*

g) *RNA Metabolism (10)*
   a. *RNA processing and modification (3)*
      i. *ATP-dependent RNA helicases, bacterial (1)*
      ii. *RNA processing and degradation, bacterial (2)*
   b. *Transcription (7)*
      i. *Transcription initiation, bacterial sigma factors (1)*
      ii. *RNA polymerase bacterial (3)*
      iii. *Transcription factors bacterial (3)*

h) *Nucleosides and Nucleotides (3)*
   a. *Ribonucleotide reduction (2)*
   b. *Hydantoin metabolism (1)*

i) *Protein Metabolism (33)*
   a. *Protein folding (3)*
      i. *GroEL GroES (3)*
   b. *Selenoproteins (1)*
      i. *Selenoprotein O (1)*
   c. *Protein biosynthesis (16)*
      i. *tRNA aminoacylation, Val (1)*

ii. tRNA aminoacylation, Cys (1)
iii. tRNA aminoacylation, Met (1)
iv. tRNA aminoacylation, His (1)
v. tRNA aminoacylation, Ile (1)
vi. tRNA aminoacylation, Arg (1)
vii. tRNA aminoacylation, Lys (1)
viii. tRNA aminoacylation, Thr (1)
ix. tRNA aminoacylation, Ala (1)
x. Translation elongation factor G family (3)
xi. tRNA aminoacylation, Glu and Gln (3)
xii. tRNA aminoacylation, Leu (1)

d. Protein processing and modification (3)
i. G3E family of P-loop GTPases (metallocenter biosynthesis) (3)

e. Protein degradation (10)
i. Aminopeptidases (EC 3.4.11.-) (3)
ii. Protein degradation (1)
iii. Metallocarboxypeptidases (EC 3.4.17.-) (1)
iv. Proteolysis in bacteria, ATP-dependent (5)

j) Motility and Chemotaxis (13)
a. Motility and Chemotaxis - no subcategory (13)
i. Bacterial Chemotaxis (13)

k) Regulation and Cell signaling (13)
a. cAMP signaling in bacteria (11)
b. Stringent Response, (p)ppGpp metabolism (2)

l) DNA Metabolism (16)
a. DNA repair (8)
i. DNA repair, bacterial RecBCD pathway (1)
ii. DNA repair, bacterial MutL-MutS system (2)
iii. DNA repair, UvrABC system (1)
iv. DNA repair system including RecA, MutS and a hypothetical protein (2)
v. DNA repair, bacterial UvrD and related helicases (2)
b. DNA Metabolism - no subcategory (3)
i. DNA ligases (3)
c. DNA replication (4)
i. DNA topoisomerases, Type I, ATP-independent (1)
ii. DNA topoisomerases, Type II, ATP-dependent (3)
d. DNA uptake, competence (1)
i. DNA processing cluster (1)

m) Fatty Acids, Lipids, and Isoprenoids (1)
a. Phospholipids (1)
i. Cardiolipin synthesis (1)

n) Nitrogen Metabolism (9)
a. Nitrate and nitrite ammonification (2)
b. Ammonia assimilation (7)

o) Respiration (16)
a. Electron accepting reactions (3)
i. Anaerobic respiratory reductases (3)
b. Electron donating reactions (13)
i. Respiratory Complex I (7)
ii. Respiratory dehydrogenases 1 (4)
iii. Succinate dehydrogenase (2)

p) *Stress Response (19)*
    a. *Osmotic stress (8)*
        i. *Synthesis of osmoregulated periplasmic glucans (2)*
        ii. *Choline and Betaine Uptake and Betaine Biosynthesis (6)*
    b. *Oxidative stress (8)*
        i. *Protection from Reactive Oxygen Species (3)*
        ii. *Oxidative stress (5)*
    c. *Stress Response - no subcategory (1)*
        i. *SigmaB stress responce regulation (1)*
    d. *Periplasmic Stress (2)*
        i. *Periplasmic Stress Response (2)*

q) *Metabolism of Aromatic Compounds (1)*
    a. *Peripheral pathways for catabolism of aromatic compounds (1)*
        i. *Benzoate catabolism (1)*

r) *Amino Acids and Derivatives (31)*
    a. *Glutamine, glutamate, aspartate, asparagine; ammonia assimilation (8)*
        i. *Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis (5)*
        ii. *Glutamate dehydrogenases (1)*
        iii. *Glutamine synthetases (2)*
    b. *Arginine; urea cycle, polyamines (15)*
        i. *Polyamine Metabolism (3)*
        ii. *Arginine and Ornithine Degradation (4)*
        iii. *Urea decomposition (4)*
        iv. *Urea carboxylase and Allophanate hydrolase cluster (4)*
    c. *Lysine, threonine, methionine, and cysteine (2)*
        i. *Threonine degradation (1)*
        ii. *Lysine degradation (1)*
    d. *Amino Acids and Derivatives - no subcategory (2)*
        i. *Creatine and Creatinine Degradation (2)*

    e. *Aromatic amino acids and derivatives (2)*
        i. *Common Pathway For Synthesis of Aromatic Compounds (DAHP synthase to chorismate) (2)*
    f. *Proline and 4-hydroxyproline (2)*
        i. *Proline, 4-hydroxyproline uptake and utilization (2)*

s) *Sulfur Metabolism (1)*
    a. *Sulfur Metabolism - no subcategory (1)*
        i. *Galactosylceramide and Sulfatide metabolism (1)*

t) *Phosphorus Metabolism (1)*
    a. *Phosphorus Metabolism - no subcategory (1)*
        i. *Carbohydrates (34)*

u) *Carbohydrates (34)*
    a. *Central carbohydrate metabolism (20)*
        i. *TCA Cycle (8)*
        ii. *Methylglyoxal Metabolism (3)*
        iii. *Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate (4)*
        iv. *Pyruvate metabolism I: anaplerotic reactions, PEP (5)*
    b. *Di- and oligosaccharides (6)*
        i. *Trehalose Biosynthesis (5)*
        ii. *Lactose utilization (1)*
    c. *Polysaccharides (4)*
        i. *Glycogen metabolism (3)*
        ii. *Alpha-Amylase locus in Streptococcoccus (1)*
    d. *Monosaccharides (4)*
        i. *Xylose utilization (3)*
        ii. *D-gluconate and ketogluconates metabolism* (1)

## Tabla 10. Genes sobre-expresados en frijol y no en maíz.

| Gen | Función Putativa |
|---|---|
| CromosomaCh24-10_draft_0112 | *Putative sugar ABC transporter 2C substrate-binding protein* |
| CromosomaCh24-10_draft_0495 | *D-mannonate oxidoreductase protein* |
| CromosomaCh24-10_draft_0515 | *Putative uncharacterized protein* |
| CromosomaCh24-10_draft_0905 | *Dihydroorotase* |
| CromosomaCh24-10_draft_1093 | *Flagellar P-ring protein* |
| CromosomaCh24-10_draft_1106 | *Chemotaxis motility protein* |
| CromosomaCh24-10_draft_1124 | *Bifunctional protein FolD1* |
| CromosomaCh24-10_draft_1150 | *Hypothetical conserved membrane protein* |
| CromosomaCh24-10_draft_1360 | *D-alanyl-D-alanine carboxypeptidase* |
| CromosomaCh24-10_draft_1559 | *Probable+ABC+transporter%2C+ATP-binding+protein* |
| CromosomaCh24-10_draft_1692 | *Probable+glutamine+synthetase+protein* |
| CromosomaCh24-10_draft_1982 | *Putative+glycosyltransferase+protein* |
| CromosomaCh24-10_draft_2009 | *Putative sensory box/GGDEF family protein* |
| CromosomaCh24-10_draft_2048 | *Hypothetical+conserved+protein* |
| CromosomaCh24-10_draft_2368 | *Lipid-A-disaccharide synthase* |
| CromosomaCh24-10_draft_2409 | *Nitrogen regulation / Two-component response regulator protein* |
| CromosomaCh24-10_draft_2450 | *Dipeptide ABC transporter 2C ATP-binding protein* |
| CromosomaCh24-10_draft_2453 | *Putative transcriptional regulator protein 2C LacI family* |
| CromosomaCh24-10_draft_2473 | *Putative periplasmic-binding protein component of ABC transporter* |
| CromosomaCh24-10_draft_2474 | *ABC transporter related* |
| CromosomaCh24-10_draft_2531 | *Putative transcriptional regulator protein* |
| CromosomaCh24-10_draft_2583 | *Adenylate cyclase protein* |
| CromosomaCh24-10_draft_2756 | *Glutathione-disulfide reductase* |
| CromosomaCh24-10_draft_2915 | *Putative hydrolase protein* |
| CromosomaCh24-10_draft_2924 | *Hypothetical protein* |
| CromosomaCh24-10_draft_2952 | *Hypothetical protein* |
| CromosomaCh24-10_draft_2961 | *Putative adenylate cyclase protein* |
| CromosomaCh24-10_draft_3355 | *Putative uncharacterized protein* |
| CromosomaCh24-10_draft_3561 | *Xylose ABC transporter permease* |
| CromosomaCh24-10_draft_3854 | *Two component sigma54 specific / Transcriptional regulator Fis family* |
| CromosomaCh24-10_draft_3927 | *Transporter permease* |
| CromosomaCh24-10_draft_3947 | *Methyl-accepting chemotaxis protein* |
| CromosomaCh24-10_draft_4155 | *ABC+transporter+related* |
| pCh24-10d_draft_0740 | *Lipoprotein YaeC family* |
| pCh24-10d_draft_0530 | *Putative sugarABC transporter / ATP-bindin protein* |
| pCh24-10d_draft_0157 | *Flavin-dependent oxidoreductase, F420-dependent methylene-tetrahydromethanopterin reductase* |

| | |
|---|---|
| pCh24-10d_draft_0475 | *Putative transcriptiona regulator protein GntR family* |
| pCh24-10d_draft_0161 | *Putative alkanesulfonate+monooxygenase* |
| pCh24-10d_draft_0159 | *Ketopantoate reductase* |
| pCh24-10d_draft_0165 | *Peptide ABC transporter* |
| pCh24-10d_draft_0665 | *Peptide ABC transporter* |
| pCh24-10d_draft_0739 | *Putative ABC transporter ATP-binding protein* |
| pCh24-10d_draft_0592 | *Putative undecaprenyl-phosphate galactose phosphotransferase* |
| pCh24-10c_draft_0148 | *Carboxylesterase* |
| pCh24-10c_draft_0161 | *Sugar ABC transporter substrate-binding protein* |
| pCh24-10c_draft_0167 | *Hydrolase* |
| pCh24-10c_draft_0248 | *Hypothetical protein* |
| pCh24-10c_draft_0118 | *Type III secretion system ATP synthase protein* |
| pCh24-10b_draft_0189 | *Sulfate ABC transporter Periplasmic sulfate-binding protein* |
| pCh24-10b_draft_0264 | *Aspartate+ammonia-lyase* |
| pCh24-10b_draft_0201 | *Probable mercury reductase protein* |
| pCh24-10b_draft_0263 | *L-asparaginase II protein* |

## Tabla 11. Genes sobre-expresados en maíz y no en frijol.

| Gen | Función putativa |
|---|---|
| CromosomaCh24-10_draft_0004 | *Protoporphyrinogen oxidase / Methyltransferase protein* |
| CromosomaCh24-10_draft_0214 | *Hypothetical conserved protein* |
| CromosomaCh24-10_draft_0301 | *RNA polymerase sigma factor* |
| CromosomaCh24-10_draft_0408 | *Transcription termination factor Rho* |
| CromosomaCh24-10_draft_0534 | *Putative integral membrane protein (Dominio: TolB N Superfamily)* |
| CromosomaCh24-10_draft_0627 | *Nicotinate phosphoribosyltransferase* |
| CromosomaCh24-10_draft_0740 | *Putative uncharacterized protein* |
| CromosomaCh24-10_draft_0755 | *DNA polymerase III beta chain* |
| CromosomaCh24-10_draft_0856 | *Putative uncharacterized protein* |
| CromosomaCh24-10_draft_0882 | *Hypothetical conserved protein* |
| CromosomaCh24-10_draft_0897 | *Putative transcriptional regulator protein TetR family* |
| CromosomaCh24-10_draft_0971 | *Hypothetical protein* |
| CromosomaCh24-10_draft_1153 | *Hypothetical+conserved+protein* |
| CromosomaCh24-10_draft_1252 | *Modulator of DNA gyrase (Dominio: TidD, PmbA)* |
| CromosomaCh24-10_draft_1463 | *Putative polyhydroxybutyrate depolymerase protein* |
| CromosomaCh24-10_draft_1543 | *Aminopeptidase* |
| CromosomaCh24-10_draft_1603 | *Putative polysaccharide deacetylase protein* |
| CromosomaCh24-10_draft_1619 | *Probable molecular chaperone small heat shock protein hsp20 family* |
| CromosomaCh24-10_draft_1651 | *Probable serine protease protein* |
| CromosomaCh24-10_draft_1881 | *Hypothetical protein* |
| CromosomaCh24-10_draft_2074 | *Sin identidad significativa* |
| CromosomaCh24-10_draft_2170 | *Hypothetical conserved protein* |
| CromosomaCh24-10_draft_2171 | *Phenylalanine-4-hydroxylase protein* |
| CromosomaCh24-10_draft_2174 | *4-hydroxyphenylpyruvate dioxygenase* |
| CromosomaCh24-10_draft_2490 | *Putative uncharacterized protein* |
| CromosomaCh24-10_draft_2604 | *GTP cyclohydrolase 1* |
| CromosomaCh24-10_draft_2625 | *Putative uncharacterized protein* |
| CromosomaCh24-10_draft_2693 | *Glutaredoxin* |
| CromosomaCh24-10_draft_3213 | *SpoVR family protein* |
| CromosomaCh24-10_draft_3399 | *Hypothetical conserved protein* |
| CromosomaCh24-10_draft_3450 | *2-isopropylmalate synthase* |
| CromosomaCh24-10_draft_3537 | *Hypothetical protein* |
| CromosomaCh24-10_draft_3729 | *Amino acid ABC transporter permease protein* |
| CromosomaCh24-10_draft_3748 | *RNA polymerase sigma factor* |
| CromosomaCh24-10_draft_3757 | *Putative esterase protein* |
| CromosomaCh24-10_draft_3764 | *Putative protein secretion protein, HlyD family (Dominio: Biotin, lipoyl)* |
| CromosomaCh24-10_draft_3834 | *Putative uncharacterized protein* |

| | |
|---|---|
| pCh24-10d_draft_0304 | *Binding-protein-dependent / Transport systems inner membrane component* |
| pCh24-10d_draft_0318 | *Putative tartrate dehydrogenase* |
| pCh24-10d_draft_0363 | *Hypothetical protein* |
| pCh24-10d_draft_0692 | *Putative hydrolase protein* |
| pCh24-10d_draft_0850 | *Putative two-component+sensor+histidine+kinase+protein* |
| pCh24-10d_draft_1036 | *Phosphoserine phosphatase* |
| pCh24-10c_draft_0009 | *Methionyl-tRNA synthetase* |
| pCh24-10c_draft_0033 | *Plasmid partitioning protein RepAb* |
| pCh24-10c_draft_0258 | *Transposase* |
| pCh24-10b_draft_0007 | *Methionine sulfoxide reductase protein* |
| pCh24-10b_draft_0025 | *Cytochrome O ubiquinol oxidase, subunit II* |
| pCh24-10b_draft_0076 | *Imidazolonepropionase* |

# 10. Anexo II. Publicaciones.

1. MG López-Guerrero, et al., *Rhizobium etli* taxonomy revised with novel genomic data and analyses, Syst. Appl. Microbiol. (2012) http://dx.doi.org/10.1016/j.syapm.2012.06.009
2. MG López-Guerrero, et al. Rhizobial extrachromosomal replicon variability, stability and expression in natural niches. Plasmid (2012) http://dx.doi.org/10.1016/j.plasmid.2012.07.002
3. López-Guerrero MG, E Martínez-Romero. Bacterias fijadoras de nitrógeno. Diversidad y uso como biofertilizantes. Ciencia y Desarrollo. CONACYT. Mayo-Junio, 201 http://www.conacyt.gob.mx/comunicacion/Revista/252/articulos/microorganismos-bacterias-fijadoras.html
4. MG López-Guerrero, Ramírez-Romero MA, Martínez-Romero ME. Rhizobial genetic repertoire to inhabit legume and non-legume rhizospheres. En: Molecular Microbial Ecology of the Rhizosphere. Frans J. de Bruijn, ed. Wiley-Blackwell Publishers. En prensa.

Systematic and Applied Microbiology 35 (2012) 353–358

Contents lists available at SciVerse ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

# *Rhizobium etli* taxonomy revised with novel genomic data and analyses

Martha G. López-Guerrero[a], Ernesto Ormeño-Orrillo[a], Encarna Velázquez[b], Marco A. Rogel[a], José Luis Acosta[a,1], Victor Gónzalez[a], Julio Martínez[a], Esperanza Martínez-Romero[a,*]

[a] *Centro de Ciencias Genómicas, UNAM, Cuernavaca, Morelos, Mexico*
[b] *Departamento de Microbiología y Genética, Universidad de Salamanca, Salamanca, Spain*

ARTICLE INFO

*Article history:*
Received 8 March 2012
Received in revised form 14 May 2012
Accepted 14 June 2012

*Keywords:*
Nitrogen fixation
Legume nodulation
*Phaseolus vulgaris*

ABSTRACT

The taxonomic position of *Phaseolus vulgaris* rhizobial strains with available sequenced genomes was examined. Phylogenetic analyses with concatenated conserved genomic fragments accounting for over half of each genome showed that *Rhizobium* strains CIAT 652, Ch24-10 (newly reported genome) and CNPAF 512 constituted a well-supported group independent from *Rhizobium etli* CFN 42$^T$. DNA–DNA hybridization results indicated that CIAT 652, Ch24-10 and CNPAF 512 could correspond to *R. etli*, although the hybridization values were at the borderline that distinguishes different species. In contrast, experimental hybridization results were higher (over 80%) with *Rhizobium phaseoli* type strain ATCC 14482$^T$ in congruence to phylogenetic and ANIm analyses. The latter criterion allowed the reclassification of *R. etli* strains 8C-3 and Brasil5 as *R. phaseoli*. It was therefore concluded, based on all the evidence, that the CIAT 652, Ch24-10, and CNPAF 512 strains should be reclassified as *R. phaseoli* in spite of several common features linking them to *R. etli*. The *R. phaseoli* and *R. etli* speciation process seems to be a more recent event than the speciation that has occurred among other sister species, such as *R. leguminosarum*–*R. etli* or *R. rhizogenes*–*R. tropici*.



## Introduction

The advent of the genomic era has provided both a plethora of molecular markers useful in taxonomy and the possibility to compare whole genomes instead of a few genes. Toward this end, novel algorithms and parameters have been proposed to compare genomes for taxonomic purposes. ANI (average nucleotide identity) has been defined as a very useful parameter to delineate different species and it correlates with DNA–DNA hybridization (DDH) [12,19,32]. Based upon a large set of experimental results from diverse bacteria, thresholds of ANI (94–96%) have been recommended for distinguishing species.

*Phaseolus vulgaris* (common bean) is the legume grain most consumed for human nutrition and, like other legumes, it forms symbiosis with nitrogen-fixing bacteria. *P. vulgaris* symbiotic bacteria have been widely studied [4,6,13,16,22,24,31,34,37,43] and this legume has become a model for studying nodule-bacterial diversity from plants grown in diverse conditions or geographical regions where bean is native or introduced. In its sites of origin and in some introduced areas, *Rhizobium etli* has been reported as the dominant *P. vulgaris* bean nodule bacterium identified on the basis of 16S rRNA gene sequences [22]. In addition, core gene sequences have been used to characterize nodule isolates, thus providing a better phylogenetic resolution and revealing that other species besides *R. etli* can also represent a significant fraction of the bean nodule occupants [4,13,34,37]. Recently, *R. etli* type strain CFN 42$^T$ was found to have a low recombination with *R. etli* CIAT 652 and other *P. vulgaris* isolates [2]. Richter and Rosselló-Móra [32] calculated the ANI of CFN 42$^T$ and CIAT 652, and reported that CIAT 652 was not a member of *R. etli*. DDH results were not available for comparison and the species affiliation of CIAT 652 was not identified. At the same time, *Rhizobium phaseoli* was re-recognized as a valid and different species from *R. etli* because it presented divergent core genes and experimental DNA–DNA relatedness values significantly lower than 70% with *R. etli* CFN 42$^T$ [30]. *R. etli* and *R. phaseoli* are sympatric species nodulating bean and both have been found recently in Ethiopia [4]. On a phylogenetic basis, using *recA*, *atpD* and *celC* partial gene sequences, CIAT 652 was recognized as belonging to *R. phaseoli* [34] but this taxonomic affiliation was not otherwise confirmed by DDH analysis. Based on the same phylogenetic analysis, the Mim2 strain (a *Mimosa affinis* isolate) was also recognized as *R. phaseoli*, however, DDH and multilocus enzyme electrophoresis (MLEE) placed it within *R. etli* [44]. Clearly, there were conflicting data. Further, it was recommended that "the taxonomic status of the strains currently named *R. etli*

* Corresponding author.
*E-mail address:* esperanzaeriksson@yahoo.com.mx (E. Martínez-Romero).
[1] Present address: Instituto Nacional de Medicina Genómica, México DF, Mexico.

0723-2020/$ – see front matter © 2012 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.syapm.2012.06.009

**Table 1**
*Rhizobium etli* and *R. phaseoli* strains used in this study.

| Strains | Host | Origin | Reference |
|---|---|---|---|
| *R. phaseoli* | | | |
| Ch24-10 | *Zea mays* and *Phaseolus vulgaris* | Puebla, Mexico | [35] |
| CIAT 652 | *P. vulgaris* | Buitrera, Colombia | [42] |
| CNPAF 512 | *P. vulgaris* | Brazil | [9,23] |
| ATCC 14482$^T$ | *P. vulgaris* | Beltsville, Maryland | ATCC |
| *R. etli* | | | |
| CFN 42$^T$ | *P. vulgaris* | Guanajuato, Mexico | [29] |

should be revised" [34]. Consequently, our aim in this study was to revise the taxonomic status of different *R. etli* strains with available sequenced genomes.

## Materials and methods

### *Strains, growth and DNA extraction*

Strains used in the DDH studies are shown in Table 1. They were grown in 5 mL liquid PY medium for DNA extraction. Ch24-10 was grown on PY plates and fresh cultures were grown in 50 mL liquid PY. DNA was extracted by the DNA Isolation Kit for Cells and Tissues (Roche, USA). *Rhizobium* strains were maintained in YM with glycerol at −70 °C.

### *Phylogenetic analysis*

*rpoB* primers and PCR conditions were as described previously [21,25]. PCR sequences were compared to sequences obtained from whole genomes. Sequence alignments were generated and edited with BioEdit 7 [14]. Percentage identity between sequences was obtained after removing all columns with gaps from the alignments. Best-fit models of sequence evolution were selected for each gene with JModelTest 0.1.1 using the Akaike information criterion [28]. Maximum likelihood (ML) and neighbor-joining (NJ) phylogenies were constructed with Mega 5 [39]. Support for tree nodes was evaluated by bootstrap analysis with 100 or 1000 pseudoreplicates for ML and NJ, respectively.

### *Genomic sequencing*

The genomic sequence from strain Ch24-10 was obtained using two platforms: the Roche 454 pyrosequencing system (350 bp from 3K long-tag paired end sequencing protocol in the Genome Sequencer FLX) using a commercial service, and the Illumina technology (*Genome Analyzer GAIIx*, paired-end protocol, 200 base pairs-inserts library, reads with 36 nucleotides in length) at the Unidad Universitaria de Secuenciación Masiva de DNA (USMDNA) of the Universidad Nacional Autónoma de México (UNAM).

### *Genome assembly*

Roche reads were *de novo* assembled using GSassembler, Newbler version 2.5.3 with default parameters. Illumina reads were assembled using TAIPAN [36] and then SSAKE [45] separately, both with default parameters. The contigs generated by these programs were reassembled using minimus2 [http://sourceforge.net/apps/mediawiki/amos/index.php?title=Minimus2].

Both assemblies, the one generated with Newbler and the one generated with minimus2 (TAIPAN + SSAKE), were reassembled again to generate a hybrid assembly with minimus2.

### *Prediction and annotation of the Ch24-10 genome*

The contigs generated with the minimus2 assembler were assembled in scaffolds with ABACAS [5] using the complete genomes of CFN 42$^T$ and CIAT 652 as anchors. Next, the pseudochromosome and pseudoplasmids were constructed for each replicon armed with ABACAS, by adding the following sequence "NNNNCATTCCATTCATTAATTAATTAATGAATGAATGNNNNN" (containing the six open reading frames) at the 5′ end of each contig [40]. Prediction and annotation of genes for each pseudoreplicon were undertaken with the CG-Pipeline program [17]. The draft genome sequence of *Rhizobium* Ch24-10 obtained by hybrid assembly using sequences derived by Illumina and 454 Roche technologies was then used. An 80× genome coverage was obtained. The results of this Whole Genome Shotgun project have been deposited at DDBJ/EMBL/GenBank under the accession AHJU00000000. The version described in this paper is the first version, AHJU01000000.

### *Genome tree construction*

Regions shared among *Rhizobium* strains CIAT 652, CFN 42$^T$, Ch24-10, and CNPAF 512, *Rhizobium leguminosarum* sv. viciae 3841 and *Sinorhizobium meliloti* 1021 were identified using Mugsy [3] with the following parameters: minlength = 30, distance = 100, duplications 1, and fullsearch refine. Each orthologous region was extracted by an *ad hoc* Perl script (homemade) and filtered for long gaps with trimAl [8] with the parameter -automated1. All regions were concatenated to build a genome tree by the ML and neighbor-net network methods. ML was performed using RaxML [38] with the GTR + I + G nucleotide substitution model, 1000 distinct randomized maximum parsimony trees and the parameters: p 12345, e 0.0000001, c 8, j STRICT, and k 1000. Finally, the neighbor-net was created with the Splits Tree 4 program [15]. Other genome comparisons were performed as described [20].

### *ANIm*

ANIm values were calculated using the JSpecies package [32].

### *DNA–DNA hybridization (DDH)*

The procedure was as described in [41,44]. DNA was quantified with NanoDrop 2000 (Thermo Scientific) and in gels. DNA was digested with the *Eco*R1 restriction enzyme and electrophoresis was performed in 1% agarose gels. Only lanes with homogeneous DNA quantities were hybridized in Southern blot experiments to total DNA from reference strains with probes labeled with [α-32P] dCTP using RediPrimeTM II (GE Healthcare). Rapid-hyb buffer was used for hybridization and washings were carried out using 2× to 1× SSC with 0.1% SDS at 65 °C.

Filters were cut and individual lanes were counted in scintillation liquid in a LS6500 multi-purpose scintillation counter (Beckman Coulter). Three independent experiments were performed with similar results.

## Results

### *Phylogenetic and similarity gene analysis*

*R. etli* and *R. phaseoli* strains had highly similar 16S rRNA genes (over 99.3% identical) but were slightly separated in a NJ phylogenetic tree (Fig. 1). Ch24-10, CNPAF 512 and CIAT 652 strains reported as *R. etli* were found to group with *R. phaseoli* in the phylogenetic analysis of housekeeping genes *recA*, *atpD* and *rpoB* (Fig. 2). Furthermore, the genomic-based phylogenetic analyses

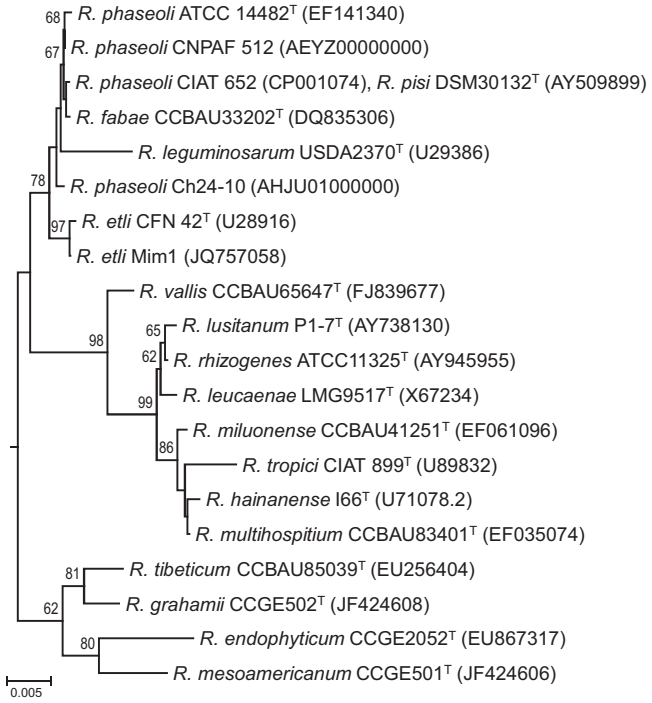



**Fig. 1.** Neighbor-joining phylogenetic tree based on nearly full length 16S rRNA gene sequences from *Rhizobium* and related genera. Only bootstrap supports larger than 50% are indicated. Bar, 5 nt substitutions per 1000 nt.

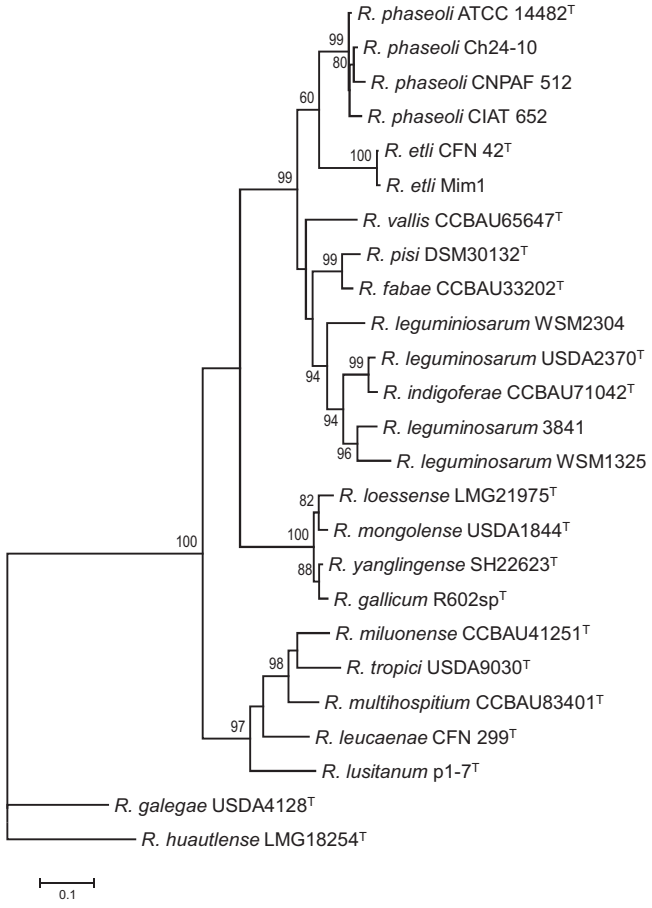



**Fig. 2.** Maximum likelihood phylogenetic tree based on partial concatenated sequences of *recA*, *atpD* and *rpoB* genes of *Rhizobium etli*, *Rhizobium phaseoli* and other rhizobia. Only bootstrap supports larger than 50% are indicated. Bar, 1 nt substitution per 10 nt.

**Table 2**
Nucleotide sequence identity for five complete core genes between pairs of sister *Rhizobium* species.

| Gene | Re/Rp | Re/Rl | Rr/Rt |
|---|---|---|---|
| *dnaK* | 95.3–95.8 | 94 | 93.4 |
| *recA* | 91.7–92.3 | 90.9 | 92.5 |
| *atpD* | 94.3–94.5 | 90.7 | 91.5 |
| *rpoB* | 95.2–95.4 | 94.3 | 93.6 |
| *celC* | 87.2–87.7 | 85.2 | 71.4 |

Re, *R. etli* CFN 42$^T$; Rp, *R. phaseoli* CIAT 652, Ch24-10 and CNPAF 512; Rl, *R. leguminosarum* 3841; Rr, *R. rhizogenes* K84; Rt, *R. tropici* CIAT 899$^T$.

clearly showed that Ch24-10, CNPAF 512 and CIAT 652 strains constituted a well-supported group separated from *R. etli* CFN 42$^T$ (Fig. 3).

Single gene comparisons showed that there were slightly less differences between each of the strains in the group (CIAT 652, CNPAF 512 or Ch24-10) and *R. etli* CFN 42$^T$ than those found for the corresponding genes between *Rhizobium* sister species, such as *R. tropici*-*R. rhizogenes* or *R. leguminosarum*-*R. tropici* (Table 2). The whole core genes compared were *dnaK*, *recA*, *atpD*, *rpoB* and *celC*. In particular, *dnaK*, which has been used as a marker to distinguish species in rhizobia [33], was 94% identical between *R. etli* and *R. leguminosarum*, 93.4% between *R. tropici* and *R. rhizogenes*, and slightly more conserved (95.3–95.8%) between each of the strains (CIAT 652, CNPAF 512 or Ch24-10) and *R. etli* CFN 42$^T$.

### *DNA–DNA hybridization*

DDH between strains CIAT 652 and CFN 42$^T$ was 68%, which was lower than the estimated value for species differentiation (Table 3). DDH between strains CNPAF 512 and CFN 42$^T$ was 67%. Results of the experimental DDH of Ch24-10 to CFN 42$^T$ were 75% (Table 3). Ch24-10 and CFN 42$^T$ showed DDH values sufficient to be considered as belonging to the same species, *R. etli*.

Our experimental DDH results showed values well over 70% for ATCC 14482$^T$ with CIAT 652 (84%), Ch24-10 (88%) and CNPAF 512 (84%), and they constituted the most important results for allocating these strains to *R. phaseoli*.

### *ANIm estimations*

The *R. etli* CFN 42$^T$ and either Ch24-10 or CNPAF 512 estimated ANIm was 90.5% (Table 4), which was in agreement with that previously estimated for CFN 42$^T$ and CIAT 652 by Richter and Roselló-Móra [32]. We estimated the *R. etli* CFN 42$^T$ and *R. leguminosarum* 3841 ANIm as 87.9%. ANIm values were also calculated for other strains with available partial genome sequences that are presently classified as *R. etli* (Table 4) [9,11]. It has been shown that sequences covering 20% of the genome size are enough to obtain reliable estimates of ANIm [32]. Strains 8C-3 and Brasil5 showed ANIm values ≥96.7% with CIAT 652, Ch24-10 and CNPAF

**Table 3**
Experimental DNA–DNA hybridization.

| | *R. etli* CFN 42$^T$ | *R. phaseoli* CNPAF 512 | *R. phaseoli* ATCC 14482$^T$ |
|---|---|---|---|
| *R. etli* CFN 42$^T$ | 100 | 67 ± 3 | 70 ± 2 |
| *R. phaseoli* CIAT 652 | 68 ± 3 | 76 ± 4 | 84 ± 3 |
| *R. phaseoli* Ch24-10 | 75 ± 6 | 78 ± 0 | 88 ± 6 |
| *R. phaseoli* CNPAF 512 | 67 ± 3 | 100 | 84 ± 2 |
| *R. phaseoli* ATCC 14482$^T$ | 69 ± 4 | 78 ± 5 | 100 |

In the experimental procedure radioactive DNA was from CFN 42$^T$, CNPAF 512 or ATCC 14482$^T$.
± standard error.

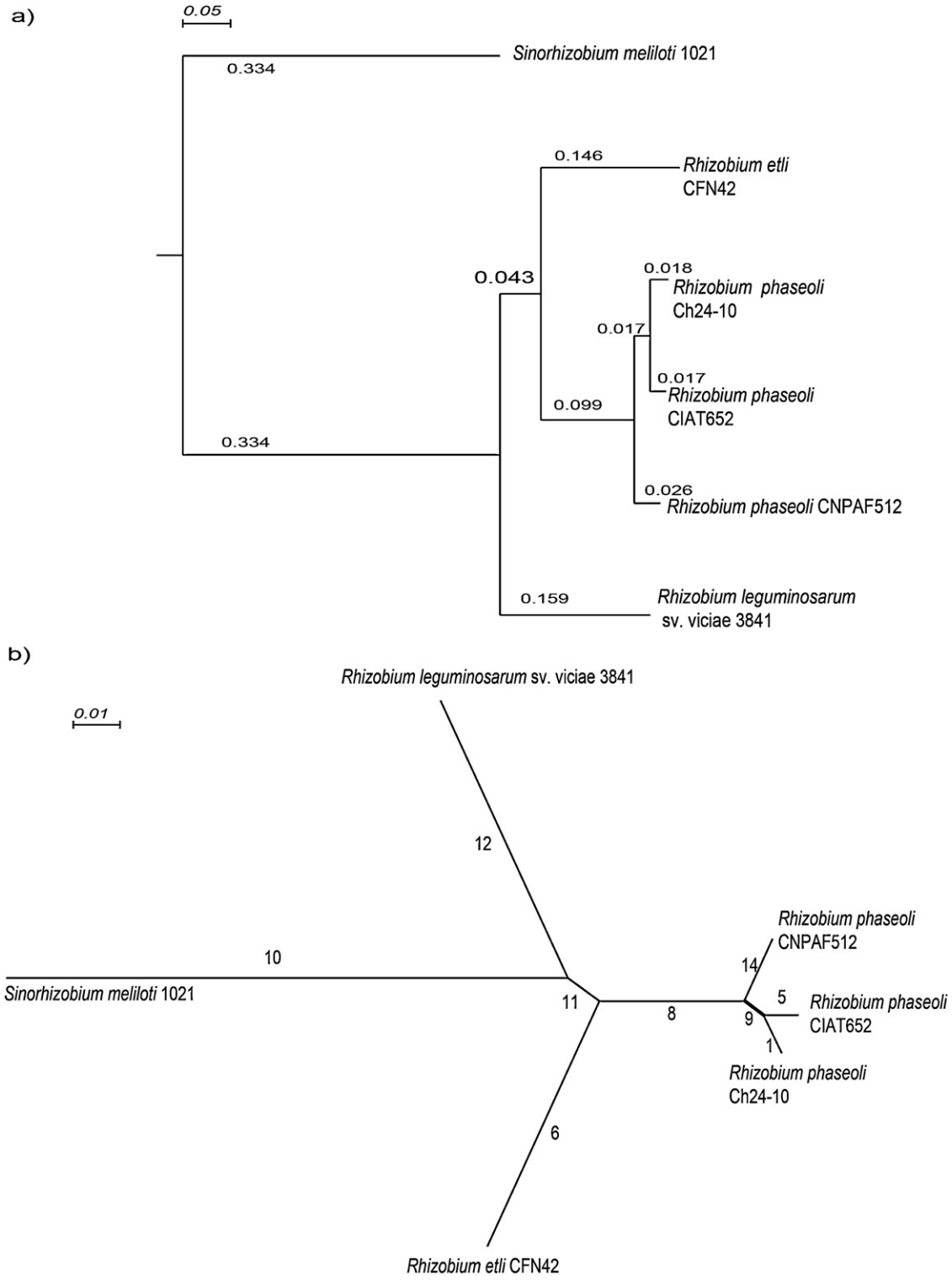



**Fig. 3.** Phylogenetic trees of concatenated genomic conserved regions (3,441,383 pb) in 1373 (coding and non-coding) fragments, on average 2.5 kb. (a) Rooted tree by maximum likelihood with *Sinorhizobium meliloti* 1021 as an outgroup. The scale bar denotes the expected number of nucleotide substitutions per site. (b) Network joining phylogeny inferred from the same concatenated fragments. The tree is unrooted and the number on each branch denotes the split decomposition according to the consensus. The scale bar denotes the number of nucleotide substitutions per site.

512 indicating that they may also belong to *R. phaseoli*. Strains IE4771 and Kim5 shared a high ANIm value between them (96.9%) and may constitute a separate species from *R. etli* and *R. phaseoli* with low ANIm values to both species (89.8–92.3). Strain GR56 may belong to another species, although it is closely related to strains IE4771 and Kim5 with ANIm values of 93–93.4%. Finally, strain CIAT 894 showed low ANI values with all strains indicating that it belonged to another independent species. The same relationships were previously noted in a genomic tree reported for the same strains [11].

**Table 4**
ANIm values between different bean nodulating *Rhizobium* strains with complete or partial genome sequences available.

| | Strain | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | CFN 42$^T$ | 100 | | | | | | | | | |
| 2 | CIAT 652 | 90.5 | 100 | | | | | | | | |
| 3 | Ch24-10[a] | 90.5 | 98.1 | 100 | | | | | | | |
| 4 | CNPAF 512[a] | 90.5 | 97.3 | 97.1 | 100 | | | | | | |
| 5 | 8C-3[a] | 90.9 | 98.2 | 97.6 | 97.1 | 100 | | | | | |
| 6 | Brasil5[a] | 90.7 | 97 | 96.8 | 97.5 | 96.7 | 100 | | | | |
| 7 | IE4771[a] | 89.8 | 91.8 | 91.7 | 91.8 | 91.8 | 91.8 | 100 | | | |
| 8 | Kim5[a] | 90.4 | 92.3 | 92.3 | 92.3 | 92.4 | 92.4 | 96.9 | 100 | | |
| 9 | GR56[a] | 90.6 | 92.8 | 92.7 | 92.9 | 92.8 | 92.7 | 93 | 93.4 | 100 | |
| 10 | CIAT 894[a] | 89.5 | 90.2 | 89.9 | 90 | 90.7 | 90.5 | 89.3 | 90 | 90.3 | 100 |

[a] Based on partial genome sequence. The minimum percentage of genome sequenced was 62.6% for strain 8C-3 [11].

## Discussion

Hybridization to *R. etli* CFN 42$^T$ showed that strains CIAT 652, Ch24-10 and CNPAF 512 are at the borderline that distinguishes different species. It is worth considering that *R. phaseoli* and *R. etli* are barely distinguished by 16S rRNA gene phylogenies and based on this marker they were all recognized as *R. etli*. A discrepancy was shown between the hybridization and the phylogenetic results in which borderline strains (in the DDH analysis) are clearly not *R. etli* but are in fact *R. phaseoli*. In other cases, borderline strains have been considered to correspond to subspecies within a single species [7]. However, there are examples of new species described with borderline values of DDH [46] and it has been recommended that bacterial species comprising distinct genetic subgroups should be split into different species [1].

The conclusion that strains Ch24-10, CIAT 652 and CNPAF 512 are not *R. etli* but belong to *R. phaseoli* is supported by the following considerations: previous [32] and new ANIm analysis, such as the low 90.5% calculated for *R. etli* CFN 42$^T$ and CIAT 652, Ch24-10 or CNPAF 512; the suggested low recombination of CIAT 652, 8C-3 and Brasil5 with *R. etli* CFN 42$^T$ [2] that indicates different evolutionary histories; and new results on the genomic-based phylogenetic analysis and the experimental DNA–DNA hybridization that showed Ch24-10, CIAT 652 and CNPAF 512 were closely related to *R. phaseoli* type strain ATCC 14482$^T$. Such a conclusion supports the previous claim that CIAT 652 corresponded to *R. phaseoli* based on partial *recA*, *atpD* and *celC* gene sequences [34]. From a practical point of view, a phylogenetically based decision with genes other than ribosomal genes seems congruent with the type of data that are normally used to identify strains in rhizobial studies.

Common genes encountered in *R. etli* CFN 42$^T$ and in different *R. phaseoli* strains (data not shown) and the close relatedness of some core genes in *R. phaseoli* and *R. etli* (Table 2) may be explained as a result of an ongoing speciation process that has not been as long as the divergence process occurring among the sister species *R. leguminosarum*–*R. etli* or *R. rhizogenes*–*R. tropici*. As speciation is occurring, many characteristics may still be common between *R. etli* and *R. phaseoli*, such as the 16S ribosomal gene sequences, the physicochemical properties of 705 syntenic ortholog products from CFN 42$^T$ and CIAT 652 that showed an almost identical pattern of polarity between the two strains [26], their sharing of lytic phages (V. Gonzalez and G. Davila, unpublished data) or their close relatedness in MALDI analyses [10] or in MLEE [44]. It is also remarkable that there are very few distinct phenotypes reported that distinguish *R. etli* CFN 42$^T$ from *R. phaseoli* ATCC 14482$^T$ [30]. *R. etli* and *R. phaseoli* speciation occurs within the same environment, since both species are sympatric and nodulate bean. Curiously, the *R. phaseoli* strains analyzed (CIAT 652, Ch24-10, CNPAF 512 and Brasil5) have a better capacity to fix nitrogen in bean than *R. etli* CFN 42$^T$ (data not shown). It seems that Mexican strains *R. phaseoli* Ch24-10 and *R. etli* CFN 42$^T$ are slightly more similar (experimental DDH values over 70%; Table 2) than the other *R. phaseoli* strains from distant geographic regions, such as Colombia (CIAT 652) or Brazil (CNPAF 512), and this may be in relation to the possibility of local exchange of genetic material, especially symbiotic plasmids. In agreement with this, the *R. phaseoli* strain Ch24-10 *nodC* gene but not the *fixL* gene is more similar to that of *R. etli* CFN 42$^T$ than to that of *R. phaseoli* CIAT 652 (data not shown).

Methods to analyze bacterial populations used in the past, such as MLEE, showed that *R. etli* was highly diverse and it was discussed that some of the most distant groups could be considered as different species (Fig. 2 in [27]). New genomic-based analysis using ANIm confirms this observation (Table 4) and indicates that lineages within *R. etli*, such as that represented by CIAT 894 from Colombia or IE4771 and Kim5 or GR56, correspond to novel undescribed species. *R. etli*, *R. phaseoli* and related lineages provide novel examples of the "sequence-discrete groups" that "reflect presumably some fundamental properties of the microbial world" [18].

## Acknowledgments

To PAPIIT IN2054 from UNAM, CONACyT CB 131499, the SUBNARGEM (Subsistema Nacional de Recursos Genéticos Microbianos) grant from SAGARPA Mexico for financial support, and Michael Dunn for kindly reviewing the manuscript. Martha G. López-Guerrero acknowledges the CONACyT fellowship during her Ph.D. studies in the Doctorado en Ciencias Biomédicas program at UNAM.

## References

[1] Achtman, M., Wagner, M. (2008) Microbial diversity and the genetic nature of microbial species. Nat. Rev. Microbiol. 6, 431–440.
[2] Acosta, J.L., Eguiarte, L.E., Santamaría, R.I., Bustos, P., Vinuesa, P., Martínez-Romero, E., Dávila, G., González, V. (2011) Genomic lineages of *Rhizobium etli* revealed by the extent of nucleotide polymorphisms and low recombination. BMC Evol. Biol. 11 (305), 11–305.
[3] Angiuoli, S.V., Salzberg, S.L. (2010) Mugsy: fast multiple alignment of closely related whole genomes. Bioinformatics 27, 334–342.
[4] Aserse, A.A., Räsänen, L.A., Assefa, F., Hailemariam, A., Lindström, K. (2012) Phylogeny and genetic diversity of native rhizobia nodulating common bean (*Phaseolus vulgaris* L.) in Ethiopia. Syst. Appl. Microbiol. 35, 120–131.
[5] Assefa, S., Keane, T.M., Otto, K.T., Newbold, C., Berriman, M. (2009) ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics 25, 1968–1969.
[6] Beyene, D., Kassa, S., Ampy, F., Asseffa, A., Gebremedhin, T., van Berkum, P. (2004) Ethiopian soils harbor natural populations of rhizobia that form symbioses with common bean (*Phaseolus vulgaris* L.). Arch. Microbiol. 181, 129–136.
[7] Borriss, R., Chen, X.H., Rueckert, C., Blom, J., Becker, A., Baumgarth, B., Fan, B., Pukall, R., Schumann, P., Spröer, C., Junge, H., Vater, J., Pühler, A., Klenk, H.P. (2011) Relationship of *Bacillus amyloliquefaciens* clades associated with strains DSM 7T and FZB42T: a proposal for *Bacillus amyloliquefaciens* subsp. *amyloliquefaciens* subsp. nov. and *Bacillus amyloliquefaciens* subsp. *plantarum* subsp. nov. based on complete genome sequence comparisons. Int. J. Syst. Evol. Microbiol. 61, 1786–1801.
[8] Capella-Gutierrez, S., Silla-Martinez, J.M. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.
[9] Fauvart, M., Sánchez-Rodríguez, A., Beullens, S., Marchal, K., Michiels, J. (2011) Genome sequence of *Rhizobium etli* CNPAF512, a nitrogen-fixing symbiont isolated from bean root nodules in Brazil. J. Bacteriol. 193, 3158–3159.

[10] Ferreira, L.L., Sánchez-Juanes, F.F., García-Fraile, P.P., Rivas, R.R., Mateos, P.F., Martínez-Molina, E., González-Buitrago, J.M., Velázquez, E. (2011) MALDI-TOF mass spectrometry is a fast and reliable platform for identification and ecological studies of species from family *Rhizobiaceae*. PloS One 6, e20223.
[11] González, V., Acosta, J.L., Santamaría, R.I., Bustos, P., Fernández, J.L., Hernández González, I.L., Díaz, R., Flores, M., Palacios, R., Mora, J., Dávila, G.I.L. (2010) Conserved symbiotic plasmid DNA sequences in the multireplicon pangenomic structure of *Rhizobium etli*. Appl. Environ. Microbiol. 76, 1604–1614.
[12] Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., Tiedje, J.M. (2007) DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. 57, 81–91.
[13] Grange, L., Hungria, M. (2004) Genetic diversity of indigenous common bean (*Phaseolus vulgaris*) rhizobia in two Brazilian ecosystems. Soil Biol. Biochem. 36, 1389–1398.
[14] Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
[15] Huson, D.H., Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267.
[16] Kaschuk, G., Hungria, M., Andrade, D.S., Campo, R.J. (2006) Genetic diversity of rhizobia associated with common bean (*Phaseolus vulgaris* L.) grown under no-tillage and conventional systems in Southern Brazil. Appl. Soil Ecol., 32210–32220.
[17] Kislyuk, A.O., Katz, L.S., Agrawal, S., Hagen, M.S., Conley, A.B., Jayaraman, P., Nelakuditi, V., Humphrey, J.C., Sammons, S.A., Govil, D., Mair, R.D., Tatti, K.M., Tondella, M.L., Harcourt, B.H., Mayer, L.W., Jordan, I.K. (2010) A computational genomics pipeline for prokaryotic sequencing projects. Bioinformatics 26, 1819–1826.
[18] Konstantinidis, K.T. (2011) Metagenomic insights into bacterial species. In: De Bruijn, F.J. (Ed.), Handbook of Molecular Microbial Ecology. II: Metagenomics in Different Habitats, John Wiley & Sons, Inc., Hoboken, NJ, USA.
[19] Konstantinidis, K.T., Ramette, A., Tiedje, J.M. (2006) The bacterial species definition in the genomic era. Philos. Trans. R. Soc. Lond. Ser. B: Biol. Sci. 361, 1929–1940.
[20] Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., Salzberg, S.L. (2004) Versatile and open software for comparing large genomes. Genome Biol. 5, R12.
[21] Lloret, L., Ormeno-Orrillo, E., Rincon, R., Martinez-Romero, J., Rogel-Hernandez, M.A., Martínez-Romero, E. (2007) *Ensifer mexicanum* sp. nov. a new species nodulating *Acacia angustissima* (Mill.) Kuntze in Mexico. Syst. Appl. Environ. 30, 280–290.
[22] Martinez-Romero, E. (2003) Diversity of *Rhizobium–Phaseolus vulgaris* symbiosis: overview and perspectives. Plant Soil 252, 11–23.
[23] Michiels, J., D'Hooghe, I., Verreth, C., Pelemans, H., Vanderleyden, J. (1994) Characterization of the *Rhizobium leguminosarum* biovar *phaseoli nifA* gene, a positive regulator of *nif* gene expression. Arch. Microbiol. 161, 404–408.
[24] Mnasri, B., Mrabet, M., Laguerre, G., Aouani, M.E., Mhamdi, R. (2007) Salt-tolerant rhizobia isolated from a Tunisian oasis that are highly effective for symbiotic $N_2$-fixation with *Phaseolus vulgaris* constitute a novel biovar (bv. *mediterranense*) of *Sinorhizobium meliloti*. Arch. Microbiol. 187, 79–85.
[25] Ormeño-Orrillo, E., Rogel-Hernández, M.A., Lloret, L., López-López, A., Martínez, J., Barois, I., Martínez-Romero, E. (2012) Change in land use alters the diversity and composition of *Bradyrhizobium* communities and led to the introduction of *Rhizobium etli* into the tropical rain forest of Los Tuxtlas (Mexico). Microb. Ecol. 63, 822–834.
[26] Peralta, H., Guerrero, G., Aguilar, A., Mora, J. (2011) Sequence variability of Rhizobiales orthologs and relationship with physico-chemical characteristics of proteins. Biol. Direct 6, 48.
[27] Pinero, D., Martinez, E., Selander, R.K. (1988) Genetic diversity and relationships among isolates of *Rhizobium leguminosarum* biovar *phaseoli*. Appl. Environ. Microbiol. 54, 2825–2832.
[28] Posada, D. (2008) jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
[29] Quinto, C., de la Vega, H., Flores, M., Fernández, L., Ballado, T., Soberón, G., Palacios, R. (1982) Reiteration of nitrogen fixation gene sequences in *Rhizobium phaseoli*. Nature 299, 724–726.
[30] Ramírez-Bahena, M.H., García-Fraile, P., Peix, A., Valverde, A., Rivas, R., Igual, J.M., Mateos, P.F., Martínez-Molina, E., Velázquez, E. (2008) Revision of the taxonomic status of the species *Rhizobium leguminosarum* (Frank 1879) Frank 1889AL *Rhizobium phaseoli* Dangeard 1926AL and *Rhizobium trifolii* Dangeard 1926AL. *R. trifolii* is a later synonym of *R. leguminosarum*. Reclassification of the strain *R. leguminosarum* DSM 30132 (=NCIMB 11478) as *Rhizobium pisi* sp. nov. Int. J. Syst. Evol. Microbiol. 58, 2484–2490.
[31] Ribeiro, R.A., Rogel, M.A., López-López, A., Ormeño-Orrillo, E., Gomes Barcellos, F., Martínez, J., Lopes Thompson, F., Martínez-Romero, E., Hungria, M. (2012) Reclassification of *Rhizobium tropici* type A strains as *Rhizobium leucaenae* sp. nov. Int. J. Syst. Evol. Microbiol. 62, 1179–1184.
[32] Richter, M., Rosselló-Móra, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. U.S.A. 106, 19126–19131.
[33] Rivas, R., Martens, M., de Lajudie, P., Willems, A. (2009) Multilocus sequence analysis of the genus *Bradyrhizobium*. Syst. Appl. Microbiol. 32, 101–110.
[34] Robledo, M., Velázquez, E., Ramírez-Bahena, M.H., García-Fraile, P., Pérez-Alonso, A., Rivas, R., Martínez-Molina, E., Mateos, P.F. (2011) The *celC* gene, a new phylogenetic marker useful for taxonomic studies in *Rhizobium*. Syst. Appl. Microbiol. 34, 393–399.
[35] Rosenblueth, M., Martinez Romero, E. (2004) *Rhizobium etli* maize populations and their competitiveness for root colonization. Arch. Microbiol. 181, 337–344.
[36] Schmidt, B., Sinha, R., Beresford-Smith, B., Puglisi, S.J. (2009) A fast hybrid short read fragment assembly algorithm. Bioinformatics 25, 2279–2280.
[37] Silva, C., Vinuesa, P., Eguiarte, L.E., Souza, V., Martinez-Romero, E. (2005) Evolutionary genetics and biogeographic structure of *Rhizobium gallicum sensu lato*, a widely distributed bacterial symbiont of diverse legumes. Mol. Ecol. 14, 4033–4050.
[38] Stamatakis, A., Ludwing, T., Meier, H. (2005) RaxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 21, 456–463.
[39] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
[40] Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., DeBoy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit, Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sullivan, S.A., Daugherty, S.C., Haft, D.H., Selengut, J., Gwinn, M.L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K.J.B., Smith, S., Utterback, T.R., White, O., Rubens, C.E., Grandi, G., Madoff, L.C., Kasper, D.L., Telford, J.L., Wessels, M.R., Rappuoli, R., Fraser, C.M. (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc. Natl. Acad. Sci. U.S.A. 102, 13950–13955.
[41] Toledo, I., Lloret, L., Martinez-Romero, E. (2003) *Sinorhizobium americanum* sp. nov., a new *Sinorhizobium* species nodulating native *Acacia* spp. in Mexico. Syst. Appl. Microbiol. 26, 54–64.
[42] Uribe, L. (1993) Evaluación de medios para la selección de cepas de *Rhizobium leguminosarum* bv. *phaseoli* tolerantes a baja concentración de fosfato en medio de cultivo. Agron. Costarric. 17, 103–109.
[43] Valverde, A., Igual, J.M., Peix, A., Cervantes, E., Velazquez, E. (2006) *Rhizobium lusitanum* sp. nov. a bacterium that nodulates *Phaseolus vulgaris*. Int. J. Syst. Evol. Microbiol. 56, 2631–2637.
[44] Wang, E.T., Rogel, M.A., Garcia-De los Santos, A., Martinez-Romero, J., Cevallos, M.A., Martínez-Romero, E. (1999) *Rhizobium etli* bv. *mimosae*, a novel biovar isolated from *Mimosa affinis*. Int. J. Syst. Bacteriol. 49, 1479–1491.
[45] Warren, R.L., Sutton, G.G., Jones, S.J.M., Holt, R.A. (2007) Assembling millions of short DNA sequences using SSAKE. Bioinformatics 23, 500–501.
[46] Wink, J.M., Kroppenstedt, R.M., Ganguli, B.N., Nadkarni, S.R. (2003) Three new antibiotic producing species of the genus *Amycolatopsis*, *Amycolatopsis balhimydna* sp. nov., *A. tolypomycina* sp. nov., *A. vancoresmycina* sp. nov., and description of *Amycolatopsis keratiniphila* subsp. *keratiniphila* subsp. nov. and *A. keratiniphila* subsp. *nogabecina* subsp. nov. Syst. Appl. Microbiol. 26, 38–46.

Plasmid 68 (2012) 149–158

Contents lists available at SciVerse ScienceDirect

Plasmid

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

Review

# Rhizobial extrachromosomal replicon variability, stability and expression in natural niches

Martha G. López-Guerrero [a], Ernesto Ormeño-Orrillo [a], José Luis Acosta [a,1], Alfredo Mendoza-Vargas [b,1], Marco A. Rogel [a], Miguel Angel Ramírez [a], Mónica Rosenblueth [a], Julio Martínez-Romero [a], Esperanza Martínez-Romero [a,*]

[a] Centro de Ciencias Genómicas, UNAM, Av. Universidad SN, Cuernavaca, Morelos, Mexico
[b] Instituto de Biotecnología, UNAM, Av. Universidad 2001, Cuernavaca, Morelos, Mexico

ARTICLE INFO

*Article history:*
Received 30 March 2012
Accepted 6 July 2012
Available online 16 July 2012
Communicated by Eva Top

*Keywords:*
Plasmids
Plasmid instability
Symbiotic plasmids
*Rhizobium*
*Sinorhizobium*
*Ensifer*

ABSTRACT

In bacteria, niche adaptation may be determined by mobile extrachromosomal elements. A remarkable characteristic of *Rhizobium* and *Ensifer* (*Sinorhizobium*) but also of *Agrobacterium* species is that almost half of the genome is contained in several large extrachromosomal replicons (ERs). They encode a plethora of functions, some of them required for bacterial survival, niche adaptation, plasmid transfer or stability. In spite of this, plasmid loss is common in rhizobia upon subculturing. Rhizobial gene-expression studies in plant rhizospheres with novel results from transcriptomic analysis of *Rhizobium phaseoli* in maize and *Phaseolus vulgaris* roots highlight the role of ERs in natural niches and allowed the identification of common extrachromosomal genes expressed in association with plant rootlets and the replicons involved.



**Contents**



* Corresponding author.
*E-mail address:* emartine@ccg.unam.mx (E. Martínez-Romero).
[1] Present address: Instituto Nacional de Medicina Genómica. Periférico Sur No. 4809, México, D.F. C.P. 14610, Mexico.

0147-619X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.plasmid.2012.07.002

## 1. Introduction

Rhizobia is a generic name to refer to several genera of α and β-Proteobacteria. Rhizobia are successful legume and non-legume rhizosphere colonizers and form nitrogen fixing nodules in legumes. Rhizobia inhabit the soil and other niches such as seeds (López-López et al., 2010; Pérez-Ramírez et al., 1998) or inside plant tissues as endophytes (Chaintreuil et al., 2000; Gutiérrez-Zamora and Martínez-Romero, 2001; Reiter et al., 2003; Yanni et al., 1997). Legumes that establish symbiosis with rhizobia can colonize nitrogen poor environments, may enrich the soil or require less chemical nitrogen fertilizers as agricultural crops.

In rhizobial research, an outstanding discovery was that symbiosis abilities resided in plasmids that could be lost or transferred among bacteria (Bánfalvi et al., 1981; Hooykaas et al., 1982; Johnston et al., 1978; Nuti et al., 1977; Nuti et al., 1979; Rosenberg et al., 1982; Sutton, 1974; Tshitinge et al., 1975; Zurkowski, 1982; Zurkowski and Lorkiewic, 1976). Symbiotic plasmids are found in *Rhizobium*, *Ensifer* = *Sinorhizobium*, in few *Mesorhizobium* species, in the β-Protobacterium *Cupriavidus taiwanensis* that forms nodules in *Mimosa* species (Amadou et al., 2008) and in *Burkholderia* sp. CCGE 1002 isolated from a nodule of *Mimosa occidentalis* collected in Tepic, Mexico (genome NCBI ID 640511). However symbiotic plasmids are not found in *Bradyrhizobium* (Cytryn et al., 2008; Hahn and Hennecke, 1987; Haugland and Verma, 1981), in *Azorhizobium caulinodans* (Lee et al., 2008) or in most *Mesorhizobium* strains (Wang et al., 1999; Xu and Murooka, 1995; Zou et al., 1997). Nitrogen fixation occurring in nodules may be considered as an ecological service. Genes involved in this process (*nif* genes) are plasmid encoded in *Rhizobium, Ensifer* (*Sinorhizobium*), few *Mesorhizobium* species, *Burkholderia* and *Cupriavidus* strains but located in chromosomes in many bacteria (reviewed in Ormeño-Orrillo et al., in press). In rhizobia, symbiosis variants (symbiovars) are recognized on the basis of host specificity and effectiveness (nitrogen fixation) mainly determined by symbiotic plasmids or islands (Rogel et al., 2011). Reviews on symbiotic plasmids (Romero and Brom, 2004) and on the bacterial and plant functions required during the symbiotic process have been published (Oldroyd et al., 2011; Peix et al., 2010).

Methods to visualize rhizobial plasmids (Eckhardt, 1978; Hirsch et al., 1980; Hynes and McGregor, 1990) were pivotal to the study of their diverse patterns, their stability and for the determination of the plasmid location of symbiosis significant genes. In addition to symbiotic plasmids, different large plasmids or extrachromosomal replicons (ER) are found in nodule forming bacteria. However, only 23% of *Bradyrhizobium japonicum* and *B. elkanii* strains from different geographical regions contained plasmids (Cytryn et al., 2008). The role of plasmids in the Rhizobiaceae focusing on interbacterial and transkingdom interactions was recently reviewed (Pappas and Cevallos, 2011). Different types of ER have been described, such as chromids (Harrison et al., 2010) as well as secondary chromosomes (Slater et al., 2009). Housekeeping and ribosomal genes that are relocated to plasmids may make them look like secondary chromosomes. ER that encode housekeeping or essential functions, stably maintained in bacteria and having a GC content similar to that of the chromosome, have been designated chromids and have been identified from genomic data in several rhizobial strains (Harrison et al., 2010). The definition of essential functions encoded in ER must be reviewed because genes may only be conditionally essential on some media or conditions. For example, a plasmid may be cured in the laboratory and thus be considered non essential but may be essential in soil or in the rhizosphere. On the other hand, use of the curing plasmid strategy to recognize essential genes may lead to erroneous conclusions if essential genes move to other replicons during the plasmid elimination (curing) process and selection of survivors. Genome sequence analysis of cured strains would reveal such events.

## 2. Extrachromosomal replicons in rhizobia, a substantial proportion of their genomes

We will focus mainly on *Rhizobium* with only some references on *Ensifer* and the related *Agrobacterium* genus that includes species forming tumors in plants. A remarkable characteristic of *Rhizobium*, *Ensifer* but also of *Agrobacterium* species is the large amount of genomic DNA contained in ER. From 30% to almost 50% of the genome may be extrachromosomal in symbiotic or pathogenic strains (Table 1). Agrobacterial plasmids were reviewed in Suzuki et al. (2009). Although ER may represent a burden for bacterial growth in some cases, this is not the case with rhizobial plasmids. On the contrary, they are important for bacterial physiology as has been shown for *Rhizobium etli* CFN 42 in which strains cured of most of the plasmids had larger duplication times (Brom et al., 1992). Furthermore, ER may contribute significantly to the phenotype and to the bacterial pangenome, the whole species genome.

Most rhizobial ERs are large and in low copy number. Rhizobial strains have several ERs (Table 1 in Romero and Brom, 2004), up to 11 in *R. leguminosarum*. Agrobacteria, *R. galegae*, *R. phaseoli*, *R. tropici* and *R. gallicum* seem to have fewer, 2–4. In rhizobia and in other α-Proteobacteria most ERs have *repABC* replication systems (Cervantes-Rivera et al., 2011; Pappas and Cevallos, 2011). A 7.2 kb plasmid with rolling circle replication was described in an *E. meliloti* strain but small size plasmids are uncommon in rhizobia

**Table 1**
Size and percent of extrachromosomal genome in rhizobia and related strains with completely sequenced genomes.

| Strain | Genome size (Mb) | Percent in extrachromosomal replicons (%) |
|---|---|---|
| *Rhizobium tropici* CIAT 899 | 6.69 | 42.6 |
| *Rhizobium etli* CFN 42 | 6.53 | 32.9 |
| *Rhizobium phaseoli* CIAT 652 | 6.44 | 30.1 |
| *Rhizobium phaseoli* Ch24-10 | 6.63 | 32.0 |
| *Rhizobium leguminosarum* 3841 | 7.79 | 34.5 |
| *Rhizobium leguminosarum* WSM1325 | 7.45 | 35.6 |
| *Rhizobium leguminosarum* WSM2304 | 6.87 | 34.0 |
| *Rhizobium rhizogenes* (*Agrobacterium radiobacter*) K84 | 7.31 | 44.7 |
| *Agrobacterium tumefaciens* C58 | 5.65 | 50.0[a] |
| *Agrobacterium vitis* S4 | 6.31 | 41.0[a] |
| *Ensifer meliloti* (*Sinorhizobium meliloti*)1021 | 6.80 | 44.9 |
| *Ensifer* sp. NGR234 (*Sinorhizobium* sp.) | 6.90 | 43.0 |
| *Ensifer medicae* WSM419 (*Sinorhizobium medicae*) | 6.82 | 44.5 |

[a] Including the secondary chromosome that has ribosomal genes but an origin of replication typical of plasmids.

(Barran et al., 2001). ER sizes in *Rhizobium* and *Ensifer* (*Sinorhizobium*) are in the range of 45 kb to around 2.5 Mb.

## 3. Rhizobial hypervariable genome is in extrachromosomal elements

Chromosomes are more conserved than ER both at the gene sequence and synteny levels (Guerrero et al., 2005). Plasmid patterns are different even within a single rhizobial species (Rosenblueth and Martínez Romero, 2004; Wang et al., 1999). This is particularly evident among *R. etli* and *R. leguminosarum* strains but less variability has been observed in *R. tropici*, *R. phaseoli* or *Ensifer* plasmid profiles (not shown). Plasmid pattern differences suggest that rhizobia may thrive in different environments.

Plasmid gene content variation has been revealed from genomic projects and mosaicism seems to be a common characteristic of plasmids (Cervantes et al., 2011) and symbiotic plasmids (Freiberg et al., 1997; González et al., 2003). Recombination was evidenced with a PCR approach in *Rhizobium etli* plasmids (Flores et al., 2005). Plasmids seem to be prone to pick up novel genes or to suffer deletions. How are plasmids assembled or disassembled? Once a successful plasmid is arranged it may be stably maintained even in distinct chromosomal backgrounds over time (Crossman et al., 2008).

Duplicated copies from chromosomal genes have been allocated to plasmids. In *R. tropici* and in *R. leucaenae* a duplicated citrate synthase gene is found in the symbiotic plasmid, conditioning nodulation (Pardo et al., 1994) and differentially regulated from the chromosomal copy (Hernández-Lucas et al., 1995). Glucosamine synthase (*nodM*) duplicated genes in plasmids (Marie et al., 1992), are needed to provide additional substrates for Nod factor production.

ER may integrate into chromosomes (Guo et al., 2003), rearrange (Brom et al., 1991; Flores et al., 1988, 2000; Soberón-Chávez et al., 1986; Zhang et al., 2001) or form cointegrates with other plasmids (Brom et al., 2004; Cervantes et al., 2011; Guo et al., 2003; Mavingui et al., 2002). Fragments of plasmids may be amplified and in some cases this leads to enhanced nodulation (Mavingui et al., 1997, 1998; Romero et al., 1991; Romero et al., 1995). Extrachromosomal location of genes is not universal and fixed in strains because some genes may be in chromosomes and in other cases in extrachromosomal elements (Crossman et al., 2008 and Fig. 1). There are clues that indicate that some plasmids may be chimeras resulting from the fusion of different plasmids (Cervantes et al., 2011; Ormeño-Orrillo et al., unpublished). Plasmid co-integrates may excise correctly or incorrectly. Plasmids seem to be more dynamic than chromosomes and equivalent genes found among distinct ER in related species are evidence of extensive plasmid rearrangements (Fig. 1, Fig. 6 in Crossman et al., 2008).

## 4. Instability and stability of extrachromosomal elements

Plasmid instability has been known for a long time and it has been recommended to avoid the practice of single colony isolation when purifying rhizobia especially for inoculant production as they may lose relevant plasmids (Weaver and Wright, 1987). Absence of symbiotic plasmids is remarkable as rhizobial natural populations without symbiotic plasmids lose their access to legume nodules, however *R. etli* strains lacking Sym plasmids seem to be very successful rhizospheric or endophytic colonizers (López-López et al., 2010; Segovia et al., 1991). As plasmids encode carbon assimilation genes, rhizobia may change phenotype in one step when losing or gaining plasmids. After *R. etli* CFN42 was resequenced to test Illumina sequencing facilities at UNAM, it was evident that plasmid pReCFN42a was lost in the cultured cells grown to extract DNA (González and Lozano, personal communication) while the original stock maintained the whole set of plasmids. Some rhizobial strains when subcultured in the lab were prone to lose their plasmids (Weaver et al., 1990). A Tn5 had to be inserted in CFN 23 symbiotic plasmid to exert a selective pressure to maintain the plasmid in this *Rhizobium* strain (Soberón-Chávez and Nájera, 1989).

Instability has also been observed in *Burkholderia* strain CCGE 1001 isolated in our laboratory from a nodule of a *Mimosa affinis* plant grown in soils from Acayuca, Veracruz.

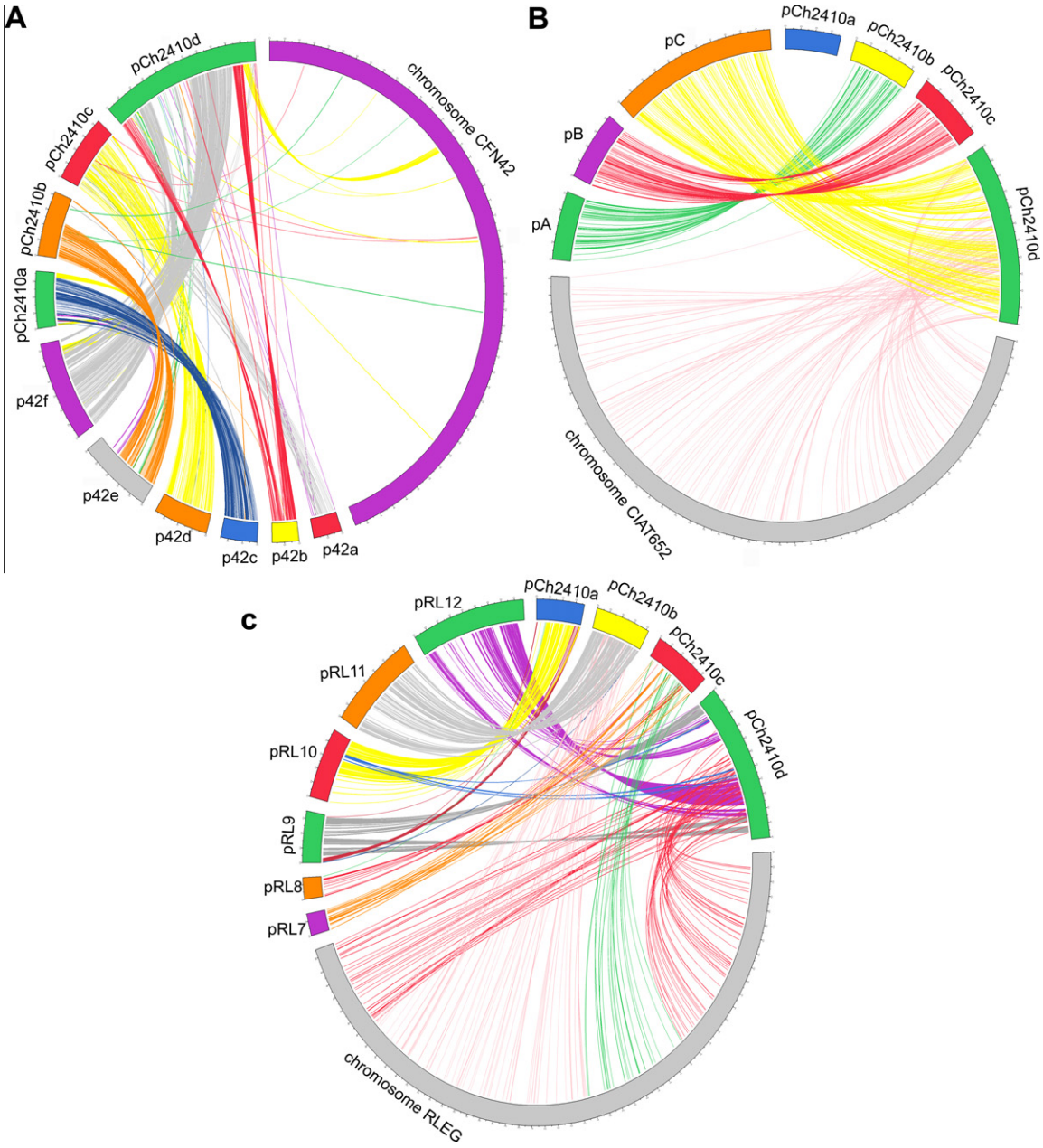



**Fig. 1.** Comparison using satsumasinteny of *R. phaseoli* Ch24-10 extrachromosomal replicons (ERs) to (A) *R. etli* CFN42, (B) *R. phaseoli* CIAT 652 and (C) *R. leguminosarum* 3841 (RLEG) chromosomes and ERs.

Upon subculturing this strain lost its symbiotic plasmid as evidenced from the whole genome analysis (NCBI ID 640510). The original strain is still capable of nodulating *Phaseolus vulgaris* and mimosa plants (unpublished). In another case, when we analyzed the transcripts from *R. phaseoli* strain Ch24-10 (see below) there were none corresponding to a 370 kb plasmid (the smallest, non-symbiotic plasmid) that was revealed in the whole genome analysis of the same strain (López-Guerrero et al., in press). We supposed that the plasmid was lost upon subculturing as the original stock has all plasmids. Our analysis of the published genome of *R. phaseoli* CNPAF512 (Fauvart et al., 2011) revealed sequences corresponding to the 370 kb plasmid from Ch24-10, however these were not found (Fig. 1B) in the published genome of another *R. phaseoli* strain, CIAT 652 (González et al., 2010). This shows that this plasmid is not homogenously conserved among *R. phaseoli* strains. It is worth mentioning that *R. phaseoli* CIAT652 is a very efficient *P. vulgaris* symbiont in spite of lacking this plasmid.

ER maintenance seems to be forced when carrying genes required for growth or survival. This is illustrated in *R. etli*

CFN 42 with pReCFN42e carrying genes needed for growth or optimal growth in rich medium such as those encoding a sensor histidine kinase/ response regulator hybrid protein and a hypothetical protein with a winged helix-turn-helix motif (Landeta et al., 2011) in addition to containing some of the genes for cobalamin biosynthesis. Both genes encoding the sensor histidine kinase/ response regulator hybrid and the hypothetical protein with a winged helix-turn-helix motif are found in *R. leguminosarum* sv. viciae 3841 (in chromid PRL11) and in sv. trifolii strains 1325 and 2304 plasmids as well as in an *R. phaseoli* CIAT 652 plasmid (pRp652a) that corresponds to pReCFN42e.

Toxin–antitoxin genes were discovered as plasmid stabilizers (Jensen and Gerdes, 1995; Ogura and Hiraga, 1983) and have been identified in many bacteria (Pandey and Gerdes, 2005; Van Melderen et al., 2009). Toxin–antitoxin genes have been found in the symbiotic plasmid of *Ensifer* sp. NGR234 (Falla and Chopra, 1999). Antitoxins are more unstable than toxins so when the antitoxin is missing due to plasmid loss, the toxin inhibits cell growth and leads to death (Jensen and Gerdes, 1995). Bacterial genetic mechanisms to ensure plasmid maintenance both in symbionts and pathogens have been reviewed (Sengupta and Austin, 2011).

## 5. Extrachromosomal replicons involved in plant–rhizobium interactions

ERs in addition to the symbiotic plasmids have roles in symbiosis with legumes (Hynes and McGregor, 1990). Curing of a cryptic plasmid in *Ensifer* (*Sinorhizobium*) *meliloti* led to a more efficient symbiosis in alfalfa (Velázquez et al., 1995). In *R. leguminosarum* an exogenous RP4 plasmid decreased symbiotic effectiveness (O'Connell et al., 1998). Enhanced nodulation competitiveness was recorded in *R. etli* strains that gained an *R. leucaenae* (185 kb) plasmid (Martínez-Romero and Rosenblueth, 1990). *A. tumefaciens* transconjugants that in addition to carrying the *nod-nif* plasmid had a 200 kb plasmid from *R. leucaenae* fixed more nitrogen than that with only the symbiotic plasmid (Martínez et al., 1987).

Non symbiotic plasmids participate in rhizobial interactions with plants (Brom et al., 2000; Chen et al., 2000; Hynes and McGregor, 1990; Pappas and Cevallos, 2011). Some *R. leguminosarum* strains capable of associating with rice promoted its growth and alleviated N deficiencies (Yanni et al., 1997), but others from clover inhibited rice root growth. Rice inhibition or promotion is plasmid dependent in *R. leguminosarum* (Perrine et al., 2001) and in *E. meliloti* (Perrine et al., 2005). Derivatives of *R. leguminosarum* sv. trifolii W14–12 lacking two plasmids were unable to grow in soil (Moënne-Loccoz and Weaver, 1995a) and different plasmids were found to contribute to growth in the clover rhizosphere (Moënne-Loccoz and Weaver, 1995b) or in saprophytic life (Moënne-Loccoz et al., 1995). The most competitive maize colonizing *R. phaseoli* strains had the most common plasmid pattern observed among many rhizospheric strains analyzed (Rosenblueth and Martínez Romero, 2004). In *R. leguminosarum* sv. viciae, a plasmid contains several genes needed and expressed by bacterial cells when colonizing the pea rhizosphere (Ramachandran et al., 2011). Similarly we found that extrachromosomal genes were expressed in *R. phaseoli* strain Ch24-10 (Rosenblueth and Martínez Romero, 2004) associated with maize and *P. vulgaris* (common bean) roots (see Section 7).

## 6. Extrachromosomal genes associated with rhizobial environmental adaptation

There is a functional bias in extrachromosomal genes, the ERs tend to contain genes implicated in processes like chemotaxis (Yost et al., 1998) and transport, and they are enriched in elements of external origin (Crossman et al., 2008). Some plasmids, megaplasmids or chromids encode many carbon assimilation genes (Baldani et al., 1992; Oresnik et al., 1998); vitamins like biotin, thiamine or pantothenate (Finan et al., 1986; Miranda-Ríos et al., 1997; Streit et al., 1996; Villaseñor et al., 2011), bacteriocin (Oresnik et al., 1999; Venter et al., 2001), melanin (Hynes et al., 1988) or autoinducer (Schripsema et al., 1996) biosynthetic pathways; and may encode chaperons and modification-restriction systems (Rochepeau et al., 1997). Quorum sensing systems that regulate plasmid transfer or expression of genes in plants may be plasmid encoded in rhizobia (Cubo et al., 1992; Edwards et al., 2009; Lithgow et al., 2000). Reviews on gene functions of plasmids (García-de los Santos and Brom, 1996; Mercado-Blanco and Toro, 1996; Pappas and Cevallos, 2011) and of megaplasmids from *Ensifer* sp. NGR234 (Mavingui, 2009) and *E. meliloti* (Barloy-Hubler and Jebbar, 2009) have been published. Only some functions that we considered important for plant niche colonization will be reviewed here.

### 6.1. Transporters in ERs

In megaplasmid pSymA but especially in pSymB of *Ensifer meliloti* 1021 there are large numbers of transporters (Mauchline et al., 2006) that may allow the bacteria to use different soil nutrients or root exudates. They are inducible by a large number of substrates (Mauchline et al., 2006). Plasmids in *R. etli*, *R. tropici*, *R. leucaenae* and *R. gallicum* sv. phaseoli carry *teu* genes that code for putative sugar ABC transporters involved in the uptake of molecules found in *P. vulgaris* and siratro exudates (Rosenblueth et al., 1998). Four of six quaternary amine transporters that were characterized are located in chromids pRL10 and pRL12 in *R. leguminosarum* 3841 (Fox et al., 2008).

### 6.2. Catabolism

In *E. meliloti putA* genes (for proline catabolism) are involved in rhizobial competitiveness (Van Dillewijn et al., 2002), *putA* is in the chromosome in *E. meliloti* and in *Ensifer* sp. NGR234. *putA* genes are in ER in *R. etli*, *R. phaseoli* and *R. leguminosarum*.

Rhamnose catabolic genes are plasmid borne and inducible (Oresnik et al., 1998). Transport and catabolism of erythriol is plasmid dependent (Geddes et al., 2010; Yost et al., 2006). *R. leguminosarum* mutants in glycerol catabo-

lism have diminished competitiveness. Glycerol uptake and catabolism is plasmid encoded (Ding et al., 2012).

Calystegine catabolism genes are plasmid borne in *E. meliloti* (Guntli et al., 1999; Tepfer et al., 1988). These genes participate in bacterial competitive colonization of non legume rhizozpheres such as those from morning glory plants. Mimosine catabolism genes are also plasmid borne (Borthakur et al., 2003). Opine uptake and catabolism genes reside in the symbiotic megaplasmid a in *E. meliloti* (Murphy et al., 1987). There are also opine catabolizing plasmids in agrobacteria (Bruce et al., 1990).

### 6.3. Surface polysaccharides

Different surface polysaccharides are needed in rhizobial attachment to roots (Downie, 2010) and genes for their biosynthesis are located in different bacterial replicons. Some lipopolysaccharide (LPS) biosynthetic genes have been found in *R. etli* plasmids (García-de los Santos and Brom, 1997). Biosynthetic genes for exopolysaccharides reside in megaplasmid b in *E. meliloti* (Finan et al., 1986) and also in megaplasmids of other rhizobia (Skorupska et al., 2006). Megaplasmid a of *Ensifer* sp. NGR234 encodes flavonoid-inducible genes required for the biosynthesis of a rhamnose-rich LPS produced only inside nodules and that is required for symbiosis (Broughton et al., 2006).

### 6.4. Hormone biosynthesis and protein secretion

Upon inspection of reported genomes we found genes that seem to be involved in gibberellin biosynthesis located in the symbiotic plasmids of *E. fredii* HH103 and *Ensifer sp.* NGR234, *R. etli* CFN42, *R. phaseoli* CIAT 652, *R. tropici* CIAT 899 and in the symbiosis islands of *B. japonicum* USDA 6, *Mesorhizobium loti* R7A, and *M. huakuii* MAFF303099. These genes were originally described in *Bradyrhizobium japonicum* USDA 110 (Morrone et al., 2009) and are not present in the reported genomes of *E. meliloti* and *R. leguminosarum* strains. Gibberellins have diverse effects on plants and its balance in relation to auxins affects plant growth (Brian, 2008). Rhizobial mutants in these genes have not been tested in their hosts. ACC deaminases that modulate ethylene levels are encoded in symbiosis islands in mesorhizobial strains R7A and MAFF303099 (Conforte et al., 2010) and in the symbiotic plasmid of *R. tropici* (Ormeño-Orrillo et al., unpublished). Genes for different auxin biosynthetic pathways are plasmidic in NGR234 (Theunis et al., 2004) and in *R. tropici* CIAT 899 and they are flavonoid inducible (Theunis et al., 2004; Ormeño-Orrillo et al., unpublished).

Rhizobia use different types of secretion systems (excellently reviewed in Downie, 2010). Type III secretion systems (T3SS) are found in several *Rhizobium* and *Ensifer* strains (Marie et al., 2001), these genes are in the symbiotic plasmid in *Ensifer* sp. strain NGR234 and mutants in this system have altered plant specificity. NGR234 T3SS genes are inducible and expressed in the presence of flavonoids (Viprey et al., 1998). A T3SS cluster is also present in the phaseoli symbiotic plasmid (González et al. 2006). Genes coding for Type 1 and 5 secretion systems are found in megaplasmids in *R. tropici* (Ormeño-Orrillo et al., unpublished).

### 6.5. Other functions

In *R. etli*, genes to tolerate polyphenols are plasmid borne (García-de los Santos et al., 2008). The only *R. etli* CFN42 catalase is located in a large ER (pReCFN42f) and is required for bacterial survival in polyphenol rich medium (García-de los Santos et al., 2008). The same replicon carries *nirK* and *norCB* genes for nitrite reduction involved in nitrite detoxification but not in nitrite respiration (Gómez-Hernández et al., 2011). Genes that encode efflux pumps (inducible with bean exudates) that eliminate plant toxic molecules or antibiotics are located in pReCFN42b (184 kb) (González-Pasayo and Martínez-Romero, 2000). The same replicon carries genes for thiamine biosynthesis (Miranda-Ríos et al., 1997).

## 7. Transcriptional profiling of rhizobial ER in natural niches such as the root environment

Are there rhizobial genomic islands or plasmids that are preferentially expressed in the environment? Many stress induced genes that could play a role in the environment are extrachromosomal in *R. etli* CFN42 (Ramírez, unpublished). Expression of symbiosis genes dependent on plant hosts and the molecules and conditions required for gene expression have been well studied and have been extensively reviewed (Cooper, 2004; Le Strange et al., 1990; Maj et al., 2010; Masson-Boivin et al., 2009). Rhizobial genes expressed under stress (Vercruysse et al., 2011), in presence of flavonoids (Perret et al., 1999; Zhang and Cheng, 2006) or in nodules have been reported (Barnett et al., 2004; Chang et al., 2007; Karunakaran et al., 2009; Tsukada et al., 2009) but less is known on genes expressed in soil or in the rhizosphere. Mutations in the *cin* and *rhi* quorum sensing systems affect rhizospheric growth (Cubo et al., 1992; Edwards et al., 2009).

### 7.1. Rhizobium leguminosarum ER rhizospheric expression

A microarray based approach to study *R. leguminosarum* gene expression in pea, alfalfa or sugar beet rhizospheres showed that many of the genes preferentially expressed in *R. leguminosarum* 3841 when inhabiting the pea rhizosphere are encoded in the conjugative 147 kb plasmid pRL8 (Ramachandran et al., 2011). From pRL8, 11 or 21 genes (depending on the threshold considered) were up regulated in pea and only 3 or 2 in alfalfa or sugar beet rhizospheres. Pea induced genes represented around 15% of all genes on pRL8. In total 138 genes were specifically up regulated in 7 day old pea plants and 106 genes were up regulated in all rhizospheres, 70 of those were hypothetical. Among genes expressed were those encoding phenylanine and tyrosine catabolism, dicarboxylate transport, *rhiABC*, *rhiI*, *cinI*, protocatechuate and shikimate uptake, xanthine, formate and other dehydrogenases, as well as some *nod* genes (Ramachandran et al., 2011).

### 7.2. Rhizobium phaseoli ER rhizoplane expression

*R. phaseoli* Ch24-10 was chosen to study gene expression in plant roots because it represents a group of dominant

bacteria in maize rhizosphere (Rosenblueth and Martínez Romero, 2004), is highly competitive to colonize maize and rice and is a very efficient bean symbiont. Bean and maize plants have been grown in association in traditional agriculture for some thousand years and rhizobial gene expression was analyzed in both hosts independently. Upper value tails of bacterial gene transcript distribution in a reported transcriptomic analysis were found to correlate to RNA polymerase occupancy meaning that transcription was occurring in those genes (Vijayan et al., 2011) and, on that basis, highly expressed genes in the Ch24-10 transcriptomic profiling were selected. The 324 extrachromosomal genes highly expressed in maize and/or bean rootlets represented 22% of pRpCh24-10b and 16% of pRpCh24-10d. pSym genes were also expressed in the rhizosphere of maize and bean (representing 13% of the plasmid). Examples of ER genes that were highly expressed in both maize and bean root samples (Supplementary Table S1) are those responsible for proline catabolism, iron uptake, thiamine and gibberellin biosynthesis, a type VI secretion system, oligopeptide or sugar transporters and extrusion pumps as well as polygalacturonase, alpha amylase and Deg protease genes. *teu* genes were not expressed in maize roots in agreement to previous results showing that they are not induced by maize exudates (Rosenblueth et al., 1998). A promoter-less *gusA* gene reporter fused to the polygalacturonase gene was found to be expressed in maize and bean exudates (unpublished) and antibiotic resistance promoter-less genes were found to be expressed in plants when fused to the extrusion pump genes *rmrAB* (González-Pasayo and Martínez-Romero, 2000) or to Deg protease genes (unpublished); this additional evidence is in agreement to the transcriptomic results presented. Furthermore, a radioactive polygalacturonase probe was found to hybridize to the 1 Mb Ch24-10 ER (not shown). Orthologues to previously reported *R. leguminosarum* genes expressed in plant rhizospheres (Ramachandran et al., 2011) were found to be highly expressed in *R. phaseoli* in maize and bean roots (Supplementary Table S1). As in *R. leguminosarum* (Ramachandran et al., 2011), many *R. phaseoli* Ch24-10 highly expressed genes were hypothetical, one of them in common to *R. leguminosarum*.

A comparison of the Ch24-10 transcripts from maize and from bean roots suggested that replicons were differentially expressed depending on the plant host colonized. ER transcripts highly expressed in bean and not in maize roots were found in the Ch24-10 symbiotic plasmid (11 out of 26 bean specific genes) and in a 400 kb ER (pRpCh24-10b with equivalent genes to pReCFN42e), with 9 out of 26 specific genes, while most of the transcripts highly expressed in maize but not in bean (11 out of 14 maize specific genes) were found in pRpCh24-10d, a 1 Mb replicon sharing genes with *R. etli* pReCFN42f. No transcripts could be assigned to a 370 kb plasmid (pRpCh24-10a) as the strain used for the transcriptomic analysis unfortunately lost this plasmid that shares genes with pRL10 and pReCFN42c.

## 8. Concluding remarks

There is still scarce knowledge of rhizobial genes that are functional in nature, in soil, rhizospheric niches or complex microbial communities. Future studies may provide more data to support that ERs, highly dynamic and variable, determine or condition fitness or survival of rhizobia in the environment. Our data extend the knowledge of root-expressed genes in *Rhizobium* and allowed the identification of some extrachromosomal genes commonly expressed in association with plants such as those for thiamine biosynthesis, oligopeptide, proline betaine, α-galactosidase and other ABC transporters, α-*N*-arabinofuranosidase, *rmrA* (González-Pasayo and Martínez-Romero, 2000) and *nod* genes.

## Acknowledgements

To PAPIIT (UNAM) grants IN200709 and IN205412. Martha López-Guerrero was a Ph.D. student at the Programa de Doctorado en Ciencias Biomédicas and had a Consejo Nacional de Ciencia y Tecnología (CONACyT) fellowship. Illumina Sequencing was performed at the Unidad Universitaria de Secuenciación Masiva de DNA (USMDNA) of Universidad Nacional Autónoma de México (UNAM). To Dr. M. Dunn for critically reading the ms.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.plasmid.2012.07.002.

## References

Amadou, C., Pascal, G., Mangenot, S., Glew, M., Bontemps, C., et al., 2008. Genome sequence of the beta-*Rhizobium Cupriavidus taiwanensis* and comparative genomics of rhizobia. Genome Res. 18, 1472–1483.

Baldani, J.I., Weaver, R.W., Hynes, M.F., Eardly, B.D., 1992. Utilization of carbon substrates, electrophoretic enzyme patterns, and symbiotic performance of plasmid-cured clover rhizobia. Appl. Environ. Microbiol. 58, 2308–2314.

Bánfalvi, Z., Sakanyan, V., Koncz, C., Kiss, A., Dusha, I., et al., 1981. Location of nodulation and nitrogen fixation genes on a high molecular weight plasmid of *R. meliloti*. Mol. Gen. Genet. 184, 318–325.

Barloy-Hubler, F., Jebbar, J., 2009. *Sinorhizobium meliloti* megaplasmids and symbiosis in *S. meliloti*. In: Schwartz, E. (Ed.), Microbiology Monographs, vol. 11. Microbial Megaplasmids, Münster, Germany, p. 91.

Barnett, M.J., Toman, C.J., Fisher, R.F., Long, S.R., 2004. A dual-genome symbiosis chip for coordinate study of signal exchange and development in a prokaryote–host interaction. Proc. Natl. Acad. Sci. USA 101, 16636–16641.

Barran, L.R., Ritchot, N., Bromfield, E.S.P., 2001. *Sinorhizobium meliloti* plasmid pRm1132f replicates by a rolling-circle mechanism. J. Bacteriol. 183, 2704–2708.

Borthakur, D., Soedarjo, M., Fox, P.M., Webb, D.T., 2003. The *mid* genes of *Rhizobium* sp. strain TAL1145 are required for degradation of mimosine into 3-hydroxy-4-pyridone and are inducible by mimosine. Microbiology 149, 537–546.

Brian, P.W., 2008. Effects of gibberellins on plant growth and development. Biol. Rev. 83, 37–77.

Brom, S., García-De los Santos, A., Cervantes, L., Palacios, R., Romero, D., 2000. In *Rhizobium etli* symbiotic plasmid transfer, nodulation competitivity and cellular growth require interaction among different replicons. Plasmid 44, 34–43.

Brom, S., García de los Santos, A.G., Girard, M.L., Dávila, G., Palacios, R., et al., 1991. High-frequency rearrangements in *Rhizobium leguminosarum* bv. phaseoli plasmids. J. Bacteriol. 173, 1344–1346.

Brom, S., García de los Santos, A.G., Stepkowsky, T., Flores, M., Dávila, G., et al., 1992. Different plasmids of *Rhizobium leguminosarum* bv. phaseoli are required for optimal symbiotic performance. J. Bacteriol. 174, 5183–5189.

Brom, S., Girard, L., Tun-Garrido, C., García-de los Santos, A., 2004. Transfer of the symbiotic plasmid of *Rhizobium etli* CFN42 requires cointegration with p42a, which may be mediated by site-specific recombination. J. Bacteriol. 186, 7538–7548.

Broughton, W.J., Hanin, M., Relić, B., Kopciñska, J., Golinowski, W., Şimşek, S., Ojanen-Reuhs, T., Reuhs, B., Marie, C., Kobayashi, H., Bordogna, B., Le Quéré, A., Jabbouri, S., Fellay, R., Perret, X., Deakin, W.J., 2006. Flavonoid-inducible modifications to rhamnan O antigens are necessary for *Rhizobium* sp. strain NGR234-legume symbioses. J. Bacteriol. 188, 3654–3663.

Bruce, G.C., Kerr, A., Jones, D.A., 1990. Characteristics of the nopaline catabolic plasmid in *Agrobacterium* strains K84 and K1026 used for biological control of crown gall disease. Plasmid 23, 126–137.

Cervantes, L., Bustos, P., Girard, L., Santamaría, R.I., Dávila, G., et al., 2011. The conjugative plasmid of a bean-nodulating *Sinorhizobium fredii* strain is assembled from sequences of two *Rhizobium* plasmids and the chromosome of a *Sinorhizobium* strain. BMC Microbiol. 11, 149.

Cervantes-Rivera, R., Pedraza-López, F., Pérez-Segura, G., Cevallos, M.A., 2011. The replication origin of a *repABC* plasmid. BMC Microbiol. 11, 158.

Chaintreuil, C., Giraud, E., Prin, Y., Lorquin, J., Ba, A., et al., 2000. Photosynthetic bradyrhizobia are natural endophytes of the African wild rice *Oryza breviligulata*. Appl. Environ. Microbiol. 66, 5437–5447.

Chang, W.S., Franck, W.L., Cytryn, E., Jeong, S., Joshi, T., Emerich, D.W., Sadowsky, M.J., Xu, D., Stacey, G., 2007. An oligonucleotide microarray resource for transcriptional profiling of *Bradyrhizobium japonicum*. Mol. Plant Microbe Interact. 20, 298–307.

Chen, H., Higgins, J., Oresnik, I.J., Hynes, M.F., Natera, S., Djordjevic, M.A., Weinman, J.J., Rolfe, B.G., 2000. Proteome analysis demonstrates complex replicon and luteolin interactions in pSyma-cured derivatives of *Sinorhizobium meliloti* strain 2011. Electrophoresis 21, 3833–3842.

Conforte, V.P., Echeverria, M., Sánchez, C., Ugalde, R.A., Menéndez, A.B., Lepek, V.C., 2010. Engineered ACC deaminase-expressing free-living cells of *Mesorhizobium loti* show increased nodulation efficiency and competitiveness on *Lotus* spp. J. Gen. Appl. Microbiol. 56, 331–338.

Cooper, J., 2004. Multiple responses of rhizobia to flavonoids during legume root infection. Adv. Bot. Res. 41, 1–62.

Crossman, L.C., Castillo-Ramírez, S., McAnnula, C., Lozano, L., Vernikos, G.S., et al., 2008. A common genomic framework for a diverse assembly of plasmids in the symbiotic nitrogen fixing bacteria. PloS One 3, e2567.

Cubo, M.T., Economou, A., Murphy, G., Johnston, A.W., Downie, J.A., 1992. Molecular characterization and regulation of the rhizosphere-expressed genes *rhiABCR* that can influence nodulation by *Rhizobium leguminosarum* biovar viciae. J. Bacteriol. 174, 4026–4035.

Cytryn, E.J., Jitacksorn, S., Giraud, E., Sadowsky, M.J., 2008. Insights learned from pBTAi1, a 229-kb accessory plasmid from *Bradyrhizobium* sp. strain BTAi1 and prevalence of accessory plasmids in other *Bradyrhizobium* sp. strains. ISME J. 2, 158–170.

Ding, H., Yip, C.B., Geddes, B.A., Oresnik, I.J., Hynes, M.F., 2012. Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. Microbiology. http://dx.doi.org/10.1099/mic.0.057281-0.

Downie, J.A., 2010. The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. FEMS Microbiol. Rev. 34, 150–170.

Eckhardt, T., 1978. A rapid method for the identification of plasmid desoxyribonucleic acid in bacteria. Plasmid 1, 584–588.

Edwards, A., Frederix, M., Wisniewski-Dyé, F., Jones, J., Zorreguieta, A., Downie, J.A., 2009. The *cin* and *rai* quorum-sensing regulatory systems in *Rhizobium leguminosarum* are coordinated by ExpR and CinS, a small regulatory protein coexpressed with CinI. J. Bacteriol. 191, 3059–3067.

Falla, T.J., Chopra, I., 1999. Stabilization of *Rhizobium* symbiosis plasmids. Microbiology 145, 515–516.

Fauvart, M., Sánchez-Rodríguez, A., Beullens, S., Marchal, K., Michiels, J., 2011. Genome sequence of *Rhizobium etli* CNPAF512, a nitrogen-fixing symbiont isolated from bean root nodules in Brazil. J. Bacteriol. 193, 3158–3159.

Finan, T.M., Kunkel, B., de Vos, G.F., Signer, E.R., 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. J. Bacteriol. 167, 66–72.

Flores, M., González, V., Pardo, M.A., Leija, A., Martínez, E., et al., 1988. Genomic instability in *Rhizobium phaseoli*. J. Bacteriol. 170, 1191–1196.

Flores, M., Mavingui, P., Perret, X., Broughton, W.J., Romero, D., et al., 2000. Prediction, identification, and artificial selection of DNA rearrangements in *Rhizobium*: toward a natural genomic design. Proc. Natl. Acad. Sci. USA 97, 9138–9143.

Flores, M., Morales, L., Ávila, A., González, V., Bustos, P., et al., 2005. Diversification of DNA sequences in the symbiotic genome of *Rhizobium etli*. J. Bacteriol. 187, 7185–7192.

Fox, M.A., Karunakaran, R., Leonard, M.E., Mouhsine, B., Williams, A., East, A.K., Downie, J.A., Poole, P.S., 2008. Characterization of the quaternary amine transporters of *Rhizobium leguminosarum* bv. viciae 3841. FEMS Microbiol. Lett. 287, 212–220.

Freiberg, C., Fellay, R., Bairoch, A., Broughton, W.J., Rosenthal, A., et al., 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. Nature 387, 394–401.

García-de los Santos, A., Brom, S., 1997. Characterization of two plasmid-borne *lps* $\beta$ loci of Rhizobium etli required for lipopolysaccharide synthesis and for optimal interaction with plants. Mol. Plant-Microbe Interact. 10, 891–902.

García-de los Santos, A., Brom, S., Romero, D., 1996. Rhizobium plasmids in bacteria–legume interactions. World J. Microbiol. Biotechnol. 12, 119–125.

García-de los Santos, A., López, E., Cubillas, C.A., Noel, K.D., Brom, S., et al., 2008. Requirement of a plasmid-encoded catalase for survival of *Rhizobium etli* CFN42 in a polyphenol-rich environment. Appl. Environ. Microbiol. 74, 2398–2403.

Geddes, B.A., Pickering, B.S., Poysti, N.J., Collins, H., Yudistira, H., et al., 2010. A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*. Microbiology 156, 2970–2981.

Gómez-Hernández, N., Reyes-González, A., Sánchez, C., Mora, Y., Delgado, M.J., Girard, L., 2011. Regulation and symbiotic role of *nirK* and *norC* expression in *Rhizobium etli*. Mol. Plant-Microbe Interact. 24, 233–245.

González, V., Acosta, J.L., Santamaría, R.I., Bustos, P., Fernández, J.L., et al., 2010. Conserved symbiotic plasmid DNA sequences in the multireplicon pangenomic structure of *Rhizobium etli*. Appl. Environ. Microbiol. 76, 1604–1614.

González, V., Bustos, P., Ramírez-Romero, M.A., Medrano-Soto, A., Salgado, H., et al., 2003. The mosaic structure of the symbiotic plasmid of *Rhizobium etli* CFN42 and its relation to other symbiotic genome compartments. Genome Biol. 4, R36.

González, V., Santamaría, R.I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S.C., Ramírez, M.A., Jiménez-Jacinto, V., Collado-Vides, J., Dávila, G., 2006. The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. Proc. Natl. Acad. Sci. USA 103, 3834–3839.

González-Pasayo, R., Martínez-Romero, E., 2000. Multiresistance genes of *Rhizobium etli* CFN42. Mol. Plant-Microbe Interact. 13, 572–577.

Guerrero, G., Peralta, H., Aguilar, A., Díaz, R., Villalobos, M.A., et al., 2005. Evolutionary, structural and functional relationships revealed by comparative analysis of syntenic genes in Rhizobiales. BMC Evol. Biol. 5, 55.

Guntli, D., Heeb, M., Moenne-Loccoz, Y., Defago, G., 1999. Contribution of calystegine catabolic plasmid to competitive colonization of the rhizosphere of calystegine-producing plants by *Sinorhizobium meliloti* Rm41. Mol. Ecol. 8, 855–863.

Guo, X., Flores, M., Mavingui, P., Fuentes, S.I., Hernández, G., et al., 2003. Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. Genome Res. 13, 1810–1817.

Gutiérrez-Zamora, M.L., Martínez-Romero, E., 2001. Natural endophytic association between *Rhizobium etli* and maize (*Zea mays* L.). J. Biotechnol. 91, 117–126.

Hahn, M., Hennecke, H., 1987. Mapping of a *Bradyrhizobium japonicum* DNA region carrying genes for symbiosis and an asymmetric accumulation of reiterated sequences. Appl. Environ. Microbiol. 53, 2247–2252.

Harrison, P.W., Lower, R.P.J., Kim, N.K.D., Young, J.P.W., 2010. Introducing the bacterial chromid: not a chromosome, not a plasmid. Trends Microbiol. 18, 141–148.

Haugland, R., Verma, D.P.S., 1981. Interspecific plasmid and genomic DNA sequence homologies and localisation of *nif*-genes in effective and ineffective strains of *Rhizobium japonicum*. J. Mol. Appl. Genet. 1, 205–217.

Hernández-Lucas, I., Pardo, M.A., Segovia, L., Miranda, J., Martínez-Romero, E., 1995. *Rhizobium tropici* chromosomal citrate synthase gene. Appl. Environ. Microbiol. 61, 3992–3997.

Hirsch, P.R., Van Montagu, M., Johnston, A.W.B., Brewin, N.J., Schell, J., 1980. Physical identification of bacteriocinogenic, nodulation and other plasmids in strains of *Rhizobium leguminosarum*. J. Gen. Microbiol. 120, 403–412.

Hooykaas, P.J., Snijdewint, F.G., Schilperoort, R.A., 1982. Identification of the Sym plasmid of *Rhizobium leguminosarum* strain 1001 and its

transfer to and expression in other rhizobia and *Agrobacterium tumefaciens*. Plasmid 8, 73–82.
Hynes, M.F., Brucksch, K., Priefer, U., 1988. Melanin production encoded by a cryptic plasmid in a *Rhizobium leguminosarum* strain. Arch. Microbiol. 150, 326–332.
Hynes, M.F., McGregor, N.F., 1990. Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. Mol. Microbiol. 4, 567–574.
Jensen, R.B., Gerdes, K., 1995. Programmed cell death in bacteria: proteic plasmid stabilization systems. Mol. Microbiol. 17, 205–210.
Johnston, A.W.B., Beynon, J.L., Buchanan-Wollaston, A.V., Setchell, S.M., Hirsch, P.R., Beringer, J.E., et al., 1978. High frequency transfer of nodulating ability between strains and species of *Rhizobium*. Nature 276, 634–636.
Karunakaran, R., Ramachandran, V.K., Seaman, J.C., East, A.K., Mouhsine, B., et al., 2009. Transcriptomic analysis of *Rhizobium leguminosarum* biovar viciae in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. J. Bacteriol. 191, 4002–4014.
Landeta, C., Dávalos, A., Cevallos, M.A., Geiger, O., Brom, S., et al., 2011. Plasmids with a chromosome-like role in rhizobia. J. Bacteriol. 193, 1317–1323.
Lee, K.B., De Backer, P., Aono, T., Liu, C.T., Suzuki, S., et al., 2008. The genome of the versatile nitrogen fixer *Azorhizobium caulinodans* ORS571. BMC Genomics 9, 271.
Le Strange, K.K., Bender, G.L., Djordjevic, M.A., Rolfe, B.G., Redmond, J.W., 1990. The *Rhizobium* strain NGR234 *nodDI* gene product responds to activation by the simple phenolic compounds vanillin and isovanillin present in wheat seedling extracts. Mol. Plant-Microbe Interact. 3, 214–220.
Lithgow, J.K., Wilkinson, A., Hardman, A., Rodelas, B., Wisniewski-Dyé, F., Williams, P., Downie, J.A., 2000. The regulatory locus *cinRI* in *Rhizobium leguminosarum* controls a network of quorum-sensing loci. Mol. Microbiol. 37, 81–97.
López-López, A., Rogel, M.A., Ormeño-Orrillo, E., Martínez-Romero, J., Martínez-Romero, E., 2010. *Phaseolus vulgaris* seed-borne endophytic community with novel bacterial species such as *Rhizobium endophyticum sp. nov*. Syst. Appl. Microbiol. 33, 322–327.
López-Guerrero, M.G., Ormeño-Orrillo, E., Velázquez, E., Rogel, M.A., Acosta, J.L., Gónzalez, V., Martínez, J., Martínez-Romero, E., 2012. *Rhizobium etli* taxonomy revised with novel genomic data and analyses. Syst. Appl. Microbiol. In Press. Available online 1 August 2012.
Maj, D., Wielbo, J., Marek-Kozaczuk, M., Skorupska, A., 2010. Response to flavonoids as a factor influencing competitiveness and symbiotic activity of *Rhizobium leguminosarum*. Microbiol. Res. 165, 50–60.
Marie, C., Barny, M.-A., Downie, J.A., 1992. *Rhizobium leguminosarum* has two glucosamine synthases, GlmS and NodM, required for nodulation and development of nitrogen-fixing nodules. Mol. Microbiol. 6, 843–851.
Marie, C., Broughton, W.J., Deakin, W.J., 2001. *Rhizobium* type III secretion systems: legume charmers or alarmers? Curr. Opinion Plant Biol. 4, 336–342.
Martínez, E., Palacios, R., Sánchez, F., 1987. Nitrogen-fixing nodules induced by *Agrobacterium tumefaciens* harboring *Rhizobium phaseoli* plasmids. J. Bacteriol. 169, 2828–2834.
Martínez-Romero, E., Rosenblueth, M., 1990. Increased bean (*Phaseolus vulgaris* L.) nodulation competitiveness of genetically modified *Rhizobium* strains. Appl. Environ. Microbiol. 56, 2384–2388.
Masson-Boivin, C., Giraud, E., Perret, X., Batut, J., 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? Trends Microbiol. 17, 458–466.
Mauchline, T.H., Fowler, J.E., East, A.K., Sartor, A.L., Zaheer, R., et al., 2006. Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. Proc. Natl. Acad. Sci. USA 103, 17933–17938.
Mavingui, P., 2009. The megaplasmid pNGR234a of *Rhizobium* sp. strain NGR234. In: Schwartz, E. (Ed.), Microbiology Monographs, vol. 11. Microbial Megaplasmids, Münster, Germany, pp. 119–132.
Mavingui, P., Flores, M., Guo, X., Davila, G., Perret, X., et al., 2002. Dynamics of genome architecture in *Rhizobium* sp. strain NGR234. J. Bacteriol. 184, 171–176.
Mavingui, P., Flores, M., Romero, D., Martínez-Romero, E., Palacios, R., 1997. Generation of *Rhizobium* strains with improved symbiotic properties by random DNA amplification (RDA). Nat. Biotechnol. 15, 564–569.
Mavingui, P., Laeremans, T., Flores, M., Romero, D., Martínez-Romero, E., et al., 1998. Genes essential for Nod factor production and nodulation are located on a symbiotic amplicon (AMPRtrCFN299pc60) in *Rhizobium tropici*. J. Bacteriol. 180, 2866–2874.
Mercado-Blanco, J., Toro, N., 1996. Plasmids in rhizobia: the role of nonsymbiotic plasmids. Mol. Plant-Microbe Interact. 9, 535–545.
Miranda-Ríos, J., Morera, C., Taboada, H., Dávalos, A., Encarnación, S., et al., 1997. Expression of thiamin biosynthetic genes (*thiCOGE*) and production of symbiotic terminal oxidase cbb sub(3) in *Rhizobium etli*. J. Bacteriol. 179, 6887–6893.
Moënne-Loccoz, Y., Weaver, R.W., 1995a. Plasmids and saprophytic growth of *Rhizobium leguminosarum* bv. trifolii W14–2 in soil. FEMS Microbiol. Ecol. 18, 139–144.
Moënne-Loccoz, Y., Weaver, R.W., 1995b. Plasmids influence growth of rhizobia in the rhizosphere of clover. Soil Biol. Biochem. 27, 1001–1004.
Moënne-Loccoz, Y., Baldani, J.I., Weaver, R.W., 1995. Sequential heat-curing of Tn*5*-Mob-*sac* labelled plasmids from *Rhizobium* to obtain derivatives with various combinations of plasmids and no plasmid. Lett. Appl. Microbiol. 20, 175–179.
Morrone, D., Chambers, J., Lowry, L., Kim, G., Anterola, A., Bender, K., Peters, R.J., 2009. Gibberellin biosynthesis in bacteria: Separate ent-copalyl diphosphate and ent-kaurene synthases in *Bradyrhizobium japonicum*. FEBS Lett. 583, 475–480.
Murphy, P.J., Heycke, N., Banfalvi, Z., Tate, M.E., de Bruijn, F., Kondorosi, A., Tempé, J., Schell, J., 1987. Genes for the catabolism and synthesis of an opine-like compound in *Rhizobium meliloti* are closely linked and on the Sym plasmid. Proc. Natl. Acad. Sci. USA 84, 493–497.
Nuti, M.P., Ledeboera, M., Lepidi, A.A., Schilperoorrt, A., 1977. Large plasmids in different *Rhizobium species*. J. Gen. Microbiol. 100, 241–248.
Nuti, M.P., Lepidi, A.A., Prakash, R.K., Schilperoort, R.A., Cannon, F.C., 1979. Evidence for nitrogen fixation (*nif*) genes on indigenous *Rhizobium* plasmids. Nature 282, 533–535.
O'Connell, M., Noel, T.C., Yeung, E.C., Hynes, M., Hynes, M.F., 1998. Decreased symbiotic effectiveness of *Rhizobium leguminosarum* strains carrying plasmid RP4. FEMS Microbiol. Lett. 161, 275–283.
Ogura, T., Hiraga, S., 1983. Mini-F plasmid genes that couple host cell division to plasmid proliferation. Proc. Natl. Acad. Sci. USA 80, 4784–4788.
Oldroyd, G.E., Murray, J.D., Poole, P.S., Downie, J.A., 2011. The rules of engagement in the legume-rhizobial symbiosis. Ann. Rev. Genet. 45, 119–144.
Oresnik, I.J., Pacarynuk, L.A., O'Brien, Sh.A.P., Yost, Ch.K., Hynes, M.F., 1998. Plasmid-encoded catabolic genes in *Rhizobium leguminosarum* bv. trifolii: evidence for a plant-inducible rhamnose locus involved in competition for nodulation. Mol. Plant-Microbe Interact. 11, 1175–1185.
Oresnik, I.J., Twelker, S., Hynes, M.F., 1999. Cloning and characterization of a *Rhizobium leguminosarum* gene encoding a bacteriocin with similarities to RTX toxins. Appl. Environ. Microbiol. 65, 2833–2840.
Ormeño-Orrillo, E., Hungria, M., Martinez-Romero, E., in press. Dinitrogen-fixing prokaryotes. In: Rosenberg, E., DeLong, E.F., Stackebrandt, E., Lory, S., Thompson, F. (Eds.), The Prokaryotes vol. 1: Symbiotic Associations, Biotechnology, Applied Microbiology. 4th ed. ISBN 978-3-642-30194-0, Springer.
Pandey, D.P., Gerdes, K., 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res. 33, 966–976.
Pappas, Cevallos, M.A., 2011. Plasmids of the Rhizobiaceae and their role in interbacterial and transkingdom interactions. In: Witzany, G. (Ed.), Soil Biology. Biocommunication in Soil Microorganisms, Austria.
Pardo, M.A., Lagúnez, J., Miranda, J., Martínez, E., 1994. Nodulating ability of *Rhizobium tropici* is conditioned by a plasmid- encoded citrate synthase. Mol. Microbiol. 11, 315–321.
Peix, A., Velázquez, E., Silva, L.R., Mateos, P.F., 2010. Key molecules involved in beneficial infection process in rhizobia–legume symbiosis. In: Khan, M.H., Zaidi, A., Musarrat, J. (Eds.), Microbes for Legume Improvement. Springer, Wien, pp. 55–80.
Pérez-Ramírez, N.O., Rogel, M.A., Wang, E., Castellanos, J.Z., Martínez-Romero, E., 1998. Seeds of *Phaseolus vulgaris* bean carry *Rhizobium etli*. FEMS Microbiol. Ecol. 26, 289–296.
Perret, X., Freiberg, C., Rosenthal, A., Broughton, W.J., Fellay, R., 1999. High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234. Mol. Microbiol. 32, 415–425.
Perrine, F.M., Hocart, C.H., Hynes, M.F., Rolfe, B.G., 2005. Plasmid-associated genes in the model micro-symbiont *Sinorhizobium meliloti* 1021 affect the growth and development of young rice seedlings. Environ. Microbiol. 7, 1826–1838.
Perrine, F.M., Prayitno, J., Weinman, J.J., Dazzo, F.B., Rolfe, B.G., 2001. *Rhizobium* plasmids are involved in the inhibition or stimulation of rice growth and development. Aust. J. Plant Physiol. 28, 923–937.
Ramachandran, V.K., East, A.K., Karunakaran, R., Downie, J.A., Poole, P.S., 2011. Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and

sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biol. 12, R106.
Reiter, B., Buergmann, H., Burg, K., Sessitsch, A., 2003. Endophytic *nifH* gene diversity in African sweet potato. Can. J. Microbiol. 49, 549–555.
Rochepeau, P., Selinger, L.B., Hynes, M.F., 1997. Transposon-like structure of a new plasmid-encoded restriction-modification system in *Rhizobium leguminosarum* VF39SM. Mol. Gen. Genet. 256, 387–396.
Rogel, M.A., Ormeño-Orrillo, E., Martínez Romero, E., 2011. Symbiovars in rhizobia reflect bacterial adaptation to legumes. Syst. Appl. Microbiol. 34, 96–104.
Romero, D., Brom, S., 2004. The symbiotic plasmids of the Rhizobiaceae: Chapter 12 in Plasmid Biology. In: Funnell, B.E., Phillips, G.J. (Eds.), ASM Press, Washington, pp. 271–290.
Romero, D., Brom, S., Martínez-Salazar, J., De, L., Girard, M., Palacios, R., et al., 1991. Amplification and deletion of a *nod-nif* region in the symbiotic plasmid of *Rhizobium phaseoli*. J. Bacteriol. 173, 2435–2441.
Romero, D., Martínez-Salazar, J., Girard, L., Brom, S., Dávila, G., et al., 1995. Discrete amplifiable regions (amplicons) in the symbiotic plasmid of *Rhizobium etli* CFN42. J. Bacteriol. 177, 973–980.
Rosenberg, C., Casse-Delbart, F., Dusha, I., David, M., Boucher, C., et al., 1982. Megaplasmids in the plant-associated bacteria *Rhizobium meliloti* and *Pseudomonas solanacearum*. J. Bacteriol. 150, 402–406.
Rosenblueth, M., Hynes, M.F., Martínez-Romero, E., 1998. *Rhizobium tropici teu* genes involved in specific uptake of *Phaseolus vulgaris* bean-exudate compounds. Mol. Gen. Genet. 258, 587–598.
Rosenblueth, M., Martínez Romero, E., 2004. *Rhizobium etli* maize populations and their competitiveness for root colonization. Arch. Microbiol. 181, 337–344.
Schripsema, J., De Rudder, K.E.E., Van Vliet, T.B., Lankhorst, P.P., De Vroom, E., et al., 1996. Bacteriocin small of *Rhizobium leguminosarum* belongs to the class of *N*-acyl-ι-homoserine lactone molecules, known as autoinducers and as quorum sensing co-transcription factors. J. Bacteriol. 178, 366–371.
Sengupta, M., Austin, S., 2011. Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. Infect. Immun. 79, 2502–2509.
Segovia, L., Piñero, D., Palacios, R., Martínez-Romero, E., 1991. Genetic structure of a soil population of nonsymbiotic *Rhizobium leguminosarum*. Appl. Environ. Microbiol. 57, 426–433.
Skorupska, A., Janczarek, M., Marczak, M., Mazur, A., Krol, J., 2006. Rhizobial exopolysaccharides: genetic control and symbiotic functions. Microbial Cell Factories 5, 7.
Slater, S.C., Goldman, B.S., Goodner, B., Setubal, J.C., Farrand, S.K., et al., 2009. Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. J. Bacteriol. 191, 2501–2511.
Soberón-Chávez, G., Nájera, R., 1989. Isolation from soil of *Rhizobium leguminosarum* lacking symbiotic information. Can. J. Microbiol. 35, 464–468.
Soberón-Chávez, G., Nájera, R., Olivera, H., Segovia, L., 1986. Genetic rearrangements of a *Rhizobium phaseoli* symbiotic plasmid. J. Bacteriol. 167, 487–491.
Streit, W.R., Joseph, C.M., Phillips, D.A., 1996. Biotin and other water-soluble vitamins are key growth factors for alfalfa root colonization by *Rhizobium meliloti* 1021. Mol. Plant-Microbe Interact. 9, 330–338.
Sutton, W.D., 1974. Some features of the DNA of *Rhizobium* bacteroids and bacteria. Biochim. Biophys. Acta 336, 1–10.
Suzuki, T., Tanaka, K., Yamamoto, S., Kiyokawa, K., Moriguchi, K., Yoshida, K., 2009. Ti and Ri plasmids. In: Schwartz, E. (Ed.), Microbiology Monographs, vol. 11. Microbial Megaplasmids, Münster, Germany, p. 133.
Tepfer, D., Goldmann, A., Pamboukdjian, N., Maille, M., Lepingle, A., et al., 1988. A plasmid of *Rhizobium meliloti* 41 encodes catabolism of two compounds from root exudate of *Calystegium sepium*. J. Bacteriol. 170, 1153–1161.
Theunis, M., Kobayashi, H., Broughton, W.J., Prinsen, E., 2004. Flavonoids, NodD1, NodD2, and nod-box NB15 modulate expression of the y4wEFG locus that is required for indole-3-acetic acid synthesis in *Rhizobium* sp. strain NGR234. Mol. Plant-Microbe Interact. 17, 1153–1161.
Tshitinge, G., Luyindula, N., Lurquin, P.F., Ledoux, L., 1975. Plasmid DNA in *Rhizobium vigna* and *Rhizobium trifolii*. Biochim. Biophys. Acta 414, 357–361.
Tsukada, S., Aono, T., Akiba, N., Lee, K.-B., Liu, C.-T., et al., 2009. Comparative genome-wide transcriptional profiling of *Azorhizobium caulinodans* ORS571 grown under free-living and symbiotic conditions. Appl. Environ. Microbiol. 75, 5037–5046.
Van Dillewijn, P., Villadas, P.J., Toro, N., 2002. Effect of a *Sinorhizobium meliloti* strain with a modified *putA* gene on the rhizosphere microbial community of alfalfa. Appl. Environ. Microbiol. 68, 4201–4208.
Van Melderen, L., Saavedra De Bast, M., Rosenberg, S.M., 2009. Bacterial toxin–antitoxin systems: more than selfish entities? PLoS Genet. 5, e1000437.
Velázquez, E., Mateos, P.F., Pedrero, P., Dazzo, F.B., Martínez-Molina, E., 1995. Attenuation of symbiotic effectiveness by *Rhizobium meliloti* SAF22 related to the presence of a cryptic plasmid. Appl. Environ. Microbiol. 61, 2033–2036.
Venter, A.P., Twelker, S., Oresnik, I.J., Hynes, M.F., 2001. Analysis of the genetic region encoding a novel rhizobiocin from *Rhizobium leguminosarum* bv. viciae strain 306. Can. J. Microbiol. 47, 495–502.
Vercruysse, M., Fauvart, M., Jans, A., Beullens, S., Braeken, K., et al., 2011. Stress response regulators identified through genome-wide transcriptome analysis of the (p)ppGpp-dependent response in *Rhizobium etli*. Genome Biol. 12, R17.
Vijayan, V., Jain, I.H., O'Shea, E.K., 2011. A high resolution map of a cyanobacterial transcriptome. Genome Biol. 12, R47.
Villaseñor, T., Brom, S., Dávalos, A., Lozano, L., Romero, D., et al., 2011. Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*. BMC Microbiol. 11, 66.
Viprey, V., Del Greco, A., Golinowski, W., Broughton, W.J., Perret, X., 1998. Symbiotic implications of type III protein secretion machinery in *Rhizobium*. Mol. Microbiol. 28, 1381–1389.
Wang, E.T., Van Berkum, P., Sui, X.H., Beyene, D., Chen, W.X., et al., 1999. Diversity of rhizobia associated with *Amorpha fruticosa* isolated from Chinese soils and description of *Mesorhizobium amorphae sp. nov*. Int. J. Syst. Bacteriol. 49, 51–65.
Weaver, R.W., Wei, G.R., Berryhill, D.L., 1990. Stability of plasmids in *Rhizobium phaseoli* during culture. Soil Biol. Biochem. 22, 465–469.
Weaver, R.W., Wright, S.F., 1987. Variability in effectiveness of rhizobia during culture and in nodules. Appl. Environ. Microbiol. 53, 2972–2974.
Xu, Y., Murooka, Y., 1995. A large plasmid isolated from *Rhizobium huakuii* bv. renge that includes genes for both nodulation of *Astragalus sinicus* cv. Japan and nitrogen fixation. J. Ferment. Bioeng. 80, 276–279.
Yanni, Y.G., Rizk, R.Y., Corich, V., Squartini, A., Ninke, K., et al., 1997. Natural endophytic association between *Rhizobium leguminosarum* bv. trifolii and rice roots and assessment of its potential to promote rice growth. Plant Soil 194, 99–114.
Yost, C.K., Rath, A.M., Noel, T.C., Hynes, M.F., 2006. Characterization of genes involved in erythritol catabolism in *Rhizobium leguminosarum* bv. viciae. Microbiology 152, 2061–2074.
Yost, C.K., Rochepeau, P., Hynes, M.F., 1998. *Rhizobium leguminosarum* contains a group of genes that appear to code for methyl-accepting chemotaxis proteins. Microbiology 144, 1945–1956.
Zhang, X.S., Cheng, H.P., 2006. Identification of *Sinorhizobium meliloti* early symbiotic genes by use of a positive functional screen. Appl. Environ. Microbiol. 72, 2738–2748.
Zhang, X., Kosier, B., Priefer, U.B., 2001. Symbiotic plasmid rearrangement in *Rhizobium leguminosarum* bv. viciae VF39SM. J. Bacteriol. 183, 2141–2144.
Zou, X., Li, F., Chen, H., 1997. Characteristics of plasmids in *Rhizobium huakuii*. Curr. Microbiol. 35, 215–220.
Zurkowski, W., 1982. Molecular mechanism for loss of nodulation properties of *Rhizobium trifolii*. J Bacteriol. 150, 999–1007.
Zurkowski, W., Lorkiewic, Z., 1976. Plasmid deoxyribonucleic acid in *Rhizobium trifolii*. J. Bacteriol. 128, 481–484.

MARTHA G. LÓPEZ-GUERRERO Y ESPERANZA MARTÍNEZ ROMERO

# Bacterias fijadoras de nitrógeno

## Diversidad y uso como biofertilizantes

**La revolución verde que transformó la agricultura y ha proporcionado alimentos a una población humana creciente fue posible gracias a la fertilización; lamentablemente, los fertilizantes aplicados a los cultivos resultan ser contaminantes de lagos y ríos.**

**En el otro extremo, tenemos que los microbios ostentan capacidades metabólicas que ningún otro ser vivo posee, pueden sintetizar los mismos fertilizantes que se producen industrialmente, a un bajo costo; particularmente, las bacterias producen los fertilizantes en simbiosis con las plantas, y éstos no se liberan al ambiente. En este capítulo revisaremos cuales son estas bacterias.**

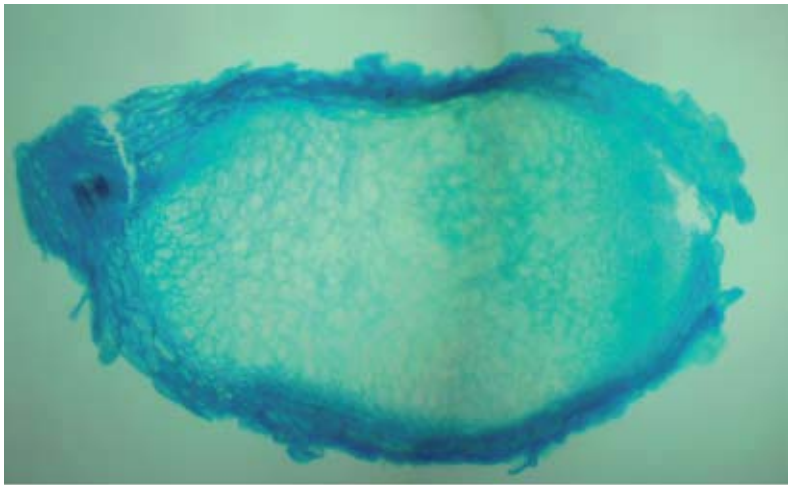
Un biofertilizante es una bacteria u hongo capaz de proporcionar a la planta minerales esenciales que se encuentran en concentración limitada en el suelo, como el nitrógeno (N o también $N_2$), el cual constituye 78% de los gases de la atmósfera. Los biofertilizantes pueden ser usados para disminuir los costos ambientales y económicos que generan el uso de fertilizantes químicos nitrogenados aplicados en agricultura. Es mediante la Fijación Biológica de Nitrógeno (FBN), proceso realizado sólo por microbios procariontes, que el $N_2$ es transformado en un compuesto asimilable; así, la FBN se lleva a cabo en asociación con plantas e insectos, en el océano y hasta en el fondo del mar.

Los rizobios[1] son proteobacterias alfa (como *Rhizobium* y *Bradyrhizobium*) que, en simbiosis, inducen la formación de estructuras especializadas (nódulos) en las que se lleva a cabo la FBN, en las raíces de leguminosas (frijol, soya, chícharo, trébol, haba, cacahuate). A la fecha, se han descrito cerca de 60 especies de rizobios, mientras que existen más de 1,800 especies de leguminosas, por lo que se estima que no se conoce la totalidad de bacterias asociadas a estas plantas. En México, hemos descubierto nuevas especies de rizobios asociadas al frijol y a las leguminosas arbóreas tropicales como mimosas, acacias y guajes.

En Brasil existe una industria exitosa que produce biofertilizantes de cepas eficientes para fijar nitrógeno a partir de la bacteria *Bradyrhizobium*. En algunas condiciones, las bacterias aportan cerca de 250 kilos de nitrógeno por hectárea por año, cantidad equivalente a la que se aplica con fertilizantes químicos. En ese país se cuenta con un control oficial de la calidad de los biofertilizantes comerciales.

En México, contamos con una gran diversidad de bacterias que viven en el suelo y pueden formar nódulos en plantas de frijol por lo que, aparentemente, no es necesario añadir más bacterias a los suelos o a las semillas; sin embargo, se ha tenido éxito utilizando algunas cepas de *Rhizobium* que son capaces de fijar más nitrógeno y que son producto de la investigación realizada en la UNAM.

Los risobios son bacterias que ayudan a configurar nódulos de las raíces de las leguminosas, en las cuales se establecen y, a partir de esta asociación, logran fijar el nitrógeno atmosférico



>> Margarita: esto es un corte transversal de un nódulo radicular formado en la raíz de una leguminosa

## FBN EN PLANTAS NO LEGUMINOSAS

Los cereales –arroz, maíz, trigo–, ocupan la mayor superficie agrícola en el ámbito mundial, y son fertilizados químicamente para generar altos rendimientos, pero lo ideal sería aplicar la FBN mediante biofertilizantes. Existen bacterias que pueden mejorar el crecimiento de los cereales. Las aportaciones de nitrógeno son pequeñas, pero como alternativa se puede aumentar su número en las plantas para incrementar la tasa de nitrógeno fijado.

Mediante enfoques modernos de análisis de todos los genomas de una comunidad (metagenómica) se descubrió que de todos los fijadores de nitrógeno asociados a caña de azúcar o arroz, son los rizobios los que expresan los genes que codifican para la nitrogenasa (enzima necesaria para fijar nitrógeno). Este resultado es sorprendente y puede impulsar el uso de rizobios como biofertilizantes.

En nuestro laboratorio descubrimos que *Rhizobium etli*, simbionte del frijol, se encuentra de manera natural dentro del maíz y es capaz de promover su crecimiento. Actualmente, estamos estudiando la expresión genética de *R. etli* en asociación con el maíz, conocimiento con el que esperamos generar herramientas que permitan, en un futuro, optimizar su uso como biofertilizante.

1. Bacterias que ayudan a configurar nódulos de las raíces de las leguminosas, en las cuales se establecen y, a partir de esta asociación logran fijar el nitrógeno atmosférico.

Un biofertilizante es una bacteria u hongo capaz de proporcionar a la planta minerales esenciales que se encuentran en concentración limitada en el suelo, como el nitrógeno

Algunas bacterias pertenecientes a los géneros *Klebsiella y Burkholderia* son capaces de promover el crecimiento vegetal, pero pueden ser patógenas de humanos, por lo que se ha recomendado que no sean utilizadas en experimentos en campo, y mucho menos en cultivos destinados para consumo humano.

De caña de azúcar, maíz y plátano se aisló *K. variicola*, bacteria fijadora de nitrógeno que promueve el crecimiento vegetal. Esta bacteria, además de asociarse a plantas, se ha encontrado en la sangre de bebés con infecciones originadas en hospitales, así como en la de pacientes inmunosuprimidos.

Antes de utilizar cualquier bacteria se requiere su identificación y caracterización para evaluar cualquier riesgo a la salud humana. Las cianobacterias que fijan nitrógeno y pudieran ser de gran utilidad como biofertilizantes pueden producir toxinas, hepatotoxinas o neurotoxinas muy dañinas para el humano y otros animales.

Estos son dos ejemplos para resaltar la importancia de realizar estudios filogenéticos y de diversidad para elegir adecuadamente las bacterias biofertilizantes, ya que pueden representar un riesgo para la salud humana.

## LOS BIOFERTILIZANTES Y LA FBN

La necesidad de desarrollar biofertilizantes eficientes se puede ver claramente en la producción de biocombustibles a partir de plantas, pues se evitaría el uso de fertilizantes químicos sintetizados a partir de la energía derivada del petróleo. Es ilógico y energéticamente poco favorable invertir en petróleo para generar biocombustibles. El proceso de fijación de nitrógeno es para el ciclo del nitrógeno como la fotosíntesis para el ciclo del carbono.

RECONOCIMIENTO

Agradecimientos: PAPIIT IN 200709, FOMIX 04-09-08. A Ernesto Ormeño, Tabita Ramírez y Mónica Rosenblueth por revisar el escrito. A Marco A. Rogel Hernández, Ivonne Toledo y Julio Martínez por su valioso apoyo técnico.

» Nódulos en las raíces de soya, conteniendo miles de millones de Bradyihizobium

SI DESEA SABER MÁS

» López-López, A. Rosenblueth, M., Martínez, J. and Martínez-Romero, E. 2010. "Rhizobial symbioses in tropical legumes and non-legumes", en "Soil Biology and Agriculture in the Tropics" Patrice Dion (ed.) Srpinger-Verlag Berlin pp. 163-184.

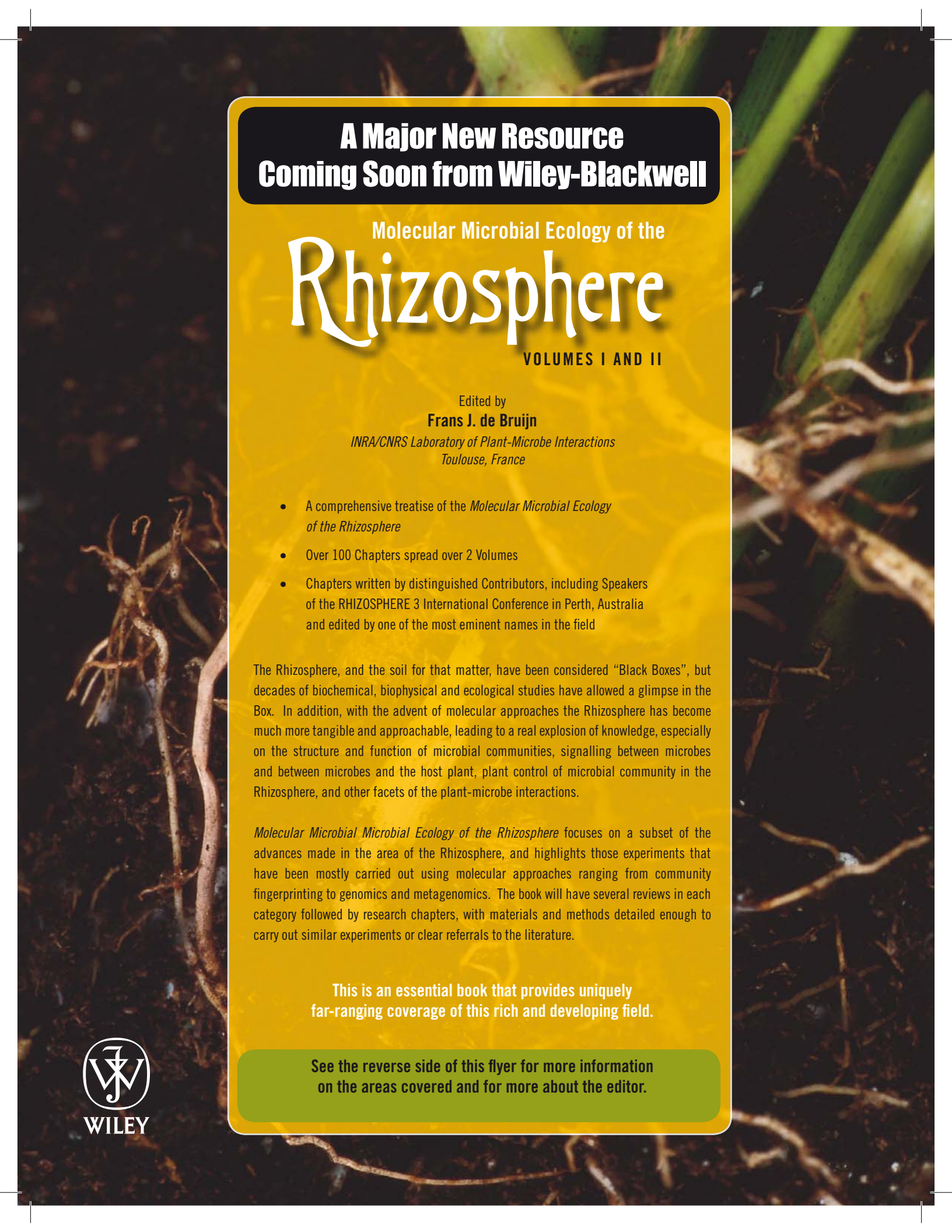
» Martínez Romero, J., Ormeño-Orrillo E., Rogel, M. A., López-López, A. and Martínez Romero, E. 2010. "Trends in rhizobial evolution and some taxonomic remarks" en "Evolutionary Biology – Concepts, Molecular and Morphological Evolution" Springer-Verlag Berlin.

» Fulweiler R. W. 2009. Fantastic Fixers. Science. 326: 377 – 378.

» Martínez Romero, E. 2008. "Detective de bacterias", en "Aportaciones Científicas e Humanísticas Mexicanas en el Siglo XX" Fondo de Cultura Económico, México, D. F. pp. 561-569.

» Lloret L., Rosenblueth, M, Toledo, I. and Martínez Romero, E. 2004. "Ecología bacteriana. Beneficios y riesgos del uso agrícola de rizobios y klebsiellas" Scientific American Latinoamérica 25: 34-35.

» Martínez Romero, E., Palacios, R. and Mora, J. 1998. "Cepas mejoradas de *Rhizobium*", Investigación y Ciencia (Edición en español de Scientific American), 265: 14-19.



# A Major New Resource Coming Soon from Wiley-Blackwell

## Molecular Microbial Ecology of the

# Rhizosphere

**VOLUMES I AND II**


Edited by

**Frans J. de Bruijn**

*INRA/CNRS Laboratory of Plant-Microbe Interactions*
*Toulouse, France*


- A comprehensive treatise of the *Molecular Microbial Ecology of the Rhizosphere*
- Over 100 Chapters spread over 2 Volumes
- Chapters written by distinguished Contributors, including Speakers of the RHIZOSPHERE 3 International Conference in Perth, Australia and edited by one of the most eminent names in the field

The Rhizosphere, and the soil for that matter, have been considered "Black Boxes", but decades of biochemical, biophysical and ecological studies have allowed a glimpse in the Box. In addition, with the advent of molecular approaches the Rhizosphere has become much more tangible and approachable, leading to a real explosion of knowledge, especially on the structure and function of microbial communities, signalling between microbes and between microbes and the host plant, plant control of microbial community in the Rhizosphere, and other facets of the plant-microbe interactions.

*Molecular Microbial Microbial Ecology of the Rhizosphere* focuses on a subset of the advances made in the area of the Rhizosphere, and highlights those experiments that have been mostly carried out using molecular approaches ranging from community fingerprinting to genomics and metagenomics. The book will have several reviews in each category followed by research chapters, with materials and methods detailed enough to carry out similar experiments or clear referrals to the literature.

**This is an essential book that provides uniquely far-ranging coverage of this rich and developing field.**


**See the reverse side of this flyer for more information on the areas covered and for more about the editor.**


WILEY

# Rhizobial genetic repertoire to inhabit legume and non-legume rhizospheres

**Martha G. López-Guerrero, Miguel Ángel Ramírez and Esperanza Martínez-Romero[1]**

**Centro de Ciencias Genómicas, UNAM. Av. Universidad SN, Chamilpa, Cuernavaca, Mor. México**

Number of words: 4, 215

Number of Figures: 1

[1]Corresponding author: E Martinez-Romero

Email: emartine@ccg.unam.mx, Telephone: +52(777)-329-1692, Fax: +52(777)317-5581

Running title: Rhizospheric rhizobial genes and functions

Keywords: Transcriptomics, *Rhizobium*, root colonization, root survival, plant exudates

**Frans J. de Bruijn (ed).**

***Molecular Microbial Ecology of the Rhizosphere, Volume 1*.**

**First Edition. Wiley-Blackwell. 1326.**

**Mayo 2013.**

**ISBN 978-1-1182-9617-2.**

## Abstract

Rhizosphere colonization seems to be an old rhizobial trick, older that nodulation and similar to the colonization processes of other rhizospheric bacteria. Besides legumes, rhizobia colonize the rhizospheres of many other plants. Root exudates have an important role in rhizosphere colonization and are inducers of rhizobial genes involved in bacterial competitive growth and survival preceded by root adhesion. Examples of bacterial functions needed at the rhizosphere are presented here as well as novel results derived from *Rhizobium* transcriptomic analyses. The need to recognize a heterogeneous bacterial physiology on roots is discussed.


## Introduction

Rhizobia are soil bacteria, best known for forming nitrogen fixing nodules on legume roots. For a long time the importance of rhizobial rhizosphere colonization to achieve nodulation has been emphasized [Hossain and Alexander, 1984]. As is the case with other bacteria, rhizobia are successful inhabitants of plant rhizospheres. *Rhizobium etli* strains may attain large numbers on plant roots, up to $10^9$ cells per gram of maize root (fresh weight) [Gutierrez-Zamora and Martínez-Romero, 2001] and *Rhizobium leguminosarum* strains promote rice, canola and lettuce growth [Noel et al., 1996; Yanni et al., 1997].

Root rhizospheres are rich in nutrients and support the growth and proliferation of diverse microbes [Andrews and Harris, 2000]. In contrast to what occurs in nodules, where rhizobia are exclusive or almost exclusive occupants, in the rhizosphere these bacteria share the habitat with many other microorganisms [Barea et al., 2005]. It is unfortunate that most rhizosphere studies are performed using single strain inoculations and not microbial communities, as proposed [Sørensen et al., 2009].

Besides mucilage and sloughed root cortex tissues, root exudates are an important source of nutrients. Root exudates composition has been analyzed from different species [Rovira, 1969]. Exudates contain, amino acids, organic acids [Nardi et al., 1997], sugars [Gransee, 2002] and vitamins [Rajamani et al., 2008]; while there are common plant exudate molecules there are others that are plant specific [reviewed in Walker et al., 2003]. Plant growth and root age affect exudate composition [Walker et al., 2003]. Exudation is different and more active in some parts of the root [Bringhurst et al., 2001]. At the root base the influx of amino acids is greater than in the other regions [Jones and Darrah, 1994]. Accordingly, microbial community may change in relation to

differences in exudate availability [Baudoin, 2002; Bais et al., 2006]. Microbial colonization on roots is not homogeneous [Watt et al., 2006].

Among root exudates flavonoids have attracted great attention for inducing rhizobial *nod* genes. Flavonoids induce other genes such as those encoding type III secretion systems [Viprey et al., 1998] or efflux pumps [Parniske et al., 1991; Gonzalez-Pasayo and Martinez-Romero, 2000] or auxin production (our unpublished results). Additionally there were reports showing that flavonoids stimulated rhizobial growth [Hartwig et al., 1991]. Flavonoids stimulate competitiveness in *R. leguminosarum* in early stages of the interaction with clover and vetch [Maj et al., 2010]. Induction of rhizobial genes *nif* and *nod* genes in the presence of host plants or in nodules has been known for quite a number of years. Besides flavonoids there are other substances in exudates that induce rhizobial *nod* genes such as phenolic compounds [Le Strange et al., 1990], jasmonates [Rosas et al., 1998; Mabood et al., 2006] and xanthones [Yuen et al., 1995].

Rhizobia have large genomes and differential gene expression seems to account for their adaptation to different niches, as occurs in *Pseudomonas* spp. colonizing rhizospheres [Rainey, 1999; Ramos-González et al., 2005]. Some rhizobial functions expressed in plant rhizospheres are reviewed here in comparison to other rhizospheric bacteria.

## Motility and root adhesion

Legume exudates are rhizobial attractants. Flavonoids, sugars, amino acids and dicarboxylic acids induce chemotaxis [Cooper, 2004]. The response towards amino acids and flavonoids has been observed in different rhizobial species [Cooper, 2007]. In *Pseudomonas fluorescens* chemotaxis is required for successful root colonization, since mutants in the *cheA* gene with reduced flagellar driven chemotaxis have diminished tomato root colonization. Malic and citric acids are attractants in the tomato rhizosphere [de Weert et al., 2002]. Using IVET (*In vivo* expression technology) to study genes expressed in the *Pisum sativa* rhizosphere, the *flgG* gene was found to be associated with chemotaxis and motility of *Rhizobium leguminosarum* sv. viciae [Barr et al 2008].

There is compelling evidence that adhesion to roots is an early *sin e qua non* step in the root colonization process for many bacteria. For *Rhizobium* and *Bradyrhizobium*, attachment is the first and the most important step in

legume plant colonization [reviewed in Albareda et al., 2006; Downie, 2010]. In the case of different bacteria, distinct molecules mediate their adhesion to plant roots [Danhorn and Fuqua, 2007].

The *pilAB* gene products are required for the synthesis of a type IV pili that is needed for the establishment of *Azoarcus* sp. strain BH72 on rice rootlets [Dörr et al., 1998] and is essential for root surface colonization [Böhm et al., 2007]. In *Pseudomonas putida* KT244, a secretion system is involved in the release of large proteins implicated in rhizosphere colonization as well as in iron uptake [Molina et al., 2006].

In the rhizosphere or in the presence of exudates, *Rhizobium* produces novel surface polysaccharides and proteins secreted by type I, III and IV secretion systems, some of which may participate in root adhesion [reviewed in Cooper, 2007; Krehenbrink and Downie, 2008].

Different phases in *Rhizobium* root attachment are recognized. First, the initial adsorption is the result of the interaction between plant lectins, bacterial surface polysaccharides and the $Ca^{2+}$ bacterial binding protein rhicadhesin [Dazzo et al., 1984]. The subsequent step involves cellulose fibrils secreted by bacteria which are responsible for irreversible binding to root surfaces [Laus et al., 2005].

## Catabolism of plant substances

Different plants contain different metabolites, most of them unknown. Plant metabolite uptake and catabolism confer on rhizobia and other bacteria an adaptative advantage to differentially colonize the rhizosphere. Plants may be engineered to produce particular metabolites to select rhizospheric species [Mansouri et al., 2002; see also Chapters 116-119].

Induction of catabolism must occur in the presence of the substance or a related substance. Inducible catabolism genes seem to be unevenly distributed among rhizobial species or rhizobial strains. Novel catabolic genes are likely to be found in rhizobia to profit from soil and plant nutrients. The transporters for such substrates may be unknown as well. It is remarkable that hundreds of ABC transporters of unknown substrate have been found in rhizobial genomes [Gonzalez et al., 2006]. Adhesion, transport and catabolism genes seem to constitute a substantial part of the rhizobial genetic repertoire to inhabit legume and non legume rhizospheres.

Rhizobia may catabolize rhamnose [Oresnik et al., 1998] and arabinogalactan which are common in mucilage [Knee et al., 2001], protocatechuate [MacLean et al., 2011] and a diversity of complex carbohydrate molecules and phenolic compunds [Parke et al., 1991]. IVET was used to identify genes that are differentially expressed in

*R. leguminosarum* bv. viciae A34 in pea rhizosphere. Induced genes are involved in membrane transport, such as those encoding for a sulphonate ABC transport system, indicating that sulphate containing compounds are available at the rhizosphere [Barr et al., 2008]. Other genes encoding transporters were also expressed. RL0362 (*araJ*), encodes a permease of the major facilitator system (MFS) family transporters. The latter probably forms an operon with RL0363, encoding a glyoxalase, and may be involved in transporting and processing arabinose polymers [Barr et al., 2008].

Homoserine is found in pea exudates and may be used by *R. leguminosarum* [van Egeraat, 1975]. Calystegines are plant secondary metabolites first found in *Calystegia sepium* (morning glory) that can be poisonous to arthropods and mammals. They are polyhydroxyl nortropame alkaloids that occur in the Solanaceae. In a particular strain of *Sinorhizobium meliloti*, plasmid genes for catabolism of calystegine contribute to competitive colonization of morning glory [Tepfer et al., 1988] and genetically modified alfalfa plants [Guntli et al., 1999a, b]. Other strains of *S. meliloti* do not metabolize calystegine [Tepfer et al., 1988]. Mimosine produced by *Leucaena* plants may be catabolized by the *mid* genes in *Rhizobium* strains that nodulate them [Borthakur et al., 2003].

Genes that are involved in bean exudate uptake (*teu*) are required in *R. etli* and *R. tropici* for competitive *Phaseolus vulgaris* nodulation. These genes encode an ABC transporter that specifically determines the uptake of a bean exudate molecule, the chemical structure of which has not been determined yet [Rosenblueth et al., 1998]. Proline catabolism genes including *putA* have a role in rhizospheric competitive colonization [van Dillewijn et al., 2002].

In non-sterile conditions, *Sinorhizobium meliloti* is able to use galactosides secreted by alfalfa, clover and barrel medic seeds. Using a biosensor constructed using the *melA* promoter fused to the *gfp* protein and induced in presence of galactosides and galactose, it was found that galactosides are present in defined zones at the lateral root initiation and around root hairs but no at the root tip [Bringhurst et al., 2001].

In contrast to the large diversity of carbon and nitrogen sources used in the rhizosphere, in nodules only few amino acids and dicarboxylic acids are the main nutrients provided by the plant to rhizobial bacteroids [Lodwig and Poole, 2003].

In *Pseudomonas stutzeri* the usage of diverse carbon sources seems to be related to rhizospheric colonization capacity [Yan et al., 2008].

## Biofilm formation and quorum sensing

*S. meliloti* cells form biofilms on alfalfa root surfaces [Bringhurst et al., 2001]. Exopolysaccharides (EPS) are needed to colonize *Arabidopsis thaliana* and *Brassica napus* roots by an unclassified *Rhizobium* strain, but are not required for biofilm formation *in vitro* [Santaella et al., 2008]. Similarly, *Rhizobium leguminosarum* biofilm formation on roots does not require the same gene functions as those needed *in vitro* [Williams et al., 2008; see also Chapters 66-70].

Bacteria regulate gene expression in relation to population density by a mechanism called quorum sensing (QS; see also Chapters 71-80). In this process bacteria produce and secrete an autoinductor which actives a transcriptional regulator which, in turn, regulates the transcription of specific genes. QS controls processes involved in the interaction with eucaryotes, like motility, biofilm formation and the production of toxins, EPS and virulence factors [González and Keshavan, 2006].

In *Pseudomonas fluorescens* 2P24 a QS related system PcoI-PcoR has been identified. A mutant in the *pcoI* gene encoding a N-acylhomoserine lactone synthase has been found to have a reduced capacity to form biofilms on non-biological surfaces and also in the wheat rhizosphere in sterilized and non-sterilized soil. The mutant had a reduced capacity to colonize root tips and rhizosphere [Wei and Zhang, 2006]. N-acyl homoserine lactone has been found to be produced by *P. putida* F117 and *S. liquefaciens* MG44 in the tomato rhizosphere [Steidle et al., 2001].

## Rhizobial protection from plant defense

Plants activate defense responses after bacterial colonization. Consequently rhizobia protect themselves by changing their gene expression at the rhizosphere. Phytoalexin resistance is induced by soybean isoflavonoids in *Bradyrhizobium japonicum* [Parniske et al., 1991]. Similarly, genes encoding a multidrug efflux systems were identified using IVET in *R. leguminosarum* bv. viciae A34 in the rhizosphere [Barr et al., 2008]. *gusA* promoterless genes were introduced as transposons into *R. etli* in a random mutagenesis procedure. *gusA* activity of bacteria grown in minimal medium (MM) or in MM with *P. vulgaris* flavonoids was examined. Insertions expressed only in presence of flavonoids were analyzed. In *R. etli* genes involved in the production of an efflux pump system were detected that were also induced by phytoalexins [Gonzalez-Pasayo and Martínez-Romero, 2000].

There are surface polysaccharides that are important for the establishment of the rhizobial-plant interaction such as lipopolysaccharides (LPS), EPS and cyclic β-glucans. They may act as physical barriers to plant defense compounds or as suppressors of plant defense responses [reviewed in Soto et al., 2009 and Downie, 2010]. In other bacteria effector molecules that are exported to plants by type III secretion modulate plant defense responses [Alfano and Collmer, 2004]. In *Sinorhizobium* strain NGR234 type III secretion genes are induced by flavonoids [Viprey et al., 1998].

## Transcriptional regulators expressed at the rhizosphere

Using IVET two transcriptional regulators from the LysR and GntR families, one sigma factor and two genes involved in environmental sensing that use as second messenger cyclic di-GMP with GGDEF and EAL domains were found to be expressed in *R. leguminosarum* in the pea rhizosphere [Barr et al., 2008]. *Pseudomonas putida* bacteria on maize roots express transcriptional regulators belonging to the AraC and TetR families [Matilla et al., 2007].

A proteomic study of the phylospheric *Methylobacterium extorquens* bacterium on *Arabidopsis thaliana* leaves revealed that a regulatory factor PhyR is required for the expression of a number of stress proteins and is a key regulator in *M. extorquens* for its adaptation to its epiphytic lifestyle [Gourion et al., 2006]. Community proteogenomics have revealed insights into the physiology of phyllospheric bacteria [Delmotte et al., 2009]. Similar approaches may be used to study rhizobial communities in the rhizosphere.

The transcriptional regulator RpoN has a role in *Sinorhizobium meliloti* [Barnett et al., 2004] and in *Rhizobium etli* [Salazar et al., 2010] legume interactions. For another example of the identification of regulatory circuits see Chapter 83.

## Genome wide transcriptomic analyses

A microarray study compared gene expression in *R. leguminosarum* strain 3841 in pea, alfalfa and sugar beet (non -legume) rhizospheres. Many genes expressed in the rhizosphere were identified and plant specific expression was shown to occur [Ramachandran et al., 2011]. Interestingly a *R. leguminosarum* plasmid contains many of the genes expressed in the pea rhizosphere.

Genes with unknown function were found to be highly expressed in *R. phaseoli* Ch24-10 recovered from maize (non-legume) rhizosphere by an RNA-Seq procedure using the Illumina sequencing platform [López-Guerrero et al, submitted]. This is similar to what occurs in *R. leguminosarum* in the pea rhizosphere where 66% of the genes expressed are of unknown function [Ramachandran et al., 2011].

In *R. phaseoli* strain Ch24-10, some genes were commonly expressed in maize and bean rhizospheres but others were plant specific. Bacteria growing on roots were not physiologically homogeneous as rhizobial transcripts reflected conflicting bacterial physiological conditions, some genes would correspond to those from bacteria in rich media while others in starvation. All transcripts obtained may not be present in a single cell. Even though young roots from 5 day old plants were used, it seems that they are not homogeneous niches for bacterial growth. Reporter-gene approaches would allow the spatial detection of gene expression on roots.

**Acknowledgements**

Partial financial support was from PAPIIT IN200709 and IN205412 from UNAM. We thank Julio Martínez for technical help.

## References


Albareda MM, Dardanelli MS, Sousa C, Megías M, Temprano F. 2006. Factors affecting the attachment of rhizospheric bacteria to bean and soybean roots. FEMS Microbiol. Lett. 259:67-73.

Alfano JR, Collmer A. 2004. Type III secretion system effector proteins: double agents in bacterial disease and plant defense. Ann. Rev. Phytopathol. 42:385-414.

Andrews JH, Harris RF. 2000. The ecology and biogeography of microorganisms on plant surfaces. Annu. Rev. Phytopathol. 38:145-180.

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM. 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. Ann. Rev. Plant Biol. 57:233-266.

Barea J-M, Pozo MJ, Azcón R, Azcón-Aguilar C. 2005. Microbial co-operation in the rhizosphere. J. Experimental Bot. 56:1761-1778.

Barnett MJ, Toman CJ, Fisher RF, Long SR. 2004. A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. Proc. Natl. Acad. Sci. USA. 101:16636-16641.

Barr M, East AK, Leonard M, Mauchline TH, Poole PS. 2008. In vivo expression technology (IVET) selection of genes of *Rhizobium leguminosarum* biovar viciae A34 expressed in the rhizosphere. FEMS Microbiol. Lett. 282:219-227.

Baudoin E. 2002. Impact of growth stage on the bacterial community structure along maize roots, as determined by metabolic and genetic fingerprinting. Appl. Soil Ecol. 19:135-145.

Böhm M, Hurek T, Reinhold-Hurek B. 2007. Twitching motility is essential for endophytic rice colonization by the N2-fixing endophyte *Azoarcus* sp. strain BH72. Mol. Plant-Microbe Interact. 20:526-533.

Borthakur D, Soedarjo M, Fox PM, Webb DT. 2003. The *mid* genes of *Rhizobium* sp. strain TAL1145 are required for degradation of mimosine into 3-hydroxy-4-pyridone and are inducible by mimosine. Microbiol. 149:537-546.

Bringhurst RM, Cardon ZG, Gage DJ. 2001. Galactosides in the rhizosphere: utilization by *Sinorhizobium meliloti* and development of a biosensor. Proc. Natl. Acad. Sci. USA 98: 4540-4545.

Cooper JE. 2004. Multiple responses of rhizobia to flavonoids during legume root infection. Adv. Bot. Res. 41: 1–62.

Cooper JE. 2007. Early interactions between legumes and rhizobia : disclosing complexity in a molecular dialogue. J. App. Microbiol. 103: 1355-1365.

Danhorn T, Fuqua C. 2007. Biofilm formation by plant-associated bacteria. Ann. Rev. Microbiol. 61:401-422.

Dazzo FB, Truchet GL, Sherwood JE, Hrabak EM, Abe M, Pankratz SH. 1984. Specific phases of root hair attachment in the *Rhizobium trifolii*-clover symbiosis. Appl. Environ. Microbiol. 48: 1140-50.

Delmotte N, Knief C, Chaffron S, Innerebner G, Roschitzki B, Schlapbach R, Von Mering C, Vorholt JuA. 2009. Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc. Natl. Acad. Sci. USA 106, 16428-16433.

de Weert S, Vermeiren H, Mulders IHM, Kuiper I, Hendrickx N, Bloemberg GV, *et al*. 2002. Flagella-driven chemotaxis towards exudate components is an important trait for tomato root colonization by *Pseudomonas fluorescens*. Mol. Plant-Microbe Interact. 15, 1173-1180.

Dörr J, Hurek T, Reinhold-Hurek B. 1998. Type IV pili are involved in plant-microbe and fungus-microbe interactions. Mol. Microbiol. 30, 7-17.

Downie JA. 2010. The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. FEMS Microbiol. Rev. 34: 150-170.

González JE, Keshavan ND. 2006. Messing with bacterial quorum sensing. Microbiol. Mol. Biol. Rev. 70:859-75.

Gonzalez V, Santamaria RI, Bustos P, Hernandez-Gonzalez I, Medrano-Soto A. 2006. The partitioned *Rhizobium etli* genome: Genetic and metabolic redundancy in seven interacting replicons. Proc. Natl. Acad. Sci. USA. 103:3834-3839.

Gonzalez-Pasayo R, Martinez-Romero E. 2000. Multiresistance genes of *Rhizobium etli* CFN42. Mol. Plant-Microbe Interactions 13:572-577.

Gourion B, Rossignol M, Vorholt JA. 2006. A proteomic study of *Methylobacterium extorquens* reveals a response regulator essential for epiphytic growth. Proc. Natl. Acad. Sci. USA 103, 13186-13191.

Gransee A. 2002. Effects of root exudates on nutrient availability in the rhizosphere. Plant Nutrition 92: 626–627.

Guntli D, Burgos S, Moeenne-Loccoz Y, Defago G. 1999a. Calystegine degradation capacities of microbial rhizosphere communities of *Zea mays* (calystegine-negative) and *Calystegia sepium* (calystegine-positive). FEMS Microbiol. Ecol. 28:75-84.

Guntli D, Heeb M, Moenne-Loccoz Y, Defago G. 1999b. Contribution of calystegine catabolic plasmid to competitive colonization of the rhizosphere of calystegine-producing plants by *Sinorhizobium meliloti* Rm41. Mol. Ecol. 8:855-863.

Gutierrez-Zamora ML, Martinez-Romero E. 2001. Natural endophytic association between *Rhizobium etli* and maize (*Zea mays* L.). J. Biotechnol. 91:117-126.

Hartwig UA, Joseph CM, Phillips DA. 1991. Flavonoids released naturally from alfalfa seeds enhance growth rate of *Rhizobium meliloti*. Plant Physiol. 95:797-803.

Hossain AK, Alexander M. 1984. Enhancing soybean rhizosphere colonization by *Rhizobium japonicum*. Appl. Environ. Microbiol. 48: 468-742.

Jones DL, Darrah PR. 1994. Amino acid influx at the soil-root interface of *Zea mays* L and its implications in the rhizosphere. Plant and Soil 163:1-12.

Krehenbrink MM, Downie JA. 2008. Identification of protein secretion systems and novel secreted proteins in *Rhizobium leguminosarum* bv. viciae. BMC Genomics 9: 55.

Knee EM, Gong F-C, Gao M, Teplitski M, Jones AR, Foxworthy A, Mort AJ, Bauer WD. 2001. Root mucilage from pea and its utilization by rhizosphere bacteria as a sole carbon source. Mol. Plant-Microbe Interact. 14:775-784.

Laus MC, van Brussel, AAN, Kijne JW. 2005. Role of cellulose fibrils and exopolysaccharides of *Rhizobium leguminosarum* in attachment to and infection of Vicia sativa root hairs. Mol. Plant-Microbe Interact. 18:533-538.

Le Strange KK, Bender GL, Djordjevic MA, Rolfe BG, Redmond JW. 1990. The *Rhizobium* strain NGR234 *nodD1* gene product responds to activation by the simple phenolic compounds vanillin and isovanillin present in wheat seedling extracts. Mol. Plant-Microbe Interact. 3:214-220.

Lodwig E, Poole P. 2003. Metabolism of *Rhizobium* bacteroids. Crit. Rev. Plant Sci. 22:37-78.

Mabood F, Souleimanov A, Khan W, Smith DL. 2006. Jasmonates induce Nod factor production by *Bradyrhizobium japonicum*. Plant Physiol. Biochem. 44:759-765.

MacLean AM, Haerty W, Golding GB, Finan TM. 2011. The LysR-type PcaQ protein regulates expression of a protocatechuate-inducible A 1 BC-type transport 2 system in *Sinorhizobium meliloti*. Microbiology doi:10.1099/mic.0.050542-0

Maj D, Wielbo J, Marek-Kozaczuk M, Skorupska A. 2010. Response to flavonoids as a factor influencing competitiveness and symbiotic activity of *Rhizobium leguminosarum*. Microbiol. Res. 165:50–60.

Mansouri H, Petit A, Oger P, Dessaux, Y. 2002. Engineered rhizosphere: the trophic bias generated by opine-producing plants is independent of the opine type, the soil origin, and the plant species. Appl. Environ. Microbiol. 68:2562-2566.

Matilla MA, Espinosa-Urgel M, Rodríguez-Herva JJ, Ramos JL, Ramos-González MI. 2007. Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. Genome Biol. 8:R179.

Molina MA, Ramos JL, Espinosa-Urgel M. 2006. A two-partner secretion system is involved in seed and root colonization and iron uptake by *Pseudomonas putida* KT2440. Environ. Microbiol. 8:639-647.

Nardi S, Reniero F, Concheri G. 1997. Soil organis matter mobilization by root exudates of three maize hybrids. Chemosphere 35:2237-2244.

Noel TC, Sheng C, Yost CK, Pharis RP, Hynes MF. 1996. *Rhizobium leguminosarum* as a plant growth-promoting rhizobacterium: Direct growth promotion of canola and lettuce. Can. J. Microbiol. 42:279-283.

Oresnik IJ, Pacarynuk LA, O'Brien SAP, Yost CK, Hynes MF. 1998. Plasmid-encoded catabolic genes in *Rhizobium leguminosarum* bv. trifolii: evidence for a plant-inducible rhamnose locus involved in competition for nodulation. Mol. Plant-Microbe Interactions 11:1175-1185.

Parke D, Rynne F, Glenn A. 1991. Regulation of phenolic catabolism in *Rhizobium leguminosarum* biovar trifolii. J. Bacteriol. 173:5546-5550.

Parniske M, Ahlborn B, Werner D. 1991. Isoflavonoid-inducible resistance to the phytoalexin glyceollin in soybean rhizobia. J. Bacteriol. 173:3432-3439.

Rainey PB. 1999. Adaptation of *Pseudomonas fluorescens* to the plant rizhosphere. Environ. Microbiol. 1:243-257.

Rajamani S, Bauer WD, Robinson JB, Farrow JM, Pesci EC, Teplitski M, *et al.* 2008. The vitamin riboflavin and its derivative lumichrome activate the LasR bacterial quorum-sensing receptor. Mol. Plant-Microbe Interact. 21: 1184-92.

Ramachandran V, East A, Karunakaran R, Downie A, Poole PS. 2011. Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biol. 12:R106.

Ramos-Gonzalez MI, Campos MJ, Ramos JL. 2005. Analysis of *Pseudomonas putida* KT2440 gene expression in the maize rhizosphere: *In vitro* expression technology capture and identification of root-activated promoters. J. Bacteriol. 187:4033-4041.

Rosas S, Soria R, Correa N, Abdala G. 1998. Jasmonic acid stimulates the expression of *nod* genes in *Rhizobium*. Plant Mol. Biol. 38:1161-1168.

Rosenblueth M, Hynes MF, Martinez-Romero E. 1998. *Rhizobium tropici teu* genes involved in specific uptake of *Phaseolus vulgaris* bean-exudate compounds. Mol. Gen. Genet. 258:587-598.

Rovira AD. 1969. Plant root exudates. Bot. Rev. 35:35-57.

Salazar E, Díaz-Mejía JJ, Moreno-Hagelsieb G, Martínez-Batallar G, Mora Y, Mora J, Encarnación S. 2010. Characterization of the NifA-RpoN regulon in *Rhizobium etli* in free life and in symbiosis with *Phaseolus vulgaris*. Appl. Environ. Microbiol. 76: 4510-4520.

Santaella C, Schue M, Berge O, Heulin T, Achouak W. 2008. The exopolysaccharide of *Rhizobium* sp. YAS34 is not necessary for biofilm formation on *Arabidopsis thaliana* and *Brassica napus* roots but contributes to root colonization. Environ. Microbiol. 10:2150-2163.

Sørensen J, Nicolaisen MH, Ron E, Simonet P. 2009. Molecular tools in rhizosphere microbiology - from single cells to whole-community analysis. Plant Soil 321:483-512.

Soto MJ, Domínguez-Ferreras A, Pérez-Mendoza D, Sanjuán J, Olivares J. 2009. Mutualism versus pathogenesis: the give-and-take in plant-bacteria interactions. Cell. Microbiol. 11:381-388.

Steidle A, Sigl K, Schuhegger R, Ihring A, Schmid M, Gantner S, *et al*. 2001. Visualization of N-Acylhomoserine lactone-mediated cell-cell communication between bacteria colonizing the tomato rhizosphere. J. Bacteriol. 67: 5761-5770.

Tepfer D, Goldmann A, Pamboukdjian N, Maille M, Lepingle A, Chevalier D, Denarie J, Rosenberg C. 1988. A plasmid of *Rhizobium meliloti* 41 encodes catabolism of two compounds from root exudate of *Calystegium sepium*. J. Bacteriol. 170:1153-1161.

van Dillewijn P, Villadas PJ, Toro N. 2002. Effect of a *Sinorhizobium meliloti* strain with a modified *putA* gene on the rhizosphere microbial community of alfalfa. Appl. Environ. Microbiol. 68:4201-4208.

van Egeraat A. 1975. The possible role of homoserine in the development of *Rhizobium leguminosarum* in the rhizosphere of pea seedlings. Plant and Soil 386: 381-386.

Viprey V, Del Greco A, Golinowski W, Broughton WJ, Perret X. 1998. Symbiotic implications of type III protein secretion machinery in *Rhizobium*. Mol. Microbiol. 28:1381-1389.

Walker TS, Bais HP, Grotewold E, Vivanco JM. 2003. Root exudation and rhizosphere biology. Plant Physiol. 132:44-51.

Watt M, Silk WK, Passioura JB. 2006. Rates of root and organism growth, soil conditions, and temporal and spatial development of the rhizosphere. Ann. Botany. 97: 839-55.

Wei HL, Zhang LQ. 2006. Quorum-sensing system influences root colonization and biological control ability in *Pseudomonas fluorescens* 2P24. Antonie Van Leeuwenhoek. 89:267-80.

Williams A, Wilkinson A, Krehenbrink M, Russo DM, Zorreguieta A. 2008. Glucomannan-mediated attachment of *Rhizobium leguminosarum* to pea root hairs is required for competitive nodule infection. J. Bacteriol. 190:4706.

Yan Y, Yang J, Dou Y, Chen M, Ping S. 2008. Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501. Proc. Natl. Acad. Sci. USA. 105:7564-7569.

Yanni YG, Rizk RY, Corich V, Squartini A, Ninke K. 1997. Natural endophytic association between *Rhizobium leguminosarum* bv. trifolii and rice roots and assessment of its potential to promote rice growth. Plant Soil 194:99-114.

Yuen JPY, Cassini ST, De Oliveira TT, Nagem TJ, Stacey G. 1995. Xanthone induction of *nod* gene expression in *Bradyrhizobium japonicum*. Symbiosis 19:131-140.

Systematic and Applied Microbiology 35 (2012) 353–358

Contents lists available at SciVerse ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

# *Rhizobium etli* taxonomy revised with novel genomic data and analyses

Martha G. López-Guerrero[a], Ernesto Ormeño-Orrillo[a], Encarna Velázquez[b], Marco A. Rogel[a], José Luis Acosta[a,1], Victor Gónzalez[a], Julio Martínez[a], Esperanza Martínez-Romero[a,*]

[a] *Centro de Ciencias Genómicas, UNAM, Cuernavaca, Morelos, Mexico*
[b] *Departamento de Microbiología y Genética, Universidad de Salamanca, Salamanca, Spain*

ARTICLE INFO

*Article history:*
Received 8 March 2012
Received in revised form 14 May 2012
Accepted 14 June 2012

*Keywords:*
Nitrogen fixation
Legume nodulation
*Phaseolus vulgaris*

ABSTRACT

The taxonomic position of *Phaseolus vulgaris* rhizobial strains with available sequenced genomes was examined. Phylogenetic analyses with concatenated conserved genomic fragments accounting for over half of each genome showed that *Rhizobium* strains CIAT 652, Ch24-10 (newly reported genome) and CNPAF 512 constituted a well-supported group independent from *Rhizobium etli* CFN 42$^T$. DNA–DNA hybridization results indicated that CIAT 652, Ch24-10 and CNPAF 512 could correspond to *R. etli*, although the hybridization values were at the borderline that distinguishes different species. In contrast, experimental hybridization results were higher (over 80%) with *Rhizobium phaseoli* type strain ATCC 14482$^T$ in congruence to phylogenetic and ANIm analyses. The latter criterion allowed the reclassification of *R. etli* strains 8C-3 and Brasil5 as *R. phaseoli*. It was therefore concluded, based on all the evidence, that the CIAT 652, Ch24-10, and CNPAF 512 strains should be reclassified as *R. phaseoli* in spite of several common features linking them to *R. etli*. The *R. phaseoli* and *R. etli* speciation process seems to be a more recent event than the speciation that has occurred among other sister species, such as *R. leguminosarum*–*R. etli* or *R. rhizogenes*–*R. tropici*.



## Introduction

The advent of the genomic era has provided both a plethora of molecular markers useful in taxonomy and the possibility to compare whole genomes instead of a few genes. Toward this end, novel algorithms and parameters have been proposed to compare genomes for taxonomic purposes. ANI (average nucleotide identity) has been defined as a very useful parameter to delineate different species and it correlates with DNA–DNA hybridization (DDH) [12,19,32]. Based upon a large set of experimental results from diverse bacteria, thresholds of ANI (94–96%) have been recommended for distinguishing species.

*Phaseolus vulgaris* (common bean) is the legume grain most consumed for human nutrition and, like other legumes, it forms symbiosis with nitrogen-fixing bacteria. *P. vulgaris* symbiotic bacteria have been widely studied [4,6,13,16,22,24,31,34,37,43] and this legume has become a model for studying nodule-bacterial diversity from plants grown in diverse conditions or geographical regions where bean is native or introduced. In its sites of origin and in some introduced areas, *Rhizobium etli* has been reported as the dominant *P. vulgaris* bean nodule bacterium identified on the basis of 16S rRNA gene sequences [22]. In addition, core gene sequences have been used to characterize nodule isolates, thus providing a better phylogenetic resolution and revealing that other species besides *R. etli* can also represent a significant fraction of the bean nodule occupants [4,13,34,37]. Recently, *R. etli* type strain CFN 42$^T$ was found to have a low recombination with *R. etli* CIAT 652 and other *P. vulgaris* isolates [2]. Richter and Rosselló-Móra [32] calculated the ANI of CFN 42$^T$ and CIAT 652, and reported that CIAT 652 was not a member of *R. etli*. DDH results were not available for comparison and the species affiliation of CIAT 652 was not identified. At the same time, *Rhizobium phaseoli* was re-recognized as a valid and different species from *R. etli* because it presented divergent core genes and experimental DNA–DNA relatedness values significantly lower than 70% with *R. etli* CFN 42$^T$ [30]. *R. etli* and *R. phaseoli* are sympatric species nodulating bean and both have been found recently in Ethiopia [4]. On a phylogenetic basis, using *recA*, *atpD* and *celC* partial gene sequences, CIAT 652 was recognized as belonging to *R. phaseoli* [34] but this taxonomic affiliation was not otherwise confirmed by DDH analysis. Based on the same phylogenetic analysis, the Mim2 strain (a *Mimosa affinis* isolate) was also recognized as *R. phaseoli*, however, DDH and multilocus enzyme electrophoresis (MLEE) placed it within *R. etli* [44]. Clearly, there were conflicting data. Further, it was recommended that "the taxonomic status of the strains currently named *R. etli*

* Corresponding author.
*E-mail address:* esperanzaeriksson@yahoo.com.mx (E. Martínez-Romero).
[1] Present address: Instituto Nacional de Medicina Genómica, México DF, Mexico.

0723-2020/$ – see front matter © 2012 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.syapm.2012.06.009

**Table 1**
*Rhizobium etli* and *R. phaseoli* strains used in this study.

| Strains | Host | Origin | Reference |
|---|---|---|---|
| *R. phaseoli* | | | |
| Ch24-10 | *Zea mays* and *Phaseolus vulgaris* | Puebla, Mexico | [35] |
| CIAT 652 | *P. vulgaris* | Buitrera, Colombia | [42] |
| CNPAF 512 | *P. vulgaris* | Brazil | [9,23] |
| ATCC 14482$^T$ | *P. vulgaris* | Beltsville, Maryland | ATCC |
| *R. etli* | | | |
| CFN 42$^T$ | *P. vulgaris* | Guanajuato, Mexico | [29] |

should be revised" [34]. Consequently, our aim in this study was to revise the taxonomic status of different *R. etli* strains with available sequenced genomes.

## Materials and methods

### *Strains, growth and DNA extraction*

Strains used in the DDH studies are shown in Table 1. They were grown in 5 mL liquid PY medium for DNA extraction. Ch24-10 was grown on PY plates and fresh cultures were grown in 50 mL liquid PY. DNA was extracted by the DNA Isolation Kit for Cells and Tissues (Roche, USA). *Rhizobium* strains were maintained in YM with glycerol at −70 °C.

### *Phylogenetic analysis*

*rpoB* primers and PCR conditions were as described previously [21,25]. PCR sequences were compared to sequences obtained from whole genomes. Sequence alignments were generated and edited with BioEdit 7 [14]. Percentage identity between sequences was obtained after removing all columns with gaps from the alignments. Best-fit models of sequence evolution were selected for each gene with JModelTest 0.1.1 using the Akaike information criterion [28]. Maximum likelihood (ML) and neighbor-joining (NJ) phylogenies were constructed with Mega 5 [39]. Support for tree nodes was evaluated by bootstrap analysis with 100 or 1000 pseudoreplicates for ML and NJ, respectively.

### *Genomic sequencing*

The genomic sequence from strain Ch24-10 was obtained using two platforms: the Roche 454 pyrosequencing system (350 bp from 3K long-tag paired end sequencing protocol in the Genome Sequencer FLX) using a commercial service, and the Illumina technology (*Genome Analyzer GAIIx*, paired-end protocol, 200 base pairs-inserts library, reads with 36 nucleotides in length) at the Unidad Universitaria de Secuenciación Masiva de DNA (USMDNA) of the Universidad Nacional Autónoma de México (UNAM).

### *Genome assembly*

Roche reads were *de novo* assembled using GSassembler, Newbler version 2.5.3 with default parameters. Illumina reads were assembled using TAIPAN [36] and then SSAKE [45] separately, both with default parameters. The contigs generated by these programs were reassembled using minimus2 [http://sourceforge.net/apps/mediawiki/amos/index.php?title=Minimus2].

Both assemblies, the one generated with Newbler and the one generated with minimus2 (TAIPAN + SSAKE), were reassembled again to generate a hybrid assembly with minimus2.

### *Prediction and annotation of the Ch24-10 genome*

The contigs generated with the minimus2 assembler were assembled in scaffolds with ABACAS [5] using the complete genomes of CFN 42$^T$ and CIAT 652 as anchors. Next, the pseudochromosome and pseudoplasmids were constructed for each replicon armed with ABACAS, by adding the following sequence "NNNNCATTCCATTCATTAATTAATTAATGAATGAATGNNNNN" (containing the six open reading frames) at the 5′ end of each contig [40]. Prediction and annotation of genes for each pseudoreplicon were undertaken with the CG-Pipeline program [17]. The draft genome sequence of *Rhizobium* Ch24-10 obtained by hybrid assembly using sequences derived by Illumina and 454 Roche technologies was then used. An 80× genome coverage was obtained. The results of this Whole Genome Shotgun project have been deposited at DDBJ/EMBL/GenBank under the accession AHJU00000000. The version described in this paper is the first version, AHJU01000000.

### *Genome tree construction*

Regions shared among *Rhizobium* strains CIAT 652, CFN 42$^T$, Ch24-10, and CNPAF 512, *Rhizobium leguminosarum* sv. viciae 3841 and *Sinorhizobium meliloti* 1021 were identified using Mugsy [3] with the following parameters: minlength = 30, distance = 100, duplications 1, and fullsearch refine. Each orthologous region was extracted by an *ad hoc* Perl script (homemade) and filtered for long gaps with trimAl [8] with the parameter -automated1. All regions were concatenated to build a genome tree by the ML and neighbor-net network methods. ML was performed using RaxML [38] with the GTR + I + G nucleotide substitution model, 1000 distinct randomized maximum parsimony trees and the parameters: p 12345, e 0.0000001, c 8, j STRICT, and k 1000. Finally, the neighbor-net was created with the Splits Tree 4 program [15]. Other genome comparisons were performed as described [20].

### *ANIm*

ANIm values were calculated using the JSpecies package [32].

### *DNA–DNA hybridization (DDH)*

The procedure was as described in [41,44]. DNA was quantified with NanoDrop 2000 (Thermo Scientific) and in gels. DNA was digested with the *Eco*R1 restriction enzyme and electrophoresis was performed in 1% agarose gels. Only lanes with homogeneous DNA quantities were hybridized in Southern blot experiments to total DNA from reference strains with probes labeled with [α-32P] dCTP using RediPrimeTM II (GE Healthcare). Rapid-hyb buffer was used for hybridization and washings were carried out using 2× to 1× SSC with 0.1% SDS at 65 °C.

Filters were cut and individual lanes were counted in scintillation liquid in a LS6500 multi-purpose scintillation counter (Beckman Coulter). Three independent experiments were performed with similar results.

## Results

### *Phylogenetic and similarity gene analysis*

*R. etli* and *R. phaseoli* strains had highly similar 16S rRNA genes (over 99.3% identical) but were slightly separated in a NJ phylogenetic tree (Fig. 1). Ch24-10, CNPAF 512 and CIAT 652 strains reported as *R. etli* were found to group with *R. phaseoli* in the phylogenetic analysis of housekeeping genes *recA*, *atpD* and *rpoB* (Fig. 2). Furthermore, the genomic-based phylogenetic analyses

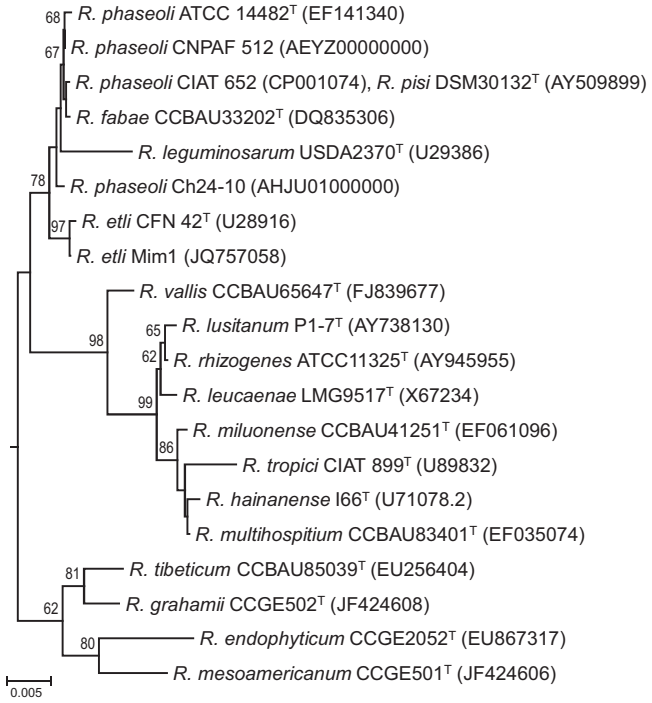



**Fig. 1.** Neighbor-joining phylogenetic tree based on nearly full length 16S rRNA gene sequences from *Rhizobium* and related genera. Only bootstrap supports larger than 50% are indicated. Bar, 5 nt substitutions per 1000 nt.

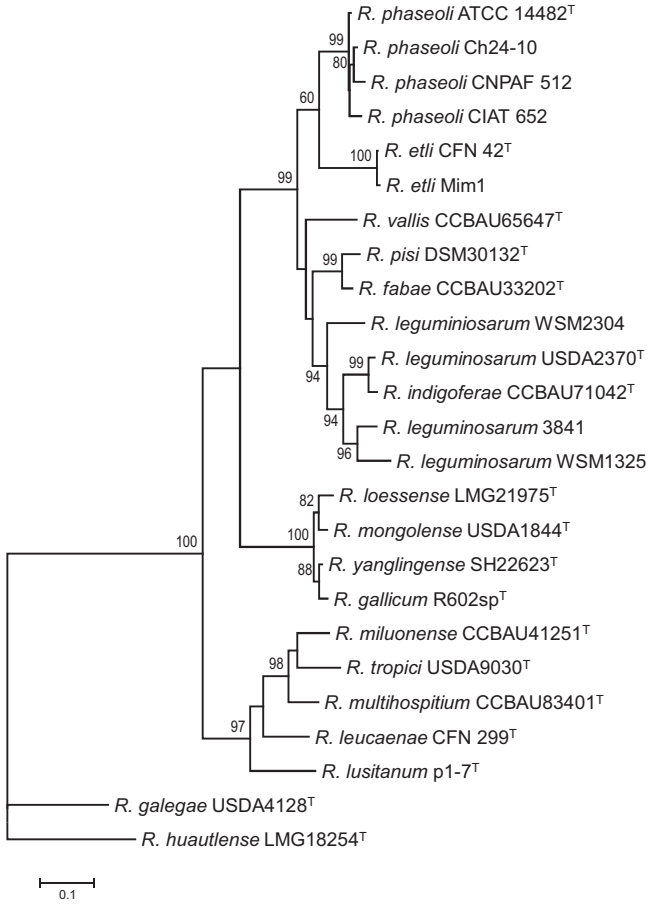



**Fig. 2.** Maximum likelihood phylogenetic tree based on partial concatenated sequences of *recA*, *atpD* and *rpoB* genes of *Rhizobium etli*, *Rhizobium phaseoli* and other rhizobia. Only bootstrap supports larger than 50% are indicated. Bar, 1 nt substitution per 10 nt.

**Table 2**
Nucleotide sequence identity for five complete core genes between pairs of sister *Rhizobium* species.

| Gene | Re/Rp | Re/Rl | Rr/Rt |
|---|---|---|---|
| *dnaK* | 95.3–95.8 | 94 | 93.4 |
| *recA* | 91.7–92.3 | 90.9 | 92.5 |
| *atpD* | 94.3–94.5 | 90.7 | 91.5 |
| *rpoB* | 95.2–95.4 | 94.3 | 93.6 |
| *celC* | 87.2–87.7 | 85.2 | 71.4 |

Re, *R. etli* CFN 42$^T$; Rp, *R. phaseoli* CIAT 652, Ch24-10 and CNPAF 512; Rl, *R. leguminosarum* 3841; Rr, *R. rhizogenes* K84; Rt, *R. tropici* CIAT 899$^T$.

clearly showed that Ch24-10, CNPAF 512 and CIAT 652 strains constituted a well-supported group separated from *R. etli* CFN 42$^T$ (Fig. 3).

Single gene comparisons showed that there were slightly less differences between each of the strains in the group (CIAT 652, CNPAF 512 or Ch24-10) and *R. etli* CFN 42$^T$ than those found for the corresponding genes between *Rhizobium* sister species, such as *R. tropici*-*R. rhizogenes* or *R. leguminosarum*-*R. tropici* (Table 2). The whole core genes compared were *dnaK*, *recA*, *atpD*, *rpoB* and *celC*. In particular, *dnaK*, which has been used as a marker to distinguish species in rhizobia [33], was 94% identical between *R. etli* and *R. leguminosarum*, 93.4% between *R. tropici* and *R. rhizogenes*, and slightly more conserved (95.3–95.8%) between each of the strains (CIAT 652, CNPAF 512 or Ch24-10) and *R. etli* CFN 42$^T$.

### *DNA–DNA hybridization*

DDH between strains CIAT 652 and CFN 42$^T$ was 68%, which was lower than the estimated value for species differentiation (Table 3). DDH between strains CNPAF 512 and CFN 42$^T$ was 67%. Results of the experimental DDH of Ch24-10 to CFN 42$^T$ were 75% (Table 3). Ch24-10 and CFN 42$^T$ showed DDH values sufficient to be considered as belonging to the same species, *R. etli*.

Our experimental DDH results showed values well over 70% for ATCC 14482$^T$ with CIAT 652 (84%), Ch24-10 (88%) and CNPAF 512 (84%), and they constituted the most important results for allocating these strains to *R. phaseoli*.

### *ANIm estimations*

The *R. etli* CFN 42$^T$ and either Ch24-10 or CNPAF 512 estimated ANIm was 90.5% (Table 4), which was in agreement with that previously estimated for CFN 42$^T$ and CIAT 652 by Richter and Roselló-Móra [32]. We estimated the *R. etli* CFN 42$^T$ and *R. leguminosarum* 3841 ANIm as 87.9%. ANIm values were also calculated for other strains with available partial genome sequences that are presently classified as *R. etli* (Table 4) [9,11]. It has been shown that sequences covering 20% of the genome size are enough to obtain reliable estimates of ANIm [32]. Strains 8C-3 and Brasil5 showed ANIm values ≥96.7% with CIAT 652, Ch24-10 and CNPAF

**Table 3**
Experimental DNA–DNA hybridization.

| | *R. etli* CFN 42$^T$ | *R. phaseoli* CNPAF 512 | *R. phaseoli* ATCC 14482$^T$ |
|---|---|---|---|
| *R. etli* CFN 42$^T$ | 100 | 67 ± 3 | 70 ± 2 |
| *R. phaseoli* CIAT 652 | 68 ± 3 | 76 ± 4 | 84 ± 3 |
| *R. phaseoli* Ch24-10 | 75 ± 6 | 78 ± 0 | 88 ± 6 |
| *R. phaseoli* CNPAF 512 | 67 ± 3 | 100 | 84 ± 2 |
| *R. phaseoli* ATCC 14482$^T$ | 69 ± 4 | 78 ± 5 | 100 |

In the experimental procedure radioactive DNA was from CFN 42$^T$, CNPAF 512 or ATCC 14482$^T$.
± standard error.

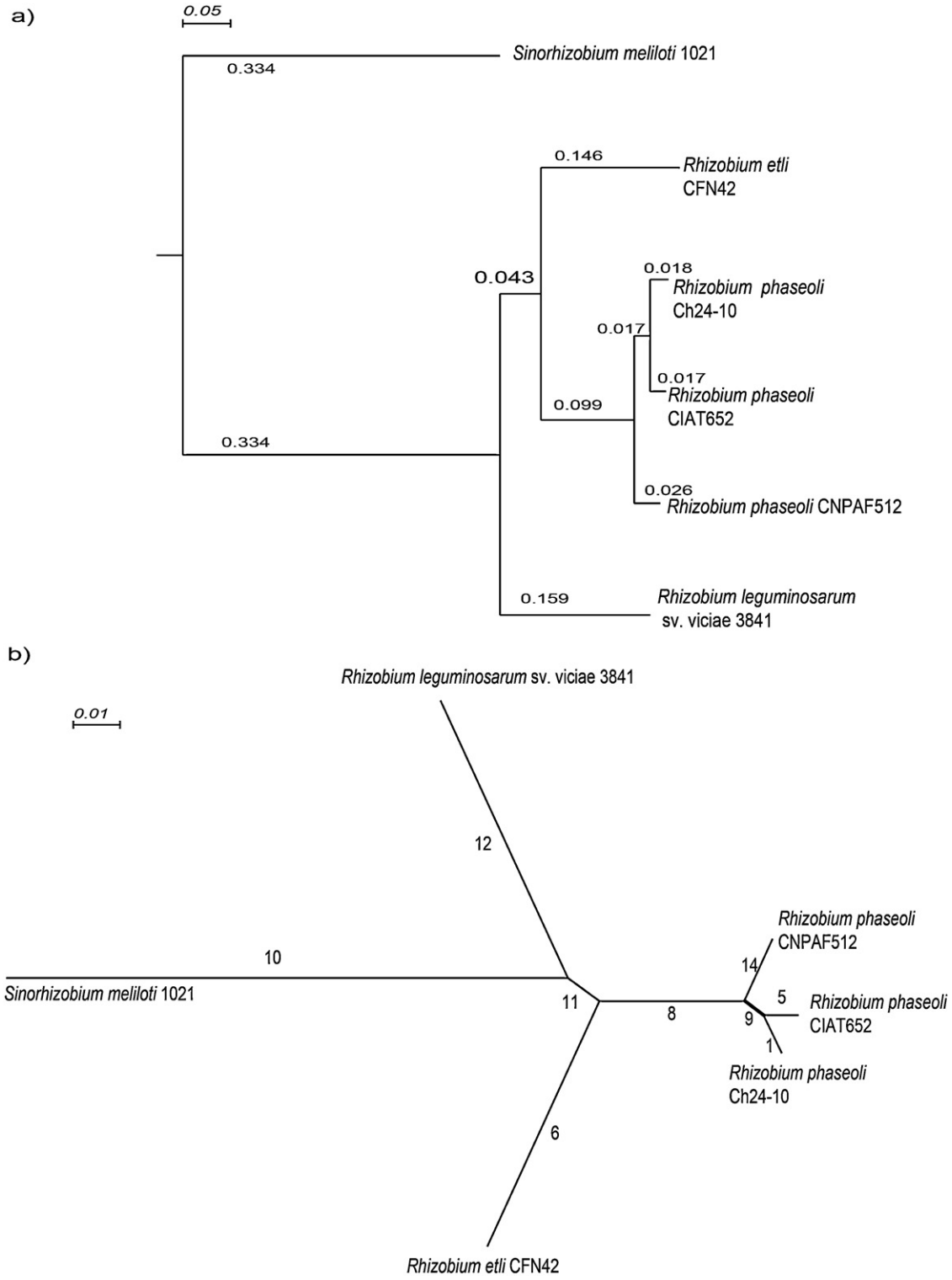



**Fig. 3.** Phylogenetic trees of concatenated genomic conserved regions (3,441,383 pb) in 1373 (coding and non-coding) fragments, on average 2.5 kb. (a) Rooted tree by maximum likelihood with *Sinorhizobium meliloti* 1021 as an outgroup. The scale bar denotes the expected number of nucleotide substitutions per site. (b) Network joining phylogeny inferred from the same concatenated fragments. The tree is unrooted and the number on each branch denotes the split decomposition according to the consensus. The scale bar denotes the number of nucleotide substitutions per site.

512 indicating that they may also belong to *R. phaseoli*. Strains IE4771 and Kim5 shared a high ANIm value between them (96.9%) and may constitute a separate species from *R. etli* and *R. phaseoli* with low ANIm values to both species (89.8–92.3). Strain GR56 may belong to another species, although it is closely related to strains IE4771 and Kim5 with ANIm values of 93–93.4%. Finally, strain CIAT 894 showed low ANI values with all strains indicating that it belonged to another independent species. The same relationships were previously noted in a genomic tree reported for the same strains [11].

**Table 4**
ANIm values between different bean nodulating *Rhizobium* strains with complete or partial genome sequences available.

| | Strain | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | CFN 42$^T$ | 100 | | | | | | | | | |
| 2 | CIAT 652 | 90.5 | 100 | | | | | | | | |
| 3 | Ch24-10[a] | 90.5 | 98.1 | 100 | | | | | | | |
| 4 | CNPAF 512[a] | 90.5 | 97.3 | 97.1 | 100 | | | | | | |
| 5 | 8C-3[a] | 90.9 | 98.2 | 97.6 | 97.1 | 100 | | | | | |
| 6 | Brasil5[a] | 90.7 | 97 | 96.8 | 97.5 | 96.7 | 100 | | | | |
| 7 | IE4771[a] | 89.8 | 91.8 | 91.7 | 91.8 | 91.8 | 91.8 | 100 | | | |
| 8 | Kim5[a] | 90.4 | 92.3 | 92.3 | 92.3 | 92.4 | 92.4 | 96.9 | 100 | | |
| 9 | GR56[a] | 90.6 | 92.8 | 92.7 | 92.9 | 92.8 | 92.7 | 93 | 93.4 | 100 | |
| 10 | CIAT 894[a] | 89.5 | 90.2 | 89.9 | 90 | 90.7 | 90.5 | 89.3 | 90 | 90.3 | 100 |

[a] Based on partial genome sequence. The minimum percentage of genome sequenced was 62.6% for strain 8C-3 [11].

## Discussion

Hybridization to *R. etli* CFN 42$^T$ showed that strains CIAT 652, Ch24-10 and CNPAF 512 are at the borderline that distinguishes different species. It is worth considering that *R. phaseoli* and *R. etli* are barely distinguished by 16S rRNA gene phylogenies and based on this marker they were all recognized as *R. etli*. A discrepancy was shown between the hybridization and the phylogenetic results in which borderline strains (in the DDH analysis) are clearly not *R. etli* but are in fact *R. phaseoli*. In other cases, borderline strains have been considered to correspond to subspecies within a single species [7]. However, there are examples of new species described with borderline values of DDH [46] and it has been recommended that bacterial species comprising distinct genetic subgroups should be split into different species [1].

The conclusion that strains Ch24-10, CIAT 652 and CNPAF 512 are not *R. etli* but belong to *R. phaseoli* is supported by the following considerations: previous [32] and new ANIm analysis, such as the low 90.5% calculated for *R. etli* CFN 42$^T$ and CIAT 652, Ch24-10 or CNPAF 512; the suggested low recombination of CIAT 652, 8C-3 and Brasil5 with *R. etli* CFN 42$^T$ [2] that indicates different evolutionary histories; and new results on the genomic-based phylogenetic analysis and the experimental DNA–DNA hybridization that showed Ch24-10, CIAT 652 and CNPAF 512 were closely related to *R. phaseoli* type strain ATCC 14482$^T$. Such a conclusion supports the previous claim that CIAT 652 corresponded to *R. phaseoli* based on partial *recA*, *atpD* and *celC* gene sequences [34]. From a practical point of view, a phylogenetically based decision with genes other than ribosomal genes seems congruent with the type of data that are normally used to identify strains in rhizobial studies.

Common genes encountered in *R. etli* CFN 42$^T$ and in different *R. phaseoli* strains (data not shown) and the close relatedness of some core genes in *R. phaseoli* and *R. etli* (Table 2) may be explained as a result of an ongoing speciation process that has not been as long as the divergence process occurring among the sister species *R. leguminosarum*–*R. etli* or *R. rhizogenes*–*R. tropici*. As speciation is occurring, many characteristics may still be common between *R. etli* and *R. phaseoli*, such as the 16S ribosomal gene sequences, the physicochemical properties of 705 syntenic ortholog products from CFN 42$^T$ and CIAT 652 that showed an almost identical pattern of polarity between the two strains [26], their sharing of lytic phages (V. Gonzalez and G. Davila, unpublished data) or their close relatedness in MALDI analyses [10] or in MLEE [44]. It is also remarkable that there are very few distinct phenotypes reported that distinguish *R. etli* CFN 42$^T$ from *R. phaseoli* ATCC 14482$^T$ [30]. *R. etli* and *R. phaseoli* speciation occurs within the same environment, since both species are sympatric and nodulate bean. Curiously, the *R. phaseoli* strains analyzed (CIAT 652, Ch24-10, CNPAF 512 and Brasil5) have a better capacity to fix nitrogen in bean than *R. etli* CFN 42$^T$ (data not shown). It seems that Mexican strains *R. phaseoli* Ch24-10 and *R. etli* CFN 42$^T$ are slightly more similar (experimental DDH values over 70%; Table 2) than the other *R. phaseoli* strains from distant geographic regions, such as Colombia (CIAT 652) or Brazil (CNPAF 512), and this may be in relation to the possibility of local exchange of genetic material, especially symbiotic plasmids. In agreement with this, the *R. phaseoli* strain Ch24-10 *nodC* gene but not the *fixL* gene is more similar to that of *R. etli* CFN 42$^T$ than to that of *R. phaseoli* CIAT 652 (data not shown).

Methods to analyze bacterial populations used in the past, such as MLEE, showed that *R. etli* was highly diverse and it was discussed that some of the most distant groups could be considered as different species (Fig. 2 in [27]). New genomic-based analysis using ANIm confirms this observation (Table 4) and indicates that lineages within *R. etli*, such as that represented by CIAT 894 from Colombia or IE4771 and Kim5 or GR56, correspond to novel undescribed species. *R. etli*, *R. phaseoli* and related lineages provide novel examples of the "sequence-discrete groups" that "reflect presumably some fundamental properties of the microbial world" [18].

## Acknowledgments

To PAPIIT IN2054 from UNAM, CONACyT CB 131499, the SUBNARGEM (Subsistema Nacional de Recursos Genéticos Microbianos) grant from SAGARPA Mexico for financial support, and Michael Dunn for kindly reviewing the manuscript. Martha G. López-Guerrero acknowledges the CONACyT fellowship during her Ph.D. studies in the Doctorado en Ciencias Biomédicas program at UNAM.

## References

[1] Achtman, M., Wagner, M. (2008) Microbial diversity and the genetic nature of microbial species. Nat. Rev. Microbiol. 6, 431–440.
[2] Acosta, J.L., Eguiarte, L.E., Santamaría, R.I., Bustos, P., Vinuesa, P., Martínez-Romero, E., Dávila, G., González, V. (2011) Genomic lineages of *Rhizobium etli* revealed by the extent of nucleotide polymorphisms and low recombination. BMC Evol. Biol. 11 (305), 11–305.
[3] Angiuoli, S.V., Salzberg, S.L. (2010) Mugsy: fast multiple alignment of closely related whole genomes. Bioinformatics 27, 334–342.
[4] Aserse, A.A., Räsänen, L.A., Assefa, F., Hailemariam, A., Lindström, K. (2012) Phylogeny and genetic diversity of native rhizobia nodulating common bean (*Phaseolus vulgaris* L.) in Ethiopia. Syst. Appl. Microbiol. 35, 120–131.
[5] Assefa, S., Keane, T.M., Otto, K.T., Newbold, C., Berriman, M. (2009) ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics 25, 1968–1969.
[6] Beyene, D., Kassa, S., Ampy, F., Asseffa, A., Gebremedhin, T., van Berkum, P. (2004) Ethiopian soils harbor natural populations of rhizobia that form symbioses with common bean (*Phaseolus vulgaris* L.). Arch. Microbiol. 181, 129–136.
[7] Borriss, R., Chen, X.H., Rueckert, C., Blom, J., Becker, A., Baumgarth, B., Fan, B., Pukall, R., Schumann, P., Spröer, C., Junge, H., Vater, J., Pühler, A., Klenk, H.P. (2011) Relationship of *Bacillus amyloliquefaciens* clades associated with strains DSM 7T and FZB42T: a proposal for *Bacillus amyloliquefaciens* subsp. *amyloliquefaciens* subsp. nov. and *Bacillus amyloliquefaciens* subsp. *plantarum* subsp. nov. based on complete genome sequence comparisons. Int. J. Syst. Evol. Microbiol. 61, 1786–1801.
[8] Capella-Gutierrez, S., Silla-Martinez, J.M. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.
[9] Fauvart, M., Sánchez-Rodríguez, A., Beullens, S., Marchal, K., Michiels, J. (2011) Genome sequence of *Rhizobium etli* CNPAF512, a nitrogen-fixing symbiont isolated from bean root nodules in Brazil. J. Bacteriol. 193, 3158–3159.

[10] Ferreira, L.L., Sánchez-Juanes, F.F., García-Fraile, P.P., Rivas, R.R., Mateos, P.F., Martínez-Molina, E., González-Buitrago, J.M., Velázquez, E. (2011) MALDI-TOF mass spectrometry is a fast and reliable platform for identification and ecological studies of species from family *Rhizobiaceae*. PloS One 6, e20223.
[11] González, V., Acosta, J.L., Santamaría, R.I., Bustos, P., Fernández, J.L., Hernández González, I.L., Díaz, R., Flores, M., Palacios, R., Mora, J., Dávila, G.I.L. (2010) Conserved symbiotic plasmid DNA sequences in the multireplicon pangenomic structure of *Rhizobium etli*. Appl. Environ. Microbiol. 76, 1604–1614.
[12] Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., Tiedje, J.M. (2007) DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. 57, 81–91.
[13] Grange, L., Hungria, M. (2004) Genetic diversity of indigenous common bean (*Phaseolus vulgaris*) rhizobia in two Brazilian ecosystems. Soil Biol. Biochem. 36, 1389–1398.
[14] Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
[15] Huson, D.H., Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267.
[16] Kaschuk, G., Hungria, M., Andrade, D.S., Campo, R.J. (2006) Genetic diversity of rhizobia associated with common bean (*Phaseolus vulgaris* L.) grown under no-tillage and conventional systems in Southern Brazil. Appl. Soil Ecol., 32210–32220.
[17] Kislyuk, A.O., Katz, L.S., Agrawal, S., Hagen, M.S., Conley, A.B., Jayaraman, P., Nelakuditi, V., Humphrey, J.C., Sammons, S.A., Govil, D., Mair, R.D., Tatti, K.M., Tondella, M.L., Harcourt, B.H., Mayer, L.W., Jordan, I.K. (2010) A computational genomics pipeline for prokaryotic sequencing projects. Bioinformatics 26, 1819–1826.
[18] Konstantinidis, K.T. (2011) Metagenomic insights into bacterial species. In: De Bruijn, F.J. (Ed.), Handbook of Molecular Microbial Ecology. II: Metagenomics in Different Habitats, John Wiley & Sons, Inc., Hoboken, NJ, USA.
[19] Konstantinidis, K.T., Ramette, A., Tiedje, J.M. (2006) The bacterial species definition in the genomic era. Philos. Trans. R. Soc. Lond. Ser. B: Biol. Sci. 361, 1929–1940.
[20] Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., Salzberg, S.L. (2004) Versatile and open software for comparing large genomes. Genome Biol. 5, R12.
[21] Lloret, L., Ormeno-Orrillo, E., Rincon, R., Martinez-Romero, J., Rogel-Hernandez, M.A., Martínez-Romero, E. (2007) *Ensifer mexicanum* sp. nov. a new species nodulating *Acacia angustissima* (Mill.) Kuntze in Mexico. Syst. Appl. Environ. 30, 280–290.
[22] Martinez-Romero, E. (2003) Diversity of *Rhizobium–Phaseolus vulgaris* symbiosis: overview and perspectives. Plant Soil 252, 11–23.
[23] Michiels, J., D'Hooghe, I., Verreth, C., Pelemans, H., Vanderleyden, J. (1994) Characterization of the *Rhizobium leguminosarum* biovar *phaseoli nifA* gene, a positive regulator of *nif* gene expression. Arch. Microbiol. 161, 404–408.
[24] Mnasri, B., Mrabet, M., Laguerre, G., Aouani, M.E., Mhamdi, R. (2007) Salt-tolerant rhizobia isolated from a Tunisian oasis that are highly effective for symbiotic $N_2$-fixation with *Phaseolus vulgaris* constitute a novel biovar (bv. *mediterranense*) of *Sinorhizobium meliloti*. Arch. Microbiol. 187, 79–85.
[25] Ormeño-Orrillo, E., Rogel-Hernández, M.A., Lloret, L., López-López, A., Martínez, J., Barois, I., Martínez-Romero, E. (2012) Change in land use alters the diversity and composition of *Bradyrhizobium* communities and led to the introduction of *Rhizobium etli* into the tropical rain forest of Los Tuxtlas (Mexico). Microb. Ecol. 63, 822–834.
[26] Peralta, H., Guerrero, G., Aguilar, A., Mora, J. (2011) Sequence variability of Rhizobiales orthologs and relationship with physico-chemical characteristics of proteins. Biol. Direct 6, 48.
[27] Pinero, D., Martinez, E., Selander, R.K. (1988) Genetic diversity and relationships among isolates of *Rhizobium leguminosarum* biovar *phaseoli*. Appl. Environ. Microbiol. 54, 2825–2832.
[28] Posada, D. (2008) jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
[29] Quinto, C., de la Vega, H., Flores, M., Fernández, L., Ballado, T., Soberón, G., Palacios, R. (1982) Reiteration of nitrogen fixation gene sequences in *Rhizobium phaseoli*. Nature 299, 724–726.
[30] Ramírez-Bahena, M.H., García-Fraile, P., Peix, A., Valverde, A., Rivas, R., Igual, J.M., Mateos, P.F., Martínez-Molina, E., Velázquez, E. (2008) Revision of the taxonomic status of the species *Rhizobium leguminosarum* (Frank 1879) Frank 1889AL *Rhizobium phaseoli* Dangeard 1926AL and *Rhizobium trifolii* Dangeard 1926AL. *R. trifolii* is a later synonym of *R. leguminosarum*. Reclassification of the strain *R. leguminosarum* DSM 30132 (=NCIMB 11478) as *Rhizobium pisi* sp. nov. Int. J. Syst. Evol. Microbiol. 58, 2484–2490.
[31] Ribeiro, R.A., Rogel, M.A., López-López, A., Ormeño-Orrillo, E., Gomes Barcellos, F., Martínez, J., Lopes Thompson, F., Martínez-Romero, E., Hungria, M. (2012) Reclassification of *Rhizobium tropici* type A strains as *Rhizobium leucaenae* sp. nov. Int. J. Syst. Evol. Microbiol. 62, 1179–1184.
[32] Richter, M., Rosselló-Móra, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. U.S.A. 106, 19126–19131.
[33] Rivas, R., Martens, M., de Lajudie, P., Willems, A. (2009) Multilocus sequence analysis of the genus *Bradyrhizobium*. Syst. Appl. Microbiol. 32, 101–110.
[34] Robledo, M., Velázquez, E., Ramírez-Bahena, M.H., García-Fraile, P., Pérez-Alonso, A., Rivas, R., Martínez-Molina, E., Mateos, P.F. (2011) The *celC* gene, a new phylogenetic marker useful for taxonomic studies in *Rhizobium*. Syst. Appl. Microbiol. 34, 393–399.
[35] Rosenblueth, M., Martinez Romero, E. (2004) *Rhizobium etli* maize populations and their competitiveness for root colonization. Arch. Microbiol. 181, 337–344.
[36] Schmidt, B., Sinha, R., Beresford-Smith, B., Puglisi, S.J. (2009) A fast hybrid short read fragment assembly algorithm. Bioinformatics 25, 2279–2280.
[37] Silva, C., Vinuesa, P., Eguiarte, L.E., Souza, V., Martinez-Romero, E. (2005) Evolutionary genetics and biogeographic structure of *Rhizobium gallicum sensu lato*, a widely distributed bacterial symbiont of diverse legumes. Mol. Ecol. 14, 4033–4050.
[38] Stamatakis, A., Ludwing, T., Meier, H. (2005) RaxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 21, 456–463.
[39] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
[40] Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., DeBoy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit, Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sullivan, S.A., Daugherty, S.C., Haft, D.H., Selengut, J., Gwinn, M.L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K.J.B., Smith, S., Utterback, T.R., White, O., Rubens, C.E., Grandi, G., Madoff, L.C., Kasper, D.L., Telford, J.L., Wessels, M.R., Rappuoli, R., Fraser, C.M. (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc. Natl. Acad. Sci. U.S.A. 102, 13950–13955.
[41] Toledo, I., Lloret, L., Martinez-Romero, E. (2003) *Sinorhizobium americanum* sp. nov., a new *Sinorhizobium* species nodulating native *Acacia* spp. in Mexico. Syst. Appl. Microbiol. 26, 54–64.
[42] Uribe, L. (1993) Evaluación de medios para la selección de cepas de *Rhizobium leguminosarum* bv. *phaseoli* tolerantes a baja concentración de fosfato en medio de cultivo. Agron. Costarric. 17, 103–109.
[43] Valverde, A., Igual, J.M., Peix, A., Cervantes, E., Velazquez, E. (2006) *Rhizobium lusitanum* sp. nov. a bacterium that nodulates *Phaseolus vulgaris*. Int. J. Syst. Evol. Microbiol. 56, 2631–2637.
[44] Wang, E.T., Rogel, M.A., Garcia-De los Santos, A., Martinez-Romero, J., Cevallos, M.A., Martínez-Romero, E. (1999) *Rhizobium etli* bv. *mimosae*, a novel biovar isolated from *Mimosa affinis*. Int. J. Syst. Bacteriol. 49, 1479–1491.
[45] Warren, R.L., Sutton, G.G., Jones, S.J.M., Holt, R.A. (2007) Assembling millions of short DNA sequences using SSAKE. Bioinformatics 23, 500–501.
[46] Wink, J.M., Kroppenstedt, R.M., Ganguli, B.N., Nadkarni, S.R. (2003) Three new antibiotic producing species of the genus *Amycolatopsis*, *Amycolatopsis balhimydna* sp. nov., *A. tolypomycina* sp. nov., *A. vancoresmycina* sp. nov., and description of *Amycolatopsis keratiniphila* subsp. *keratiniphila* subsp. nov. and *A. keratiniphila* subsp. *nogabecina* subsp. nov. Syst. Appl. Microbiol. 26, 38–46.

Plasmid 68 (2012) 149–158

Contents lists available at SciVerse ScienceDirect

Plasmid

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

Review

# Rhizobial extrachromosomal replicon variability, stability and expression in natural niches

Martha G. López-Guerrero [a], Ernesto Ormeño-Orrillo [a], José Luis Acosta [a,1], Alfredo Mendoza-Vargas [b,1], Marco A. Rogel [a], Miguel Angel Ramírez [a], Mónica Rosenblueth [a], Julio Martínez-Romero [a], Esperanza Martínez-Romero [a,*]

[a] Centro de Ciencias Genómicas, UNAM, Av. Universidad SN, Cuernavaca, Morelos, Mexico
[b] Instituto de Biotecnología, UNAM, Av. Universidad 2001, Cuernavaca, Morelos, Mexico

ARTICLE INFO

*Article history:*
Received 30 March 2012
Accepted 6 July 2012
Available online 16 July 2012
Communicated by Eva Top

*Keywords:*
Plasmids
Plasmid instability
Symbiotic plasmids
*Rhizobium*
*Sinorhizobium*
*Ensifer*

ABSTRACT

In bacteria, niche adaptation may be determined by mobile extrachromosomal elements. A remarkable characteristic of *Rhizobium* and *Ensifer* (*Sinorhizobium*) but also of *Agrobacterium* species is that almost half of the genome is contained in several large extrachromosomal replicons (ERs). They encode a plethora of functions, some of them required for bacterial survival, niche adaptation, plasmid transfer or stability. In spite of this, plasmid loss is common in rhizobia upon subculturing. Rhizobial gene-expression studies in plant rhizospheres with novel results from transcriptomic analysis of *Rhizobium phaseoli* in maize and *Phaseolus vulgaris* roots highlight the role of ERs in natural niches and allowed the identification of common extrachromosomal genes expressed in association with plant rootlets and the replicons involved.



**Contents**



* Corresponding author.
*E-mail address:* emartine@ccg.unam.mx (E. Martínez-Romero).
[1] Present address: Instituto Nacional de Medicina Genómica. Periférico Sur No. 4809, México, D.F. C.P. 14610, Mexico.

0147-619X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.plasmid.2012.07.002

## 1. Introduction

Rhizobia is a generic name to refer to several genera of α and β-Proteobacteria. Rhizobia are successful legume and non-legume rhizosphere colonizers and form nitrogen fixing nodules in legumes. Rhizobia inhabit the soil and other niches such as seeds (López-López et al., 2010; Pérez-Ramírez et al., 1998) or inside plant tissues as endophytes (Chaintreuil et al., 2000; Gutiérrez-Zamora and Martínez-Romero, 2001; Reiter et al., 2003; Yanni et al., 1997). Legumes that establish symbiosis with rhizobia can colonize nitrogen poor environments, may enrich the soil or require less chemical nitrogen fertilizers as agricultural crops.

In rhizobial research, an outstanding discovery was that symbiosis abilities resided in plasmids that could be lost or transferred among bacteria (Bánfalvi et al., 1981; Hooykaas et al., 1982; Johnston et al., 1978; Nuti et al., 1977; Nuti et al., 1979; Rosenberg et al., 1982; Sutton, 1974; Tshitinge et al., 1975; Zurkowski, 1982; Zurkowski and Lorkiewic, 1976). Symbiotic plasmids are found in *Rhizobium*, *Ensifer* = *Sinorhizobium*, in few *Mesorhizobium* species, in the β-Protobacterium *Cupriavidus taiwanensis* that forms nodules in *Mimosa* species (Amadou et al., 2008) and in *Burkholderia* sp. CCGE 1002 isolated from a nodule of *Mimosa occidentalis* collected in Tepic, Mexico (genome NCBI ID 640511). However symbiotic plasmids are not found in *Bradyrhizobium* (Cytryn et al., 2008; Hahn and Hennecke, 1987; Haugland and Verma, 1981), in *Azorhizobium caulinodans* (Lee et al., 2008) or in most *Mesorhizobium* strains (Wang et al., 1999; Xu and Murooka, 1995; Zou et al., 1997). Nitrogen fixation occurring in nodules may be considered as an ecological service. Genes involved in this process (*nif* genes) are plasmid encoded in *Rhizobium, Ensifer* (*Sinorhizobium*), few *Mesorhizobium* species, *Burkholderia* and *Cupriavidus* strains but located in chromosomes in many bacteria (reviewed in Ormeño-Orrillo et al., in press). In rhizobia, symbiosis variants (symbiovars) are recognized on the basis of host specificity and effectiveness (nitrogen fixation) mainly determined by symbiotic plasmids or islands (Rogel et al., 2011). Reviews on symbiotic plasmids (Romero and Brom, 2004) and on the bacterial and plant functions required during the symbiotic process have been published (Oldroyd et al., 2011; Peix et al., 2010).

Methods to visualize rhizobial plasmids (Eckhardt, 1978; Hirsch et al., 1980; Hynes and McGregor, 1990) were pivotal to the study of their diverse patterns, their stability and for the determination of the plasmid location of symbiosis significant genes. In addition to symbiotic plasmids, different large plasmids or extrachromosomal replicons (ER) are found in nodule forming bacteria. However, only 23% of *Bradyrhizobium japonicum* and *B. elkanii* strains from different geographical regions contained plasmids (Cytryn et al., 2008). The role of plasmids in the Rhizobiaceae focusing on interbacterial and transkingdom interactions was recently reviewed (Pappas and Cevallos, 2011). Different types of ER have been described, such as chromids (Harrison et al., 2010) as well as secondary chromosomes (Slater et al., 2009). Housekeeping and ribosomal genes that are relocated to plasmids may make them look like secondary chromosomes. ER that encode housekeeping or essential functions, stably maintained in bacteria and having a GC content similar to that of the chromosome, have been designated chromids and have been identified from genomic data in several rhizobial strains (Harrison et al., 2010). The definition of essential functions encoded in ER must be reviewed because genes may only be conditionally essential on some media or conditions. For example, a plasmid may be cured in the laboratory and thus be considered non essential but may be essential in soil or in the rhizosphere. On the other hand, use of the curing plasmid strategy to recognize essential genes may lead to erroneous conclusions if essential genes move to other replicons during the plasmid elimination (curing) process and selection of survivors. Genome sequence analysis of cured strains would reveal such events.

## 2. Extrachromosomal replicons in rhizobia, a substantial proportion of their genomes

We will focus mainly on *Rhizobium* with only some references on *Ensifer* and the related *Agrobacterium* genus that includes species forming tumors in plants. A remarkable characteristic of *Rhizobium*, *Ensifer* but also of *Agrobacterium* species is the large amount of genomic DNA contained in ER. From 30% to almost 50% of the genome may be extrachromosomal in symbiotic or pathogenic strains (Table 1). Agrobacterial plasmids were reviewed in Suzuki et al. (2009). Although ER may represent a burden for bacterial growth in some cases, this is not the case with rhizobial plasmids. On the contrary, they are important for bacterial physiology as has been shown for *Rhizobium etli* CFN 42 in which strains cured of most of the plasmids had larger duplication times (Brom et al., 1992). Furthermore, ER may contribute significantly to the phenotype and to the bacterial pangenome, the whole species genome.

Most rhizobial ERs are large and in low copy number. Rhizobial strains have several ERs (Table 1 in Romero and Brom, 2004), up to 11 in *R. leguminosarum*. Agrobacteria, *R. galegae*, *R. phaseoli*, *R. tropici* and *R. gallicum* seem to have fewer, 2–4. In rhizobia and in other α-Proteobacteria most ERs have *repABC* replication systems (Cervantes-Rivera et al., 2011; Pappas and Cevallos, 2011). A 7.2 kb plasmid with rolling circle replication was described in an *E. meliloti* strain but small size plasmids are uncommon in rhizobia

**Table 1**
Size and percent of extrachromosomal genome in rhizobia and related strains with completely sequenced genomes.

| Strain | Genome size (Mb) | Percent in extrachromosomal replicons (%) |
|---|---|---|
| *Rhizobium tropici* CIAT 899 | 6.69 | 42.6 |
| *Rhizobium etli* CFN 42 | 6.53 | 32.9 |
| *Rhizobium phaseoli* CIAT 652 | 6.44 | 30.1 |
| *Rhizobium phaseoli* Ch24-10 | 6.63 | 32.0 |
| *Rhizobium leguminosarum* 3841 | 7.79 | 34.5 |
| *Rhizobium leguminosarum* WSM1325 | 7.45 | 35.6 |
| *Rhizobium leguminosarum* WSM2304 | 6.87 | 34.0 |
| *Rhizobium rhizogenes* (*Agrobacterium radiobacter*) K84 | 7.31 | 44.7 |
| *Agrobacterium tumefaciens* C58 | 5.65 | 50.0[a] |
| *Agrobacterium vitis* S4 | 6.31 | 41.0[a] |
| *Ensifer meliloti* (*Sinorhizobium meliloti*)1021 | 6.80 | 44.9 |
| *Ensifer* sp. NGR234 (*Sinorhizobium* sp.) | 6.90 | 43.0 |
| *Ensifer medicae* WSM419 (*Sinorhizobium medicae*) | 6.82 | 44.5 |

[a] Including the secondary chromosome that has ribosomal genes but an origin of replication typical of plasmids.

(Barran et al., 2001). ER sizes in *Rhizobium* and *Ensifer* (*Sinorhizobium*) are in the range of 45 kb to around 2.5 Mb.

## 3. Rhizobial hypervariable genome is in extrachromosomal elements

Chromosomes are more conserved than ER both at the gene sequence and synteny levels (Guerrero et al., 2005). Plasmid patterns are different even within a single rhizobial species (Rosenblueth and Martínez Romero, 2004; Wang et al., 1999). This is particularly evident among *R. etli* and *R. leguminosarum* strains but less variability has been observed in *R. tropici*, *R. phaseoli* or *Ensifer* plasmid profiles (not shown). Plasmid pattern differences suggest that rhizobia may thrive in different environments.

Plasmid gene content variation has been revealed from genomic projects and mosaicism seems to be a common characteristic of plasmids (Cervantes et al., 2011) and symbiotic plasmids (Freiberg et al., 1997; González et al., 2003). Recombination was evidenced with a PCR approach in *Rhizobium etli* plasmids (Flores et al., 2005). Plasmids seem to be prone to pick up novel genes or to suffer deletions. How are plasmids assembled or disassembled? Once a successful plasmid is arranged it may be stably maintained even in distinct chromosomal backgrounds over time (Crossman et al., 2008).

Duplicated copies from chromosomal genes have been allocated to plasmids. In *R. tropici* and in *R. leucaenae* a duplicated citrate synthase gene is found in the symbiotic plasmid, conditioning nodulation (Pardo et al., 1994) and differentially regulated from the chromosomal copy (Hernández-Lucas et al., 1995). Glucosamine synthase (*nodM*) duplicated genes in plasmids (Marie et al., 1992), are needed to provide additional substrates for Nod factor production.

ER may integrate into chromosomes (Guo et al., 2003), rearrange (Brom et al., 1991; Flores et al., 1988, 2000; Soberón-Chávez et al., 1986; Zhang et al., 2001) or form cointegrates with other plasmids (Brom et al., 2004; Cervantes et al., 2011; Guo et al., 2003; Mavingui et al., 2002). Fragments of plasmids may be amplified and in some cases this leads to enhanced nodulation (Mavingui et al., 1997, 1998; Romero et al., 1991; Romero et al., 1995). Extrachromosomal location of genes is not universal and fixed in strains because some genes may be in chromosomes and in other cases in extrachromosomal elements (Crossman et al., 2008 and Fig. 1). There are clues that indicate that some plasmids may be chimeras resulting from the fusion of different plasmids (Cervantes et al., 2011; Ormeño-Orrillo et al., unpublished). Plasmid co-integrates may excise correctly or incorrectly. Plasmids seem to be more dynamic than chromosomes and equivalent genes found among distinct ER in related species are evidence of extensive plasmid rearrangements (Fig. 1, Fig. 6 in Crossman et al., 2008).

## 4. Instability and stability of extrachromosomal elements

Plasmid instability has been known for a long time and it has been recommended to avoid the practice of single colony isolation when purifying rhizobia especially for inoculant production as they may lose relevant plasmids (Weaver and Wright, 1987). Absence of symbiotic plasmids is remarkable as rhizobial natural populations without symbiotic plasmids lose their access to legume nodules, however *R. etli* strains lacking Sym plasmids seem to be very successful rhizospheric or endophytic colonizers (López-López et al., 2010; Segovia et al., 1991). As plasmids encode carbon assimilation genes, rhizobia may change phenotype in one step when losing or gaining plasmids. After *R. etli* CFN42 was resequenced to test Illumina sequencing facilities at UNAM, it was evident that plasmid pReCFN42a was lost in the cultured cells grown to extract DNA (González and Lozano, personal communication) while the original stock maintained the whole set of plasmids. Some rhizobial strains when subcultured in the lab were prone to lose their plasmids (Weaver et al., 1990). A Tn5 had to be inserted in CFN 23 symbiotic plasmid to exert a selective pressure to maintain the plasmid in this *Rhizobium* strain (Soberón-Chávez and Nájera, 1989).

Instability has also been observed in *Burkholderia* strain CCGE 1001 isolated in our laboratory from a nodule of a *Mimosa affinis* plant grown in soils from Acayuca, Veracruz.

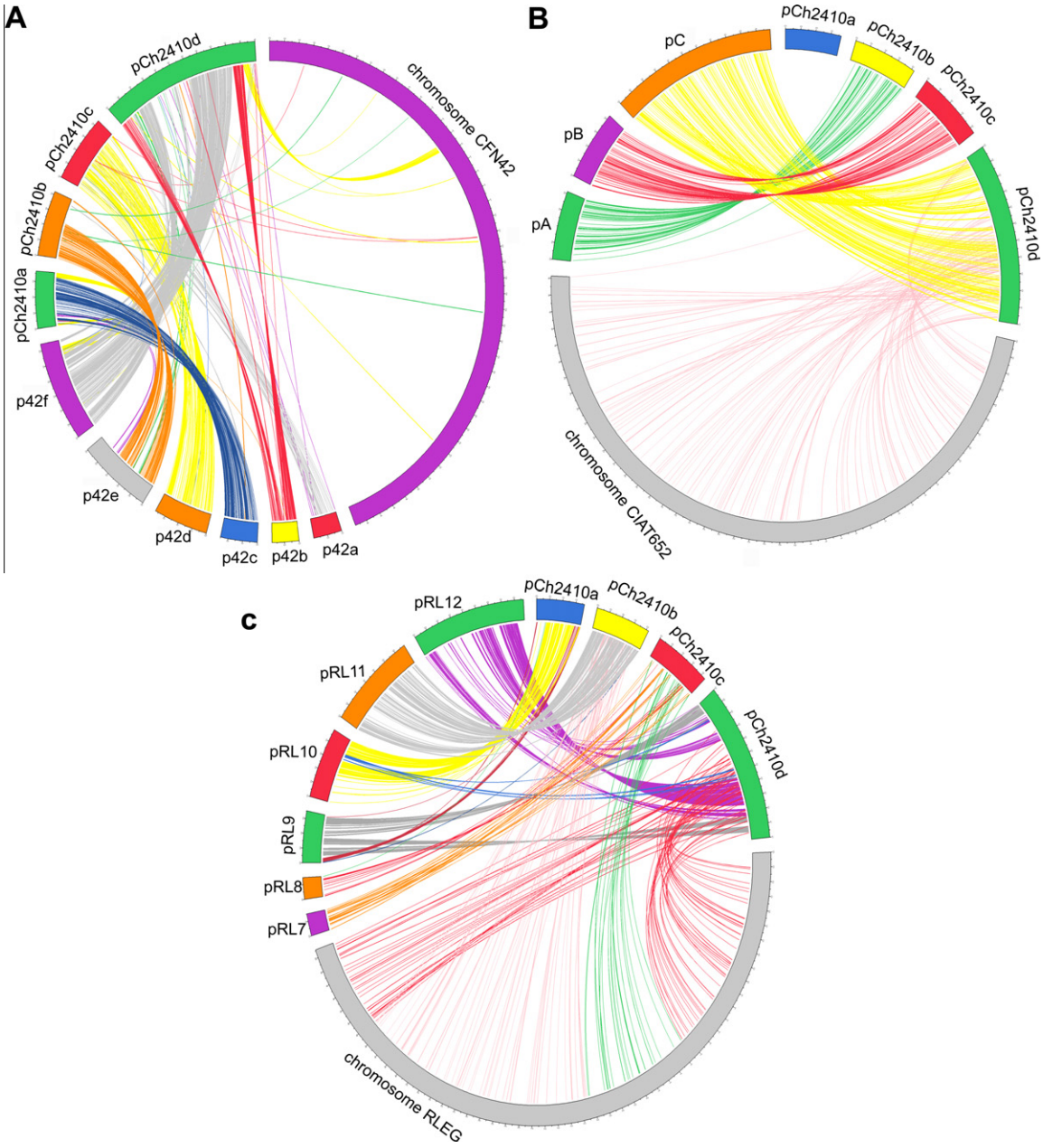



**Fig. 1.** Comparison using satsumasinteny of *R. phaseoli* Ch24-10 extrachromosomal replicons (ERs) to (A) *R. etli* CFN42, (B) *R. phaseoli* CIAT 652 and (C) *R. leguminosarum* 3841 (RLEG) chromosomes and ERs.

Upon subculturing this strain lost its symbiotic plasmid as evidenced from the whole genome analysis (NCBI ID 640510). The original strain is still capable of nodulating *Phaseolus vulgaris* and mimosa plants (unpublished). In another case, when we analyzed the transcripts from *R. phaseoli* strain Ch24-10 (see below) there were none corresponding to a 370 kb plasmid (the smallest, non-symbiotic plasmid) that was revealed in the whole genome analysis of the same strain (López-Guerrero et al., in press). We supposed that the plasmid was lost upon subculturing as the original stock has all plasmids. Our analysis of the published genome of *R. phaseoli* CNPAF512 (Fauvart et al., 2011) revealed sequences corresponding to the 370 kb plasmid from Ch24-10, however these were not found (Fig. 1B) in the published genome of another *R. phaseoli* strain, CIAT 652 (González et al., 2010). This shows that this plasmid is not homogenously conserved among *R. phaseoli* strains. It is worth mentioning that *R. phaseoli* CIAT652 is a very efficient *P. vulgaris* symbiont in spite of lacking this plasmid.

ER maintenance seems to be forced when carrying genes required for growth or survival. This is illustrated in *R. etli*

CFN 42 with pReCFN42e carrying genes needed for growth or optimal growth in rich medium such as those encoding a sensor histidine kinase/ response regulator hybrid protein and a hypothetical protein with a winged helix-turn-helix motif (Landeta et al., 2011) in addition to containing some of the genes for cobalamin biosynthesis. Both genes encoding the sensor histidine kinase/ response regulator hybrid and the hypothetical protein with a winged helix-turn-helix motif are found in *R. leguminosarum* sv. viciae 3841 (in chromid PRL11) and in sv. trifolii strains 1325 and 2304 plasmids as well as in an *R. phaseoli* CIAT 652 plasmid (pRp652a) that corresponds to pReCFN42e.

Toxin–antitoxin genes were discovered as plasmid stabilizers (Jensen and Gerdes, 1995; Ogura and Hiraga, 1983) and have been identified in many bacteria (Pandey and Gerdes, 2005; Van Melderen et al., 2009). Toxin–antitoxin genes have been found in the symbiotic plasmid of *Ensifer* sp. NGR234 (Falla and Chopra, 1999). Antitoxins are more unstable than toxins so when the antitoxin is missing due to plasmid loss, the toxin inhibits cell growth and leads to death (Jensen and Gerdes, 1995). Bacterial genetic mechanisms to ensure plasmid maintenance both in symbionts and pathogens have been reviewed (Sengupta and Austin, 2011).

## 5. Extrachromosomal replicons involved in plant–rhizobium interactions

ERs in addition to the symbiotic plasmids have roles in symbiosis with legumes (Hynes and McGregor, 1990). Curing of a cryptic plasmid in *Ensifer* (*Sinorhizobium*) *meliloti* led to a more efficient symbiosis in alfalfa (Velázquez et al., 1995). In *R. leguminosarum* an exogenous RP4 plasmid decreased symbiotic effectiveness (O'Connell et al., 1998). Enhanced nodulation competitiveness was recorded in *R. etli* strains that gained an *R. leucaenae* (185 kb) plasmid (Martínez-Romero and Rosenblueth, 1990). *A. tumefaciens* transconjugants that in addition to carrying the *nod-nif* plasmid had a 200 kb plasmid from *R. leucaenae* fixed more nitrogen than that with only the symbiotic plasmid (Martínez et al., 1987).

Non symbiotic plasmids participate in rhizobial interactions with plants (Brom et al., 2000; Chen et al., 2000; Hynes and McGregor, 1990; Pappas and Cevallos, 2011). Some *R. leguminosarum* strains capable of associating with rice promoted its growth and alleviated N deficiencies (Yanni et al., 1997), but others from clover inhibited rice root growth. Rice inhibition or promotion is plasmid dependent in *R. leguminosarum* (Perrine et al., 2001) and in *E. meliloti* (Perrine et al., 2005). Derivatives of *R. leguminosarum* sv. trifolii W14–12 lacking two plasmids were unable to grow in soil (Moënne-Loccoz and Weaver, 1995a) and different plasmids were found to contribute to growth in the clover rhizosphere (Moënne-Loccoz and Weaver, 1995b) or in saprophytic life (Moënne-Loccoz et al., 1995). The most competitive maize colonizing *R. phaseoli* strains had the most common plasmid pattern observed among many rhizospheric strains analyzed (Rosenblueth and Martínez Romero, 2004). In *R. leguminosarum* sv. viciae, a plasmid contains several genes needed and expressed by bacterial cells when colonizing the pea rhizosphere (Ramachandran et al., 2011). Similarly we found that extrachromosomal genes were expressed in *R. phaseoli* strain Ch24-10 (Rosenblueth and Martínez Romero, 2004) associated with maize and *P. vulgaris* (common bean) roots (see Section 7).

## 6. Extrachromosomal genes associated with rhizobial environmental adaptation

There is a functional bias in extrachromosomal genes, the ERs tend to contain genes implicated in processes like chemotaxis (Yost et al., 1998) and transport, and they are enriched in elements of external origin (Crossman et al., 2008). Some plasmids, megaplasmids or chromids encode many carbon assimilation genes (Baldani et al., 1992; Oresnik et al., 1998); vitamins like biotin, thiamine or pantothenate (Finan et al., 1986; Miranda-Ríos et al., 1997; Streit et al., 1996; Villaseñor et al., 2011), bacteriocin (Oresnik et al., 1999; Venter et al., 2001), melanin (Hynes et al., 1988) or autoinducer (Schripsema et al., 1996) biosynthetic pathways; and may encode chaperons and modification-restriction systems (Rochepeau et al., 1997). Quorum sensing systems that regulate plasmid transfer or expression of genes in plants may be plasmid encoded in rhizobia (Cubo et al., 1992; Edwards et al., 2009; Lithgow et al., 2000). Reviews on gene functions of plasmids (García-de los Santos and Brom, 1996; Mercado-Blanco and Toro, 1996; Pappas and Cevallos, 2011) and of megaplasmids from *Ensifer* sp. NGR234 (Mavingui, 2009) and *E. meliloti* (Barloy-Hubler and Jebbar, 2009) have been published. Only some functions that we considered important for plant niche colonization will be reviewed here.

### 6.1. Transporters in ERs

In megaplasmid pSymA but especially in pSymB of *Ensifer meliloti* 1021 there are large numbers of transporters (Mauchline et al., 2006) that may allow the bacteria to use different soil nutrients or root exudates. They are inducible by a large number of substrates (Mauchline et al., 2006). Plasmids in *R. etli*, *R. tropici*, *R. leucaenae* and *R. gallicum* sv. phaseoli carry *teu* genes that code for putative sugar ABC transporters involved in the uptake of molecules found in *P. vulgaris* and siratro exudates (Rosenblueth et al., 1998). Four of six quaternary amine transporters that were characterized are located in chromids pRL10 and pRL12 in *R. leguminosarum* 3841 (Fox et al., 2008).

### 6.2. Catabolism

In *E. meliloti putA* genes (for proline catabolism) are involved in rhizobial competitiveness (Van Dillewijn et al., 2002), *putA* is in the chromosome in *E. meliloti* and in *Ensifer* sp. NGR234. *putA* genes are in ER in *R. etli*, *R. phaseoli* and *R. leguminosarum*.

Rhamnose catabolic genes are plasmid borne and inducible (Oresnik et al., 1998). Transport and catabolism of erythriol is plasmid dependent (Geddes et al., 2010; Yost et al., 2006). *R. leguminosarum* mutants in glycerol catabo-

lism have diminished competitiveness. Glycerol uptake and catabolism is plasmid encoded (Ding et al., 2012).

Calystegine catabolism genes are plasmid borne in *E. meliloti* (Guntli et al., 1999; Tepfer et al., 1988). These genes participate in bacterial competitive colonization of non legume rhizozpheres such as those from morning glory plants. Mimosine catabolism genes are also plasmid borne (Borthakur et al., 2003). Opine uptake and catabolism genes reside in the symbiotic megaplasmid a in *E. meliloti* (Murphy et al., 1987). There are also opine catabolizing plasmids in agrobacteria (Bruce et al., 1990).

### 6.3. Surface polysaccharides

Different surface polysaccharides are needed in rhizobial attachment to roots (Downie, 2010) and genes for their biosynthesis are located in different bacterial replicons. Some lipopolysaccharide (LPS) biosynthetic genes have been found in *R. etli* plasmids (García-de los Santos and Brom, 1997). Biosynthetic genes for exopolysaccharides reside in megaplasmid b in *E. meliloti* (Finan et al., 1986) and also in megaplasmids of other rhizobia (Skorupska et al., 2006). Megaplasmid a of *Ensifer* sp. NGR234 encodes flavonoid-inducible genes required for the biosynthesis of a rhamnose-rich LPS produced only inside nodules and that is required for symbiosis (Broughton et al., 2006).

### 6.4. Hormone biosynthesis and protein secretion

Upon inspection of reported genomes we found genes that seem to be involved in gibberellin biosynthesis located in the symbiotic plasmids of *E. fredii* HH103 and *Ensifer sp.* NGR234, *R. etli* CFN42, *R. phaseoli* CIAT 652, *R. tropici* CIAT 899 and in the symbiosis islands of *B. japonicum* USDA 6, *Mesorhizobium loti* R7A, and *M. huakuii* MAFF303099. These genes were originally described in *Bradyrhizobium japonicum* USDA 110 (Morrone et al., 2009) and are not present in the reported genomes of *E. meliloti* and *R. leguminosarum* strains. Gibberellins have diverse effects on plants and its balance in relation to auxins affects plant growth (Brian, 2008). Rhizobial mutants in these genes have not been tested in their hosts. ACC deaminases that modulate ethylene levels are encoded in symbiosis islands in mesorhizobial strains R7A and MAFF303099 (Conforte et al., 2010) and in the symbiotic plasmid of *R. tropici* (Ormeño-Orrillo et al., unpublished). Genes for different auxin biosynthetic pathways are plasmidic in NGR234 (Theunis et al., 2004) and in *R. tropici* CIAT 899 and they are flavonoid inducible (Theunis et al., 2004; Ormeño-Orrillo et al., unpublished).

Rhizobia use different types of secretion systems (excellently reviewed in Downie, 2010). Type III secretion systems (T3SS) are found in several *Rhizobium* and *Ensifer* strains (Marie et al., 2001), these genes are in the symbiotic plasmid in *Ensifer* sp. strain NGR234 and mutants in this system have altered plant specificity. NGR234 T3SS genes are inducible and expressed in the presence of flavonoids (Viprey et al., 1998). A T3SS cluster is also present in the phaseoli symbiotic plasmid (González et al. 2006). Genes coding for Type 1 and 5 secretion systems are found in megaplasmids in *R. tropici* (Ormeño-Orrillo et al., unpublished).

### 6.5. Other functions

In *R. etli*, genes to tolerate polyphenols are plasmid borne (García-de los Santos et al., 2008). The only *R. etli* CFN42 catalase is located in a large ER (pReCFN42f) and is required for bacterial survival in polyphenol rich medium (García-de los Santos et al., 2008). The same replicon carries *nirK* and *norCB* genes for nitrite reduction involved in nitrite detoxification but not in nitrite respiration (Gómez-Hernández et al., 2011). Genes that encode efflux pumps (inducible with bean exudates) that eliminate plant toxic molecules or antibiotics are located in pReCFN42b (184 kb) (González-Pasayo and Martínez-Romero, 2000). The same replicon carries genes for thiamine biosynthesis (Miranda-Ríos et al., 1997).

## 7. Transcriptional profiling of rhizobial ER in natural niches such as the root environment

Are there rhizobial genomic islands or plasmids that are preferentially expressed in the environment? Many stress induced genes that could play a role in the environment are extrachromosomal in *R. etli* CFN42 (Ramírez, unpublished). Expression of symbiosis genes dependent on plant hosts and the molecules and conditions required for gene expression have been well studied and have been extensively reviewed (Cooper, 2004; Le Strange et al., 1990; Maj et al., 2010; Masson-Boivin et al., 2009). Rhizobial genes expressed under stress (Vercruysse et al., 2011), in presence of flavonoids (Perret et al., 1999; Zhang and Cheng, 2006) or in nodules have been reported (Barnett et al., 2004; Chang et al., 2007; Karunakaran et al., 2009; Tsukada et al., 2009) but less is known on genes expressed in soil or in the rhizosphere. Mutations in the *cin* and *rhi* quorum sensing systems affect rhizospheric growth (Cubo et al., 1992; Edwards et al., 2009).

### 7.1. Rhizobium leguminosarum ER rhizospheric expression

A microarray based approach to study *R. leguminosarum* gene expression in pea, alfalfa or sugar beet rhizospheres showed that many of the genes preferentially expressed in *R. leguminosarum* 3841 when inhabiting the pea rhizosphere are encoded in the conjugative 147 kb plasmid pRL8 (Ramachandran et al., 2011). From pRL8, 11 or 21 genes (depending on the threshold considered) were up regulated in pea and only 3 or 2 in alfalfa or sugar beet rhizospheres. Pea induced genes represented around 15% of all genes on pRL8. In total 138 genes were specifically up regulated in 7 day old pea plants and 106 genes were up regulated in all rhizospheres, 70 of those were hypothetical. Among genes expressed were those encoding phenylanine and tyrosine catabolism, dicarboxylate transport, *rhiABC*, *rhiI*, *cinI*, protocatechuate and shikimate uptake, xanthine, formate and other dehydrogenases, as well as some *nod* genes (Ramachandran et al., 2011).

### 7.2. Rhizobium phaseoli ER rhizoplane expression

*R. phaseoli* Ch24-10 was chosen to study gene expression in plant roots because it represents a group of dominant

bacteria in maize rhizosphere (Rosenblueth and Martínez Romero, 2004), is highly competitive to colonize maize and rice and is a very efficient bean symbiont. Bean and maize plants have been grown in association in traditional agriculture for some thousand years and rhizobial gene expression was analyzed in both hosts independently. Upper value tails of bacterial gene transcript distribution in a reported transcriptomic analysis were found to correlate to RNA polymerase occupancy meaning that transcription was occurring in those genes (Vijayan et al., 2011) and, on that basis, highly expressed genes in the Ch24-10 transcriptomic profiling were selected. The 324 extrachromosomal genes highly expressed in maize and/or bean rootlets represented 22% of pRpCh24-10b and 16% of pRpCh24-10d. pSym genes were also expressed in the rhizosphere of maize and bean (representing 13% of the plasmid). Examples of ER genes that were highly expressed in both maize and bean root samples (Supplementary Table S1) are those responsible for proline catabolism, iron uptake, thiamine and gibberellin biosynthesis, a type VI secretion system, oligopeptide or sugar transporters and extrusion pumps as well as polygalacturonase, alpha amylase and Deg protease genes. *teu* genes were not expressed in maize roots in agreement to previous results showing that they are not induced by maize exudates (Rosenblueth et al., 1998). A promoter-less *gusA* gene reporter fused to the polygalacturonase gene was found to be expressed in maize and bean exudates (unpublished) and antibiotic resistance promoter-less genes were found to be expressed in plants when fused to the extrusion pump genes *rmrAB* (González-Pasayo and Martínez-Romero, 2000) or to Deg protease genes (unpublished); this additional evidence is in agreement to the transcriptomic results presented. Furthermore, a radioactive polygalacturonase probe was found to hybridize to the 1 Mb Ch24-10 ER (not shown). Orthologues to previously reported *R. leguminosarum* genes expressed in plant rhizospheres (Ramachandran et al., 2011) were found to be highly expressed in *R. phaseoli* in maize and bean roots (Supplementary Table S1). As in *R. leguminosarum* (Ramachandran et al., 2011), many *R. phaseoli* Ch24-10 highly expressed genes were hypothetical, one of them in common to *R. leguminosarum*.

A comparison of the Ch24-10 transcripts from maize and from bean roots suggested that replicons were differentially expressed depending on the plant host colonized. ER transcripts highly expressed in bean and not in maize roots were found in the Ch24-10 symbiotic plasmid (11 out of 26 bean specific genes) and in a 400 kb ER (pRpCh24-10b with equivalent genes to pReCFN42e), with 9 out of 26 specific genes, while most of the transcripts highly expressed in maize but not in bean (11 out of 14 maize specific genes) were found in pRpCh24-10d, a 1 Mb replicon sharing genes with *R. etli* pReCFN42f. No transcripts could be assigned to a 370 kb plasmid (pRpCh24-10a) as the strain used for the transcriptomic analysis unfortunately lost this plasmid that shares genes with pRL10 and pReCFN42c.

## 8. Concluding remarks

There is still scarce knowledge of rhizobial genes that are functional in nature, in soil, rhizospheric niches or complex microbial communities. Future studies may provide more data to support that ERs, highly dynamic and variable, determine or condition fitness or survival of rhizobia in the environment. Our data extend the knowledge of root-expressed genes in *Rhizobium* and allowed the identification of some extrachromosomal genes commonly expressed in association with plants such as those for thiamine biosynthesis, oligopeptide, proline betaine, α-galactosidase and other ABC transporters, α-*N*-arabinofuranosidase, *rmrA* (González-Pasayo and Martínez-Romero, 2000) and *nod* genes.

## Acknowledgements

To PAPIIT (UNAM) grants IN200709 and IN205412. Martha López-Guerrero was a Ph.D. student at the Programa de Doctorado en Ciencias Biomédicas and had a Consejo Nacional de Ciencia y Tecnología (CONACyT) fellowship. Illumina Sequencing was performed at the Unidad Universitaria de Secuenciación Masiva de DNA (USMDNA) of Universidad Nacional Autónoma de México (UNAM). To Dr. M. Dunn for critically reading the ms.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.plasmid.2012.07.002.

## References

Amadou, C., Pascal, G., Mangenot, S., Glew, M., Bontemps, C., et al., 2008. Genome sequence of the beta-*Rhizobium Cupriavidus taiwanensis* and comparative genomics of rhizobia. Genome Res. 18, 1472–1483.

Baldani, J.I., Weaver, R.W., Hynes, M.F., Eardly, B.D., 1992. Utilization of carbon substrates, electrophoretic enzyme patterns, and symbiotic performance of plasmid-cured clover rhizobia. Appl. Environ. Microbiol. 58, 2308–2314.

Bánfalvi, Z., Sakanyan, V., Koncz, C., Kiss, A., Dusha, I., et al., 1981. Location of nodulation and nitrogen fixation genes on a high molecular weight plasmid of *R. meliloti*. Mol. Gen. Genet. 184, 318–325.

Barloy-Hubler, F., Jebbar, J., 2009. *Sinorhizobium meliloti* megaplasmids and symbiosis in *S. meliloti*. In: Schwartz, E. (Ed.), Microbiology Monographs, vol. 11. Microbial Megaplasmids, Münster, Germany, p. 91.

Barnett, M.J., Toman, C.J., Fisher, R.F., Long, S.R., 2004. A dual-genome symbiosis chip for coordinate study of signal exchange and development in a prokaryote–host interaction. Proc. Natl. Acad. Sci. USA 101, 16636–16641.

Barran, L.R., Ritchot, N., Bromfield, E.S.P., 2001. *Sinorhizobium meliloti* plasmid pRm1132f replicates by a rolling-circle mechanism. J. Bacteriol. 183, 2704–2708.

Borthakur, D., Soedarjo, M., Fox, P.M., Webb, D.T., 2003. The *mid* genes of *Rhizobium* sp. strain TAL1145 are required for degradation of mimosine into 3-hydroxy-4-pyridone and are inducible by mimosine. Microbiology 149, 537–546.

Brian, P.W., 2008. Effects of gibberellins on plant growth and development. Biol. Rev. 83, 37–77.

Brom, S., García-De los Santos, A., Cervantes, L., Palacios, R., Romero, D., 2000. In *Rhizobium etli* symbiotic plasmid transfer, nodulation competitivity and cellular growth require interaction among different replicons. Plasmid 44, 34–43.

Brom, S., García de los Santos, A.G., Girard, M.L., Dávila, G., Palacios, R., et al., 1991. High-frequency rearrangements in *Rhizobium leguminosarum* bv. phaseoli plasmids. J. Bacteriol. 173, 1344–1346.

Brom, S., García de los Santos, A.G., Stepkowsky, T., Flores, M., Dávila, G., et al., 1992. Different plasmids of *Rhizobium leguminosarum* bv. phaseoli are required for optimal symbiotic performance. J. Bacteriol. 174, 5183–5189.

Brom, S., Girard, L., Tun-Garrido, C., García-de los Santos, A., 2004. Transfer of the symbiotic plasmid of *Rhizobium etli* CFN42 requires cointegration with p42a, which may be mediated by site-specific recombination. J. Bacteriol. 186, 7538–7548.

Broughton, W.J., Hanin, M., Relić, B., Kopciñska, J., Golinowski, W., Şimşek, S., Ojanen-Reuhs, T., Reuhs, B., Marie, C., Kobayashi, H., Bordogna, B., Le Quéré, A., Jabbouri, S., Fellay, R., Perret, X., Deakin, W.J., 2006. Flavonoid-inducible modifications to rhamnan O antigens are necessary for *Rhizobium* sp. strain NGR234-legume symbioses. J. Bacteriol. 188, 3654–3663.

Bruce, G.C., Kerr, A., Jones, D.A., 1990. Characteristics of the nopaline catabolic plasmid in *Agrobacterium* strains K84 and K1026 used for biological control of crown gall disease. Plasmid 23, 126–137.

Cervantes, L., Bustos, P., Girard, L., Santamaría, R.I., Dávila, G., et al., 2011. The conjugative plasmid of a bean-nodulating *Sinorhizobium fredii* strain is assembled from sequences of two *Rhizobium* plasmids and the chromosome of a *Sinorhizobium* strain. BMC Microbiol. 11, 149.

Cervantes-Rivera, R., Pedraza-López, F., Pérez-Segura, G., Cevallos, M.A., 2011. The replication origin of a *repABC* plasmid. BMC Microbiol. 11, 158.

Chaintreuil, C., Giraud, E., Prin, Y., Lorquin, J., Ba, A., et al., 2000. Photosynthetic bradyrhizobia are natural endophytes of the African wild rice *Oryza breviligulata*. Appl. Environ. Microbiol. 66, 5437–5447.

Chang, W.S., Franck, W.L., Cytryn, E., Jeong, S., Joshi, T., Emerich, D.W., Sadowsky, M.J., Xu, D., Stacey, G., 2007. An oligonucleotide microarray resource for transcriptional profiling of *Bradyrhizobium japonicum*. Mol. Plant Microbe Interact. 20, 298–307.

Chen, H., Higgins, J., Oresnik, I.J., Hynes, M.F., Natera, S., Djordjevic, M.A., Weinman, J.J., Rolfe, B.G., 2000. Proteome analysis demonstrates complex replicon and luteolin interactions in pSyma-cured derivatives of *Sinorhizobium meliloti* strain 2011. Electrophoresis 21, 3833–3842.

Conforte, V.P., Echeverria, M., Sánchez, C., Ugalde, R.A., Menéndez, A.B., Lepek, V.C., 2010. Engineered ACC deaminase-expressing free-living cells of *Mesorhizobium loti* show increased nodulation efficiency and competitiveness on *Lotus* spp. J. Gen. Appl. Microbiol. 56, 331–338.

Cooper, J., 2004. Multiple responses of rhizobia to flavonoids during legume root infection. Adv. Bot. Res. 41, 1–62.

Crossman, L.C., Castillo-Ramírez, S., McAnnula, C., Lozano, L., Vernikos, G.S., et al., 2008. A common genomic framework for a diverse assembly of plasmids in the symbiotic nitrogen fixing bacteria. PloS One 3, e2567.

Cubo, M.T., Economou, A., Murphy, G., Johnston, A.W., Downie, J.A., 1992. Molecular characterization and regulation of the rhizosphere-expressed genes *rhiABCR* that can influence nodulation by *Rhizobium leguminosarum* biovar viciae. J. Bacteriol. 174, 4026–4035.

Cytryn, E.J., Jitacksorn, S., Giraud, E., Sadowsky, M.J., 2008. Insights learned from pBTAi1, a 229-kb accessory plasmid from *Bradyrhizobium* sp. strain BTAi1 and prevalence of accessory plasmids in other *Bradyrhizobium* sp. strains. ISME J. 2, 158–170.

Ding, H., Yip, C.B., Geddes, B.A., Oresnik, I.J., Hynes, M.F., 2012. Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. Microbiology. http://dx.doi.org/10.1099/mic.0.057281-0.

Downie, J.A., 2010. The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. FEMS Microbiol. Rev. 34, 150–170.

Eckhardt, T., 1978. A rapid method for the identification of plasmid desoxyribonucleic acid in bacteria. Plasmid 1, 584–588.

Edwards, A., Frederix, M., Wisniewski-Dyé, F., Jones, J., Zorreguieta, A., Downie, J.A., 2009. The *cin* and *rai* quorum-sensing regulatory systems in *Rhizobium leguminosarum* are coordinated by ExpR and CinS, a small regulatory protein coexpressed with CinI. J. Bacteriol. 191, 3059–3067.

Falla, T.J., Chopra, I., 1999. Stabilization of *Rhizobium* symbiosis plasmids. Microbiology 145, 515–516.

Fauvart, M., Sánchez-Rodríguez, A., Beullens, S., Marchal, K., Michiels, J., 2011. Genome sequence of *Rhizobium etli* CNPAF512, a nitrogen-fixing symbiont isolated from bean root nodules in Brazil. J. Bacteriol. 193, 3158–3159.

Finan, T.M., Kunkel, B., de Vos, G.F., Signer, E.R., 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. J. Bacteriol. 167, 66–72.

Flores, M., González, V., Pardo, M.A., Leija, A., Martínez, E., et al., 1988. Genomic instability in *Rhizobium phaseoli*. J. Bacteriol. 170, 1191–1196.

Flores, M., Mavingui, P., Perret, X., Broughton, W.J., Romero, D., et al., 2000. Prediction, identification, and artificial selection of DNA rearrangements in *Rhizobium*: toward a natural genomic design. Proc. Natl. Acad. Sci. USA 97, 9138–9143.

Flores, M., Morales, L., Ávila, A., González, V., Bustos, P., et al., 2005. Diversification of DNA sequences in the symbiotic genome of *Rhizobium etli*. J. Bacteriol. 187, 7185–7192.

Fox, M.A., Karunakaran, R., Leonard, M.E., Mouhsine, B., Williams, A., East, A.K., Downie, J.A., Poole, P.S., 2008. Characterization of the quaternary amine transporters of *Rhizobium leguminosarum* bv. viciae 3841. FEMS Microbiol. Lett. 287, 212–220.

Freiberg, C., Fellay, R., Bairoch, A., Broughton, W.J., Rosenthal, A., et al., 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. Nature 387, 394–401.

García-de los Santos, A., Brom, S., 1997. Characterization of two plasmid-borne *lps* $\beta$ loci of Rhizobium etli required for lipopolysaccharide synthesis and for optimal interaction with plants. Mol. Plant-Microbe Interact. 10, 891–902.

García-de los Santos, A., Brom, S., Romero, D., 1996. Rhizobium plasmids in bacteria–legume interactions. World J. Microbiol. Biotechnol. 12, 119–125.

García-de los Santos, A., López, E., Cubillas, C.A., Noel, K.D., Brom, S., et al., 2008. Requirement of a plasmid-encoded catalase for survival of *Rhizobium etli* CFN42 in a polyphenol-rich environment. Appl. Environ. Microbiol. 74, 2398–2403.

Geddes, B.A., Pickering, B.S., Poysti, N.J., Collins, H., Yudistira, H., et al., 2010. A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*. Microbiology 156, 2970–2981.

Gómez-Hernández, N., Reyes-González, A., Sánchez, C., Mora, Y., Delgado, M.J., Girard, L., 2011. Regulation and symbiotic role of *nirK* and *norC* expression in *Rhizobium etli*. Mol. Plant-Microbe Interact. 24, 233–245.

González, V., Acosta, J.L., Santamaría, R.I., Bustos, P., Fernández, J.L., et al., 2010. Conserved symbiotic plasmid DNA sequences in the multireplicon pangenomic structure of *Rhizobium etli*. Appl. Environ. Microbiol. 76, 1604–1614.

González, V., Bustos, P., Ramírez-Romero, M.A., Medrano-Soto, A., Salgado, H., et al., 2003. The mosaic structure of the symbiotic plasmid of *Rhizobium etli* CFN42 and its relation to other symbiotic genome compartments. Genome Biol. 4, R36.

González, V., Santamaría, R.I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S.C., Ramírez, M.A., Jiménez-Jacinto, V., Collado-Vides, J., Dávila, G., 2006. The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. Proc. Natl. Acad. Sci. USA 103, 3834–3839.

González-Pasayo, R., Martínez-Romero, E., 2000. Multiresistance genes of *Rhizobium etli* CFN42. Mol. Plant-Microbe Interact. 13, 572–577.

Guerrero, G., Peralta, H., Aguilar, A., Díaz, R., Villalobos, M.A., et al., 2005. Evolutionary, structural and functional relationships revealed by comparative analysis of syntenic genes in Rhizobiales. BMC Evol. Biol. 5, 55.

Guntli, D., Heeb, M., Moenne-Loccoz, Y., Defago, G., 1999. Contribution of calystegine catabolic plasmid to competitive colonization of the rhizosphere of calystegine-producing plants by *Sinorhizobium meliloti* Rm41. Mol. Ecol. 8, 855–863.

Guo, X., Flores, M., Mavingui, P., Fuentes, S.I., Hernández, G., et al., 2003. Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. Genome Res. 13, 1810–1817.

Gutiérrez-Zamora, M.L., Martínez-Romero, E., 2001. Natural endophytic association between *Rhizobium etli* and maize (*Zea mays* L.). J. Biotechnol. 91, 117–126.

Hahn, M., Hennecke, H., 1987. Mapping of a *Bradyrhizobium japonicum* DNA region carrying genes for symbiosis and an asymmetric accumulation of reiterated sequences. Appl. Environ. Microbiol. 53, 2247–2252.

Harrison, P.W., Lower, R.P.J., Kim, N.K.D., Young, J.P.W., 2010. Introducing the bacterial chromid: not a chromosome, not a plasmid. Trends Microbiol. 18, 141–148.

Haugland, R., Verma, D.P.S., 1981. Interspecific plasmid and genomic DNA sequence homologies and localisation of *nif*-genes in effective and ineffective strains of *Rhizobium japonicum*. J. Mol. Appl. Genet. 1, 205–217.

Hernández-Lucas, I., Pardo, M.A., Segovia, L., Miranda, J., Martínez-Romero, E., 1995. *Rhizobium tropici* chromosomal citrate synthase gene. Appl. Environ. Microbiol. 61, 3992–3997.

Hirsch, P.R., Van Montagu, M., Johnston, A.W.B., Brewin, N.J., Schell, J., 1980. Physical identification of bacteriocinogenic, nodulation and other plasmids in strains of *Rhizobium leguminosarum*. J. Gen. Microbiol. 120, 403–412.

Hooykaas, P.J., Snijdewint, F.G., Schilperoort, R.A., 1982. Identification of the Sym plasmid of *Rhizobium leguminosarum* strain 1001 and its

transfer to and expression in other rhizobia and *Agrobacterium tumefaciens*. Plasmid 8, 73–82.
Hynes, M.F., Brucksch, K., Priefer, U., 1988. Melanin production encoded by a cryptic plasmid in a *Rhizobium leguminosarum* strain. Arch. Microbiol. 150, 326–332.
Hynes, M.F., McGregor, N.F., 1990. Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. Mol. Microbiol. 4, 567–574.
Jensen, R.B., Gerdes, K., 1995. Programmed cell death in bacteria: proteic plasmid stabilization systems. Mol. Microbiol. 17, 205–210.
Johnston, A.W.B., Beynon, J.L., Buchanan-Wollaston, A.V., Setchell, S.M., Hirsch, P.R., Beringer, J.E., et al., 1978. High frequency transfer of nodulating ability between strains and species of *Rhizobium*. Nature 276, 634–636.
Karunakaran, R., Ramachandran, V.K., Seaman, J.C., East, A.K., Mouhsine, B., et al., 2009. Transcriptomic analysis of *Rhizobium leguminosarum* biovar viciae in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. J. Bacteriol. 191, 4002–4014.
Landeta, C., Dávalos, A., Cevallos, M.A., Geiger, O., Brom, S., et al., 2011. Plasmids with a chromosome-like role in rhizobia. J. Bacteriol. 193, 1317–1323.
Lee, K.B., De Backer, P., Aono, T., Liu, C.T., Suzuki, S., et al., 2008. The genome of the versatile nitrogen fixer *Azorhizobium caulinodans* ORS571. BMC Genomics 9, 271.
Le Strange, K.K., Bender, G.L., Djordjevic, M.A., Rolfe, B.G., Redmond, J.W., 1990. The *Rhizobium* strain NGR234 *nodDI* gene product responds to activation by the simple phenolic compounds vanillin and isovanillin present in wheat seedling extracts. Mol. Plant-Microbe Interact. 3, 214–220.
Lithgow, J.K., Wilkinson, A., Hardman, A., Rodelas, B., Wisniewski-Dyé, F., Williams, P., Downie, J.A., 2000. The regulatory locus *cinRI* in *Rhizobium leguminosarum* controls a network of quorum-sensing loci. Mol. Microbiol. 37, 81–97.
López-López, A., Rogel, M.A., Ormeño-Orrillo, E., Martínez-Romero, J., Martínez-Romero, E., 2010. *Phaseolus vulgaris* seed-borne endophytic community with novel bacterial species such as *Rhizobium endophyticum sp. nov*. Syst. Appl. Microbiol. 33, 322–327.
López-Guerrero, M.G., Ormeño-Orrillo, E., Velázquez, E., Rogel, M.A., Acosta, J.L., Gónzalez, V., Martínez, J., Martínez-Romero, E., 2012. *Rhizobium etli* taxonomy revised with novel genomic data and analyses. Syst. Appl. Microbiol. In Press. Available online 1 August 2012.
Maj, D., Wielbo, J., Marek-Kozaczuk, M., Skorupska, A., 2010. Response to flavonoids as a factor influencing competitiveness and symbiotic activity of *Rhizobium leguminosarum*. Microbiol. Res. 165, 50–60.
Marie, C., Barny, M.-A., Downie, J.A., 1992. *Rhizobium leguminosarum* has two glucosamine synthases, GlmS and NodM, required for nodulation and development of nitrogen-fixing nodules. Mol. Microbiol. 6, 843–851.
Marie, C., Broughton, W.J., Deakin, W.J., 2001. *Rhizobium* type III secretion systems: legume charmers or alarmers? Curr. Opinion Plant Biol. 4, 336–342.
Martínez, E., Palacios, R., Sánchez, F., 1987. Nitrogen-fixing nodules induced by *Agrobacterium tumefaciens* harboring *Rhizobium phaseoli* plasmids. J. Bacteriol. 169, 2828–2834.
Martínez-Romero, E., Rosenblueth, M., 1990. Increased bean (*Phaseolus vulgaris* L.) nodulation competitiveness of genetically modified *Rhizobium* strains. Appl. Environ. Microbiol. 56, 2384–2388.
Masson-Boivin, C., Giraud, E., Perret, X., Batut, J., 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? Trends Microbiol. 17, 458–466.
Mauchline, T.H., Fowler, J.E., East, A.K., Sartor, A.L., Zaheer, R., et al., 2006. Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. Proc. Natl. Acad. Sci. USA 103, 17933–17938.
Mavingui, P., 2009. The megaplasmid pNGR234a of *Rhizobium* sp. strain NGR234. In: Schwartz, E. (Ed.), Microbiology Monographs, vol. 11. Microbial Megaplasmids, Münster, Germany, pp. 119–132.
Mavingui, P., Flores, M., Guo, X., Davila, G., Perret, X., et al., 2002. Dynamics of genome architecture in *Rhizobium* sp. strain NGR234. J. Bacteriol. 184, 171–176.
Mavingui, P., Flores, M., Romero, D., Martínez-Romero, E., Palacios, R., 1997. Generation of *Rhizobium* strains with improved symbiotic properties by random DNA amplification (RDA). Nat. Biotechnol. 15, 564–569.
Mavingui, P., Laeremans, T., Flores, M., Romero, D., Martínez-Romero, E., et al., 1998. Genes essential for Nod factor production and nodulation are located on a symbiotic amplicon (AMPRtrCFN299pc60) in *Rhizobium tropici*. J. Bacteriol. 180, 2866–2874.
Mercado-Blanco, J., Toro, N., 1996. Plasmids in rhizobia: the role of nonsymbiotic plasmids. Mol. Plant-Microbe Interact. 9, 535–545.
Miranda-Ríos, J., Morera, C., Taboada, H., Dávalos, A., Encarnación, S., et al., 1997. Expression of thiamin biosynthetic genes (*thiCOGE*) and production of symbiotic terminal oxidase cbb sub(3) in *Rhizobium etli*. J. Bacteriol. 179, 6887–6893.
Moënne-Loccoz, Y., Weaver, R.W., 1995a. Plasmids and saprophytic growth of *Rhizobium leguminosarum* bv. trifolii W14–2 in soil. FEMS Microbiol. Ecol. 18, 139–144.
Moënne-Loccoz, Y., Weaver, R.W., 1995b. Plasmids influence growth of rhizobia in the rhizosphere of clover. Soil Biol. Biochem. 27, 1001–1004.
Moënne-Loccoz, Y., Baldani, J.I., Weaver, R.W., 1995. Sequential heat-curing of Tn*5*-Mob-*sac* labelled plasmids from *Rhizobium* to obtain derivatives with various combinations of plasmids and no plasmid. Lett. Appl. Microbiol. 20, 175–179.
Morrone, D., Chambers, J., Lowry, L., Kim, G., Anterola, A., Bender, K., Peters, R.J., 2009. Gibberellin biosynthesis in bacteria: Separate ent-copalyl diphosphate and ent-kaurene synthases in *Bradyrhizobium japonicum*. FEBS Lett. 583, 475–480.
Murphy, P.J., Heycke, N., Banfalvi, Z., Tate, M.E., de Bruijn, F., Kondorosi, A., Tempé, J., Schell, J., 1987. Genes for the catabolism and synthesis of an opine-like compound in *Rhizobium meliloti* are closely linked and on the Sym plasmid. Proc. Natl. Acad. Sci. USA 84, 493–497.
Nuti, M.P., Ledeboera, M., Lepidi, A.A., Schilperoorrt, A., 1977. Large plasmids in different *Rhizobium species*. J. Gen. Microbiol. 100, 241–248.
Nuti, M.P., Lepidi, A.A., Prakash, R.K., Schilperoort, R.A., Cannon, F.C., 1979. Evidence for nitrogen fixation (*nif*) genes on indigenous *Rhizobium* plasmids. Nature 282, 533–535.
O'Connell, M., Noel, T.C., Yeung, E.C., Hynes, M., Hynes, M.F., 1998. Decreased symbiotic effectiveness of *Rhizobium leguminosarum* strains carrying plasmid RP4. FEMS Microbiol. Lett. 161, 275–283.
Ogura, T., Hiraga, S., 1983. Mini-F plasmid genes that couple host cell division to plasmid proliferation. Proc. Natl. Acad. Sci. USA 80, 4784–4788.
Oldroyd, G.E., Murray, J.D., Poole, P.S., Downie, J.A., 2011. The rules of engagement in the legume-rhizobial symbiosis. Ann. Rev. Genet. 45, 119–144.
Oresnik, I.J., Pacarynuk, L.A., O'Brien, Sh.A.P., Yost, Ch.K., Hynes, M.F., 1998. Plasmid-encoded catabolic genes in *Rhizobium leguminosarum* bv. trifolii: evidence for a plant-inducible rhamnose locus involved in competition for nodulation. Mol. Plant-Microbe Interact. 11, 1175–1185.
Oresnik, I.J., Twelker, S., Hynes, M.F., 1999. Cloning and characterization of a *Rhizobium leguminosarum* gene encoding a bacteriocin with similarities to RTX toxins. Appl. Environ. Microbiol. 65, 2833–2840.
Ormeño-Orrillo, E., Hungria, M., Martinez-Romero, E., in press. Dinitrogen-fixing prokaryotes. In: Rosenberg, E., DeLong, E.F., Stackebrandt, E., Lory, S., Thompson, F. (Eds.), The Prokaryotes vol. 1: Symbiotic Associations, Biotechnology, Applied Microbiology. 4th ed. ISBN 978-3-642-30194-0, Springer.
Pandey, D.P., Gerdes, K., 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res. 33, 966–976.
Pappas, Cevallos, M.A., 2011. Plasmids of the Rhizobiaceae and their role in interbacterial and transkingdom interactions. In: Witzany, G. (Ed.), Soil Biology. Biocommunication in Soil Microorganisms, Austria.
Pardo, M.A., Lagúnez, J., Miranda, J., Martínez, E., 1994. Nodulating ability of *Rhizobium tropici* is conditioned by a plasmid- encoded citrate synthase. Mol. Microbiol. 11, 315–321.
Peix, A., Velázquez, E., Silva, L.R., Mateos, P.F., 2010. Key molecules involved in beneficial infection process in rhizobia–legume symbiosis. In: Khan, M.H., Zaidi, A., Musarrat, J. (Eds.), Microbes for Legume Improvement. Springer, Wien, pp. 55–80.
Pérez-Ramírez, N.O., Rogel, M.A., Wang, E., Castellanos, J.Z., Martínez-Romero, E., 1998. Seeds of *Phaseolus vulgaris* bean carry *Rhizobium etli*. FEMS Microbiol. Ecol. 26, 289–296.
Perret, X., Freiberg, C., Rosenthal, A., Broughton, W.J., Fellay, R., 1999. High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234. Mol. Microbiol. 32, 415–425.
Perrine, F.M., Hocart, C.H., Hynes, M.F., Rolfe, B.G., 2005. Plasmid-associated genes in the model micro-symbiont *Sinorhizobium meliloti* 1021 affect the growth and development of young rice seedlings. Environ. Microbiol. 7, 1826–1838.
Perrine, F.M., Prayitno, J., Weinman, J.J., Dazzo, F.B., Rolfe, B.G., 2001. *Rhizobium* plasmids are involved in the inhibition or stimulation of rice growth and development. Aust. J. Plant Physiol. 28, 923–937.
Ramachandran, V.K., East, A.K., Karunakaran, R., Downie, J.A., Poole, P.S., 2011. Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and

sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biol. 12, R106.
Reiter, B., Buergmann, H., Burg, K., Sessitsch, A., 2003. Endophytic *nifH* gene diversity in African sweet potato. Can. J. Microbiol. 49, 549–555.
Rochepeau, P., Selinger, L.B., Hynes, M.F., 1997. Transposon-like structure of a new plasmid-encoded restriction-modification system in *Rhizobium leguminosarum* VF39SM. Mol. Gen. Genet. 256, 387–396.
Rogel, M.A., Ormeño-Orrillo, E., Martínez Romero, E., 2011. Symbiovars in rhizobia reflect bacterial adaptation to legumes. Syst. Appl. Microbiol. 34, 96–104.
Romero, D., Brom, S., 2004. The symbiotic plasmids of the Rhizobiaceae: Chapter 12 in Plasmid Biology. In: Funnell, B.E., Phillips, G.J. (Eds.), ASM Press, Washington, pp. 271–290.
Romero, D., Brom, S., Martínez-Salazar, J., De, L., Girard, M., Palacios, R., et al., 1991. Amplification and deletion of a *nod-nif* region in the symbiotic plasmid of *Rhizobium phaseoli*. J. Bacteriol. 173, 2435–2441.
Romero, D., Martínez-Salazar, J., Girard, L., Brom, S., Dávila, G., et al., 1995. Discrete amplifiable regions (amplicons) in the symbiotic plasmid of *Rhizobium etli* CFN42. J. Bacteriol. 177, 973–980.
Rosenberg, C., Casse-Delbart, F., Dusha, I., David, M., Boucher, C., et al., 1982. Megaplasmids in the plant-associated bacteria *Rhizobium meliloti* and *Pseudomonas solanacearum*. J. Bacteriol. 150, 402–406.
Rosenblueth, M., Hynes, M.F., Martínez-Romero, E., 1998. *Rhizobium tropici teu* genes involved in specific uptake of *Phaseolus vulgaris* bean-exudate compounds. Mol. Gen. Genet. 258, 587–598.
Rosenblueth, M., Martínez Romero, E., 2004. *Rhizobium etli* maize populations and their competitiveness for root colonization. Arch. Microbiol. 181, 337–344.
Schripsema, J., De Rudder, K.E.E., Van Vliet, T.B., Lankhorst, P.P., De Vroom, E., et al., 1996. Bacteriocin small of *Rhizobium leguminosarum* belongs to the class of *N*-acyl-ι-homoserine lactone molecules, known as autoinducers and as quorum sensing co-transcription factors. J. Bacteriol. 178, 366–371.
Sengupta, M., Austin, S., 2011. Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. Infect. Immun. 79, 2502–2509.
Segovia, L., Piñero, D., Palacios, R., Martínez-Romero, E., 1991. Genetic structure of a soil population of nonsymbiotic *Rhizobium leguminosarum*. Appl. Environ. Microbiol. 57, 426–433.
Skorupska, A., Janczarek, M., Marczak, M., Mazur, A., Krol, J., 2006. Rhizobial exopolysaccharides: genetic control and symbiotic functions. Microbial Cell Factories 5, 7.
Slater, S.C., Goldman, B.S., Goodner, B., Setubal, J.C., Farrand, S.K., et al., 2009. Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. J. Bacteriol. 191, 2501–2511.
Soberón-Chávez, G., Nájera, R., 1989. Isolation from soil of *Rhizobium leguminosarum* lacking symbiotic information. Can. J. Microbiol. 35, 464–468.
Soberón-Chávez, G., Nájera, R., Olivera, H., Segovia, L., 1986. Genetic rearrangements of a *Rhizobium phaseoli* symbiotic plasmid. J. Bacteriol. 167, 487–491.
Streit, W.R., Joseph, C.M., Phillips, D.A., 1996. Biotin and other water-soluble vitamins are key growth factors for alfalfa root colonization by *Rhizobium meliloti* 1021. Mol. Plant-Microbe Interact. 9, 330–338.
Sutton, W.D., 1974. Some features of the DNA of *Rhizobium* bacteroids and bacteria. Biochim. Biophys. Acta 336, 1–10.
Suzuki, T., Tanaka, K., Yamamoto, S., Kiyokawa, K., Moriguchi, K., Yoshida, K., 2009. Ti and Ri plasmids. In: Schwartz, E. (Ed.), Microbiology Monographs, vol. 11. Microbial Megaplasmids, Münster, Germany, p. 133.
Tepfer, D., Goldmann, A., Pamboukdjian, N., Maille, M., Lepingle, A., et al., 1988. A plasmid of *Rhizobium meliloti* 41 encodes catabolism of two compounds from root exudate of *Calystegium sepium*. J. Bacteriol. 170, 1153–1161.
Theunis, M., Kobayashi, H., Broughton, W.J., Prinsen, E., 2004. Flavonoids, NodD1, NodD2, and nod-box NB15 modulate expression of the y4wEFG locus that is required for indole-3-acetic acid synthesis in *Rhizobium* sp. strain NGR234. Mol. Plant-Microbe Interact. 17, 1153–1161.
Tshitinge, G., Luyindula, N., Lurquin, P.F., Ledoux, L., 1975. Plasmid DNA in *Rhizobium vigna* and *Rhizobium trifolii*. Biochim. Biophys. Acta 414, 357–361.
Tsukada, S., Aono, T., Akiba, N., Lee, K.-B., Liu, C.-T., et al., 2009. Comparative genome-wide transcriptional profiling of *Azorhizobium caulinodans* ORS571 grown under free-living and symbiotic conditions. Appl. Environ. Microbiol. 75, 5037–5046.
Van Dillewijn, P., Villadas, P.J., Toro, N., 2002. Effect of a *Sinorhizobium meliloti* strain with a modified *putA* gene on the rhizosphere microbial community of alfalfa. Appl. Environ. Microbiol. 68, 4201–4208.
Van Melderen, L., Saavedra De Bast, M., Rosenberg, S.M., 2009. Bacterial toxin–antitoxin systems: more than selfish entities? PLoS Genet. 5, e1000437.
Velázquez, E., Mateos, P.F., Pedrero, P., Dazzo, F.B., Martínez-Molina, E., 1995. Attenuation of symbiotic effectiveness by *Rhizobium meliloti* SAF22 related to the presence of a cryptic plasmid. Appl. Environ. Microbiol. 61, 2033–2036.
Venter, A.P., Twelker, S., Oresnik, I.J., Hynes, M.F., 2001. Analysis of the genetic region encoding a novel rhizobiocin from *Rhizobium leguminosarum* bv. viciae strain 306. Can. J. Microbiol. 47, 495–502.
Vercruysse, M., Fauvart, M., Jans, A., Beullens, S., Braeken, K., et al., 2011. Stress response regulators identified through genome-wide transcriptome analysis of the (p)ppGpp-dependent response in *Rhizobium etli*. Genome Biol. 12, R17.
Vijayan, V., Jain, I.H., O'Shea, E.K., 2011. A high resolution map of a cyanobacterial transcriptome. Genome Biol. 12, R47.
Villaseñor, T., Brom, S., Dávalos, A., Lozano, L., Romero, D., et al., 2011. Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*. BMC Microbiol. 11, 66.
Viprey, V., Del Greco, A., Golinowski, W., Broughton, W.J., Perret, X., 1998. Symbiotic implications of type III protein secretion machinery in *Rhizobium*. Mol. Microbiol. 28, 1381–1389.
Wang, E.T., Van Berkum, P., Sui, X.H., Beyene, D., Chen, W.X., et al., 1999. Diversity of rhizobia associated with *Amorpha fruticosa* isolated from Chinese soils and description of *Mesorhizobium amorphae sp. nov*. Int. J. Syst. Bacteriol. 49, 51–65.
Weaver, R.W., Wei, G.R., Berryhill, D.L., 1990. Stability of plasmids in *Rhizobium phaseoli* during culture. Soil Biol. Biochem. 22, 465–469.
Weaver, R.W., Wright, S.F., 1987. Variability in effectiveness of rhizobia during culture and in nodules. Appl. Environ. Microbiol. 53, 2972–2974.
Xu, Y., Murooka, Y., 1995. A large plasmid isolated from *Rhizobium huakuii* bv. renge that includes genes for both nodulation of *Astragalus sinicus* cv. Japan and nitrogen fixation. J. Ferment. Bioeng. 80, 276–279.
Yanni, Y.G., Rizk, R.Y., Corich, V., Squartini, A., Ninke, K., et al., 1997. Natural endophytic association between *Rhizobium leguminosarum* bv. trifolii and rice roots and assessment of its potential to promote rice growth. Plant Soil 194, 99–114.
Yost, C.K., Rath, A.M., Noel, T.C., Hynes, M.F., 2006. Characterization of genes involved in erythritol catabolism in *Rhizobium leguminosarum* bv. viciae. Microbiology 152, 2061–2074.
Yost, C.K., Rochepeau, P., Hynes, M.F., 1998. *Rhizobium leguminosarum* contains a group of genes that appear to code for methyl-accepting chemotaxis proteins. Microbiology 144, 1945–1956.
Zhang, X.S., Cheng, H.P., 2006. Identification of *Sinorhizobium meliloti* early symbiotic genes by use of a positive functional screen. Appl. Environ. Microbiol. 72, 2738–2748.
Zhang, X., Kosier, B., Priefer, U.B., 2001. Symbiotic plasmid rearrangement in *Rhizobium leguminosarum* bv. viciae VF39SM. J. Bacteriol. 183, 2141–2144.
Zou, X., Li, F., Chen, H., 1997. Characteristics of plasmids in *Rhizobium huakuii*. Curr. Microbiol. 35, 215–220.
Zurkowski, W., 1982. Molecular mechanism for loss of nodulation properties of *Rhizobium trifolii*. J Bacteriol. 150, 999–1007.
Zurkowski, W., Lorkiewic, Z., 1976. Plasmid deoxyribonucleic acid in *Rhizobium trifolii*. J. Bacteriol. 128, 481–484.

MARTHA G. LÓPEZ-GUERRERO Y ESPERANZA MARTÍNEZ ROMERO

# Bacterias fijadoras de nitrógeno

## Diversidad y uso como biofertilizantes

**La revolución verde que transformó la agricultura y ha proporcionado alimentos a una población humana creciente fue posible gracias a la fertilización; lamentablemente, los fertilizantes aplicados a los cultivos resultan ser contaminantes de lagos y ríos.**

**En el otro extremo, tenemos que los microbios ostentan capacidades metabólicas que ningún otro ser vivo posee, pueden sintetizar los mismos fertilizantes que se producen industrialmente, a un bajo costo; particularmente, las bacterias producen los fertilizantes en simbiosis con las plantas, y éstos no se liberan al ambiente. En este capítulo revisaremos cuales son estas bacterias.**

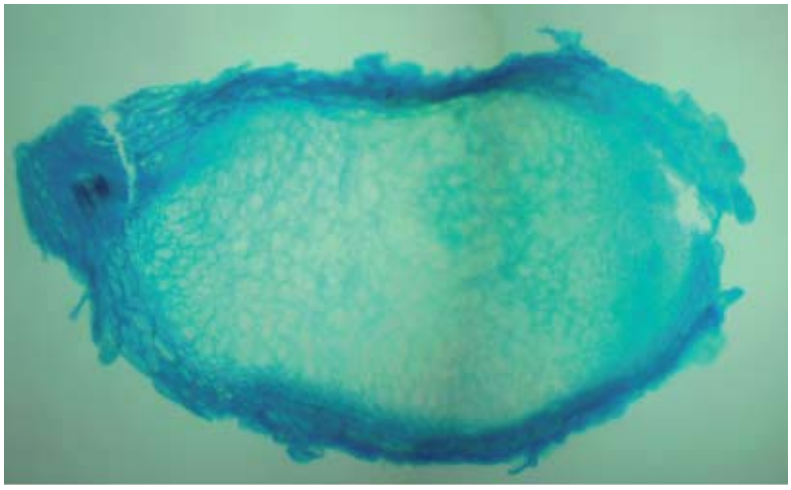
Un biofertilizante es una bacteria u hongo capaz de proporcionar a la planta minerales esenciales que se encuentran en concentración limitada en el suelo, como el nitrógeno (N o también $N_2$), el cual constituye 78% de los gases de la atmósfera. Los biofertilizantes pueden ser usados para disminuir los costos ambientales y económicos que generan el uso de fertilizantes químicos nitrogenados aplicados en agricultura. Es mediante la Fijación Biológica de Nitrógeno (FBN), proceso realizado sólo por microbios procariontes, que el $N_2$ es transformado en un compuesto asimilable; así, la FBN se lleva a cabo en asociación con plantas e insectos, en el océano y hasta en el fondo del mar.

Los rizobios[1] son proteobacterias alfa (como *Rhizobium* y *Bradyrhizobium*) que, en simbiosis, inducen la formación de estructuras especializadas (nódulos) en las que se lleva a cabo la FBN, en las raíces de leguminosas (frijol, soya, chícharo, trébol, haba, cacahuate). A la fecha, se han descrito cerca de 60 especies de rizobios, mientras que existen más de 1,800 especies de leguminosas, por lo que se estima que no se conoce la totalidad de bacterias asociadas a estas plantas. En México, hemos descubierto nuevas especies de rizobios asociadas al frijol y a las leguminosas arbóreas tropicales como mimosas, acacias y guajes.

En Brasil existe una industria exitosa que produce biofertilizantes de cepas eficientes para fijar nitrógeno a partir de la bacteria *Bradyrhizobium*. En algunas condiciones, las bacterias aportan cerca de 250 kilos de nitrógeno por hectárea por año, cantidad equivalente a la que se aplica con fertilizantes químicos. En ese país se cuenta con un control oficial de la calidad de los biofertilizantes comerciales.

En México, contamos con una gran diversidad de bacterias que viven en el suelo y pueden formar nódulos en plantas de frijol por lo que, aparentemente, no es necesario añadir más bacterias a los suelos o a las semillas; sin embargo, se ha tenido éxito utilizando algunas cepas de *Rhizobium* que son capaces de fijar más nitrógeno y que son producto de la investigación realizada en la UNAM.

Los risobios son bacterias que ayudan a configurar nódulos de las raíces de las leguminosas, en las cuales se establecen y, a partir de esta asociación, logran fijar el nitrógeno atmosférico



>> Margarita: esto es un corte transversal de un nódulo radicular formado en la raíz de una leguminosa

## FBN EN PLANTAS NO LEGUMINOSAS

Los cereales –arroz, maíz, trigo–, ocupan la mayor superficie agrícola en el ámbito mundial, y son fertilizados químicamente para generar altos rendimientos, pero lo ideal sería aplicar la FBN mediante biofertilizantes. Existen bacterias que pueden mejorar el crecimiento de los cereales. Las aportaciones de nitrógeno son pequeñas, pero como alternativa se puede aumentar su número en las plantas para incrementar la tasa de nitrógeno fijado.

Mediante enfoques modernos de análisis de todos los genomas de una comunidad (metagenómica) se descubrió que de todos los fijadores de nitrógeno asociados a caña de azúcar o arroz, son los rizobios los que expresan los genes que codifican para la nitrogenasa (enzima necesaria para fijar nitrógeno). Este resultado es sorprendente y puede impulsar el uso de rizobios como biofertilizantes.

En nuestro laboratorio descubrimos que *Rhizobium etli*, simbionte del frijol, se encuentra de manera natural dentro del maíz y es capaz de promover su crecimiento. Actualmente, estamos estudiando la expresión genética de *R. etli* en asociación con el maíz, conocimiento con el que esperamos generar herramientas que permitan, en un futuro, optimizar su uso como biofertilizante.

1. Bacterias que ayudan a configurar nódulos de las raíces de las leguminosas, en las cuales se establecen y, a partir de esta asociación logran fijar el nitrógeno atmosférico.

Un biofertilizante es una bacteria u hongo capaz de proporcionar a la planta minerales esenciales que se encuentran en concentración limitada en el suelo, como el nitrógeno

Algunas bacterias pertenecientes a los géneros *Klebsiella y Burkholderia* son capaces de promover el crecimiento vegetal, pero pueden ser patógenas de humanos, por lo que se ha recomendado que no sean utilizadas en experimentos en campo, y mucho menos en cultivos destinados para consumo humano.

De caña de azúcar, maíz y plátano se aisló *K. variicola*, bacteria fijadora de nitrógeno que promueve el crecimiento vegetal. Esta bacteria, además de asociarse a plantas, se ha encontrado en la sangre de bebés con infecciones originadas en hospitales, así como en la de pacientes inmunosuprimidos.

Antes de utilizar cualquier bacteria se requiere su identificación y caracterización para evaluar cualquier riesgo a la salud humana. Las cianobacterias que fijan nitrógeno y pudieran ser de gran utilidad como biofertilizantes pueden producir toxinas, hepatotoxinas o neurotoxinas muy dañinas para el humano y otros animales.

Estos son dos ejemplos para resaltar la importancia de realizar estudios filogenéticos y de diversidad para elegir adecuadamente las bacterias biofertilizantes, ya que pueden representar un riesgo para la salud humana.

## LOS BIOFERTILIZANTES Y LA FBN

La necesidad de desarrollar biofertilizantes eficientes se puede ver claramente en la producción de biocombustibles a partir de plantas, pues se evitaría el uso de fertilizantes químicos sintetizados a partir de la energía derivada del petróleo. Es ilógico y energéticamente poco favorable invertir en petróleo para generar biocombustibles. El proceso de fijación de nitrógeno es para el ciclo del nitrógeno como la fotosíntesis para el ciclo del carbono.

RECONOCIMIENTO

Agradecimientos: PAPIIT IN 200709, FOMIX 04-09-08. A Ernesto Ormeño, Tabita Ramírez y Mónica Rosenblueth por revisar el escrito. A Marco A. Rogel Hernández, Ivonne Toledo y Julio Martínez por su valioso apoyo técnico.

» Nódulos en las raíces de soya, conteniendo miles de millones de Bradyihizobium

SI DESEA SABER MÁS

» López-López, A. Rosenblueth, M., Martínez, J. and Martínez-Romero, E. 2010. "Rhizobial symbioses in tropical legumes and non-legumes", en "Soil Biology and Agriculture in the Tropics" Patrice Dion (ed.) Srpinger-Verlag Berlin pp. 163-184.

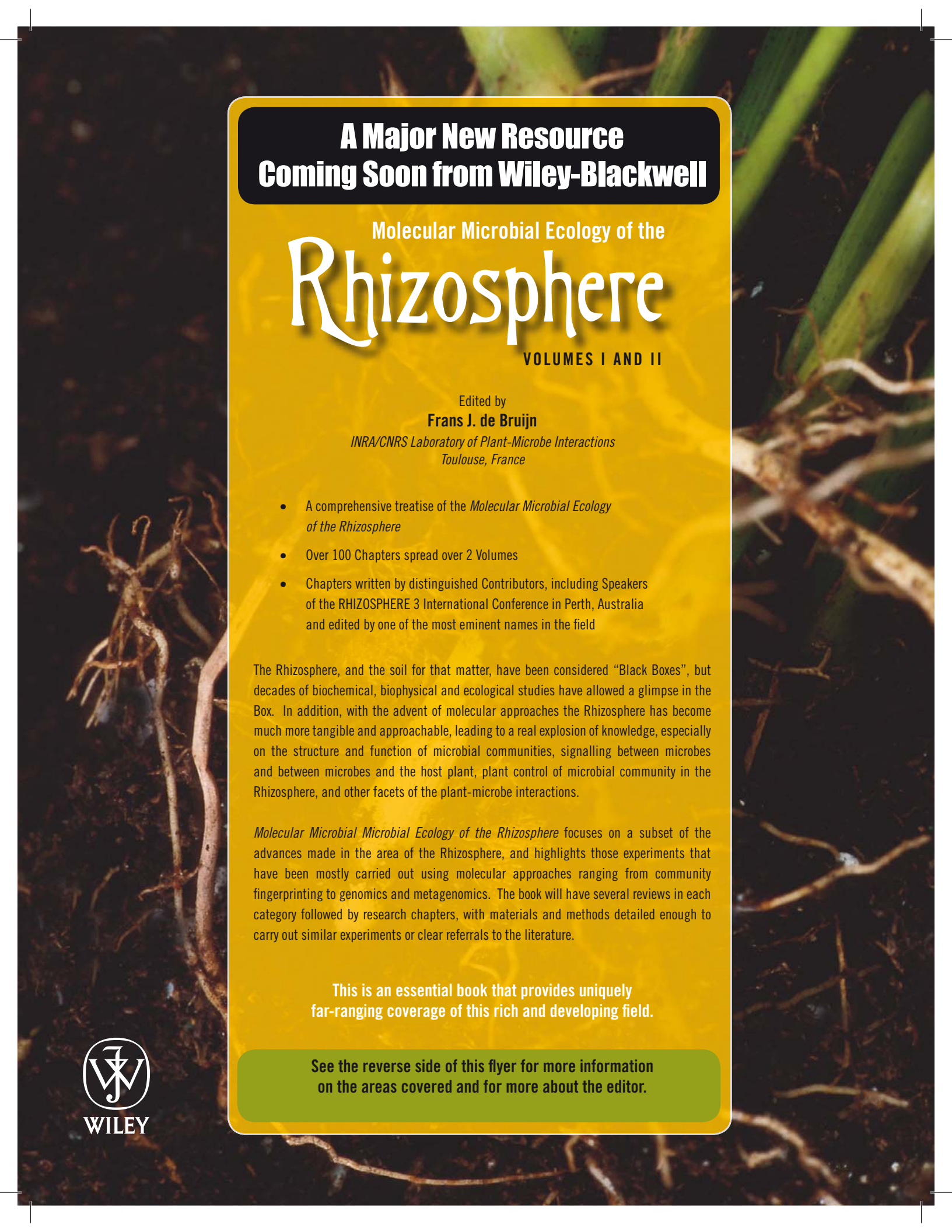
» Martínez Romero, J., Ormeño-Orrillo E., Rogel, M. A., López-López, A. and Martínez Romero, E. 2010. "Trends in rhizobial evolution and some taxonomic remarks" en "Evolutionary Biology – Concepts, Molecular and Morphological Evolution" Springer-Verlag Berlin.

» Fulweiler R. W. 2009. Fantastic Fixers. Science. 326: 377 – 378.

» Martínez Romero, E. 2008. "Detective de bacterias", en "Aportaciones Científicas e Humanísticas Mexicanas en el Siglo XX" Fondo de Cultura Económico, México, D. F. pp. 561-569.

» Lloret L., Rosenblueth, M, Toledo, I. and Martínez Romero, E. 2004. "Ecología bacteriana. Beneficios y riesgos del uso agrícola de rizobios y klebsiellas" Scientific American Latinoamérica 25: 34-35.

» Martínez Romero, E., Palacios, R. and Mora, J. 1998. "Cepas mejoradas de *Rhizobium*", Investigación y Ciencia (Edición en español de Scientific American), 265: 14-19.



# A Major New Resource Coming Soon from Wiley-Blackwell

## Molecular Microbial Ecology of the

# Rhizosphere

**VOLUMES I AND II**


Edited by

**Frans J. de Bruijn**

*INRA/CNRS Laboratory of Plant-Microbe Interactions*
*Toulouse, France*


- A comprehensive treatise of the *Molecular Microbial Ecology of the Rhizosphere*
- Over 100 Chapters spread over 2 Volumes
- Chapters written by distinguished Contributors, including Speakers of the RHIZOSPHERE 3 International Conference in Perth, Australia and edited by one of the most eminent names in the field

The Rhizosphere, and the soil for that matter, have been considered "Black Boxes", but decades of biochemical, biophysical and ecological studies have allowed a glimpse in the Box. In addition, with the advent of molecular approaches the Rhizosphere has become much more tangible and approachable, leading to a real explosion of knowledge, especially on the structure and function of microbial communities, signalling between microbes and between microbes and the host plant, plant control of microbial community in the Rhizosphere, and other facets of the plant-microbe interactions.

*Molecular Microbial Microbial Ecology of the Rhizosphere* focuses on a subset of the advances made in the area of the Rhizosphere, and highlights those experiments that have been mostly carried out using molecular approaches ranging from community fingerprinting to genomics and metagenomics. The book will have several reviews in each category followed by research chapters, with materials and methods detailed enough to carry out similar experiments or clear referrals to the literature.

**This is an essential book that provides uniquely far-ranging coverage of this rich and developing field.**


**See the reverse side of this flyer for more information on the areas covered and for more about the editor.**


WILEY

# Rhizobial genetic repertoire to inhabit legume and non-legume rhizospheres

**Martha G. López-Guerrero, Miguel Ángel Ramírez and Esperanza Martínez-Romero[1]**

**Centro de Ciencias Genómicas, UNAM. Av. Universidad SN, Chamilpa, Cuernavaca, Mor. México**

Number of words: 4, 215

Number of Figures: 1

[1]Corresponding author: E Martinez-Romero

Email: emartine@ccg.unam.mx, Telephone: +52(777)-329-1692, Fax: +52(777)317-5581

Running title: Rhizospheric rhizobial genes and functions

Keywords: Transcriptomics, *Rhizobium*, root colonization, root survival, plant exudates

**Frans J. de Bruijn (ed).**

***Molecular Microbial Ecology of the Rhizosphere, Volume 1*.**

**First Edition. Wiley-Blackwell. 1326.**

**Mayo 2013.**

**ISBN 978-1-1182-9617-2.**

## Abstract

Rhizosphere colonization seems to be an old rhizobial trick, older that nodulation and similar to the colonization processes of other rhizospheric bacteria. Besides legumes, rhizobia colonize the rhizospheres of many other plants. Root exudates have an important role in rhizosphere colonization and are inducers of rhizobial genes involved in bacterial competitive growth and survival preceded by root adhesion. Examples of bacterial functions needed at the rhizosphere are presented here as well as novel results derived from *Rhizobium* transcriptomic analyses. The need to recognize a heterogeneous bacterial physiology on roots is discussed.


## Introduction

Rhizobia are soil bacteria, best known for forming nitrogen fixing nodules on legume roots. For a long time the importance of rhizobial rhizosphere colonization to achieve nodulation has been emphasized [Hossain and Alexander, 1984]. As is the case with other bacteria, rhizobia are successful inhabitants of plant rhizospheres. *Rhizobium etli* strains may attain large numbers on plant roots, up to $10^9$ cells per gram of maize root (fresh weight) [Gutierrez-Zamora and Martínez-Romero, 2001] and *Rhizobium leguminosarum* strains promote rice, canola and lettuce growth [Noel et al., 1996; Yanni et al., 1997].

Root rhizospheres are rich in nutrients and support the growth and proliferation of diverse microbes [Andrews and Harris, 2000]. In contrast to what occurs in nodules, where rhizobia are exclusive or almost exclusive occupants, in the rhizosphere these bacteria share the habitat with many other microorganisms [Barea et al., 2005]. It is unfortunate that most rhizosphere studies are performed using single strain inoculations and not microbial communities, as proposed [Sørensen et al., 2009].

Besides mucilage and sloughed root cortex tissues, root exudates are an important source of nutrients. Root exudates composition has been analyzed from different species [Rovira, 1969]. Exudates contain, amino acids, organic acids [Nardi et al., 1997], sugars [Gransee, 2002] and vitamins [Rajamani et al., 2008]; while there are common plant exudate molecules there are others that are plant specific [reviewed in Walker et al., 2003]. Plant growth and root age affect exudate composition [Walker et al., 2003]. Exudation is different and more active in some parts of the root [Bringhurst et al., 2001]. At the root base the influx of amino acids is greater than in the other regions [Jones and Darrah, 1994]. Accordingly, microbial community may change in relation to

differences in exudate availability [Baudoin, 2002; Bais et al., 2006]. Microbial colonization on roots is not homogeneous [Watt et al., 2006].

Among root exudates flavonoids have attracted great attention for inducing rhizobial *nod* genes. Flavonoids induce other genes such as those encoding type III secretion systems [Viprey et al., 1998] or efflux pumps [Parniske et al., 1991; Gonzalez-Pasayo and Martinez-Romero, 2000] or auxin production (our unpublished results). Additionally there were reports showing that flavonoids stimulated rhizobial growth [Hartwig et al., 1991]. Flavonoids stimulate competitiveness in *R. leguminosarum* in early stages of the interaction with clover and vetch [Maj et al., 2010]. Induction of rhizobial genes *nif* and *nod* genes in the presence of host plants or in nodules has been known for quite a number of years. Besides flavonoids there are other substances in exudates that induce rhizobial *nod* genes such as phenolic compounds [Le Strange et al., 1990], jasmonates [Rosas et al., 1998; Mabood et al., 2006] and xanthones [Yuen et al., 1995].

Rhizobia have large genomes and differential gene expression seems to account for their adaptation to different niches, as occurs in *Pseudomonas* spp. colonizing rhizospheres [Rainey, 1999; Ramos-González et al., 2005]. Some rhizobial functions expressed in plant rhizospheres are reviewed here in comparison to other rhizospheric bacteria.

## Motility and root adhesion

Legume exudates are rhizobial attractants. Flavonoids, sugars, amino acids and dicarboxylic acids induce chemotaxis [Cooper, 2004]. The response towards amino acids and flavonoids has been observed in different rhizobial species [Cooper, 2007]. In *Pseudomonas fluorescens* chemotaxis is required for successful root colonization, since mutants in the *cheA* gene with reduced flagellar driven chemotaxis have diminished tomato root colonization. Malic and citric acids are attractants in the tomato rhizosphere [de Weert et al., 2002]. Using IVET (*In vivo* expression technology) to study genes expressed in the *Pisum sativa* rhizosphere, the *flgG* gene was found to be associated with chemotaxis and motility of *Rhizobium leguminosarum* sv. viciae [Barr et al 2008].

There is compelling evidence that adhesion to roots is an early *sin e qua non* step in the root colonization process for many bacteria. For *Rhizobium* and *Bradyrhizobium*, attachment is the first and the most important step in

legume plant colonization [reviewed in Albareda et al., 2006; Downie, 2010]. In the case of different bacteria, distinct molecules mediate their adhesion to plant roots [Danhorn and Fuqua, 2007].

The *pilAB* gene products are required for the synthesis of a type IV pili that is needed for the establishment of *Azoarcus* sp. strain BH72 on rice rootlets [Dörr et al., 1998] and is essential for root surface colonization [Böhm et al., 2007]. In *Pseudomonas putida* KT244, a secretion system is involved in the release of large proteins implicated in rhizosphere colonization as well as in iron uptake [Molina et al., 2006].

In the rhizosphere or in the presence of exudates, *Rhizobium* produces novel surface polysaccharides and proteins secreted by type I, III and IV secretion systems, some of which may participate in root adhesion [reviewed in Cooper, 2007; Krehenbrink and Downie, 2008].

Different phases in *Rhizobium* root attachment are recognized. First, the initial adsorption is the result of the interaction between plant lectins, bacterial surface polysaccharides and the $Ca^{2+}$ bacterial binding protein rhicadhesin [Dazzo et al., 1984]. The subsequent step involves cellulose fibrils secreted by bacteria which are responsible for irreversible binding to root surfaces [Laus et al., 2005].

## Catabolism of plant substances

Different plants contain different metabolites, most of them unknown. Plant metabolite uptake and catabolism confer on rhizobia and other bacteria an adaptative advantage to differentially colonize the rhizosphere. Plants may be engineered to produce particular metabolites to select rhizospheric species [Mansouri et al., 2002; see also Chapters 116-119].

Induction of catabolism must occur in the presence of the substance or a related substance. Inducible catabolism genes seem to be unevenly distributed among rhizobial species or rhizobial strains. Novel catabolic genes are likely to be found in rhizobia to profit from soil and plant nutrients. The transporters for such substrates may be unknown as well. It is remarkable that hundreds of ABC transporters of unknown substrate have been found in rhizobial genomes [Gonzalez et al., 2006]. Adhesion, transport and catabolism genes seem to constitute a substantial part of the rhizobial genetic repertoire to inhabit legume and non legume rhizospheres.

Rhizobia may catabolize rhamnose [Oresnik et al., 1998] and arabinogalactan which are common in mucilage [Knee et al., 2001], protocatechuate [MacLean et al., 2011] and a diversity of complex carbohydrate molecules and phenolic compunds [Parke et al., 1991]. IVET was used to identify genes that are differentially expressed in

*R. leguminosarum* bv. viciae A34 in pea rhizosphere. Induced genes are involved in membrane transport, such as those encoding for a sulphonate ABC transport system, indicating that sulphate containing compounds are available at the rhizosphere [Barr et al., 2008]. Other genes encoding transporters were also expressed. RL0362 (*araJ*), encodes a permease of the major facilitator system (MFS) family transporters. The latter probably forms an operon with RL0363, encoding a glyoxalase, and may be involved in transporting and processing arabinose polymers [Barr et al., 2008].

Homoserine is found in pea exudates and may be used by *R. leguminosarum* [van Egeraat, 1975]. Calystegines are plant secondary metabolites first found in *Calystegia sepium* (morning glory) that can be poisonous to arthropods and mammals. They are polyhydroxyl nortropame alkaloids that occur in the Solanaceae. In a particular strain of *Sinorhizobium meliloti*, plasmid genes for catabolism of calystegine contribute to competitive colonization of morning glory [Tepfer et al., 1988] and genetically modified alfalfa plants [Guntli et al., 1999a, b]. Other strains of *S. meliloti* do not metabolize calystegine [Tepfer et al., 1988]. Mimosine produced by *Leucaena* plants may be catabolized by the *mid* genes in *Rhizobium* strains that nodulate them [Borthakur et al., 2003].

Genes that are involved in bean exudate uptake (*teu*) are required in *R. etli* and *R. tropici* for competitive *Phaseolus vulgaris* nodulation. These genes encode an ABC transporter that specifically determines the uptake of a bean exudate molecule, the chemical structure of which has not been determined yet [Rosenblueth et al., 1998]. Proline catabolism genes including *putA* have a role in rhizospheric competitive colonization [van Dillewijn et al., 2002].

In non-sterile conditions, *Sinorhizobium meliloti* is able to use galactosides secreted by alfalfa, clover and barrel medic seeds. Using a biosensor constructed using the *melA* promoter fused to the *gfp* protein and induced in presence of galactosides and galactose, it was found that galactosides are present in defined zones at the lateral root initiation and around root hairs but no at the root tip [Bringhurst et al., 2001].

In contrast to the large diversity of carbon and nitrogen sources used in the rhizosphere, in nodules only few amino acids and dicarboxylic acids are the main nutrients provided by the plant to rhizobial bacteroids [Lodwig and Poole, 2003].

In *Pseudomonas stutzeri* the usage of diverse carbon sources seems to be related to rhizospheric colonization capacity [Yan et al., 2008].

## Biofilm formation and quorum sensing

*S. meliloti* cells form biofilms on alfalfa root surfaces [Bringhurst et al., 2001]. Exopolysaccharides (EPS) are needed to colonize *Arabidopsis thaliana* and *Brassica napus* roots by an unclassified *Rhizobium* strain, but are not required for biofilm formation *in vitro* [Santaella et al., 2008]. Similarly, *Rhizobium leguminosarum* biofilm formation on roots does not require the same gene functions as those needed *in vitro* [Williams et al., 2008; see also Chapters 66-70].

Bacteria regulate gene expression in relation to population density by a mechanism called quorum sensing (QS; see also Chapters 71-80). In this process bacteria produce and secrete an autoinductor which actives a transcriptional regulator which, in turn, regulates the transcription of specific genes. QS controls processes involved in the interaction with eucaryotes, like motility, biofilm formation and the production of toxins, EPS and virulence factors [González and Keshavan, 2006].

In *Pseudomonas fluorescens* 2P24 a QS related system PcoI-PcoR has been identified. A mutant in the *pcoI* gene encoding a N-acylhomoserine lactone synthase has been found to have a reduced capacity to form biofilms on non-biological surfaces and also in the wheat rhizosphere in sterilized and non-sterilized soil. The mutant had a reduced capacity to colonize root tips and rhizosphere [Wei and Zhang, 2006]. N-acyl homoserine lactone has been found to be produced by *P. putida* F117 and *S. liquefaciens* MG44 in the tomato rhizosphere [Steidle et al., 2001].

## Rhizobial protection from plant defense

Plants activate defense responses after bacterial colonization. Consequently rhizobia protect themselves by changing their gene expression at the rhizosphere. Phytoalexin resistance is induced by soybean isoflavonoids in *Bradyrhizobium japonicum* [Parniske et al., 1991]. Similarly, genes encoding a multidrug efflux systems were identified using IVET in *R. leguminosarum* bv. viciae A34 in the rhizosphere [Barr et al., 2008]. *gusA* promoterless genes were introduced as transposons into *R. etli* in a random mutagenesis procedure. *gusA* activity of bacteria grown in minimal medium (MM) or in MM with *P. vulgaris* flavonoids was examined. Insertions expressed only in presence of flavonoids were analyzed. In *R. etli* genes involved in the production of an efflux pump system were detected that were also induced by phytoalexins [Gonzalez-Pasayo and Martínez-Romero, 2000].

There are surface polysaccharides that are important for the establishment of the rhizobial-plant interaction such as lipopolysaccharides (LPS), EPS and cyclic β-glucans. They may act as physical barriers to plant defense compounds or as suppressors of plant defense responses [reviewed in Soto et al., 2009 and Downie, 2010]. In other bacteria effector molecules that are exported to plants by type III secretion modulate plant defense responses [Alfano and Collmer, 2004]. In *Sinorhizobium* strain NGR234 type III secretion genes are induced by flavonoids [Viprey et al., 1998].

## Transcriptional regulators expressed at the rhizosphere

Using IVET two transcriptional regulators from the LysR and GntR families, one sigma factor and two genes involved in environmental sensing that use as second messenger cyclic di-GMP with GGDEF and EAL domains were found to be expressed in *R. leguminosarum* in the pea rhizosphere [Barr et al., 2008]. *Pseudomonas putida* bacteria on maize roots express transcriptional regulators belonging to the AraC and TetR families [Matilla et al., 2007].

A proteomic study of the phylospheric *Methylobacterium extorquens* bacterium on *Arabidopsis thaliana* leaves revealed that a regulatory factor PhyR is required for the expression of a number of stress proteins and is a key regulator in *M. extorquens* for its adaptation to its epiphytic lifestyle [Gourion et al., 2006]. Community proteogenomics have revealed insights into the physiology of phyllospheric bacteria [Delmotte et al., 2009]. Similar approaches may be used to study rhizobial communities in the rhizosphere.

The transcriptional regulator RpoN has a role in *Sinorhizobium meliloti* [Barnett et al., 2004] and in *Rhizobium etli* [Salazar et al., 2010] legume interactions. For another example of the identification of regulatory circuits see Chapter 83.

## Genome wide transcriptomic analyses

A microarray study compared gene expression in *R. leguminosarum* strain 3841 in pea, alfalfa and sugar beet (non -legume) rhizospheres. Many genes expressed in the rhizosphere were identified and plant specific expression was shown to occur [Ramachandran et al., 2011]. Interestingly a *R. leguminosarum* plasmid contains many of the genes expressed in the pea rhizosphere.

Genes with unknown function were found to be highly expressed in *R. phaseoli* Ch24-10 recovered from maize (non-legume) rhizosphere by an RNA-Seq procedure using the Illumina sequencing platform [López-Guerrero et al, submitted]. This is similar to what occurs in *R. leguminosarum* in the pea rhizosphere where 66% of the genes expressed are of unknown function [Ramachandran et al., 2011].

In *R. phaseoli* strain Ch24-10, some genes were commonly expressed in maize and bean rhizospheres but others were plant specific. Bacteria growing on roots were not physiologically homogeneous as rhizobial transcripts reflected conflicting bacterial physiological conditions, some genes would correspond to those from bacteria in rich media while others in starvation. All transcripts obtained may not be present in a single cell. Even though young roots from 5 day old plants were used, it seems that they are not homogeneous niches for bacterial growth. Reporter-gene approaches would allow the spatial detection of gene expression on roots.

**Acknowledgements**

Partial financial support was from PAPIIT IN200709 and IN205412 from UNAM. We thank Julio Martínez for technical help.

## References


Albareda MM, Dardanelli MS, Sousa C, Megías M, Temprano F. 2006. Factors affecting the attachment of rhizospheric bacteria to bean and soybean roots. FEMS Microbiol. Lett. 259:67-73.

Alfano JR, Collmer A. 2004. Type III secretion system effector proteins: double agents in bacterial disease and plant defense. Ann. Rev. Phytopathol. 42:385-414.

Andrews JH, Harris RF. 2000. The ecology and biogeography of microorganisms on plant surfaces. Annu. Rev. Phytopathol. 38:145-180.

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM. 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. Ann. Rev. Plant Biol. 57:233-266.

Barea J-M, Pozo MJ, Azcón R, Azcón-Aguilar C. 2005. Microbial co-operation in the rhizosphere. J. Experimental Bot. 56:1761-1778.

Barnett MJ, Toman CJ, Fisher RF, Long SR. 2004. A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. Proc. Natl. Acad. Sci. USA. 101:16636-16641.

Barr M, East AK, Leonard M, Mauchline TH, Poole PS. 2008. In vivo expression technology (IVET) selection of genes of *Rhizobium leguminosarum* biovar viciae A34 expressed in the rhizosphere. FEMS Microbiol. Lett. 282:219-227.

Baudoin E. 2002. Impact of growth stage on the bacterial community structure along maize roots, as determined by metabolic and genetic fingerprinting. Appl. Soil Ecol. 19:135-145.

Böhm M, Hurek T, Reinhold-Hurek B. 2007. Twitching motility is essential for endophytic rice colonization by the N2-fixing endophyte *Azoarcus* sp. strain BH72. Mol. Plant-Microbe Interact. 20:526-533.

Borthakur D, Soedarjo M, Fox PM, Webb DT. 2003. The *mid* genes of *Rhizobium* sp. strain TAL1145 are required for degradation of mimosine into 3-hydroxy-4-pyridone and are inducible by mimosine. Microbiol. 149:537-546.

Bringhurst RM, Cardon ZG, Gage DJ. 2001. Galactosides in the rhizosphere: utilization by *Sinorhizobium meliloti* and development of a biosensor. Proc. Natl. Acad. Sci. USA 98: 4540-4545.

Cooper JE. 2004. Multiple responses of rhizobia to flavonoids during legume root infection. Adv. Bot. Res. 41: 1–62.

Cooper JE. 2007. Early interactions between legumes and rhizobia : disclosing complexity in a molecular dialogue. J. App. Microbiol. 103: 1355-1365.

Danhorn T, Fuqua C. 2007. Biofilm formation by plant-associated bacteria. Ann. Rev. Microbiol. 61:401-422.

Dazzo FB, Truchet GL, Sherwood JE, Hrabak EM, Abe M, Pankratz SH. 1984. Specific phases of root hair attachment in the *Rhizobium trifolii*-clover symbiosis. Appl. Environ. Microbiol. 48: 1140-50.

Delmotte N, Knief C, Chaffron S, Innerebner G, Roschitzki B, Schlapbach R, Von Mering C, Vorholt JuA. 2009. Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc. Natl. Acad. Sci. USA 106, 16428-16433.

de Weert S, Vermeiren H, Mulders IHM, Kuiper I, Hendrickx N, Bloemberg GV, *et al*. 2002. Flagella-driven chemotaxis towards exudate components is an important trait for tomato root colonization by *Pseudomonas fluorescens*. Mol. Plant-Microbe Interact. 15, 1173-1180.

Dörr J, Hurek T, Reinhold-Hurek B. 1998. Type IV pili are involved in plant-microbe and fungus-microbe interactions. Mol. Microbiol. 30, 7-17.

Downie JA. 2010. The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. FEMS Microbiol. Rev. 34: 150-170.

González JE, Keshavan ND. 2006. Messing with bacterial quorum sensing. Microbiol. Mol. Biol. Rev. 70:859-75.

Gonzalez V, Santamaria RI, Bustos P, Hernandez-Gonzalez I, Medrano-Soto A. 2006. The partitioned *Rhizobium etli* genome: Genetic and metabolic redundancy in seven interacting replicons. Proc. Natl. Acad. Sci. USA. 103:3834-3839.

Gonzalez-Pasayo R, Martinez-Romero E. 2000. Multiresistance genes of *Rhizobium etli* CFN42. Mol. Plant-Microbe Interactions 13:572-577.

Gourion B, Rossignol M, Vorholt JA. 2006. A proteomic study of *Methylobacterium extorquens* reveals a response regulator essential for epiphytic growth. Proc. Natl. Acad. Sci. USA 103, 13186-13191.

Gransee A. 2002. Effects of root exudates on nutrient availability in the rhizosphere. Plant Nutrition 92: 626–627.

Guntli D, Burgos S, Moeenne-Loccoz Y, Defago G. 1999a. Calystegine degradation capacities of microbial rhizosphere communities of *Zea mays* (calystegine-negative) and *Calystegia sepium* (calystegine-positive). FEMS Microbiol. Ecol. 28:75-84.

Guntli D, Heeb M, Moenne-Loccoz Y, Defago G. 1999b. Contribution of calystegine catabolic plasmid to competitive colonization of the rhizosphere of calystegine-producing plants by *Sinorhizobium meliloti* Rm41. Mol. Ecol. 8:855-863.

Gutierrez-Zamora ML, Martinez-Romero E. 2001. Natural endophytic association between *Rhizobium etli* and maize (*Zea mays* L.). J. Biotechnol. 91:117-126.

Hartwig UA, Joseph CM, Phillips DA. 1991. Flavonoids released naturally from alfalfa seeds enhance growth rate of *Rhizobium meliloti*. Plant Physiol. 95:797-803.

Hossain AK, Alexander M. 1984. Enhancing soybean rhizosphere colonization by *Rhizobium japonicum*. Appl. Environ. Microbiol. 48: 468-742.

Jones DL, Darrah PR. 1994. Amino acid influx at the soil-root interface of *Zea mays* L and its implications in the rhizosphere. Plant and Soil 163:1-12.

Krehenbrink MM, Downie JA. 2008. Identification of protein secretion systems and novel secreted proteins in *Rhizobium leguminosarum* bv. viciae. BMC Genomics 9: 55.

Knee EM, Gong F-C, Gao M, Teplitski M, Jones AR, Foxworthy A, Mort AJ, Bauer WD. 2001. Root mucilage from pea and its utilization by rhizosphere bacteria as a sole carbon source. Mol. Plant-Microbe Interact. 14:775-784.

Laus MC, van Brussel, AAN, Kijne JW. 2005. Role of cellulose fibrils and exopolysaccharides of *Rhizobium leguminosarum* in attachment to and infection of Vicia sativa root hairs. Mol. Plant-Microbe Interact. 18:533-538.

Le Strange KK, Bender GL, Djordjevic MA, Rolfe BG, Redmond JW. 1990. The *Rhizobium* strain NGR234 *nodD1* gene product responds to activation by the simple phenolic compounds vanillin and isovanillin present in wheat seedling extracts. Mol. Plant-Microbe Interact. 3:214-220.

Lodwig E, Poole P. 2003. Metabolism of *Rhizobium* bacteroids. Crit. Rev. Plant Sci. 22:37-78.

Mabood F, Souleimanov A, Khan W, Smith DL. 2006. Jasmonates induce Nod factor production by *Bradyrhizobium japonicum*. Plant Physiol. Biochem. 44:759-765.

MacLean AM, Haerty W, Golding GB, Finan TM. 2011. The LysR-type PcaQ protein regulates expression of a protocatechuate-inducible A 1 BC-type transport 2 system in *Sinorhizobium meliloti*. Microbiology doi:10.1099/mic.0.050542-0

Maj D, Wielbo J, Marek-Kozaczuk M, Skorupska A. 2010. Response to flavonoids as a factor influencing competitiveness and symbiotic activity of *Rhizobium leguminosarum*. Microbiol. Res. 165:50–60.

Mansouri H, Petit A, Oger P, Dessaux, Y. 2002. Engineered rhizosphere: the trophic bias generated by opine-producing plants is independent of the opine type, the soil origin, and the plant species. Appl. Environ. Microbiol. 68:2562-2566.

Matilla MA, Espinosa-Urgel M, Rodríguez-Herva JJ, Ramos JL, Ramos-González MI. 2007. Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. Genome Biol. 8:R179.

Molina MA, Ramos JL, Espinosa-Urgel M. 2006. A two-partner secretion system is involved in seed and root colonization and iron uptake by *Pseudomonas putida* KT2440. Environ. Microbiol. 8:639-647.

Nardi S, Reniero F, Concheri G. 1997. Soil organis matter mobilization by root exudates of three maize hybrids. Chemosphere 35:2237-2244.

Noel TC, Sheng C, Yost CK, Pharis RP, Hynes MF. 1996. *Rhizobium leguminosarum* as a plant growth-promoting rhizobacterium: Direct growth promotion of canola and lettuce. Can. J. Microbiol. 42:279-283.

Oresnik IJ, Pacarynuk LA, O'Brien SAP, Yost CK, Hynes MF. 1998. Plasmid-encoded catabolic genes in *Rhizobium leguminosarum* bv. trifolii: evidence for a plant-inducible rhamnose locus involved in competition for nodulation. Mol. Plant-Microbe Interactions 11:1175-1185.

Parke D, Rynne F, Glenn A. 1991. Regulation of phenolic catabolism in *Rhizobium leguminosarum* biovar trifolii. J. Bacteriol. 173:5546-5550.

Parniske M, Ahlborn B, Werner D. 1991. Isoflavonoid-inducible resistance to the phytoalexin glyceollin in soybean rhizobia. J. Bacteriol. 173:3432-3439.

Rainey PB. 1999. Adaptation of *Pseudomonas fluorescens* to the plant rizhosphere. Environ. Microbiol. 1:243-257.

Rajamani S, Bauer WD, Robinson JB, Farrow JM, Pesci EC, Teplitski M, *et al.* 2008. The vitamin riboflavin and its derivative lumichrome activate the LasR bacterial quorum-sensing receptor. Mol. Plant-Microbe Interact. 21: 1184-92.

Ramachandran V, East A, Karunakaran R, Downie A, Poole PS. 2011. Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biol. 12:R106.

Ramos-Gonzalez MI, Campos MJ, Ramos JL. 2005. Analysis of *Pseudomonas putida* KT2440 gene expression in the maize rhizosphere: *In vitro* expression technology capture and identification of root-activated promoters. J. Bacteriol. 187:4033-4041.

Rosas S, Soria R, Correa N, Abdala G. 1998. Jasmonic acid stimulates the expression of *nod* genes in *Rhizobium*. Plant Mol. Biol. 38:1161-1168.

Rosenblueth M, Hynes MF, Martinez-Romero E. 1998. *Rhizobium tropici teu* genes involved in specific uptake of *Phaseolus vulgaris* bean-exudate compounds. Mol. Gen. Genet. 258:587-598.

Rovira AD. 1969. Plant root exudates. Bot. Rev. 35:35-57.

Salazar E, Díaz-Mejía JJ, Moreno-Hagelsieb G, Martínez-Batallar G, Mora Y, Mora J, Encarnación S. 2010. Characterization of the NifA-RpoN regulon in *Rhizobium etli* in free life and in symbiosis with *Phaseolus vulgaris*. Appl. Environ. Microbiol. 76: 4510-4520.

Santaella C, Schue M, Berge O, Heulin T, Achouak W. 2008. The exopolysaccharide of *Rhizobium* sp. YAS34 is not necessary for biofilm formation on *Arabidopsis thaliana* and *Brassica napus* roots but contributes to root colonization. Environ. Microbiol. 10:2150-2163.

Sørensen J, Nicolaisen MH, Ron E, Simonet P. 2009. Molecular tools in rhizosphere microbiology - from single cells to whole-community analysis. Plant Soil 321:483-512.

Soto MJ, Domínguez-Ferreras A, Pérez-Mendoza D, Sanjuán J, Olivares J. 2009. Mutualism versus pathogenesis: the give-and-take in plant-bacteria interactions. Cell. Microbiol. 11:381-388.

Steidle A, Sigl K, Schuhegger R, Ihring A, Schmid M, Gantner S, *et al*. 2001. Visualization of N-Acylhomoserine lactone-mediated cell-cell communication between bacteria colonizing the tomato rhizosphere. J. Bacteriol. 67: 5761-5770.

Tepfer D, Goldmann A, Pamboukdjian N, Maille M, Lepingle A, Chevalier D, Denarie J, Rosenberg C. 1988. A plasmid of *Rhizobium meliloti* 41 encodes catabolism of two compounds from root exudate of *Calystegium sepium*. J. Bacteriol. 170:1153-1161.

van Dillewijn P, Villadas PJ, Toro N. 2002. Effect of a *Sinorhizobium meliloti* strain with a modified *putA* gene on the rhizosphere microbial community of alfalfa. Appl. Environ. Microbiol. 68:4201-4208.

van Egeraat A. 1975. The possible role of homoserine in the development of *Rhizobium leguminosarum* in the rhizosphere of pea seedlings. Plant and Soil 386: 381-386.

Viprey V, Del Greco A, Golinowski W, Broughton WJ, Perret X. 1998. Symbiotic implications of type III protein secretion machinery in *Rhizobium*. Mol. Microbiol. 28:1381-1389.

Walker TS, Bais HP, Grotewold E, Vivanco JM. 2003. Root exudation and rhizosphere biology. Plant Physiol. 132:44-51.

Watt M, Silk WK, Passioura JB. 2006. Rates of root and organism growth, soil conditions, and temporal and spatial development of the rhizosphere. Ann. Botany. 97: 839-55.

Wei HL, Zhang LQ. 2006. Quorum-sensing system influences root colonization and biological control ability in *Pseudomonas fluorescens* 2P24. Antonie Van Leeuwenhoek. 89:267-80.

Williams A, Wilkinson A, Krehenbrink M, Russo DM, Zorreguieta A. 2008. Glucomannan-mediated attachment of *Rhizobium leguminosarum* to pea root hairs is required for competitive nodule infection. J. Bacteriol. 190:4706.

Yan Y, Yang J, Dou Y, Chen M, Ping S. 2008. Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501. Proc. Natl. Acad. Sci. USA. 105:7564-7569.

Yanni YG, Rizk RY, Corich V, Squartini A, Ninke K. 1997. Natural endophytic association between *Rhizobium leguminosarum* bv. trifolii and rice roots and assessment of its potential to promote rice growth. Plant Soil 194:99-114.

Yuen JPY, Cassini ST, De Oliveira TT, Nagem TJ, Stacey G. 1995. Xanthone induction of *nod* gene expression in *Bradyrhizobium japonicum*. Symbiosis 19:131-140.