



---

---

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**CENTRO DE CIENCIAS GENÓMICAS**

**PROGRAMA DE INGENIERÍA GENÓMICA**

**ANÁLISIS FUNCIONAL DEL PLÁSMIDO**

**p42f DE *Rhizobium etli* CFN42**

**TESIS**

**QUE PARA OBTENER EL GRADO DE:**

**DOCTOR EN CIENCIAS BIOMÉDICAS**

**PRESENTA:**

**BIÓL. TOMÁS VILLASEÑOR TOLEDO**

**DIRECTOR DE TESIS: DR. ALEJANDRO GARCÍA DE LOS**

**SANTOS**

**CUERNAVACA, MOR.**

**2011**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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



El presente trabajo se realizó bajo la asesoría del Dr. Alejandro García de los Santos en el Programa de Ingeniería Genómica del Centro de Ciencias Genómicas de la Universidad Nacional Autónoma de México.

El comité tutorial estuvo conformado por:

Dr. Alejandro García de los Santos

Dr. David René Romero Camarena

Dra. Gloria Soberón Chavéz

Dra. Susana Brom Klaner

El Jurado de examen de grado estuvo integrado por:

Presidente: Dra. Esperanza Martínez Romero

Secretario: Dr. Alejandro García de los Santos

Vocal: Dra. Bertha María González Pedrajo

Vocal: Dr. Juan Miranda Ríos

Vocal: Dr. Jesús Silva Sánchez.

## CONTENIDO

<b>RESUMEN</b>	<b>4</b>
<b>ABSTRACT</b>	<b>4</b>
<b>INTRODUCCIÓN</b>	<b>6</b>
<b>ANTECEDENTES</b>	<b>28</b>
<b>JUSTIFICACIÓN Y PLANTEAMIENTO DEL PROBLEMA</b>	<b>32</b>
<b>OBJETIVO GENERAL</b>	<b>34</b>
<b>OBJETIVOS ESPECÍFICOS</b>	<b>34</b>
<b>ESTRATEGIA EXPERIMENTAL</b>	<b>35</b>
<b>RESULTADOS</b>	<b>38</b>
<b>Artículo</b>	<b>39</b>
<b>Resultados Adicionales</b>	<b>50</b>
<b>DISCUSIÓN</b>	<b>61</b>
<b>CONCLUSIONES</b>	<b>65</b>
<b>PERSPECTIVAS</b>	<b>66</b>
<b>BIBLIOGRAFÍA</b>	<b>68</b>

## RESUMEN.

Un concepto tradicional de genética bacteriana establece que las funciones metabólicas primarias (Procesos químicos que intervienen en forma directa en la supervivencia, crecimiento y reproducción de las bacterias) están codificadas en el cromosoma, mientras que los plásmidos, codifican funciones que permiten a las bacterias contender con condiciones medio ambientales adversas. Excepciones a este concepto se han generado a través de la eliminación sistemática de plásmidos en cepas de *Rhizobium etli* y *Rhizobium leguminosarum*. En estas bacterias, la pérdida de plásmidos causa una deficiencia de crecimiento en medio rico y medio mínimo (MM), sugiriendo que algunas vías metabólicas primarias, o parte de ellas, están presentes en estos replicones. El presente trabajo muestra que en el plásmido p42f de *R. etli* CFN42 están localizados los genes *panC* y *panB* que codifican para las enzimas pantotenato sintetasa y cetopantoato hidroximetiltransferasa, respectivamente. Estas enzimas son indispensables para la síntesis de pantotenato, precursor esencial en la biosíntesis de coenzima A (CoA), molécula clave en el metabolismo central de la bacteria. Determinamos que la localización de los genes *panCB* en plásmido solo está conservada entre cepas de *R. etli* y *R. leguminosarum*, en el resto de los rhizobios los genes *panCB* están localizados en el cromosoma I. La filogenia de 16 rhizobiales inferida con los genes *panCB* es congruente con la filogenia de las especies inferida con 10 genes del genoma "core" de localización cromosomal, lo anterior es indicativo de que la presencia de los genes *panCB* en plásmido se debe a un evento de transferencia intragenómica y no provienen de transferencia horizontal. También encontramos que una cepa curada del plásmido p42f, complementada con los genes *panCB*, *katG* y *oxyR* no restaura en su totalidad el crecimiento en MM, lo cual implica que, además de la síntesis de pantotenato, existen más funciones codificadas en el plásmido p42f requeridas para el crecimiento en MM. El fenotipo de mutantes con diferentes deleciones del plásmido p42f nos dan indicios de que en diferentes regiones del plásmido p42f hay funciones importantes para el crecimiento en MM.

Los resultados de este estudio confirman la presencia de funciones metabólicas primarias en los plásmidos de algunos rhizobios, apoyan la hipótesis de la cooperación funcional entre replicones y refuerzan la hipótesis de que la presencia de genes esenciales en los plásmidos podría ser parte de la evolución de estos replicones hacia la formación de cromosomas adicionales.

La búsqueda del resto de los genes involucrados en la síntesis de pantotenato, nos llevó a descubrir que la síntesis del precursor esencial del pantotenato, la  $\beta$ -alanina, se realiza mediante un mecanismo diferente en los rhizobios, contrario a lo que se ha establecido principalmente en enterobacterias.

## ABSTRACT.

A traditional concept in bacterial genetics states that housekeeping genes, those involved in basic metabolic functions needed for maintenance of the cell, are encoded in the chromosome, whereas genes required for dealing with challenging environmental conditions are located in plasmids. In contrast with this concept, the absence of several plasmids in *Rhizobium etli* and *R. leguminosarum* strains cause a growth deficiency in rich and minimal medium (MM) suggesting the presence of primary metabolic functions encoded on these plasmids. The complete genomic sequence of *R. etli* CFN42 predicts the presence of two genes: *panC*, *panB* located in the 642-kb plasmid p42f of this bacterium, which encode putative pantothenate biosynthesis enzymes, a crucial vitamin for coenzyme A (CoA) synthesis. The present study demonstrated that plasmid-borne *panCB* genes are essential for pantothenate biosynthesis. *panC* and *panB* mutants were pantothenate auxotroph and the complementation with wild-type *panCB* genes fully restored the ability of these mutants to grow in MM. A screening of the location of *panCB* genes among members of the *Rhizobiales* order indicates that the occurrence of these genes in plasmid is conserved only among different strains of the *R. etli* and *R. leguminosarum* species. The PanCB phylogeny indicated a common origin for chromosomal and plasmid-born PanCB sequences, discarded the role of horizontal gene transfer as the origin of plasmid-borne *panCB* genes, and suggested that these genes may have arrived at plasmids of *R. etli* and *R. leguminosarum* species through an intragenomic transfer event from chromosome to plasmid.

Interestingly, *panCB* genes could not totally restore the ability of plasmid-cured strain CFNX186 to grow in MM, which indicates that other important functions for growth in MM are encoded in this plasmid. Generation of site-specific deletions of p42f revealed regions required for growth of *R. etli* in MM other than *panCB* genes.

Our results confirmed the presence of primary metabolic functions encoded on plasmid p42f, support the hypothesis of cooperation among different replicons for basic cellular functions and reinforce the hypothesis that state the origin of secondary chromosomes from plasmid harboring essential genes.

In the course of the present research, we noticed that *panD* gene, encoding the aspartate 1-decarboxylase enzyme, is absent from the genome of several members of the *Rhizobiales* including *R. etli* CFN42. This enzyme catalyzes the decarboxylation of aspartate to produce  $\beta$ -alanine an essential precursor of pantothenate. Experiments conducted to understand how these bacteria synthesize  $\beta$ -alanine suggests a novel mechanism of synthesis in some rhizobias.

## INTRODUCCIÓN.

### Generalidades de los genomas multi-replicón.

En décadas pasadas, las bacterias fueron consideradas como el grupo de organismos más simples y uniformes. Este concepto incluía la organización de su genoma, el cual se consideraba invariablemente constituido por un solo cromosoma y algunos plásmidos ambos de estructura circular. Este concepto también afirmaba que en el cromosoma estaban codificadas las funciones esenciales para la vida de la bacteria y en los plásmidos las funciones para la adaptación a condiciones medio ambientales adversas. Sin embargo, la continua investigación mediante el uso de nuevas técnicas genéticas y bioquímicas, así como la relativamente nueva perspectiva genómica hacia el entendimiento de la biología microbiana, ha cambiado radicalmente nuestro concepto acerca de la estructura y organización de los genomas procariontes. El desarrollo de la técnica de electroforesis de campos pulsados y su aplicación al análisis de los genomas de diversos grupos de bacterias y arqueas, permitió el descubrimiento de novedosas estructuras genómicas como la presencia de dos cromosomas circulares en *Brucella melitensis* (Michaux *et al.*, 1993), *Rhodobacter sphaeroides* (Suwanto y Kaplan, 1989) y *Leptospira interrogans* (Zuerner *et al.*, 1993); de un cromosomas lineal en *Borrelia burgdorferi* (Ferdows y Barbour, 1989), *Rhodococcus fascians* (Crespi *et al.*, 1992) y *Streptomyces lividans* (Lin *et al.*, 1993); de un cromosoma lineal y uno circular en *Agrobacterium tumefaciens* (Allardet-Servent *et al.*, 1993); de 21 plásmidos: 12 lineales y 9 circulares en *Borrelia burgdorferi* (Barbour y Garon, 1987) así como la presencia de mega-plásmidos, mayores a una megabase, en *Sinorhizobium meliloti* (Sobral *et al.*, 1991). La secuenciación masiva de genomas, así como la invaluable herramienta que es la base de datos del NCBI, han contribuido de manera importante a ampliar nuestra visión de la diversidad genómica bacteriana y nos han enseñado que mientras

los simbioses y parásitos intracelulares obligados tienden a reducir su genoma (Moran, 2002), las bacterias de vida libre tienden a expandirlo y a organizarlo, en muchos casos, en múltiples replicones (Jordan *et al.*, 2001). Los análisis tanto *in silico* como *in vitro* de las funciones codificadas en esos replicones, indican que los genes que codifican para ARNr y ARNt así como para algunas enzimas clave de vías metabólicas primarias, es decir, aquellas que intervienen en forma directa en la supervivencia, crecimiento y reproducción de las bacterias, están distribuidos en diferentes replicones. Este hallazgo ha llevado a redefinir los conceptos de cromosoma y plásmido, a incluir nuevos conceptos como “crómidos” y a proponer hipótesis sobre la evolución de los genomas multi-replicón.

## **Cromosoma I**

A partir del descubrimiento de bacterias que poseen más de un cromosoma, se ha adoptado la nomenclatura de cromosoma I, cromosoma II y cromosoma III, y así sucesivamente, para distinguirlos. El cromosoma I es, invariablemente, el replicón más grande de la célula en todas las bacterias actualmente secuenciadas. Una característica muy conservada en los cromosomas I y que los distingue de otros replicones es el origen de replicación (*oriC*), así como los genes que flanquean esta región. Dichos genes son conocidos como *dnaA*, el cual codifica una proteína involucrada en el inicio de la replicación; *gidAB*, que codifica proteínas que participan en la división; *thdF*, el cual codifica una proteína involucrada en la partición, así como el gene *dnaN*, que codifica para una subunidad de la ADN polimerasa. Adicionalmente, el *oriC* incluye sitios de pegado para la proteína IHF (integration host factor) que dobla el ADN y estimula la actividad de desenrollamiento de la proteína DnaA. También están presentes varios sitios de metilación por la DNA adenina metiltransferasa (Dam) (MacLellan *et al.*, 2004). Los complejos nucleoproteicos (nucleoides)



estudiados en organismos modelo como *Escherichia coli*, *Bacillus subtilis*, *Caulobacter crescentus* y *Vibrio cholerae*, contienen tanto un origen de replicación (*oriC*) como una región de término de la replicación (*terC*) (Ghosh *et al.*, 2006; Toro y Shapiro, 2010).

Casi todos los genomas de las bacterias conocidas actualmente presentan una estructura circular, excepto en algunas especies de *Borrelia* (*B. afzelii*, *B. burgdorferi*, *B. hermsii*, *B. recurrentis*, *B. garinii* y *B. turicatae*) y todos los *Streptomyces* secuenciados (*S. avermitilis*, *S. coelicolor*, *S. griseus* y *S. scabiei*), en donde se ha encontrado un cromosoma I lineal. Dado que la estructura circular es predominante entre las bacterias se sugiere que esta estructura fue ancestral y que se ha mantenido como forma principal en las bacterias, sin la necesidad de mecanismos de replicación requeridos por la bacteria, para conservar los extremos en una molécula lineal (Marri *et al.*, 2008). Sin embargo, el hecho de que los cromosomas lineales hayan surgido al menos en tres ocasiones durante el curso de la evolución bacteriana y estos se encuentren propagados entre los eucariontes, sugiere que esta configuración, una vez establecida, puede también conferir alguna o algunas ventajas tales como facilitar la recombinación o integración del ADN a partir de las regiones teloméricas (Marri *et al.*, 2008; Chaconas y Kobryn, 2010). Para demostrar si esta hipótesis es cierta, Marri y colaboradores estudiaron el efecto de la geometría del cromosoma sobre la diversidad genética y la recombinación. Analizaron la variación de secuencias en 18 cepas de *A. tumefaciens* biovar I que poseen dos cromosomas, uno lineal y uno circular. Los resultados de esa investigación indicaron que los cromosomas lineales no presentan niveles más altos de diversidad genética y recombinación (Marri *et al.*, 2008).

Las moléculas de ADN lineal también pueden proveer estabilidad a las secuencias de ADN que son inestables en moléculas superenrolladas (Chaconas y Kobryn, 2010).

En los cromosomas I de todas las bacterias secuenciadas a la fecha están presentes la mayoría de los genes esenciales para la vida de la bacteria, es decir aquellos genes indispensables bajo cualquier condición ambiental como los involucrados en la replicación, transcripción y traducción del ADN. En este replicón también están codificadas la mayoría de las enzimas de las vías metabólicas centrales.

Designar y definir genes esenciales no siempre es una tarea fácil, pues dicha definición es un tema que se encuentra en constante debate, dado que las condiciones en las cuales se crecen los microorganismos para su estudio son condiciones de laboratorio plenamente controladas. Debido a que los organismos en la naturaleza, a diferencia de las condiciones de laboratorio, se encuentran expuestos a diferentes presiones y estreses tanto ambientales como nutricionales, el concepto de gene esencial se ha ampliado para incluir diferentes condiciones ambientales (Gerdes *et al.*, 2006; D' Elia *et al.*, 2009). El hecho de que los cromosomas I sean el reservorio principal de genes esenciales, no significa que todos los genes que posee lo sean. Para identificar los genes esenciales en un organismo se han seguido, principalmente, dos estrategias. Una de ellas involucra la mutagénesis en cada uno de los genes potenciales del organismo y esto lleva a la definición de un gene esencial como aquel en donde no es posible obtener una mutación nula en el medio definido para el organismo. Con los resultados de estos análisis a lo largo de varios genomas en diferentes organismos ha sido posible la creación de una base de datos de genes esenciales (Zhang *et al.*, 2004; Zhang y Lin, 2009) conocida como DEG, por sus siglas en inglés (Database of Essential Genes). La segunda estrategia que se ha empleado, aunque en menos organismos, consiste en deleciones programadas realizadas sobre diferentes regiones del genoma. Con esta técnica, ha sido posible reducir el cromosoma de por lo menos tres organismos diferentes: *E. coli*, *B. subtilis* y *Corynebacterium glutamicum*, en un porcentaje que va desde

el 5 hasta el 30% del genoma; sin embargo, la redundancia funcional entre los genes eliminados ha limitado la obtención de deleciones mayores (Suzuki *et al.*, 2008).

Por lo general, cada célula bacteriana posee una sola copia del cromosoma I. Sin embargo, en algunos casos se ha reportado poliploidía, pero esto es debido a la fase de crecimiento en las cuales se encuentra el organismo en condiciones de laboratorio. Por ejemplo, *Deinococcus radiodurans* posee de 4 a 10 copias de su cromosoma dependiendo de la fase de crecimiento (White *et al.*, 1999) y *Azotobacter vinelandii* que puede tener más de 40 copias por célula durante la fase exponencial tardía y más de 80 cromosomas por célula durante la fase estacionaria temprana (Maldonado *et al.*, 1994). Aun cuando los datos anteriores parezcan sorprendentes, la bacteria, Gram-positiva *Epulopiscium fishelsoni*, que tiene una relación simbiótica con peces de la familia *Acanthuridae*, presenta una ploidía que puede variar de 1 a 2000, dependiendo del ciclo circadiano. Este ciclo consiste en lo siguiente: en la mañana una célula bacteriana madre con dos descendientes internas y pequeñas, las cuales incrementan su tamaño durante el día hasta llenar el citoplasma de la célula madre y finalmente emergen de la madre (Bresler *et al.*, 1998; Mendell *et al.*, 2008). El significado de ésta poliploidía puede ser considerada con respecto a dos atributos de esta bacteria: su biología nutricional como un simbiote, y su extraordinario tamaño (Mendell *et al.*, 2008). La poliploidía proporciona recursos para soportar un rápido crecimiento y división, especialización celular y adaptación, y puede aumentar la reparación de lesiones en el genoma (Rosenberg y Hastings, 2004). En algunas bacterias, la poliploidía va de la mano tanto con la adaptación metabólica, como con la diferenciación simbiote (Komaki e Ishikawa, 2000). *Buchnera aphidicola*, la proteobacteria simbiote de áfidos contiene, en promedio, 120 copias de su cromosoma, y ésta expansión de recursos genéticos se relaciona con la demanda del áfido por biomoléculas esenciales proporcionadas por *Buchnera*

(Komami e Ishikawa, 2000; Latorre *et al.*, 2005). Por lo tanto, un factor que puede favorecer la poliploidía en *Epulopiscium* es la presión de selección por la evolución de una relación simbiótica que contribuye al metabolismo del hospedero (Mendell *et al.*, 2008). Esta variación en ploidía observada en *Epulopiscium*, supera a las demás observadas en cualquier organismo, aun en los mismos eucariontes (Bendich y Drlica, 2000).

El genoma bacteriano refleja las estrategias de adaptación a través de la tendencia en el uso de nucleótidos encontrados en la composición del cromosoma. La selección en contra del alto contenido de G + C es debida al alto costo energético en la síntesis de GTP y CTP, lo cual resulta en una tendencia hacia AT ya que esto brinda una ventaja selectiva. (Mann y Chen, 2010). Por lo tanto, el contenido de G + C esta íntimamente relacionado con el estilo de vida, así como con los nichos ecológicos de los microorganismos. Existe una observación general, en la cual, los genomas de mayor tamaño en organismos de vida libre poseen un alto contenido de G + C, tal vez como resultado de ambientes más variados y complejos que los genomas de menor tamaño, que se encuentran en organismos que habitan en ambientes de escasos nutrientes con la finalidad de reducir el gasto energético que implica la replicación (Mann y Chen, 2010).

La hipótesis evolutiva mejor apoyada respecto al origen del cromosoma I, fue formulada por Thomas en 2004, en ella se propone que en el proceso de generación del cromosoma bacteriano, diferentes replicones se separarían y fusionarían repetidamente dando como resultado una combinación de genes. Aquellas combinaciones que contenían genes esenciales o cuya pérdida llevaba a la muerte de la bacteria, se convirtieron en cromosomas, mientras que los otros replicones que sobrevivieron se convirtieron en plásmidos. La competencia entre las bacterias parece haber favorecido la generación de uno o unos pocos cromosomas que incluían poco ADN no codificante (Thomas, 2004). Sin embargo, lo que

hoy en día podemos observar con la secuenciación de distintos genomas, es que, a pesar de que existe colinearidad (sintenia) entre los cromosomas de bacterias relacionadas, también existen diferentes patrones de la organización del genoma bacteriano.

### **Cromosomas II y III**

Por décadas se pensó que todos los genomas bacterianos poseían un solo cromosoma indispensable y plásmidos dispensables (Krawiec y Riley, 1990). Sin embargo, hace dos décadas, aproximadamente, un segundo replicón con genes esenciales fue descubierto en la alfa-proteobacteria *Rhodobacter sphaeroides* (Suwanto y Kaplan, 1989). Este replicón fue llamado cromosoma II debido a la presencia de genes esenciales para el crecimiento: dos copias de los genes que codifican para ARNr (*rrnB* y *rrnC*), el gene para la gliceraldehído-3-fosfato deshidrogenasa (*gapB*) y el gene para la ácido  $\delta$ -aminolevulínico sintasa (*hemT*). Desde entonces, se han identificado replicones con propiedades similares en un gran número de bacterias principalmente en  $\alpha$ ,  $\beta$  y  $\gamma$  proteobacterias (Tabla 1). *Agrobacterium tumefaciens* es un caso excepcional ya que posee un cromosoma I circular y un cromosoma II lineal además de dos plásmidos (Allardet-Servent *et al.*, 1993). *Sinorhizobium meliloti* posee un cromosoma circular y dos plásmidos de gran tamaño (Sobral *et al.*, 1991). Uno de ellos (pSymB) pudiera ser un cromosoma II ya que contiene la única copia del gen que codifica para el *tARN* específico para el codón de arginina, el segundo más frecuentemente utilizado para este aminoácido (MacLellan *et al.*, 2004). En bacterias del género *Vibrio* se han encontrado dos cromosomas circulares, siendo el cromosoma II el más variado en cuanto al tamaño y contenido de genes (Okada *et al.*, 2005). Se ha observado que cada cromosoma de *V. cholerae* tiene un posible papel que actúa de acuerdo a las condiciones de crecimiento, ya que existe una expresión más alta de genes del cromosoma I para el crecimiento en

condiciones de laboratorio. Por el contrario, los genes que se encuentran en el cromosoma II se expresan a niveles mas altos cuando *V. cholerae* se encuentra en condiciones de crecimiento en el intestino del conejo (Xu *et al.*, 2003).

A diferencia de los organismos eucariontes, los cuales inician la replicación de sus cromosomas por un solo mecanismo, los procariontes multi-replicón poseen mecanismos diferentes de replicación, los cuales pueden ser independientes o dependientes de DnaA como proteína de inicio de la replicación. Como se mencionó anteriormente, la replicación del cromosoma I es muy parecida al modelo propuesto para *E. coli* la cual inicia en una región denominada *oriC* y conserva sitios para el pegado de DnaA, IHF y Dam. En contraste, la replicación del cromosoma II de la mayoría de las bacterias se lleva a cabo mediante los productos de los genes *repABC*, los cuales, recordemos, son característicos de la replicación de plásmidos. En el sistema *repABC*, la proteína de inicio de la replicación es RepC, mientras que RepAB se encargan de la segregación; a cargo de la regulación de este proceso se encuentra un ARN pequeño codificado entre *repB* y *repC* (Wood *et al.*, 2001; Cevallos *et al.*, 2008). De esta manera el tener mecanismos diferentes de replicación entre ambos cromosomas evita la competencia por la replicación entre ellos mismos (Egan *et al.*, 2005). Un ejemplo, el cual es dependiente de DnaA se encuentra en el cromosoma II de *V. cholerae*, el cual posee un origen de replicación distinto al *oriC* de *E. coli* y se replica a través de dos genes *rctAB*, siendo RctB la proteína que se une al origen de replicación; y RctA el ARN que controla la replicación y adicionalmente una región *inc* que actúa en la incompatibilidad de este sistema. Aunque este sistema es de tipo plásmido, no es independiente de la replicación del cromosoma I ya que se necesita la presencia de DnaA y Dam (DNA adenina metiltransferasa) para su replicación, ambos factores modulan la autorepresión de RctB, así, existe una coordinación en la replicación de ambos cromosomas (Egan y Waldor, 2003;

Duigou *et al.*, 2006). Al parecer, en organismos multicromosomales, resulta favorable poseer dos sistemas diferentes de replicación, ya que permite un cierto grado de coordinación del proceso de replicación, y se minimiza la competencia e incompatibilidad entre los dos replicones (Egan y Waldor, 2003).

La segregación de los cromosomas I y II, se lleva a cabo mediante mecanismos diferentes. Se ha observado en *V. cholerae* que al término de la replicación de ambos cromosomas, cada uno se encuentra en posiciones subcelulares diferentes. Ambos cromosomas son segregados mediante los productos de los genes *parAB*. El cromosoma I sigue una segregación asimétrica desde un polo de la célula hacia el otro. Mientras que el cromosoma II sigue una segregación simétrica del centro de la célula madre al centro de las células hijas. El proceso se realiza de manera independiente y el cromosoma II se agrega más tarde durante el ciclo celular (Yamaichi *et al.*, 2007; Thanbichler y Shapiro, 2008). En el caso de la segregación de replicones que poseen un sistema *repABC*, como el de *A. tumefaciens*, se ha observado que a diferencia de otros sistemas tipo plásmido, estos son segregados hacia los polos de la célula de igual manera que el cromosoma I asegurando así su permanencia en la célula (Kahng y Shapiro, 2003).

Con la secuenciación y anotación de genomas completos se hace evidente la cooperación funcional de ambos replicones, lo cual lleva a una interdependencia metabólica de ambos, pues los genes que codifican una vía metabólica completa se distribuyen entre los dos replicones (MacLellan *et al.*, 2004). Sin embargo, la mayoría de los genes esenciales para los procesos básicos de la vida están principalmente localizados en el cromosoma I. Los cromosomas II poseen unos cuantos genes esenciales, así como genes que pudieran ser parte de vías o procesos para adecuación a diferentes ambientes y nichos (Egan *et al.*, 2005).

Tanto en el cromosoma I como en el II de *V. cholerae*, existen sistemas toxina-antitoxina (TA). Existe controversia en cuanto a la fisiología de estos sistemas en los cromosomas, ya que estos sistemas presentes en plásmidos aseguran la permanencia del replicón. Sin embargo, se ha sugerido que estos sistemas podrían ser herramientas apoptóticas bajo condiciones de privación de nutrientes, en donde la muerte de una parte de los individuos puede proveer alimentación para los individuos que sobreviven (Hu *et al.*, 2010). También, estos sistemas pueden evitar que se extienda la infección de fagos y mantener la estabilidad genómica a través de la eliminación de las células mutantes (Hu *et al.*, 2010). Algunos estudios realizados con los sistemas TA (codificados por los genes *parDE*) presentes en el cromosoma II de *V. cholerae*, demuestran que la toxina (*parE*) provoca daño al ADN, induce la respuesta SOS y bloquea la división celular. Finalmente la toxina degrada al cromosoma I en aquellas células que han perdido el cromosoma II; de esta manera, se perpetúa la presencia de ambos cromosomas en el genoma de *V. cholerae* (Yuan *et al.*, 2011).

La composición nucleotídica de los cromosomas II no difiere con respecto a las de los cromosomas I. Se ha observado que la diferencia de contenido de G + C entre cromosomas I y II es de aproximadamente el 2%, esto indica que existe una relación que ha perdurado a lo largo del tiempo entre ambos cromosomas (Bavishi *et al.*, 2010). Dicha asociación evolutiva entre cromosoma I y II podría sugerir, no solo que ambos contienen genes esenciales, sino que también el cromosoma II puede tener un papel importante más allá de lo que se podía considerar tiempo atrás (Bavishi *et al.*, 2010). Al igual que la estructura de los cromosomas I, los cuales son, en su mayoría circulares y algunos lineales, los cromosomas secundarios también son circulares en su mayoría con solo dos casos lineales (Tabla 1).



Se han propuesto dos ideas acerca del origen de los cromosomas II. La primera sugiere que se originaron de un plásmido que ganó genes esenciales. El segundo escenario, sugiere que los cromosomas secundarios se originaron de un cromosoma ancestral el cual se dividió y dio lugar a una segunda molécula con un origen de replicación (Jumas-Bilak *et al.*, 1998; Egan *et al.*, 2005). Para el primer escenario, existen algunas evidencias que lo soportan, esto es debido a que miembros de las alfaproteobacterias poseen sistemas de replicación tipo plásmido en los cromosomas II, además existen eventos de translocación del cromosoma I al cromosoma secundario (Slater *et al.*, 2009; Wattam *et al.*, 2009). De acuerdo a este modelo, un plásmido ancestral con sistema de replicación tipo *repABC*, adquiere grupos de genes esenciales mediante transferencia intragenómica, del cromosoma I; o intergenómica, mediante transferencia horizontal de genes. Esta información adquirida por cualquiera de los medios antes descritos, puede reintegrarse al cromosoma I o permanecer como segundo replicón, estableciéndose con el tiempo como un cromosoma II (Figura 1, Slater *et al.*, 2009). El segundo escenario tiene como soporte las duplicaciones de fragmentos del genoma que ocurren entre secuencias repetidas. Un segundo evento de recombinación intracromosomal dentro de la región duplicada, podría resultar en la formación de dos replicones estables si ambos contienen un origen de replicación, uno de los cuales podría haber sido adquirido por transferencia lateral (Jumas-Bilak *et al.*, 1998 a y b). En apoyo a esta hipótesis es importante mencionar la división artificial del cromosoma de *Bacillus subtilis* en dos replicones estables (Itaya y Tanaka, 1997). En este estudio, mediante recombinación homóloga entre dos secuencias repetidas de 560 bp que fueron introducidas al cromosoma junto con un origen de replicación, se encontró que un segmento de DNA genómico de esta bacteria puede ser cortado y se replica como un replicón independiente, y

su presencia es importante para el crecimiento celular de la bacteria que lo porta (Itaya y Tanaka. 1997).

La evolución de los cromosomas II ha sido recientemente explorada (Bavishi *et al.*, 2010). En este trabajo, los autores emplean numerosos enfoques independientes con la finalidad de medir el grado de divergencia del cromosoma I y II dentro y entre diferentes especies bacterianas. Aunque el cromosoma II, es más pequeño que el cromosoma I en todos los organismos evaluados, su tamaño es variable en proporción al cromosoma I. Incluso entre cepas, el tamaño del cromosoma I y II varía de manera substancial. La diferencia mínima en cuanto al tamaño del cromosoma I y II entre los organismos utilizados en este análisis ocurre en *B. suis* ATCC 23445, con una diferencia de ~523 kb, siendo el cromosoma I ~57.8% del total del genoma. La diferencia máxima del tamaño entre el cromosoma I y II ocurre en *R. sphaeroides*, con una diferencia de ~2.24 Mb, siendo el cromosoma I ~77.2% del genoma total (Bavishi *et al.*, 2010). La divergencia de los cromosomas depende de la tasa de mutación y recombinación, de la transferencia horizontal de genes (HGT) y de la selección ecológica entre las bacterias (Bavishi *et al.*, 2010).

La ventaja de poseer más de un cromosoma no está del todo clara, pues no existe una relación obvia entre el tamaño del genoma y la presencia de cromosomas múltiples, ya que los genomas de procariontes varían de manera considerable en tamaño sin relación con la estructura cromosomal. Una posibilidad es que el tiempo de replicación del genoma es menor al tener dos cromosomas que al tener solo uno, pues se ha observado una reducción en el crecimiento en un medio mínimo de una cepa de *S. meliloti* la cual posee un solo cromosoma, resultado de la integración de sus tres replicones, esto es debido quizás al aumento en el tiempo de replicación de un solo cromosoma. La identificación de un segundo cromosoma puede realizarse determinando si este replicón se requiere para el crecimiento normal de la

célula, lo que implica que dicho replicón contiene uno o más genes necesarios para la supervivencia de la célula bacteriana (Krawiec y Riley, 1990). Los genes identificados han sido principalmente genes que codifican para ARNs ribosomales y de transferencia; en algunos casos, los cuales representan la minoría, se han identificado genes que codifican para otras funciones implicadas en metabolismo y que probablemente sean requeridos para la viabilidad de la célula (Gao y Chen, 2010).

Recientemente, con la secuenciación de miembros del complejo *Burkholderia cepacia* (BCC, por sus siglas en inglés), que incluyen cepas de biorremediación y biocontrol, así como patógenos de plantas, animales y humanos (Mahenthiralingam *et al.*, 2005), se ha observado que dichos miembros poseen tres cromosomas los cuales han sido nombrados como tal, ya que contienen tanto tARN's, como rARN's (Holden *et al.*, 2009). En el caso de *B. cenocepacia* J2315, el genoma está compuesto por tres cromosomas circulares I, II, y III de 3.87, 3.21 y 0.875 Mb, respectivamente, cuyo contenido de G + C es similar entre ellos. Los cromosomas CI, CII y CIII contienen 4, 1 y 1 rARN's en cada cromosoma, por otro lado, CI, CII y CIII contienen 66, 6 y 2 tARN's, respectivamente (Holden *et al.*, 2009).

### **Plásmidos.**

Los genomas procariontes, por lo general, incluyen elementos extracromosomales o plásmidos, los cuales tienen la capacidad de replicarse de forma autónoma, así como de beneficiar a la bacteria en la adaptación a ambientes especializados, además de conferir resistencia a antibióticos y poseer genes de virulencia, resistencia a metales pesados, o vías metabólicas que le proporcionan la capacidad de poblar distintos nichos (del Solar *et al.*, 1998). Por lo general no contienen genes esenciales y se consideran dispensables para la vida celular bajo condiciones controladas de crecimiento (Thomas, 2004).

Tener ciertos genes en plásmidos podría permitir que las bacterias que portan dichos elementos sean capaces de sobrevivir en una mayor diversidad de ambientes sin la carga de un cromosoma más grande, es decir, las bacterias deben ser capaces de dividirse rápidamente bajo ciertas condiciones para obtener una ventaja selectiva, el poseer cromosomas más pequeños podría permitir que estas bacterias se repliquen con mayor rapidez que si tuvieran cromosomas mayores. Aunado a esto, los plásmidos pueden ser distribuidos entre diferentes miembros de la población ya que son más fáciles de dispersar de una cepa a otra, permitiendo a las bacterias habitar muchos más nichos ecológicos, lo que contribuye a que dichas bacterias y al mismo tiempo el o los plásmidos prevalezcan con mayor éxito (Thomas, 2004).

La evolución de los plásmidos no está separada de la historia evolutiva de los cromosomas, ya que los replicones que no se integraron al cromosoma se mantuvieron como plásmidos. Estos elementos, de manera ideal, no deben ser eliminados por segregación, y además, deben de ser capaces de dispersarse entre otras bacterias. Algunos plásmidos pueden tener una alta estabilidad segregacional, la cual depende de un sistema eficiente de replicación, un mecanismo de control del número de copias, asegurando la segregación de manera eficaz. Debido a lo anterior, algunos de los genes que se encuentran en los plásmidos codifican proteínas involucradas en la resolución de multímeros y sistemas de partición activa, los cuales son claves para la estabilidad del plásmido (Thomas, 2004).

Contrario a lo que sucede con el cromosoma, la replicación de los plásmidos no está acoplada al ciclo celular ni a la replicación del cromosoma, sino es más bien una replicación controlada de manera autónoma por mecanismos que responden al número de copias. A diferencia del cromosoma, la replicación de un plásmido puede iniciarse varias veces durante el ciclo celular de la bacteria (del Solar *et al.*, 1998). La segregación de plásmidos asegura la permanencia del plásmido en la célula. Los procesos de segregación de plásmidos se llevan a

cabo mediante la intervención de proteínas que son de tipo actina, tipo tubulina y parA, las cuales llevan a cabo un proceso análogo a la segregación de los cromosomas eucariontes durante la mitosis (Ghosh *et al.*, 2006; Gerdes *et al.*, 2010).

En general, la estrategia que usan las bacterias para mantener los plásmidos es a través de los sistemas de estabilidad post-segregacional, llamados sistemas toxina-antitoxina (TA), los cuales involucran un proteína estable (toxina) y una antitoxina inestable, la cual puede ser una proteína o un ARN antisentido. De tal manera que si la célula pierde el plásmido y con ello los genes que codifican para la toxina y antitoxina, la biosíntesis de ambos se detiene; al ser la antitoxina de vida media más corta no es capaz de neutralizar la toxina de vida larga, lo que provoca la muerte celular. Existe una amplia variación de los sistemas TA, ya que estos pueden actuar como mediadores de la respuesta a estrés y de muerte celular programada, en la generación de células tolerantes conocidas como “persisters”, las cuales son variantes en estado latente que son altamente resistentes a antibióticos sin que posean los genes para la resistencia a estos (Lewis, 2010). El sistema TA también está involucrado en desarrollo celular y estabilización de regiones dispensables (Hu *et al.*, 2010). Los sistemas TA no son exclusivos de plásmidos, ya que también se encuentran en cromosomas, donde muestran un perfil filogenético mucho más complejo y su papel funcional no está del todo entendido (Hu *et al.*, 2010; Yuan *et al.*, 2011).

### **Megaplásmidos y crómidos.**

Con mayor frecuencia encontramos en la literatura el término megaplásmido para hacer referencia al o los componentes genómicos de ADN que forman parte del genoma total, pero que su tamaño es menor al de un cromosoma I. Si un replicón debe recibir el sobrenombre megaplásmido no va más allá del deseo de armonizar la terminología. Algunos

megaplásmidos llegan a medir mucho más que los cromosomas primarios de algunas bacterias con genomas reducidos. Sin embargo, una connotación asociada con el término megaplásmido es que el replicón tiene una contribución genética y bioquímica, que no necesariamente es esencial, a la fisiología de un organismo (MacLellan *et al.*, 2004).

Hoy en día, se conocen un gran número de organismos que poseen megaplásmidos (Tabla 2, \* El tamaño considerado para megaplásmidos es  $\geq 0.25$  Mb). Los genes que se encuentran en los megaplásmidos contribuyen a expandir la versatilidad catabólica o pueden afectar la adecuación de algún organismo. Los genes que codifican productos tales como sistemas de transporte, reguladores transcripcionales y aquellos que podrían participar en la degradación de compuestos que se encuentran en la naturaleza se incluyen en esta categoría (MacLellan *et al.*, 2004). Aproximadamente, en 82 genomas secuenciados se reporta la presencia de megaplásmidos, las bacterias que presentan estas estructuras genómicas tienen estilos de vida diferentes y cerca de un 15% de las bacterias cuyos genomas contienen megaplásmidos se incluyen en aquellas que guardan una estrecha relación con diferentes especies eucariontes (Harrison *et al.*, 2010). De igual manera, en estos elementos genéticos encontramos genes relacionados a simbiosis y patogénesis, así como genes que confieren, en general, adaptación a diferentes nichos especializados (MacLellan *et al.*, 2004).

En el caso de los rizobios, se han encontrado numerosos genes considerados como parte del “genoma core”, los cuales se localizan en alguno o algunos de los megaplásmidos que forman parte del genoma de dichas bacterias. Algunos de estos genes son copias de genes que se localizan también en el cromosoma, con funciones redundantes, las cuales son dispensables para un crecimiento normal, por ejemplo, las copias múltiples de los genes que codifican para el sistema de chaperonas *groEL/groES* (Bittner *et al.*, 2007; Rodríguez-Quiñones *et al.*, 2005), dos copias de genes *tpiA*; los cuales codifican una probable enzima

triosa fosfato isomerasa, una enzima clave del metabolismo de carbono (Young *et al.*, 2006; González *et al.*, 2006; Yost *et al.*, 2006); así como dos probables enzimas asparagino sintetasas (*asnO* y *asnB*), en el plásmido pSymB de *S. meliloti*, las cuales podrían estar involucradas en la síntesis de asparagina a partir de aspartato mediante la amidación dependiente de ATP (Finan *et al.*, 2001).

Recientemente se ha propuesto el término “crómido” (Harrison *et al.*, 2010), con la finalidad de describir a los cromosomas II, III y algunos plásmidos con genes esenciales y reafirmar que las bacterias tienen solo un cromosoma. Harrison y colaboradores proponen tres criterios, que en combinación, según los autores, distinguen a los crómidos de los cromosomas y plásmidos: Los crómidos tienen sistemas de replicación y de segregación del tipo de los plásmidos, y tienen una composición nucleotídica similar a la del cromosoma de la bacteria. Los crómidos portan algunos genes que por lo general tienen una localización cromosomal en otras especies, y son, por lo general, más grandes que los plásmidos que forman parte del genoma, pero son más pequeños que el cromosoma, y portan muchos genes con una distribución filogenética restringida o esporádica (Harrison *et al.*, 2010). A la fecha se reconocen 82 genomas con crómidos, los cuales pertenecen principalmente al filo Proteobacteria (alfa: *Agrobacterium*, *Brucella*, *Ochrobactrum*, *Paracoccus*, *Rhizobium*, *Rhodobacter*, *Silicibacter*, *Sinorhizobium*; beta: *Burkholderia*, *Cupriavidus*, *Ralstonia*, *Variovorax*; gama: *Aliivibrio*, *Photobacterium*, *Pseudoalteromonas*, *Vibrio*), sin embargo, hay organismos representativos de Deinococcus-Thermus (*Deinococcus*), Spirochaetes (*Leptospira*) y Cianobacteria (*Cyanothece*) (Harrison *et al.*, 2010). Estos organismos tienen estilos de vida totalmente diferentes, aunque la mayoría están asociados con animales o plantas, como patógenos o como simbioses. Algunos crómidos adicionales de menor tamaño

están presentes en los genomas de *Paracoccus denitrificans*, en cuatro especies de *Rhizobium* secuenciadas y algunas *Burkholderia* (Harrison *et al.*, 2010).

Sin embargo, queda claro que la dificultad para asignar un nombre a los replicones bacterianos nos refleja la alta complejidad de la organización del genoma en estos organismos.

**TABLA 1.** Cromosomas adicionales en Bacterias y Archaea.

Grupo	Organismo	Tamaño (Mb)	Sistema de replicación	rARN	tARN	
CROMOSOMAS II						
α-Proteobacteria	<i>Agrobacterium tumefaciens</i> C58	2.07 lineal	<i>repABC</i>	6	15	
	<i>Agrobacterium vitis</i>	1.28	<i>repABC</i>	6	7	
	<i>Asticcacaulis exentricus</i> CB 48	1.31	<i>repA/dnaA</i>	6	27	
	<i>Brucella abortus</i> S19	1.16	<i>repC</i>	3	14	
	<i>Brucella abortus</i> bv. 1 str. 9-941	1.16	<i>repABC</i>	3	14	
	<i>Brucella canis</i> ATCC 23365	1.2	<i>repAC</i>	4	14	
	<i>Brucella melitensis</i> ATCC23457	1.18	<i>repC</i>	4	14	
	<i>Brucella melitensis</i> bv. Abortus 2308	1.15	<i>repC</i>	3	14	
	<i>Brucella melitensis</i> bv. 1 str. 16M	1.17	<i>repABC</i>	4	14	
	<i>Brucella microti</i> CCM 4915	1.22	<i>repABC</i>	3	14	
	<i>Brucella ovis</i> ATCC 25840	1.16	<i>repABC</i>	3	14	
	<i>Brucella suis</i> 1330	1.2	<i>repABC</i>	3	14	
	<i>Brucella suis</i> ATCC 23445	1.4	<i>parB-like</i>	7	19	
	<i>Ochrobactrum anthropi</i> ATCC 49188	1.89	<i>repABC</i>	6	16	
	<i>Paracoccus denitrificans</i> PD1222	1.73	<i>dnaA, repA</i>	3	0	
	β-Proteobacteria	<i>Rhodobacter sphaeroides</i> 2.4.1	0.94	<i>parAB, repA</i>	6	12
		<i>Rhodobacter sphaeroides</i> ATCC 17029	1.21	<i>dnaA, repA</i>	9	17
<i>Rhodobacter sphaeroides</i> KD 131		1.29	<i>dnaA, repA</i>	9	16	
<i>Sphingobium japonicum</i> UT26S		0.68	<i>repA</i>	6	4	
<i>Burkholderia cepacia</i> AMMD		2.64	<i>parB-like, parF</i>	3	6	
<i>Burkholderia ambifaria</i> MC40-6		2.76	<i>parB-like</i>	3	6	
<i>Burkholderia cenocepacia</i> AU 1054		2.78	<i>parB-like, parF</i>	3	5	
<i>Burkholderia cenocepacia</i> HI2424		2.99	<i>parB-like</i>	3	5	
<i>Burkholderia cenocepacia</i> J2315		3.21	<i>parAB</i>	3	6	
<i>Burkholderia cenocepacia</i> MC0-3		3.21	<i>parB-like</i>	3	5	
<i>Burkholderia glumae</i> BGR1		2.82	<i>repAC</i>	6	8	
<i>Burkholderia mallei</i> ATCC 23344		2.32	<i>repC</i>	6	9	
<i>Burkholderia mallei</i> NCTC 10229		2.28	<i>repC</i>	6	9	
<i>Burkholderia mallei</i> NCTC 10247		2.35	<i>repC</i>	6	10	
<i>Burkholderia mallei</i> SAVP1		1.73	<i>repC</i>	4	6	
<i>Burkholderia multivorans</i> ATCC 17616		2.47	<i>parAB, repA</i>	3	5	
<i>Burkholderia phymatum</i> STM815		2.69	<i>parB, repC</i>	4	5	
<i>Burkholderia phytofirmans</i> Pajñ		3.62	<i>parAB</i>	9	8	
<i>Burkholderia pseudomallei</i> 1106		3.1	<i>repC</i>	3	7	
<i>Burkholderia pseudomallei</i> 1710		3.18	<i>parAB, repC</i>	3	7	
<i>Burkholderia pseudomallei</i> 668		3.12	<i>repC</i>	3	7	
<i>Burkholderia pseudomallei</i> K96234		3.17	<i>parAB, repC</i>	3	8	
<i>Burkholderia</i> sp 383		3.58	<i>parB-like, parF</i>	3	5	
<i>Burkholderia</i> sp CCGE1002		2.59	<i>parB-like, repB</i>	12	15	
<i>Burkholderia</i> sp CCGE 1003		2.96	<i>parB-like, repB</i>	9	8	
<i>Burkholderia thailandensis</i> E264		2.91	<i>parAB, repC</i>	4	6	
<i>Burkholderia vietnamensis</i> G4		2.41	<i>parB-like</i>	3	4	
<i>Burkholderia xenovorans</i> LB400	3.36	<i>parBF, rep</i>	9	8		
<i>Cupriavidus taiwanensis</i> str. LMG19424	2.5	<i>repA2</i>	6	6		
<i>Ralstonia eutropha</i> H16	2.91	<i>repA</i>	6	7		
<i>Ralstonia eutropha</i> JMP134	2.72	<i>repA</i>	9	11		
<i>Ralstonia pickettii</i> 12D	1.32	<i>repA</i>	4	4		
<i>Ralstonia pickettii</i> 12J	1.3	<i>repA</i>	3	4		



β-Proteobacteria (continuación)	<i>Variovorax paradoxus</i> S110	1.12	<i>rep3</i>	0	15
	<i>Allivibrio salmonocida</i> LFI1238	1.2	<i>parA, sopB</i>	3	12
γ-Proteobacteria	<i>Photobacterium profundum</i> SS9	2.23	-	4	19
	<i>Vibrio cholerae</i> M66-2	1.04	<i>parB</i>	0	4
	<i>Vibrio cholerae</i> MJ-1236	1.08	<i>rstAB</i>	0	4
	<i>Vibrio cholerae</i> O1 bv eltor str. N16961	1.07	<i>parB</i>	0	4
	<i>Vibrio cholerae</i> O395	3.02	<i>dnaA, cri-1</i>	25	92
	<i>Vibrio fischeri</i> ES114	1.33	<i>rstAB</i>	3	8
	<i>Vibrio fischeri</i> MJ11	1.41	<i>parAB</i>	3	11
	<i>Vibrio harveyi</i> ATCC BAA-1116	2.20	<i>rstA</i>	3	16
	<i>Vibrio parahaemolyticus</i> RIMD 2210633	1.87	<i>rstA</i>	3	14
	<i>Vibrio sp</i> Ex25	1.82	<i>rstA</i>	3	13
	<i>Vibrio splendidus</i> LGP32	1.67	<i>parAB</i>	4	17
	<i>Vibrio vulnificus</i> CMCP6	1.84	<i>parAB</i>	3	13
<i>Vibrio vulnificus</i> YJ016	1.85	<i>parAB</i>	3	12	
Firmicutes	<i>Butyrivibrio proteoclasticus</i> B316	0.3	<i>repB1</i>	6	2
Deinococcus- Thermus	<i>Deinococcus radiodurans</i> R1	0.41	<i>parAB</i>	0	1
Bacteroidetes	<i>Prevotella melaninogenica</i> ATCC 25845	1.37	<i>dnaA</i>	4	12
Cloroflexi	<i>Sphaerobacter thermophilus</i> DSM 20745	1.25	<i>parAB</i>	3	3
Arquea	<i>Haloarcula marismortui</i> ATCC 43049	0.288	<i>parA5, parA3</i>	3	1
	<i>Halorubrum lacusprofundi</i>	0.525	<i>orc1/cdc6</i>	3	1
CROMOSOMAS					
III					
β-Proteobacteria	<i>Burkholderia cepacia</i> AMMD	1.28	<i>parB-like</i>	3	2
	<i>Burkholderia ambifaria</i> MC40-6	1.12	<i>parB-like</i>	3	2
	<i>Burkholderia cenocepacia</i> AU 1054	1.19	<i>parB-like, repN</i>	3	16
	<i>Burkholderia cenocepacia</i> HI2424	1.05	<i>parB-like</i>	3	2
	<i>Burkholderia cenocepacia</i> J2315	0.87	<i>parA-like</i>	3	2
	<i>Burkholderia cenocepacia</i> MC0-3	1.22	<i>parB-like</i>	3	2
	<i>Burkholderia multivorans</i> ATCC 17616	0.91	<i>parAB, repA</i>	3	2
	<i>Burkholderia sp</i> 383	1.39	<i>parBF</i>	3	2
	<i>Burkholderia sp</i> CCGE1002	1.28	<i>parB</i>	0	2
<i>Burkholderia vietnamensis</i> G4	1.24	<i>parBF</i>	3	2	

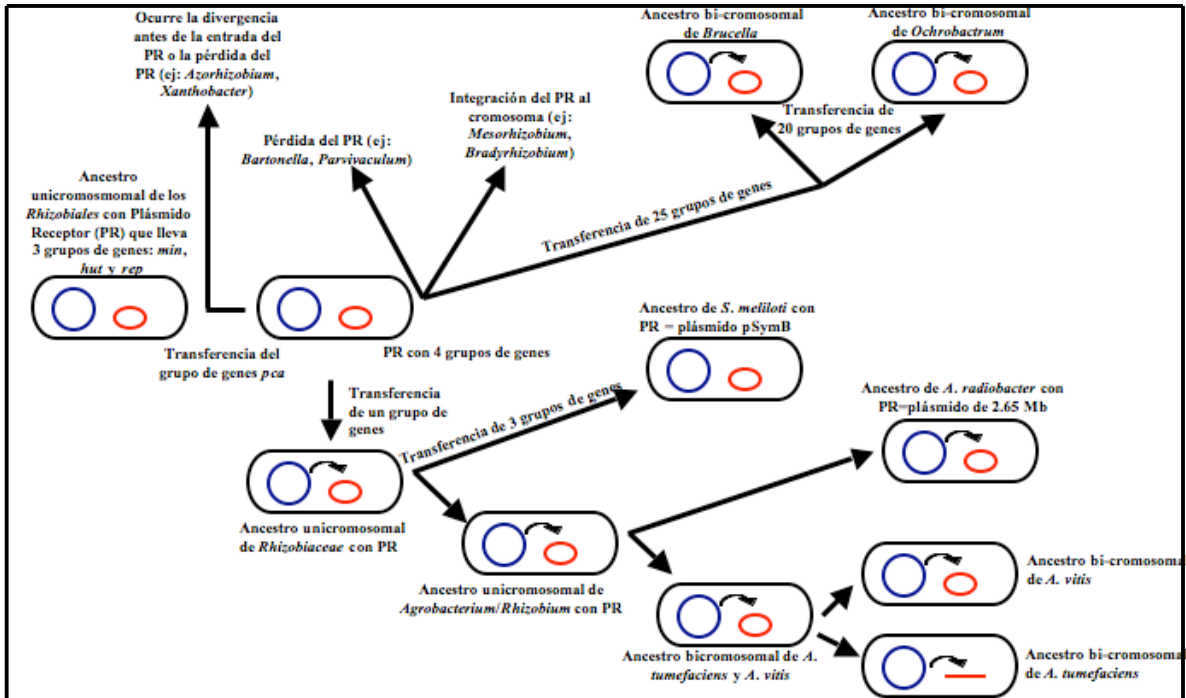
**TABLA 2.** Megaplásmidos en bacterias y arqueas \*

Grupo	Organismo	Plásmido	Tamaño Kb	rARN	tARN
$\alpha$ -Proteobacteria	<i>Agrobacterium radiobacter</i> K84	pATK84c	338.69	-	-
	<i>Agrobacterium tumefaciens</i> C58	pAt	542.86	-	-
	<i>Agrobacterium vitis</i>	pATS4e	631.77	-	-
		pTiS4	258.82	-	1
	<i>Azospirillum sp</i> B510	pAB510a	1451.1	11	14
		pAB510b	723.77	3	2
		pAB510c	681.72	3	3
		pAB510d	628.83	-	6
		pAB510e	537.29	3	9
		pAB510f	261.59	-	-
	<i>Chelativorans sp</i> BNC1	pI	343.93	-	4
	<i>Ketogulonicigenium vulgare</i> Y25	pYP1	268.67	3	3
	<i>Mesorhizobium ciceri</i> WSM1271	pMESCI01	425.53	-	-
	<i>Mesorhizobium loti</i> MAF 303099	pMLa	351.91	-	-
	<i>Methylobacterium chlorometanicum</i> CM4	pMCHL01	380.2	-	-
	<i>Methylobacterium extorquens</i> AM1	Megaplasmid	1261.46	-	6
	<i>Methylobacterium nodulans</i> ORS 2060	pMNOD01	487.73	-	2
		pMNOD02	458.07	-	-
	<i>Methylobacterium radiotolerans</i> JCM 2831	pMRAD01	586.16	6	1
	<i>Nitrobacter hamburgensis</i> X14	pI	294.82	-	-
	<i>Novosphingobium aromaticivorans</i> DSM12444	pNL2	487.26	-	-
	<i>Paracoccus denitrificans</i> PD 1222	pI	653.81	-	-
	<i>Phenylobacterium zucineum</i> HLK1	Megaplasmid	382.97	-	-
	<i>Rhizobium etli</i> CFN42	p42c	250.94	-	-
		p42e	505.33	-	-
		p42f	642.51	-	-
		p42d	371.25	-	-
	<i>Rhizobium etli</i> CIAT652	pA	414.09	-	-
		pB	429.11	-	-
		pC	1091.52	-	-
	<i>Rhizobium leguminosarum</i> bv trifolii WSM1325	pR132501	828.62	-	-
		pR132502	660.97	-	-
		pR132503	516.08	-	-
		pR132504	350.31	-	-
		pR132505	294.78	-	-
	<i>Rhizobium leguminosarum</i> bv trifolii WSM2304	pRLG202	501.94	-	-
		pRLG203	308.74	-	-
		pRLG204	257.95	-	-
	<i>Rhizobium leguminosarum</i> bv viciae 3841	pRL10	488.135	-	-
		pRL11	684.2	-	-
		pRL12	870.02	-	-
		pRL9	352.78	-	-
	<i>Rhodobacter sphaeroides</i> ATCC 17025	pRSPA01	877.87	6	11
		pRSPA02	289.48	3	3
	<i>Ruegeria pomeroyi</i> DSS-3	Megaplasmid	491.61	-	2
	<i>Ruegeria sp.</i> TM1040	Megaplasmid	821.78	9	20
	<i>Ruminococcus albus</i> 7	pRUMAL01	420.7	-	14
pRUMAL02		352.64	-	-	
<i>Sinorhizobium fredii</i> NGR234	pNGR234a	536.16	-	1	
	pNGR234b	2430.03	-	-	
<i>Sinorhizobium medicae</i> WSM419	pSMED01	1570.95	-	1	
	pSMED02	1245.4	-	-	
<i>Sinorhizobium meliloti</i> 1021	pSymA	1354.22	-	2	
	pSymB	1683.33	-	1	
<i>Sphingomonas sp.</i> KA1	pCAR3	254.79	-	-	
<i>Sphingomonas wittichii</i> RW1	pSWIT01	310.22	-	-	
<i>Xanthobacter autotrophicus</i> Py2	pXAUT01	316.16	-	-	
$\beta$ -Proteobacteria	<i>Burkholderia ambifaria</i> MC40-6	pBMC401	301.59	-	-
	<i>Burkholderia phymatum</i> STM 15	pBPHY01	1904.89	-	1
pBPHY02		595.1	-	-	

β-Proteobacteria (continuación)	<i>Burkholderia sp.</i> CCGE1002	pBC201	489.13	-	-
	<i>Burkholderia vietnamensis</i> G4	pBVIE01	397.86	-	-
		pBVIE02	265.61	-	-
	<i>Butyrivibrio proteoclasticus</i> B316	pCY360	361.69	-	-
	<i>Methylibium petroleiphilum</i> PM1	pRPME01	599.44	-	27
	<i>Polaromonas naphthalenivorans</i> CJ2	pNAP01	353.29	-	-
	<i>Polaromonas sp.</i> JS666	p1	360.4	-	-
		p2	338	-	-
	<i>Ralstonia eutropha</i> H16	pHG1	452.15	-	-
	<i>Ralstonia eutropha</i> JMP134	Megaplasmid	634.91	-	1
	<i>Ralstonia pickettii</i> 12D	pRp12D01	389.77	-	-
		pRp12D01	273.13	-	-
	<i>Ralstonia solanacearum</i> CFBP2957	pRCFBPv3	2163.37	-	4
	<i>Ralstonia solanacearum</i> GMI1000	pGMI1000MP	1094.5	3	3
<i>Rhodiferax ferrireducens</i> T118	p1	257.44	-	1	
γ-Proteobacteria	<i>Klebsiella pneumoniae</i>	pK29	269.67	-	-
	<i>Pantoea sp.</i> At-9b	pPAT9B01	793.95	-	-
		pPAT9B02	394.05	-	1
		pPAT9B03	321.08	-	-
		pPAT9B04	318.11	-	-
	<i>Pantoea vegans</i> C9-1	pPag3	529.67	-	-
<i>Pseudomonas fluorescens</i> SBW25	pQBR103	425.09	-	-	
Firmicutes	<i>Rahnella sp.</i> Y9602	pRAHAQ01	616.54	-	-
	<i>Serratia marcesens</i>	R478	274.76	-	-
	<i>Shewanella sp.</i> ANA-3	p1	278.94	-	-
	<i>Bacillus cereus</i>	pCER270	270.08	-	-
		pPER272	272.14	-	-
	<i>Bacillus cereus</i> AH187	pAH187_270	270.08	-	4
	<i>Bacillus cereus</i> AH820	pAH820_272	272.14	-	-
	<i>Bacillus cereus</i> E33L	pE33L466	466.7	-	-
	<i>Bacillus cereus</i> H3081.97	pH308197258	258.48	-	-
	<i>Bacillus pseudofirmus</i> OF4	pBpOF4-01	285.22	-	-
	<i>Bacillus thuringiensis</i> BMB171	pBMB171	312.96	-	-
	<i>Bacillus weihenstephanensis</i> KBAB4	pBWB401	417.05	-	-
	<i>Clostridium botulinum</i> A3 str. Loch Maree	pCLK	266.78	-	2
	<i>Clostridium botulinum</i> Ba4 str. 657	pCLJ	270.02	-	1
	<i>Cupriavidus metallidurans</i> CH34	Megaplasmid	2580.08	-	8
	<i>Cupriavidus taiwanensis</i>	pRALTA	557.2	-	-
	<i>Eubacterium eligens</i> ATCC 27750	Megaplasmid	626.74	6	-
	<i>Paenibacillus polymyxa</i> SC2	pSC2	510.115	-	44
	Cyanobacteria	<i>Acaryochloris marina</i> MBIC11017	pREB1	374.16	-
pREB2			356.08	-	-
pREB3			273.12	-	-
<i>Anabaena variabilis</i> ATCC 29413		pA	366.35	-	-
		pC	300.75	-	1
<i>Cyanothece sp.</i> PCC 7424		pP742401	328.63	-	-
<i>Cyanothece so.</i> PCC 7822		Cy782201	879.54	-	-
		Cy782202	473.84	-	-
		Cy782203	291.99	-	-
<i>Nostoc punctiforme</i> PCC 73102		pNPUN01	354.56	-	-
	pNPUN02	254.91	-	-	
<i>Nostoc sp.</i> PCC 7120	pPCC7120α	408.1	-	-	
Acidobacteria	<i>Acidobacterium sp.</i> MP5ACTX9	pACIX901	475.65	-	-
		pACIX902	300.29	-	-
Actinobacteria	<i>Arthrobacter aureus</i> TC1	pTC1	328.23	-	-
		pTC2	300.72	-	-
	<i>Arthrobacter chlorophenicus</i>	pACHL01	426.85	-	39
	<i>Streptomyces coelicolor</i> A3(2)	pSCP1	356.02	-	-
	<i>Mycobacterium gilvum</i> PYR-GCK	pMFLV01	321.253	-	1
<i>Mycobacterium sp.</i> KMS	pMKMS01	302.08	-	-	
Deinococcus Thermus	<i>Deinococcus deserti</i> VCD115	p1	324.71	-	-
		p2	314.31	-	-
		p3	396.45	-	-
	<i>Deinococcus geothermlis</i> DSM11300	pDGEO01	574.127	3	-
<i>Enterobacter cloacae</i>	pEC-IMP	318.78	-	-	

Deinococcus Thermus	<i>Miothermus silvanus</i> DSM 5159	pMESIL01	347.85	-	-
	<i>Thermus thermophilus</i> HB8	pTT27	256.99	-	-
Cloroflexi	<i>Herpetosiphon aurantiacus</i> ATCC 23779	pHPPC4	339.63	-	-
	<i>Thermomicrobium roseum</i> DSM 5159	Megaplasmid	917.73	-	-
Arquea	<i>Halalkalicoccus jeotgali</i> B3	1	406.28	-	1
		2	363.53	-	-
	<i>Haloarcula marismortui</i> ATCC 43049	pNG700	410.55	1	-
	<i>Halobacterium salinarum</i> R1	pHS3	284.33	-	-
	<i>Halobacterium sp.</i> NRC-1	pNRC200	365.42	-	-
	<i>Haloferax volcanii</i> DS2	pHV3	437.9	-	-
		pHV4	635.78	-	-
	<i>Halogeometricum borinquense</i> DSM 11551	pHBOR01	362.19	-	-
		pHBOR02	339.01	-	-
	<i>Halomicrobium mukohataei</i> DSM 12286	pHmurk01	221.86	4	-
	<i>Halorubrum lacusprofundi</i> ATCC 49239	pHLAC01	431.33	-	-
	<i>Haloterrigena turkmenica</i> DSM 5511	pHTUR01	698.49	-	-
		pHTUR02	413.64	-	-
	<i>Natrialba magadii</i> ATCC 43099	pNMAG02	254.95	-	-
		pNMAG01	378.34	6	1

\* El tamaño considerado para megaplásmidos fue  $\geq 0.25$  Mb



**Figura 1.** Posible origen de cromosomas secundarios y megaplásmidos en  $\alpha$ -Proteobacterias. La formación de estas estructuras probablemente se dio mediante la transferencia intragenómica de algunas regiones del cromosoma primario ancestral hacia un plásmido ancestral que poseía un sistema de replicación tipo *repABC* (Plásmido Receptor, PR) en el grupo *Agrobacterium/Rhizobium*, este PR pudo haberse cointegrado nuevamente al cromosoma primario en grupos como *Mesorhizobium* y *Bradyrhizobium*; también pudo haberse escindido en los grupos de *Ochrobactrum* y *Brucella*. En aquellos grupos en donde PR se mantuvo independiente del cromosoma primario, pudo haber adquirido más genes y así dar origen a un cromosoma secundario (Modificado de Slater *et al.*, 2009).

## ANTECEDENTES.

El género *Rhizobium*, comprende a un grupo de bacterias Gram-negativas que se ubican taxonómicamente dentro de las alfa proteobacterias, en el orden rhizobiales y en la familia *Rhizobiaceae*. La estructura genómica de los rhizobiales consta de genomas multi-replicón, constituidos por un cromosoma I y distintos replicones como cromosomas II, plásmidos, megaplásmidos y recientemente, según Harrison y colaboradores, algunos crómidos ([www.ncbi.nlm.nih.gov/Taxonomy](http://www.ncbi.nlm.nih.gov/Taxonomy)). El hábitat de estas bacterias puede ser la rizósfera, una zona rica en moléculas orgánicas alrededor de las raíces de las plantas, o debido a su capacidad para establecer asociaciones mutualistas, también se les puede encontrar en el interior de las raíces de sus plantas hospedadas. La asociación mutualista

mejor estudiada en este grupo bacteriano es la fijación simbiótica de nitrógeno con las raíces de plantas leguminosas, pero algunos estudios indican que estas bacterias también pueden vivir como endófitos en las raíces de plantas no leguminosas (Rosenblueth y Martínez-Romero, 2004).

*Rhizobium etli* cepa CFN42 se aisló en México de los nódulos formados en las raíces de plantas de frijol (*Phaseolus vulgaris*). Es una de las bacterias mejor caracterizadas en México. Su genoma, totalmente secuenciado en nuestro país, se encuentra constituido por un cromosoma circular de 4.3 Mb y seis plásmidos circulares denominados, p42a (194.2 kb), p42b (184.3 kb), p42c (250.9 kb), p42d (371.2 kb), p42e (505.3 kb) y p42f (642.5 kb) (González *et al.*, 2006). En suma, los plásmidos representan el 32% de su genoma total. De acuerdo a los datos aportados por la secuencia genómica de *R. etli*, se sabe que el cromosoma (4 Mb) codifica la mayoría de las funciones necesarias para el crecimiento celular, mientras que pocos genes esenciales o vías metabólicas completas están codificadas en plásmidos (González *et al.*, 2006). Para conocer la participación y función de los plásmidos en este organismo, se obtuvieron cepas carentes de cada uno de los plásmidos que forman parte del genoma de *R. etli* CFN42 (Brom *et al.*, 1992). El análisis fenotípico de estas cepas derivadas permitió asignar diferentes funciones codificadas en los plásmidos (Tabla 3)

**Tabla 3.** Funciones asignadas a los plásmidos de acuerdo a cepas curadas en cada uno de los plásmidos

<b>Plásmido</b>	<b>Función asignada</b>	<b>Referencia</b>
p42a	Transferencia conjugativa .	Brom <i>et al.</i> , 2000 Tun-Garrido <i>et al.</i> , 2003
p42b	Biosíntesis de lipopolisacáridos,  Biosíntesis de tiamina y degradación de glicerol	Brom <i>et al.</i> , 1992 García de los Santos y Brom, 1997 Miranda-Ríos <i>et al.</i> , 1997 Brom <i>et al.</i> , 2000

	Resistencia a compuestos tóxicos exudados por las raíces de frijol.	González-Pasayo <i>et al.</i> , 2000
p42c	Metabolismo de trigonelina y dulcitol	Brom <i>et al.</i> , 2000
p42d	Funciones simbióticas y fijación de nitrógeno.	Brom <i>et al.</i> , 2000
p42e	Degradación de melibiosa Funciones de regulación que al parecer son esenciales para el crecimiento en medio rico. Resistencia a cobre Síntesis de NAD. Morfología y división celular	Brom <i>et al.</i> , 2000 Landeta <i>et al.</i> , 2011
p42f	Deficiencia de crecimiento en medio mínimo. Actividad dual del la única catalasa-peroxidasa ( <i>katG</i> ).	Brom <i>et al.</i> , 1992 Vargas <i>et al.</i> , 2003.

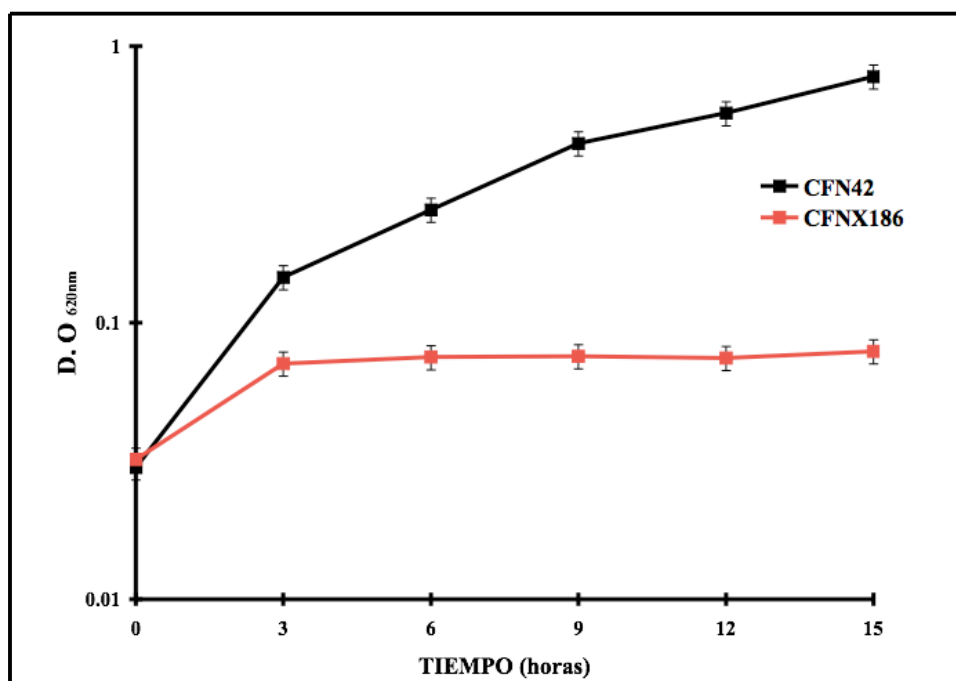
Aun cuando el trabajo de Brom *et al.*, 1992, reporta la obtención de cepas curadas en casi todos los plásmidos, de manera sorpresiva resultó imposible obtener una cepa carente de p42e (505 Kb) en su totalidad a pesar del uso de sistemas de selección positiva particularmente eficientes (Brom *et al.*, 1992). Únicamente se obtuvo una cepa con una deleción de aproximadamente 210 Kb en dicho plásmido, la cual fue designada como p42e<sub>Δ</sub> (CFNX185, Brom *et al.*, 1992). Recientemente, en un trabajo publicado en nuestro laboratorio (Landeta *et al.*, 2011) en el cual se analiza al plásmido p42e mediante herramientas bioinformáticas, mutagénesis y deleciones programadas, los autores encuentran que aproximadamente el 11% del plásmido p42e podría estar involucrado en el metabolismo, participando en funciones biosintéticas: vitaminas B1 y B12, cardiolipina y biosíntesis de NAD; funciones de degradación: ácido protocatéquico e histidina; y formación del septo de división con la participación de los genes *minCDE*. Además, el análisis de deleciones programadas sobre este plásmido, permitió la identificación de dos genes (*RHE\_PE00001* y

*RHE\_PE00024*) los cuales codifican, respectivamente, una proteína hipotética con un probable motivo HTH de unión a ADN, y una proteína híbrida entre una probable histidina-cinasa y un probable regulador de respuesta, que al parecer son esenciales para el crecimiento en medio rico (Landeta *et al.*, 2011).



## JUSTIFICACIÓN Y PLANTEAMIENTO DEL PROBLEMA.

Estudios previos a la secuenciación del genoma de *R. etli* CFN42 demostraron que la mutante CFNX186, curada del plásmido p42f, no puede crecer en medio mínimo (MM) con succinato y amonio como fuentes de carbono y nitrógeno respectivamente, así como sales minerales ( $\text{CaCl}_2$ ,  $\text{MgSO}_4$ ,  $\text{K}_2\text{HPO}_4$ , y  $\text{FeCl}_3$ ) (Figura 2). Este dato llevó a los autores a proponer como hipótesis que en este plásmido están codificadas enzimas relacionadas con algún proceso metabólico primario (Brom *et al.*, 1992). Esta posible auxotrofia asociada a la pérdida del plásmido p42f, así como la secuencia completa del plásmido p42f se convirtieron en nuestro modelo de estudio para contribuir a la identificación de funciones esenciales en el genoma extracromosomal diferentes de los ARNt's y ARNr's previamente descritos en cromosomas II y III y en algunos plásmidos.



**Figura 2.** Crecimiento de la cepa Silvestre CFN42 y de la derivada curada de p24f, CFNX186 en medio mínimo. Luego de 15 horas de crecimiento, la cepa CFNX186 no alcanza los niveles de crecimiento como lo hace la cepa silvestre.

La relevancia de este trabajo radica en proporcionar un ejemplo respecto a la importancia de genes que se encuentran en plásmidos, así como en ampliar el concepto funcional y evolutivo de los plásmidos. Tradicionalmente, se ha considerado que las funciones esenciales para el crecimiento bacteriano están codificadas en el cromosoma y no en los plásmidos. La presencia de genes del metabolismo primario en plásmidos es de gran relevancia ya que amplía el papel biológico que se ha asignado a estos replicones. Desde el punto de vista de la evolución del genoma bacteriano, apoya la hipótesis de que los plásmidos podrían ser los precursores de los cromosomas adicionales presentes en los genomas multireplicón.

### **OBJETIVO GENERAL:**

Identificar y caracterizar las funciones codificadas en el plásmido p42f necesarias para que *R. etli* pueda crecer en medio mínimo.

### **OBJETIVOS ESPECÍFICOS:**

1. Identificar *in silico* la presencia de genes esenciales en copia única codificados en el plásmido p42f, cuya función sea necesaria para crecer en medio mínimo.
2. Demostrar que mutaciones en esos genes causan deficiencia de crecimiento en medio mínimo.
3. Evaluar la posibilidad de que los genes responsables del crecimiento hayan llegado al p42f por transferencia horizontal o mediante migración intragenómica del cromosoma a dicho plásmido.
4. Determinar si los genes esenciales presentes en el plásmido p42f también están localizados en plásmidos de otros miembros del orden *Rhizobiales*.

## ESTRATEGIA EXPERIMENTAL.

Para realizar cada uno de los objetivos planteados en este proyecto se diseñaron las siguientes estrategias:

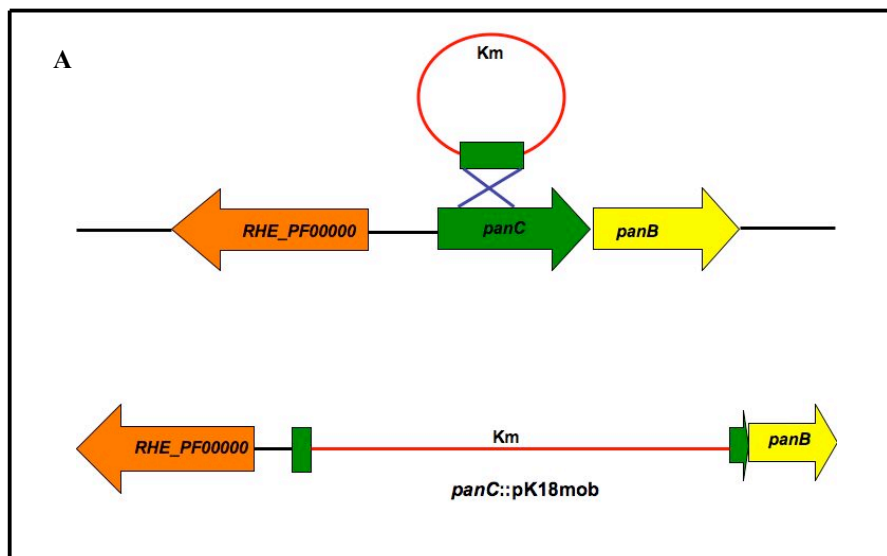
1. Búsqueda de enzimas del metabolismo primario, de copia única, codificadas en el plásmido p42f. Realizamos una re- anotación de las proteínas codificadas en el plásmido p42f de *R. etli* CFN42 con la herramienta de KEGG (Kyoto Encyclopedia of Genes and Genomes) denominada KAAS (KEGG Automatic Annotation Server and Pathway Reconstruction Server) (Moriya *et al.*, 2007). Dicho servidor proporciona la anotación funcional de genes por medio de comparaciones tipo BLAST contra la base de datos manualmente curada “KEGG GENES”. Al resultado se le asigna un valor KO (KEGG Orthology) y automáticamente se genera un mapa gráfico que contiene las vías metabólicas KEGG.

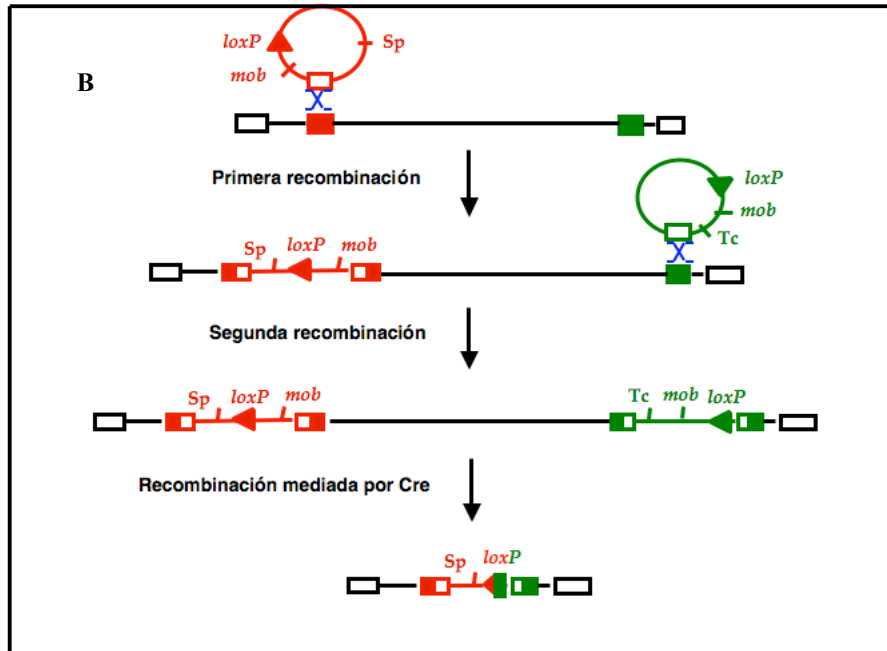
Posteriormente, seleccionamos posibles enzimas de vías metabólicas primarias, y determinamos si estas enzimas seleccionadas estaban en más de una copia en el genoma, por lo cual se realizaron comparaciones tipo BLAST-P de cada uno de las enzimas seleccionadas contra todo el genoma, mediante “Bidirectional Best Hit” (BBH).

2. Mutagénesis por inserción de vector. Para evaluar la contribución de las proteínas seleccionadas *in silico* al crecimiento en MM de *R. etli* CFN42, se mutagenizaron los respectivos genes por la técnica de inserción de vector mediante recombinación homóloga. Brevemente, un segmento interno del gene de interés, amplificado por PCR, se inserta en un plásmido (pK18mob-Km<sup>r</sup>) no replicativo en *R. etli*. Introduciendo dicho plásmido a *R. etli* mediante conjugación, y seleccionando transconjugantes con resistencia a Km, se obtienen mutantes que integraron el vector por recombinación homóloga en el gene de interés provocando la pérdida de su función (Figura

3A). La correcta interrupción del gene se verificó mediante hibridaciones tipo Southern, además de evaluar el fenotipo de cada mutante creciéndola en un medio definido.

3. Adicionalmente, se llevaron a cabo deleciones programadas a lo largo del plásmido, mediante un sistema de deleción sitio específico (Ayres *et al.*, 1993) el cual emplea sitios *loxP* y la recombinasa Cre. El sistema ha sido adaptado para su uso en *R. etli*. Brevemente, los sitios *loxP*, consisten en dos secuencias palindrómicas de 13 pares de bases (pb) flanqueado una secuencia de 8 pb, las cuales son reconocidos por la recombinasa Cre.





**Figura 3.** Estrategia experimental utilizada en este trabajo. (A) Mutagénesis por inserción de un vector realizado por un evento de recombinación. (B) Sistema de delecciones programadas mediante sitios *loxP* y la recombinasa sitio-específica Cre (Modificado de Ayres et al., 1993).

Para hacer este sistema funcional, incorporamos dos sitios *loxP* flanqueando la zona que nos interesa deletar mediante recombinación sitio específica, como se muestra en la figura 2B, se utilizan dos vectores con marcadores de resistencia diferentes. Una vez incorporados los sitios *loxP*, se expresa la recombinasa Cre en *trans*, la cual reconoce esta secuencia y corta en cada uno de los sitios, lo cual permite obtener la delección, posteriormente, el vector que lleva Cre se elimina por crecimientos a 30 grados. La correcta eliminación del fragmento del plásmido se verificó por el cambio en la movilidad electroforética del replicón y por la ausencia de genes localizados en la región deletada los cuales no pueden ser amplificados por PCR.

## **RESULTADOS.**

Parte de los resultados que se obtuvieron durante este estudio fueron publicados en el artículo: **Villaseñor, T.**, S. Brom, A. Dávalos, L. Lozano, D. Romero, and A. García de los Santos. 2011. Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*. *BMC Microbiology*. 11:66

RESEARCH ARTICLE

Open Access

# Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*

Tomás Villaseñor<sup>1</sup>, Susana Brom<sup>1</sup>, Araceli Dávalos<sup>1</sup>, Luis Lozano<sup>2</sup>, David Romero<sup>1</sup> and Alejandro García-de los Santos<sup>1\*</sup>

## Abstract

**Background:** A traditional concept in bacterial genetics states that housekeeping genes, those involved in basic metabolic functions needed for maintenance of the cell, are encoded in the chromosome, whereas genes required for dealing with challenging environmental conditions are located in plasmids. Exceptions to this rule have emerged from genomic sequence data of bacteria with multipartite genomes. The genome sequence of *R. etli* CFN42 predicts the presence of *panC* and *panB* genes clustered together on the 642 kb plasmid p42f and a second copy of *panB* on plasmid p42e. They encode putative pantothenate biosynthesis enzymes (pantoate- $\beta$ -alanine ligase and 3-methyl-2-oxobutanoate hydroxymethyltransferase, respectively). Due to their ubiquitous distribution and relevance in the central metabolism of the cell, these genes are considered part of the core genome; thus, their occurrence in a plasmid is noteworthy. In this study we investigate the contribution of these genes to pantothenate biosynthesis, examine whether their presence in plasmids is a prevalent characteristic of the *Rhizobiales* with multipartite genomes, and assess the possibility that the *panCB* genes may have reached plasmids by horizontal gene transfer.

**Results:** Analysis of mutants confirmed that the *panC* and *panB* genes located on plasmid p42f are indispensable for the synthesis of pantothenate. A screening of the location of *panCB* genes among members of the *Rhizobiales* showed that only *R. etli* and *R. leguminosarum* strains carry *panCB* genes in plasmids. The *panCB* phylogeny attested a common origin for chromosomal and plasmid-borne *panCB* sequences, suggesting that the *R. etli* and *R. leguminosarum* *panCB* genes are orthologs rather than xenologs. The *panCB* genes could not totally restore the ability of a strain cured of plasmid p42f to grow in minimal medium.

**Conclusions:** This study shows experimental evidence that core *panCB* genes located in plasmids of *R. etli* and *R. leguminosarum* are indispensable for the synthesis of pantothenate. The unusual presence of *panCB* genes in plasmids of *Rhizobiales* may be due to an intragenomic transfer from chromosome to plasmid. Plasmid p42f encodes other functions required for growth in minimal medium. Our results support the hypothesis of cooperation among different replicons for basic cellular functions in multipartite rhizobia genomes.

## Background

Multipartite genomes are common among members of the  $\alpha$ -proteobacteria [1]. Most symbiotic nitrogen-fixing bacteria belonging to the genera *Rhizobium*, *Sinorhizobium*, *Mesorhizobium* and *Bradyrhizobium* possess multipartite genomes organized as a single circular chromosome and a variable number of large plasmids [2]. In some species plasmids can represent, in terms of size, up to 40% of

the total genome. In *Rhizobium* and *Sinorhizobium* species one plasmid (pSym) concentrates most of the genes required for nodulation and nitrogen fixation [3]. The complete genome sequences of different rhizobia have revealed that plasmids harbor mainly accessory genes and that most encode predicted transport systems and a variety of catabolic pathways that may contribute to the adaptation of rhizobia to the heterogeneous soil and nodule environments [2,4]. These genes are absent from closely related genomes, lack synteny and their G+C composition differs from that of the core genes. The core genes are mainly

\* Correspondence: alex@ccg.unam.mx

Full list of author information is available at the end of the article



located on chromosomes, have essential functions in cell maintenance and have orthologs in related species [5,6]. In spite of this evidently biased distribution of core genes in the chromosome and accessory genes in plasmids, it is important to highlight the fact that there are interesting exceptions to this genomic rule: several typical core genes have been found encoded on rhizobia plasmids. Some are copies of genes located on chromosomes, with redundant functions that are totally dispensable for normal growth. Examples of these genes are the multiple copies of chaperonin-encoding genes *groEL/groES* [7,8], two *tpiA* genes encoding putative triose phosphate isomerase, a key enzyme of central carbon metabolism [4,6,9], and two putative *S. meliloti* asparagine synthetases (*asnB* and *asnO*), which may have a role in asparagine synthesis from aspartate by ATP-dependent amidation [10]. In contrast to these reiterated genes, a few single copy core genes have also been localized in plasmids. The *tRNA* specific for the second most frequently used arginine codon, CCG, is located on pSymB in *S. meliloti* [10]. Since this gene lies within a region of pSymB that could not be deleted [11], it is assumed to be essential for cell viability. The single copy of the *minCDE* genes, conceivably involved in proper cell division, have also been found in plasmids of *S. meliloti*, *R. leguminosarum* and *R. etli* [4,6,10]. Studies in *S. meliloti* have demonstrated that even though these genes are expressed in free-living cells and within nodules they are nonessential for cell division, since their deletion did not produce the small chromosomeless minicells observed in *E. coli* and *Bacillus subtilis* [12].

A recent bioinformatic study revealed that approximately ten percent of the 897 complete bacterial genomes available in 2009 carry some core genes on extrachromosomal replicons [13]. However, very few of these genes have been functionally characterized and so their real contribution to bacterial metabolism is still an open question.

The complete genome sequence of *R. etli* CFN42 predicts that two putative “housekeeping” genes, *panC* and *panB*, which may be involved in pantothenate biosynthesis, are clustered together on plasmid p42f. Pantothenate is an essential precursor of coenzyme A (CoA), a key molecule in many metabolic reactions including the synthesis of phospholipids, synthesis and degradation of fatty acids, and the operation of the tricarboxylic acid cycle [14]. The *R. etli panC* gene is predicted to encode the sole pantoate- $\beta$ -alanine ligase (PBAL), also known as pantothenate synthetase (PS) (EC 6.3.2.1), present in the *R. etli* genome. The function of this enzyme is the ATP-dependent condensation of D-pantoate with  $\beta$ -alanine to form pantothenate, the last step of the pantothenate biosynthesis pathway. The *panB* gene encodes the putative 3-methyl-2-oxobutanoate hydroxymethyltransferase (MOHMT) (EC 2.1.2.11), also known as ketopantoate hydroxymethyltransferase (KPHMT), the

first enzyme of the pathway, responsible for the formation of  $\alpha$ -ketopantoate by the transfer of a methyl group from 5,10-methylenetetrahydrofolate to  $\alpha$ -ketoisovalerate. The complete genome sequence of *R. etli* CFN42 predicts that a second putative MOHMT enzyme (RHE\_PE00443), similar to the product of *panB*, is encoded on plasmid p42e.

In this work we describe the isolation and use of *panC* and *panB* mutants to analyze the involvement of these plasmid-encoded genes in pantothenate biosynthesis. A survey of the localization of *panCB* genes among members of the *Rhizobiales* with multipartite genomes allowed us to infer a *panCB* phylogeny and to establish the probable chromosomal origin of these plasmid-borne genes. We also report that the *panCB* genes could not totally restore the growth in minimal medium (MM) of a strain cured of plasmid p42f, suggesting that other functions essential for growth in MM are encoded in this plasmid.

## Results

### Functional characterization of plasmid p42f encoded *panCB* genes

The predicted function of the product of *panC* (RHE\_PF00001) annotated as PBAL, is the catalysis of the last step of pantothenate synthesis. This PBAL (298 amino acids) showed 43% identity and 62% similarity over 279 amino acids with the functionally characterized PBAL of *E. coli* K12 (284 amino acids). A search for conserved domains (CD-search) at NCBI-CDD revealed the presence of a typical pantoate-binding site. The *panB* gene (RHE\_PF00002) is located immediately downstream of *panC*. The four nucleotide overlap between the *panC* TGA codon and *panB* ATG codon suggest that these genes might be transcribed as an operon. The *panB* gene encodes a putative MOHMT, the first enzyme of the pantothenate pathway. A BlastP comparison between the functionally characterized MOHMT of *E. coli* K12 (264 amino acids) and the putative MOHMT encoded on plasmid p42f of *R. etli* CFN42 (273 amino acids) showed 37% identity and 56% similarity over a length of 240 amino acids. A CD-search indicated that in the putative MOHMT of *R. etli* CFN42 the magnesium binding and active site domains are conserved. Additionally, Paralog Search (KEGG SSDB) and pathway tools predicted a second probable MOHMT, encoded on plasmid p42e (locus tag RHE\_PE00443). Both proteins are similar in length (273 and 270 aa for the products encoded by *panB* and RHE\_PE00443, respectively). However, a BlastP comparison of these sequences showed only 36% identity and 56% similarity over a tract of 140 amino acids. A CD-search revealed that only 5 of 12 of the invariable residues present in the active site domain are conserved in RHE\_PE00443. The metal binding domain could not be detected by the

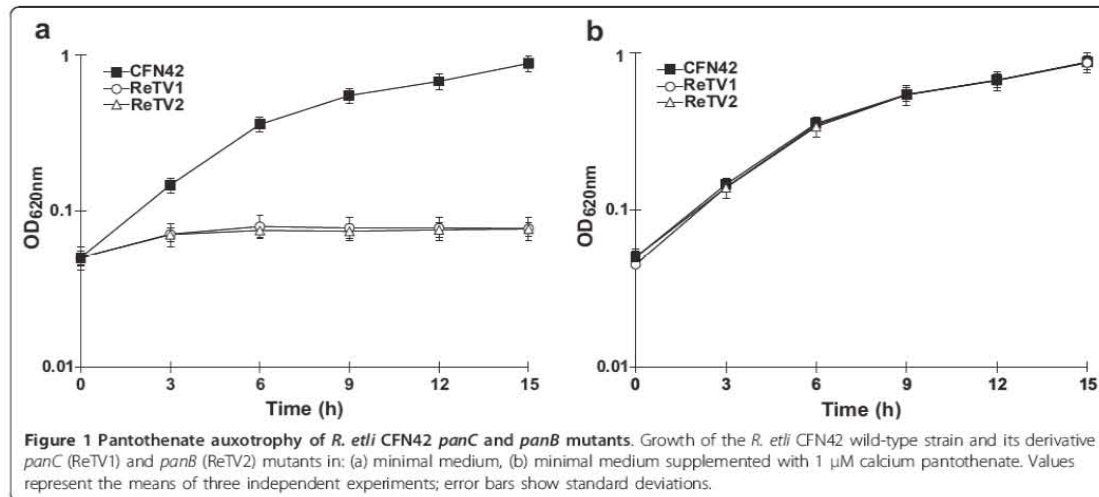
CD-search. To determine whether the *panC* and *panB* genes located on plasmid p42f are required for pantothenate synthesis, mutations in these genes were generated by site-directed vector integration mutagenesis via a single cross-over recombination (see details in Material and Methods and Table 1). Mutants ReTV1 (*panC*<sup>-</sup>) and ReTV2 (*panB*<sup>-</sup>) were unable to grow in minimal medium (MM) lacking calcium pantothenate (Figure 1a). Supplementation of MM with 1 μM calcium pantothenate allowed the *panC* and *panB* mutants to recover their wild-type growth rate (Figure 1b). The pantothenate auxotrophy displayed by the *panB* mutant ReTV2 allowed us to

discard a functional role of the putative MOHMT encoded by RHE\_PE00443 in pantothenate biosynthesis. Moreover, a pBBRMCS3 clone constitutively expressing RHE\_PE00443 (pTV7) was unable to complement the pantothenate auxotrophy of the *panB* mutant (data not shown).

Plasmid pTV4, harboring the *panC* and *panB* genes, as well as plasmids pTV5 and pTV6, carrying only *panC* or *panB* respectively, were introduced into mutant strains ReTV1 and ReTV2 and the growth phenotype of each construction was evaluated in MM. The *panC* mutant ReTV1 complemented with the *panCB* genes

**Table 1 Bacterial strains and plasmid.**

Strain or plasmid	Relevant genotype	Reference or source
<b><i>Rhizobium etli</i></b>		
CFN42	Wild type; Nal <sup>r</sup>	[6]
ReTV1	CFN42 <i>panC</i> ::pTV1; Km <sup>r</sup>	This study
ReTV1-4	CFN42 <i>panC</i> ::pTV1 complemented with pTV4; Tcr Km <sup>r</sup>	This study
ReTV1-5	CFN42 <i>panC</i> ::pTV1 complemented with pTV5; Tcr Km <sup>r</sup>	This study
ReTV2	CFN42 <i>panB</i> ::pTV2; Km <sup>r</sup>	This study
ReTV2 -4	CFN42 <i>panB</i> ::pTV2 complemented with pTV4; Tc <sup>r</sup> Km <sup>r</sup>	This study
ReTV2 -6	CFN42 <i>panB</i> ::pTV2 complemented with pTV6; Tc <sup>r</sup> Km <sup>r</sup>	This study
ReTV2 -7	CFN42 <i>panB</i> ::pTV2 complemented with pTV7; Tc <sup>r</sup> Km <sup>r</sup>	This study
ReTV3	CFN42 <i>argE</i> ::pTV3; Km <sup>r</sup>	This study
CFNX186	CFN42 cured of plasmid p42f; Nal <sup>r</sup>	[18]
CFNX186-4	CFNX186 complemented with pTV4; Tc <sup>r</sup>	This study
CFNX186-24	CFNX186 complemented with pCos24; Tc <sup>r</sup>	[30]
CIAT 652	Wild type; Nal <sup>r</sup>	[38]
CIAT 894	Wild type; Nal <sup>r</sup>	[38]
Kim5	Wild type; Nal <sup>r</sup>	J. Handelsman, University of Wisconsin, MD
IE4771	Wild type; Nal <sup>r</sup>	[15]
<b><i>Escherichia coli</i></b>		
DH5α	Host for recombinant plasmids; Nal <sup>r</sup>	Stratagene
S17-1	C600:RP4-2(Tc:Mu) (Km:Tn7) Donor for conjugation	[39]
<b>Plasmids</b>		
pBC	pBluescript II SK(+) phagemid vector; Cm <sup>r</sup>	Stratagene.
pK18mob	pK18, derivative mob; Km <sup>r</sup>	[29]
pRK7813	Broad-host-range cosmid vector; Mob, IncP, Tc <sup>r</sup>	[40]
pBBRMCS3	Broad-host-range cloning vector; Mob; Tc <sup>r</sup>	[41]
pBC1	pBC harboring a 400-bp <i>Bam</i> H- <i>Xba</i> I PCR fragment of <i>panC</i> ; Cm <sup>r</sup>	This study
pBC2	pBC harboring a 400-bp <i>Bam</i> H- <i>Xba</i> I PCR fragment of <i>panB</i> ; Cm <sup>r</sup>	This study
pTV1	pK18mob harboring a 400-bp <i>Kpn</i> I- <i>Xba</i> I PCR fragment of <i>panC</i> ; Km <sup>r</sup>	This study
pTV2	pK18mob harboring a 400-bp <i>Kpn</i> I- <i>Xba</i> I PCR fragment of <i>panB</i> ; Km <sup>r</sup>	This study
pTV3	pK18mob harboring a 400-bp <i>Kpn</i> I- <i>Xba</i> I PCR fragment of <i>argE</i> ; Km <sup>r</sup>	This study
pTV4	pRK7813 harboring a 3.1 kb <i>Eco</i> RI fragment of pCos24 containing <i>panC</i> and <i>panB</i> ; Tc <sup>r</sup>	This study
pTV5	pBBRMCS3 harboring a 1.2 kb <i>Kpn</i> I- <i>Xba</i> I PCR fragment containing <i>panC</i> ; Tc <sup>r</sup>	This study
pTV6	pBBRMCS3 harboring a 1 kb <i>Kpn</i> I- <i>Xba</i> I PCR fragment containing <i>panB</i> ; Tc <sup>r</sup>	This study
pTV7	pBBRMCS3 harboring a 1 kb <i>Kpn</i> I- <i>Xba</i> I PCR fragment containing RHE_PE00443; Tc <sup>r</sup>	This study
pcos24	20 Kb <i>Eco</i> RI fragment of plasmid p42f cloned in pLAFR1 containing <i>panC</i> , <i>panB</i> , <i>oxyR</i> and <i>katG</i> ; Tc <sup>r</sup>	[30]



**Figure 1** Pantothenate auxotrophy of *R. etli* CFN42 *panC* and *panB* mutants. Growth of the *R. etli* CFN42 wild-type strain and its derivative *panC* (ReTV1) and *panB* (ReTV2) mutants in: (a) minimal medium, (b) minimal medium supplemented with 1 μM calcium pantothenate. Values represent the means of three independent experiments; error bars show standard deviations.

(ReTV1-4) recovered wild type growth in MM. In contrast, when complemented only with *panC* (strain ReTV1-5) no growth occurred in the absence of pantothenate. These results strongly suggest that the *panCB* genes form a single transcriptional unit. As expected, wild type growth of *panB* mutant ReTV2 was recovered by complementation with the *panCB* genes or with the *panB* gene (strains ReTV2-4 and ReTV2-6 respectively).

**The occurrence of *panCB* genes in plasmids is highly conserved among *R. etli* and *R. leguminosarum* strains but not in other members of the *Rhizobiales* with multipartite genomes**

To investigate whether the presence of the *panCB* genes in plasmids is a common characteristic of the *Rhizobiales*, we examined the location of *panCB* genes in 22 members of the *Rhizobiales* having fully sequenced multipartite genomes (Table 2). To date, the genomes of seven *R. etli* strains, in addition to CFN42, have been totally sequenced [15]. However, with the exception of strain CIAT 652, the genomes were released as draft assemblies, precluding *panCB* localization. We experimentally determined the localization of *panCB* genes in the genome of four of these *R. etli* strains (CIAT 894, Kim5, 8C-3, and IE4771) by hybridization of their plasmid profiles with [<sup>32</sup>P]dCTP-labelled *panC* and *panB* genes from CFN42 under high stringency conditions. Both probes produced intense hybridization signals on the same plasmid of each strain, indicating that the *panCB* genes are also plasmid-borne in these *R. etli* strains (Table 2). Coincidentally, in the three *R. leguminosarum* strains with fully sequenced genomes reported in the NCBI database, the *panCB* genes are assigned to

plasmids. In contrast, in other species of *Rhizobiales* with multipartite genomes, the *panCB* genes are always confined to the chromosome, or to chromosome I in those species harboring two chromosomes, with exception of *Agrobacterium tumefaciens* C58 which carries *panCB* on the linear chromosome II and *Methylobacterium nodulans* ORS2060 that carries *panC* on their single chromosome and *panB* on plasmid pMNOD02 (Table 2).

**Phylogenetic analysis of rhizobial *panCB* genes indicates a common origin of chromosomal and plasmid-borne sequences**

Two possible hypotheses were considered to explain the presence of *panCB* genes in plasmids of *R. etli* and *R. leguminosarum* strains: (1) an intragenomic rearrangement of *panCB* genes from chromosome to plasmid, which must have occurred in the last common ancestor of both species; (2) by xenologous gene displacement, that is, a horizontal transfer event in which a gene is displaced by a horizontally transferred ortholog acquired from another lineage [16]. In the latter hypothesis we assume that the presence of these xenolog genes in plasmids conferred a selective advantage that may have eventually led to the loss of the chromosome-located *panCB* genes. To test these hypotheses the phylogeny of 16 rhizobial species inferred from ten orthologous single copy housekeeping genes (*fusA*, *guaA*, *ileS*, *infB*, *recA*, *rplB*, *rpoB*, *rpoC*, *secY* and *valS*) located on primary chromosomes, was compared with the phylogeny of the same rhizobial species inferred from the *panCB* genes located on plasmids and chromosomes. The rationale for this comparison was that if the plasmid-borne

**Table 2 Localization of the *panCB* genes in representative members of the *Rhizobiales* with multipartite genomes**

Strain	Localization of			
	Genome number Chr	Structure of Plasmids	<i>panC</i>	<i>panB</i>
<i>Brucella abortus</i> bv. 1 str. 9-941	2	0	ChrI	ChrI
<i>B. melitensis</i> 16M	2	0	ChrI	ChrI
<i>B. ovis</i> ATCC 25840	2	0	ChrI	ChrI
<i>Sinorhizobium meliloti</i> 1021	1	2	Chr	Chr
<i>S. medicae</i> WSM419	1	3	Chr	Chr
<i>Ochrobactrum anthropi</i> ATCC 49188	2	4	ChrI	ChrI
<i>Agrobacterium radiobacter</i> K84	2	3	ChrI	ChrI
<i>A. vitis</i> S4	2	5	ChrI	ChrI
<i>A. tumefaciens</i> C58	2	2	ChrII	ChrII
<i>Rhizobium etli</i> CFN42	1	6	p42f	p42f
<i>R. etli</i> CIAT 652	1	3	pc	pc
<i>R. etli</i> CIAT 894*	1	4	pd	pd
<i>R. etli</i> Kim5*	1	4	pc <sup>†</sup> /pd <sup>†</sup>	pc <sup>†</sup> /pd <sup>†</sup>
<i>R. etli</i> IE4771*	1	4	pd	pd
<i>R. etli</i> 8C-3*	1	3	pc	pc
<i>R. leguminosarum</i> bv. viciae 3841	1	6	pRL12	pRL12
<i>R. leguminosarum</i> WSM1325	1	5	pR132501	pR132501
<i>R. leguminosarum</i> WSM2304	1	4	pRLG201	pRLG201
<i>Rhizobium</i> sp. NGR234	1	2	Chr	Chr
<i>Mesorhizobium loti</i> MAFF303099	1	2	Chr	Chr
<i>M. sp.</i> BNC1	1	3	Chr	Chr
<i>Methylobacterium extorquens</i> AM1	1	4	Chr	Chr
<i>M. radiotolerans</i> JCM2831	1	8	Chr	Chr
<i>M. nodulans</i> ORS2060	1	7	Chr	pMNOD2
<i>Bradyrhizobium</i> sp. BTAi1	1	1	Chr	Chr
<i>Nitrobacter hamburgensis</i> X14	1	3	Chr	Chr
<i>Xantobacter autotrophicus</i> Ry2	1	1	Chr	Chr

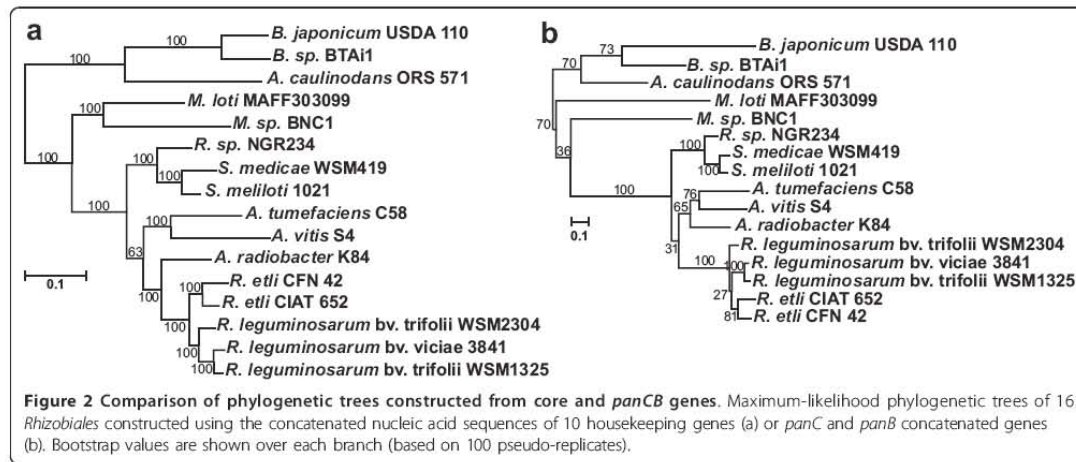
Abbreviations are as follows: Chr, chromosome of those *Rhizobiales* with one chromosome; Chr I and Chr II, chromosome I and chromosome II respectively in those *Rhizobiales* harboring two chromosomes; p, plasmid. \**Rhizobium* species in which localization of *panCB* genes was done by Southern blot hybridization of plasmid profiles. †Plasmids with very similar electrophoretic mobility gave as result ambiguous plasmid localization of *panC* and *panB* homologous sequences.

*panCB* phylogeny agrees with the current phylogeny of the *Rhizobiales*, inferred from the housekeeping genes, it would support the hypothesis of intragenomic transfer of the *panCB* genes. On the other hand, if both phylogenies are incongruent, it would favor the hypothesis of horizontal transfer of the *panCB* genes. Concatenated nucleic acids multiple alignments were used to infer both phylogenies with the maximum likelihood method described in materials and methods. The resulting phylogenetic trees are shown in Figure 2. The housekeeping genes inferred tree (Figure 2a) was consistent with the recently reported phylogeny of 19 *Rhizobiales* performed on a data set of 507 homologous proteins from the primary chromosome [17]. Both trees are in close agreement with the phylogeny inferred from the *panCB* genes (Figure 2b). Thus the phylogeny of *R. etli* and *R. leguminosarum* inferred from plasmid-encoded *panCB* genes is consistent with the phylogeny deduced

from their housekeeping genes supporting the hypothesis of a chromosomal origin for the plasmid-encoded *panCB* genes.

#### The *panCB* genes do not fully complement the growth deficiency of a *R. etli* CFN42 p42f cured derivative in MM

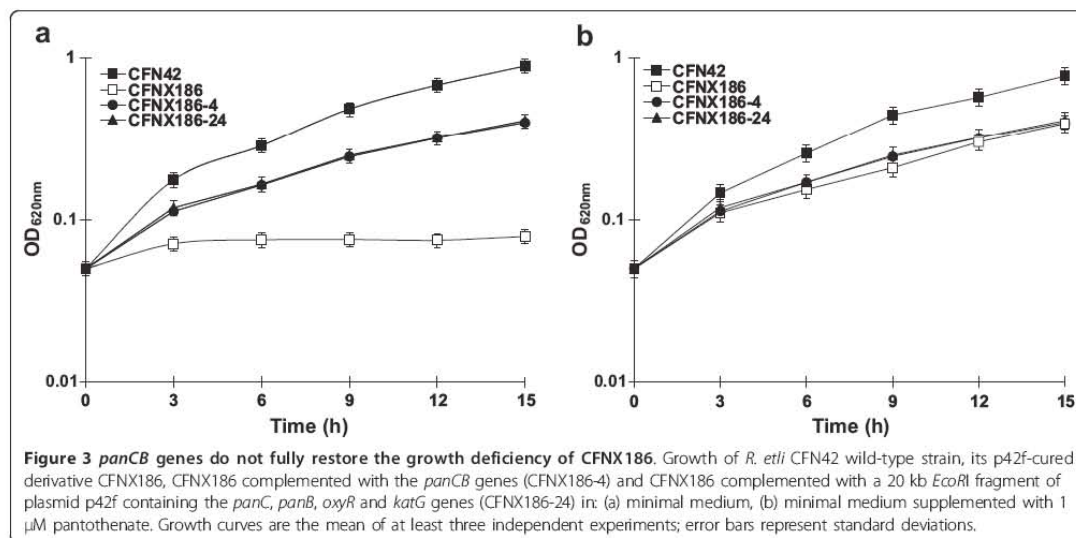
It was reported previously that *R. etli* CFNX186, a p42f-cured derivative of *R. etli* CFN42, is unable to grow in MM [18]. To assess if the growth deficiency of strain CFNX186 in MM was due to the absence of the *panC* and *panB* genes, plasmid pTV4 (*panCB*) was introduced into strain CFNX186. The growth of the transconjugant (CFNX186-4) after 15 hours of culture in MM was only 50% that of the WT strain grown under the same conditions (Figure 3a). The growth of CFNX186-4 did not improve even after 72 h in culture (data not shown). Interestingly, strain CFNX186-4 had the same growth rate as strain CFNX186 cultured in MM supplemented



with 1  $\mu$ M calcium pantothenate (Figure 3b). This shows that the growth deficiency of CFNX186 is only partly due to the absence of the *panCB* genes and indicates that other functions encoded in plasmid p42f are required for growth in MM.

Previous studies have demonstrated that the *katG* gene, which encodes the sole catalase-peroxidase expressed in free-living growth conditions, is located on plasmid p42f of *R. etli* CFN42. These studies also revealed that the growth rate of a *katG* mutant in MM was significantly reduced in comparison with that of the wild-type parental strain [19]. On plasmid p42f *katG*, as well as its putative transcriptional regulator protein

encoded by *oxyR*, are located 80 bp downstream of the *panCB* genes. We speculated that introduction of the *panCB* genes together with the *katG* and *oxyR* genes might improve the growth of CFNX186 in MM. To test this hypothesis, we used pCos24, which contains a 20 kb fragment of p42f carrying *panCB*, *katG* and *oxyR* (see Material and Methods). pCos24 was introduced into CFNX186 and the resulting transconjugant (CFNX186-24) grown in MM. Figure 3 shows that after 15 hours of culture there was no significant difference between the growth rate of CFNX186 complemented only with *panCB* (CFNX186-4), and CFNX186 complemented with cosmid pCos24 (CFNX186-24). Furthermore, the



growth of CFNX186-24 did not increase even after 72 h of culture (data not shown) indicating that *katG* and *oxyR* did not improve the growth rate of *panCB* complemented CFNX186 in MM. We also tested the possibility that arginine might improve the growth of strain CFNX186-24 due to the presence of a putative *N*-acetylornithinase (EC 3.5.1.16) encoded in the plasmid p42f. In the *Enterobacteriaceae* this enzyme catalyzes the conversion of *N*-acetylornithine to ornithine, a key step in the arginine biosynthesis pathway [20]. However, the growth deficiency of strain CFNX186-24 in MM was not corrected by the addition of 1, 5, 10 or 15 mM arginine (data not shown). Furthermore, we constructed an *argE* mutant strain (ReTV3, Table 1) that was able to grow in MM without exogenous arginine at the same rate as parental strain CFN42 (data not shown), confirming that this gene is not essential for arginine synthesis.

## Discussion

Seminal studies on the phenotypic characterisation of plasmid-cured strains of *R. leguminosarum* and *R. etli* revealed that the absence of several plasmids cause a growth deficiency in rich and minimal medium [18,21]. These findings suggested that undefined metabolic traits are present on rhizobial plasmids. The bioinformatic analysis of 897 bacterial genomes performed by Harrison *et al* [13] revealed the presence of extrachromosomal core genes in 82 genomes mainly belonging to the Proteobacteria. In contrast with these *in silico* data, there is little experimental information on the contribution of these core genes to bacterial metabolism or cellular process. The few genes that have been functionally characterized encode redundant functions and are totally dispensable for the cell [7-9,12]. Our study provides experimental evidence that the enzymes MOHMT (EC 2.1.2.11) and PBAL (EC 6.3.2.1) encoded on plasmid p42f are indispensable for the synthesis of pantothenate. Moreover, our results showed that the cluster of *panCB*, *katG* and *oxyR* genes was insufficient to restore full growth capacity to the p42f cured derivative CFNX186, implying that in addition to pantothenate synthesis, there are more functions encoded on plasmid p42f required for growth in MM. Obvious candidates for these functions could not be identified *a priori* among the 567 proteins encoded in p42f even though their predicted functions were recently updated with KAAS (KEGG Automatic Annotation Server and Pathway Reconstruction Server). We discarded arginine limitation as the cause for the growth deficiency of strain CFNX186-24. The arginine prototrophy displayed by a mutation in the p42f encoded *argE* suggests that in *R. etli* the conversion of *N*-acetylornithine to ornithine is catalyzed by the chromosome-encoded ArgJ, an ornithine acetyltransferase (OATase, EC 2.3.1.35), which transfers the acetyl group of

*N*-acetylornithine to glutamate to produce ornithine and *N*-acetylglutamate. Functional OATases have been found in the majority of bacteria [20].

Also, we have demonstrated that plasmid-localization of *panCB* in *R. etli* CFN42 is not unique to this strain. A screening of the location of *panCB* genes among members of the *Rhizobiales*, showed that the occurrence of these genes in plasmids is a highly conserved trait among *R. etli* and *R. leguminosarum* strains. Furthermore, the synteny of the *panCB*, *oxyR*, *katG* genes in *R. etli* CFN42 is conserved in *R. etli* CIAT652 and in *R. leguminosarum* strains 3841, WSM1325 and WSM2304. In contrast, genomes of *Rhizobium sp.*, *Sinorhizobium*, *Bradyrhizobium* and *Mesorhizobium* species carried chromosomal *panCB* genes. Only in *A. tumefaciens* C58 the *panCB* genes are localized in the linear chromosome, whereas in all other *Rhizobiales* harboring secondary chromosomes the *panCB* genes were located in chromosome I. A bioinformatic analysis with MicrobesOnline operon predictions [22] indicates that *panCB* genes are organized as possible operons in most of the *Rhizobiales* examined in this work: all these predicted operons conserve the four nucleotide overlap between the *panC* TGA codon and the *panB* ATG codon observed in *R. etli* CFN42 (data not shown). In the genomes of *Bradyrhizobium sp.* BTAi1, *Nitrobacter hamburgensis* X14, *Methylobacterium extorquens* AM1, *Methylobacterium radiotolerans* JCM2831 and *Xantobacter autotrophicus* Ry2, *panC* and *panB* are encoded in separate chromosomal loci, whereas in *Methylobacterium nodulans* ORS2060 *panC* is located in the chromosome and *panB* in plasmid pMNOD02.

The *Rhizobiales* phylogeny inferred from concatenated *panC* and *panB* genes was consistent with the phylogeny deduced from 10 concatenated housekeeping genes. The low bootstrap values obtained for some nodes of the *panCB* phylogeny might be due to the small number of informative characters in the alignments of only two genes (1 977 nucleotides). This is consistent with previous reports that state that trees from longer alignments obtained by the concatenation of genes encoding multiple-protein families have higher bootstrap support than trees inferred from genes encoding single proteins [23]. The phylogenetic relationships among *Rhizobium* species carrying *panCB* genes in plasmids with their closest relatives, *Agrobacterium* and *Sinorhizobium* species, harboring *panCB* genes in the chromosome was also observed in neighbor-joining trees inferred from single *panC* and *panB* genes (data not shown). These data agree with the hypothesis that plasmid-encoded *panCB* genes are orthologs of the *panCB* genes located in chromosome. From these results, we propose that the presence of the *panCB* genes in plasmids in *R. etli* and *R. leguminosarum* species may be due to an intragenomic transfer event from chromosome to plasmid. The mechanism leading to the

transfer of core genes from chromosome to plasmids could involve cointegration and excision events between the replicons, similar to rearrangements that have been visualized in *S. meliloti* [24]. The translocation of genes from chromosome to plasmids may be part of the complex evolution of multipartite genomes. A study based on the analysis of clusters of syntenic genes shared among plasmids and secondary chromosomes of bacteria with multipartite genomes suggested that secondary chromosomes may have originated from an ancestral plasmid to which genes had been transferred from a primary chromosome [17].

Our pioneering work on plasmid-encoded functions in *R. etli* CFN42 established that a functional relationship among different replicons is required for symbiotic and free-living functions [18,25]. More recently, a functional connectivity among most of the proteins encoded in the replicons of *R. etli* CFN42 was predicted *in silico* [6]. Our results demonstrated that the putative MOHMT encoded by RHE\_PE00443 is not functional under the conditions studied and provides evidence of functional cooperation between p42f and chromosomally encoded proteins for pantothenate biosynthesis.

## Conclusions

Our study shows that the presence of the core *panCB* genes in a plasmid is a characteristic conserved in *R. etli* and *R. leguminosarum* strains but not in other *Rhizobiales*. The phylogenetic approach used in this study suggests that the unusual presence of *panCB* in plasmids may be due to an intragenomic transfer event from chromosome to plasmid rather than a xenologous gene displacement. Using *R. etli* CFN42 as a model, we showed that the plasmid-encoded core *panCB* genes were indispensable for the synthesis of pantothenate. The *panCB* genes could not totally restore growth of a strain cured of plasmid p42f in minimal medium, suggesting that other functions essential for growth in this medium are encoded in this plasmid. Our results support the hypothesis of functional cooperation among different replicons for basic cellular functions in multipartite rhizobial genomes.

## Methods

### Bacterial strains, media and growth conditions

The bacterial strains and plasmids used are listed in Table 1. *Rhizobium* strains were grown at 30°C in three different media: a) PY rich medium [26], b) Minimal medium (MM) [27] and c) Minimal medium plus 1 µM calcium pantothenate (MMP). MM was prepared as follows: a solution containing 10 mM succinate as carbon source, 10 mM NH<sub>4</sub>Cl as nitrogen source, 1.26 mM K<sub>2</sub>HPO<sub>4</sub>, 0.83 mM MgSO<sub>4</sub>, was adjusted to pH 6.8 and sterilized. After sterilization the following components

were added to the final concentration indicated: 0.0184 mM FeCl<sub>3</sub> 6H<sub>2</sub>O (filter sterilized), 1.49 mM CaCl<sub>2</sub> 2H<sub>2</sub>O (autoclaved separately), 10 µg ml<sup>-1</sup> biotin and 10 µg ml<sup>-1</sup> thiamine (both filter sterilized). MMP contains the same components plus 1 µM calcium pantothenate. To determine growth rates on MM or MMP, *Rhizobium* strains were grown to saturation in PY medium, the cells were harvested by centrifugation, washed twice with sterile deionized water and diluted to an initial optical density of 0.05 at 600 nm (OD<sub>600</sub>) when added to 30 ml of MM. These cultures were grown for 24 h in 125 ml Erlenmeyer flasks to deplete any endogenous pantothenate. Cells were then harvested and washed as described above and added to fresh MM or MMP in the same manner as for the first inoculation and cultured for 15 hours. Bacterial growth was quantified by measuring optical density at 600 nm (OD<sub>600</sub>) every 3 hours. Antibiotics were used at the following concentrations (in µg ml<sup>-1</sup>): chloramphenicol (Cm), 30; tetracycline (Tc), 10; kanamycin (Km), 30; gentamicin (Gm), 30; spectinomycin (Sp), 100; nalidixic acid (Nal), 20. *E. coli* transformants harboring recombinant plasmids (β-galactosidase-positive) were identified by growth on LB plates with 30 µg ml<sup>-1</sup> 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal).

### DNA manipulations

Standard techniques described by Sambrook *et al.* [28] were used for plasmid and total DNA isolation, restriction, cloning, transformations, and agarose gel electrophoresis. Plasmid mobilization from *E. coli* to *Rhizobium* was done by conjugation performed on PY plates at 30°C by using overnight cultures grown to stationary phase. Donors (*E. coli* strain S17-1) and recipients (*R. etli* CFN42 wild type and mutant strains) were mixed at a 1:2 ratio, and suitable markers were used for transconjugant selection.

### Mutagenesis of the *panC* and *panB* genes and genetic complementation of mutant strains

Mutants were generated by site-directed vector integration mutagenesis. Internal 400 bp DNA fragments of *panC* and *panB* were amplified by PCR with primers A and B; C and D, respectively (Table 3). PCR fragments of *panC* and *panB* were cloned in vector pBC as 400 bp *Bam*HI-*Xba*I fragments, generating pBC1 and pBC2 respectively, and then subcloned as *Kpn*I-*Xba*I fragments into suicide vector pK18mob [29] to form plasmids pTV1 and pTV2, respectively. These plasmids were mobilized into *R. etli* CFN42 by conjugation and single crossover recombinants selected on PY plates containing Km and Nal. The disruption of the *panC* and *panB* genes was confirmed by Southern blot analysis using a 400-bp PCR internal fragment of each gene as a probe (data not shown). The resultant mutants were named ReTV1 and ReTV2 respectively. To complement the phenotype of

**Table 3 Primers used in this work.**

Primer	Sequence (5'- 3')
A	GCGGATCCGAAGACCTCAGCAAATACCCGC
B	CGGAGGATCCGCGCCACGACGACCGACCCGCC
C	CGGGTCTAGACTCGGCATGGTCTATGGCA
D	GACGTCTAGAGCTTGAATCGTTGAAGAGCCC
E	TGATGGTACCTTGACGGATGGGGCAATAGCCG
F	GGCGCTCTAGAATCCGATGGCGCTCATTTCC
G	GCGGGCGGTACCAGCCGGGAAAGGGAGTG
H	AAGCGTCTAGAGCCTTCGTCTTACGGCCG
I	CGTCAAGGTACCATCCCTTCTGACCGCTG
J	CCCCCTCTAGACGCTGGGAGAAGGGACTC
K	GCTGTGGTACCAGCCGTCGCGGCACTCGCG
L	ACCTTCTAGATGCCGACCTGGAGGGAGG

The restriction sites are indicated in bold.

the *panC* and *panB* mutants, plasmids pTV4, pTV5, pTV6 and pTV7 were constructed as follows: a 3.1 kb *EcoRI* fragment from cosmid vector pCos24, isolated from a genomic library of *R. etli* CFN42 [30] and containing the *panC* and *panB* genes, was subcloned in broad-host-range vector pRK7813, generating plasmid pTV4. To construct plasmid pTV5, a 1.2 kb fragment containing only *panC* (894 bp) was amplified by PCR with primers E and F and cloned in the *KpnI-XbaI* sites in the broad-host-range vector pBBRMSC3 so that the gene would be constitutively expressed from the vector's *lacZ* promoter. Primers G and H (Table 3) were used to amplify a 1 kb PCR fragment containing only the *panB* gene (822 bp). This DNA fragment was cloned in plasmid pBBRMSC3 in the *KpnI-XbaI* restriction sites, generating plasmid pTV6. Plasmid pTV7 contains the second *panB* gene (RHE\_PE00443), encoded on *R. etli* plasmid p42e, this gene was amplified with primers I and J. The resultant 1 kb PCR fragment was cloned in the *KpnI-XbaI* sites of plasmid pBBRMSC3. To complement the growth deficiency of strain CFNX186, a derivative of *R. etli* CFN42 cured of plasmid p42f, plasmid pTV4 and cosmid vector pCos24 were introduced by conjugation. The complemented strains obtained were named CFNX186-4 and CFNX186-24 respectively. The *argE* gene was disrupted as described above. Briefly, an internal 400 bp PCR fragment of *argE* amplified with primers K and L was cloned directly in pK18mob using the *KpnI* and *XbaI* sites to give pTV3 (Table 1). This recombinant suicide plasmid was mobilized into *R. etli* CFN42 and the resultant mutant named ReTV3 (Table 1).

#### Filter blots hybridization and plasmid visualization

For Southern-type hybridizations, genomic DNA was digested with appropriate restriction enzymes, electrophoresed in 1% (w/v) agarose gels, blotted onto nylon membranes, and hybridized under stringent conditions,

as previously reported by [31], using Rapid-hyb buffer. To use the *panC* and *panB* genes as probes, both genes were amplified by PCR, separated on a 1% agarose and purified by a PCR purification kit (QIAquick). They were labeled with [ $\alpha$ -<sup>32</sup>P]dCTP using a Rediprime DNA labeling system. Plasmid profiles were visualized by the Eckhardt technique as modified by [21], and hybridized in a similar manner.

#### Identification of orthologous proteins, multiple sequence alignments and phylogenetic analysis

All genomic sequences analyzed in this study were obtained from the Integrated Microbial Genomes System of the DOE Joint Genome Institute <http://img.jgi.doe.gov/>. We obtained protein and gene sequences of *panB*, *panC* and 10 chromosomal housekeeping genes (*fusA*, *guaA*, *ileS*, *infB*, *recA*, *rplB*, *rpoB*, *rpoC*, *secY* and *valS*) from 16 rhizobial species. Accession numbers for these sequences and the species list are shown in Table S1 (see Additional file 1). An orthologous data set for each gene was constructed using Blast [32] and the bidirectional best hit method applying the criteria reported by Poggio *et al* [33]. Multiple alignments of putative orthologous proteins were performed using the MUSCLE program [34] with default settings. After removing poorly conserved regions two concatenated protein alignments were obtained, one for the 10 chromosomal housekeeping genes (8469 amino acids) and the other for *panB* and *panC* (659 amino acids). Both concatenated protein multiple alignments were used to generate nucleic acids multiple alignments of their respective genes with the Tralign program of the EMBOSS suit <http://emboss.sourceforge.net/>. Nucleic acids multiple alignments were used to obtain two phylogenies with the maximum likelihood method implemented in PHYML [35] with HKY as substitution model [36]. The phylogenetic reconstruction was carried out with a nonparametric bootstrap analysis of 100 replicates for each alignment. TreeDyn program [37] was used to visualize and edit both phylogenies.

#### Additional material

**Additional file 1: Table S1.** Rhizobial species list and accession numbers of housekeeping and *panCB* genes used for phylogenetic analysis.

#### List of abbreviations

PBAL: pantoate- $\beta$ -alanine ligase; MOHMT: 3-methyl-2-oxobutanoate hydroxymethyltransferase; MM: minimal medium.

#### Acknowledgements

We are grateful to Laura Cervantes and Javier Rivera for their excellent technical assistance. We acknowledge Michael F. Dunn for critically reviewing the manuscript. This work was supported by DGAPA-PAPIIT-UNAM grant IN200309-2. Tomás Villaseñor was supported by a Ph. D. scholarship



(204725) from CONACYT México during his Ph. D. studies at UNAM, Programa de Doctorado en Ciencias Biomédicas.

#### Author details

<sup>1</sup>Programa de Ingeniería Genómica, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Apdo. Postal 565-A. Cuernavaca, Morelos, México. <sup>2</sup>Programa de Genómica Evolutiva, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Apdo. Postal 565-A. Cuernavaca, Morelos, México.

#### Authors' contributions

TV designed and constructed all the mutants, did all the experiments for genetic complementation of the mutants, performed growth experiments and Southern blot hybridizations and helped to draft the manuscript. SB provided intellectual guidance and contributed to writing the manuscript. AD performed Eckhardt gels and Southern blot to localize *panCB* homologues in plasmids of *R. etli* strains and assisted in DNA cloning. LL carried out the phylogenetic analysis and the discussion of results. DR participated in the experimental design and in the discussion of results. AGS conceived the study, supervised the experimental work and wrote the manuscript. All authors read and approved the final manuscript.

Received: 17 November 2010 Accepted: 5 April 2011

Published: 5 April 2011

#### References

1. Jumas-Bilak E, Michaux-Charachon S, Bourg G, Ramuz M, Allardet-Servent A: Unconventional genomic organization in the alpha subgroup of the Proteobacteria. *J Bacteriol* 1998, **180**:2749-2755.
2. MacLean AM, Finan TM, Sadowsky MJ: Genomes of the symbiotic nitrogen-fixing bacteria of legumes. *Plant Physiol* 2007, **144**:615-622.
3. Romero D, Brom S: The symbiotic plasmids of the *Rhizobiaceae*. In *Plasmid biology*. Edited by: Phillips G, Funnell BE. Washington, D.C: American Society for Microbiology; 2004:271-290.
4. Young JP, Crossman LC, Johnston AW, Thomson NR, Ghazoui ZF, Hull KH, Wexler M, Curson AR, Todd JD, Poole PS, Mauchline TH, East AK, Quail MA, Churcher C, Arrowsmith C, Cherevach I, Chillingworth T, Clarke K, Cronin A, Davis P, Fraser A, Hance Z, Hauser H, Jagels K, Moule S, Murgall K, Norbertczak H, Rabinowitsch E, Sanders M, Simmonds M, Whitehead S, Parkhill J: The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol* 2006, **7**:R34.
5. Crossman LC, Castillo-Ramírez S, McAnnula C, Lozano L, Vernikos GS, Acosta JL, Ghazoui ZF, Hernández-González I, Meakin G, Walker AW, Hynes MF, Young JPW, Downie JA, Romero D, Johnston AWB, Dávila G, Parkhill J, González V: A common genomic framework for a diverse assembly of plasmids in the symbiotic nitrogen fixing bacteria. *PLoS ONE* 2007, **3**:e2567.
6. González V, Santamaría RI, Bustos P, Hernández-González I, Medrano-Soto A, Moreno-Hagelsieb G, Janga SC, Ramírez MA, Jiménez-Jacinto V, Collado-Vides J, Dávila G: The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc Natl Acad Sci USA* 2006, **103**:3834-3839.
7. Bittner AN, Foltz A, Oke V: Only one of five *groEL* genes is required for viability and successful symbiosis in *Sinorhizobium meliloti*. *J Bacteriol* 2007, **189**:1884-1889.
8. Rodríguez-Quiriones F, Maguire M, Wallington EJ, Gould PS, Yerko V, Downie JA, Lund PA: Two of the three *groEL* homologues in *Rhizobium leguminosarum* are dispensable for normal growth. *Arch Microbiol* 2005, **183**:253-265.
9. Yost CK, Rath AM, Noel TC, Hynes MF: Characterization of genes involved in erythritol catabolism in *Rhizobium leguminosarum* bv. *viciae*. *Microbiology* 2006, **152**:2061-2074.
10. Finan TM, Weidner S, Wong K, Buhmester J, Chain P, Vorhölter FJ, Hernandez-Lucas I, Becker A, Cowie A, Gouzy J, Golding B, Pühler A: The complete sequence of the 1,683 kb pSymb megaplasmid from the N2-fixing endosymbiont *Sinorhizobium meliloti*. *Proc Natl Acad Sci USA* 2001, **98**:9889-9894.
11. Charles TC, Finan TM: Analysis of a 1600-Kilobase *Rhizobium meliloti* megaplasmid using defined deletions generated *in vivo*. *Genetics* 1991, **127**:5-20.
12. Cheng J, Sibley CD, Zaheer R, Finan TM: A *Sinorhizobium meliloti* minE mutant has an altered morphology and exhibits defects in legume symbiosis. *Microbiology* 2007, **153**:375-387.
13. Harrison PW, Lower RPJ, Kim NKD, Young JPW: Introducing the bacterial "chromid": not a chromosome, not a plasmid. *Trends Microbiol* 2010, **18**:141-148.
14. Jackowski S: Biosynthesis of pantothenic acid and coenzyme A. In *Escherichia coli and Salmonella: cellular and molecular biology*. Edited by: Neidhardt FC, Curtiss R III, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umberger HE. Washington, DC: ASM Press; 1996:1310-1324.
15. González V, Acosta JL, Santamaría RI, Bustos P, Fernández JL, Hernández IL, Díaz R, Flores M, Palacios R, Mora J, Dávila G: Conserved symbiotic plasmid DNA sequences in the multireplicon pangenomic structure of *Rhizobium etli*. *Appl Environ Microbiol* 2010, **76**:1604-1614.
16. Koonin EV, Makarova KS, Aravind L: Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol* 2003, **55**:709-742.
17. Slater SC, Goldman BS, Goodner B, Setubal JC, Farrand SK, Nester EW, Burr TJ, Banta L, Dickerman AW, Paulsen I, Otten L, Suen G, Welch R, Almeida NF, Arnold F, Burton OT, Du Z, Swing A, Godoy E, Heisel S, Houmliel KL, Jhaveri J, Lu J, Miller NM, Norton S, Chen Q, Phoolcharoen W, Ohlin V, Ondrusek D, Pride N, Stricklin SL, Sun J, Wheeler C, Wilson L, Zhu H, Wood DW: Genome sequences of three *Agrobacterium biovars* help elucidate the evolution of multichromosome genomes in bacteria. *J Bacteriol* 2009, **191**:2501-2511.
18. Brom S, Garcia-de los Santos A, Stepkowski T, Flores M, Dávila G, Romero D, Palacios R: Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. *J Bacteriol* 1992, **174**:5183-5189.
19. Vargas MC, Encarnación S, Davalos A, Reyes-Perez A, Mora Y, Garcia-de los Santos A, Brom S, Mora J: Only one catalase, KatG, is detectable in *Rhizobium etli*, and is encoded along with the regulator OxyR on a plasmid replicon. *Microbiology* 2003, **149**:1165-1176.
20. Xu Y, Labeledan B, Glansdorff N: Surprising arginine biosynthesis: a reappraisal of the enzymology and evolution of the pathway in microorganisms. *Microbiol Mol Biol Rev* 2007, **71**:36-47.
21. Hynes MF, McGregor NF: Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. *Mol Microbiol* 1990, **4**:567-574.
22. Dehal PS, Joachimiak MP, Price MN, Bates JT, Baumohl JK, Chivian D, Friedland GD, Huang KH, Keller K, Novichkov PS, Dubchak IL, Alm EJ, Adam PA: MicrobesOnline: an integrated portal for comparative and functional genomics. *Nucleic Acids Res* 2010, **38**:396-400.
23. Williams KP, Sobral BW, Dickerman AW: A robust species tree for the *Alphaproteobacteria*. *J Bacteriol* 2007, **189**:4578-4586.
24. Guo X, Flores M, Mavingui P, Fuentes SI, Hernández G, Dávila G, Palacios R: Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. *Genome Res* 2003, **13**:1810-1817.
25. García-de los Santos A, Brom S: Characterization of two plasmid-borne *lpsB* loci of *Rhizobium etli* required for lipopolysaccharide synthesis and for optimal interactions with plants. *Mol Plant Microbe Interact* 1997, **10**:891-902.
26. Noel KD, Sánchez A, Fernández L, Leemans J, Cevallos MA: *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. *J Bacteriol* 1984, **158**:148-155.
27. Encarnación S, Willms K, Mora J: Fermentative and aerobic metabolism in *Rhizobium etli*. *J Bacteriol* 1995, **177**:3058-3066.
28. Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual* Cold Spring Harbor, Cold Spring Harbor Press; 1989.
29. Schafer A, Tauch A, Jager W, Kalinowski J, Thierbach G, Pühler A: Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* 1994, **145**:69-73.
30. García-de los Santos A, López E, Cubillas CA, Noel KD, Brom S, Romero D: Requirement of a plasmid-encoded catalase for survival of *Rhizobium etli* CFN42 in a polyphenol-rich environment. *Appl Environ Microbiol* 2008, **74**:2398-2403.
31. Flores M, González V, Brom S, Martínez E, Piñero D, Romero D, Dávila G, Palacios R: Reiterated DNA sequences in *Rhizobium* and *Agrobacterium* spp. *J Bacteriol* 1987, **169**:5782-5788.

32. Altschul S, Madden T, Schaffer A, Zhang J, Zhang Z, Millar W, Lipman D: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997, **25**:3389-3402.
33. Poggio S, Abreu-Goodger C, Fabela S, Osorio A, Dreyfus G, Vinuesa P, Camarena L: A complete set of flagellar genes acquired by horizontal transfer coexists with the endogenous flagellar system in *Rhodobacter sphaeroides*. *J Bacteriol* 2007, **189**:3208-3216.
34. Edgar RC: MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 2004, **5**:113.
35. Guindon S, Gascuel O: Efficient biased estimation of evolutionary distances when substitution rates vary across sites. *Mol Biol Evol* 2002, **4**:534-543.
36. Hasegawa M, Kishino H, Yano T: Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 1985, **22**:160-174.
37. Chevenet F, Brun C, Banuls AL, Jacq B, Chisten R: TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 2006, **7**:439.
38. Piñero D, Martínez E, Selander RK: Genetic diversity and relationships among isolates of *Rhizobium leguminosarum* bicvar *phaseoli*. *Appl Environ Microbiol* 1988, **54**:2825-2832.
39. Simon R: High frequency mobilization of gram-negative bacterial replicons by the *in vitro* constructed Tn5-mob transposon. *Mol Gen Genet* 1984, **196**:413-420.
40. Jones JDG, Gutterston N: An efficient mobilizable cosmid vector, pRK7813, and its use in a rapid method for marker exchange in *Pseudomonas fluorescens* strain HV37a. *Gene* 1987, **61**:299-306.
41. Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, Roop RM II, Peterson KM: Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 1995, **166**:175-176.

doi:10.1186/1471-2180-11-66

Cite this article as: Villaseñor *et al.*: Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*. *BMC Microbiology* 2011 **11**:66.

Submit your next manuscript to BioMed Central  
and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)



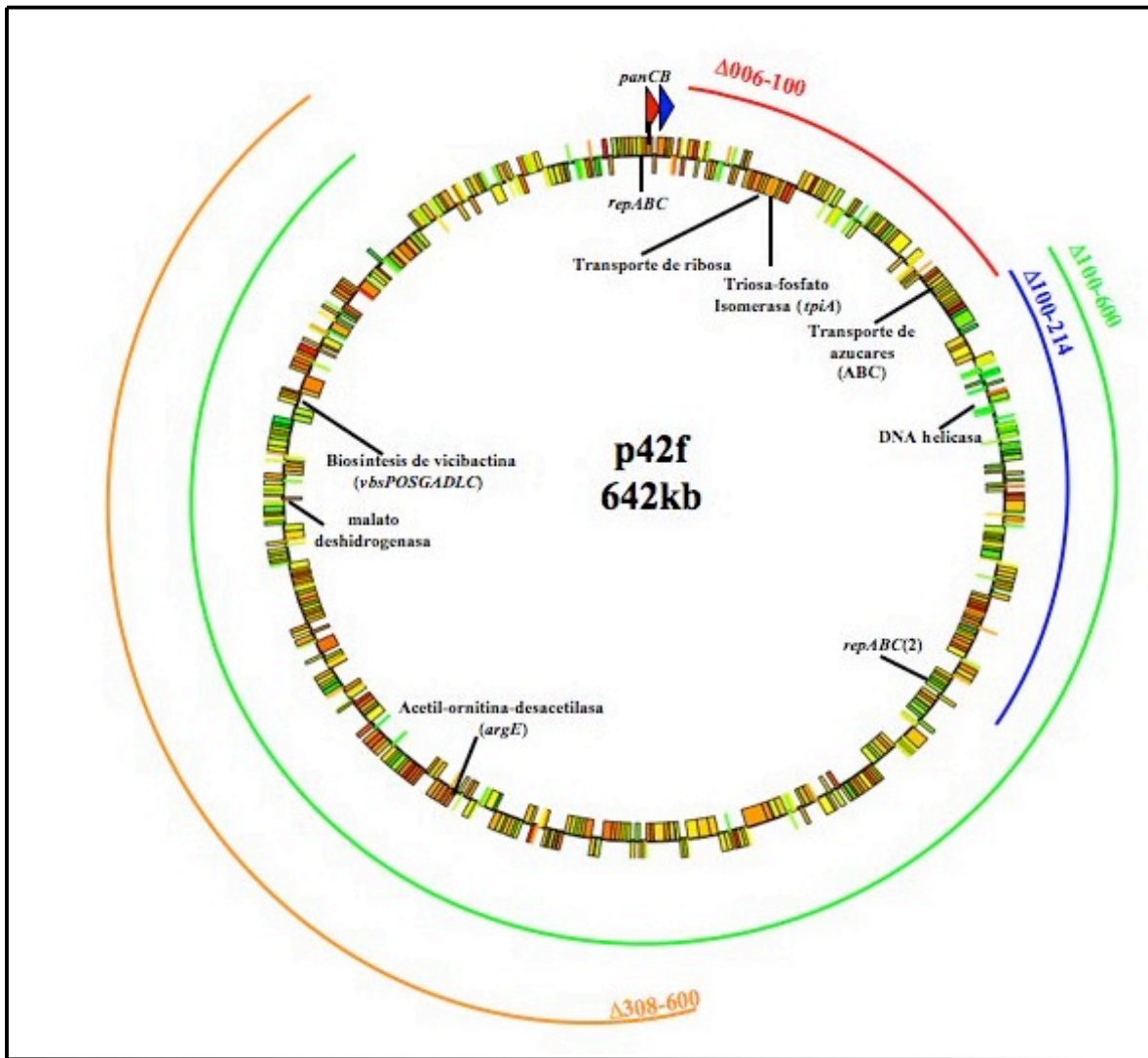
## Resultados Adicionales.

### Diseño, generación y evaluación de deleciones programadas en el plásmido p42f.

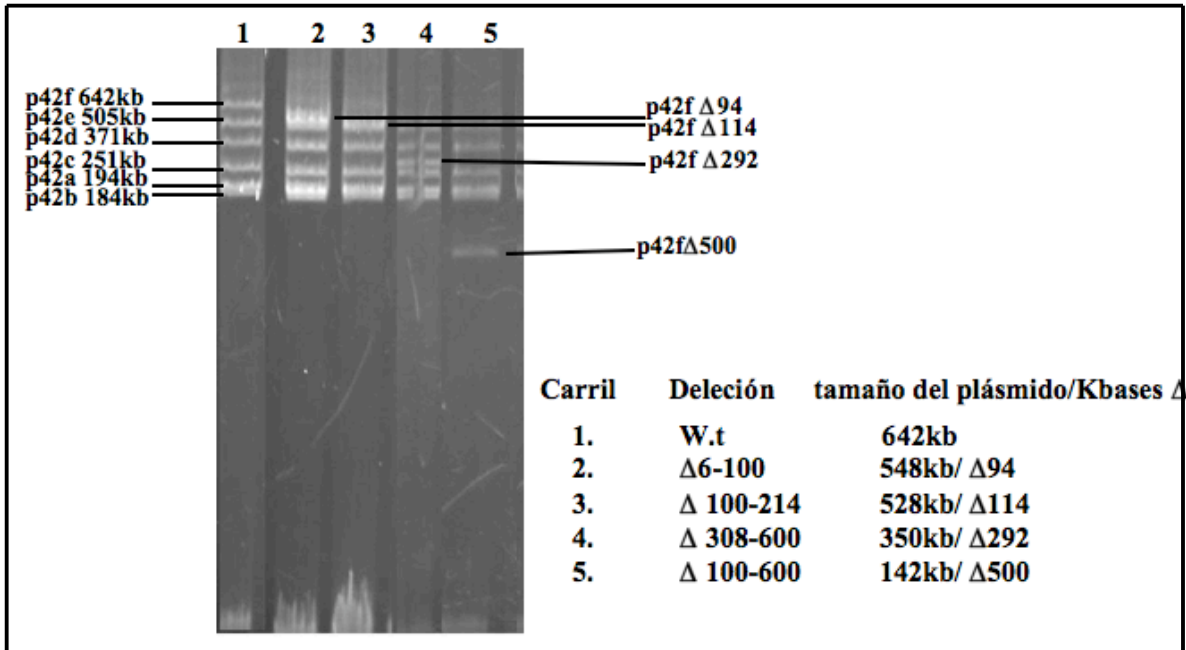
Para explicar la incapacidad de las cepas CFNX186-4 y CFNX186-24 para recuperar total crecimiento en MM proponemos como hipótesis que además de las enzimas para síntesis de pantotenato, otras enzimas codificadas en el plásmido p42f son necesarias para crecimiento en MM. Estas enzimas no pudieron ser identificadas a partir de la re- anotación del genoma con KAAS. Para comprobar la hipótesis realizamos una serie de deleciones sitio-dirigidas del plásmido mediante el sistema de recombinación sitio específica Cre-*loxP*, del fago P1 (Figura 3B). En este momento, hemos diseñado y realizado un total de 4 deleciones sobre el p42f (Figura 4), estas son, deleciones en diferentes regiones de dicho plásmido, así como deleciones sobrelapantes (Figura 4). Las deleciones generadas nos han permitido obtener diferentes tamaños del plásmido p42f, los cuales son fácilmente visibles mediante técnicas de electroforesis tipo Eckhardt, la cual consiste en una lisis suave de la bacteria dentro del gel de agarosa, que permite la visualización de plásmidos de gran tamaño (Eckhardt, 1978) (Figura 5). En todas las deleciones se observa un cambio en la movilidad electroforética con respecto al p42f silvestre (Figura 5). Las deleciones con las que actualmente contamos, se han nombrado de la siguiente manera:  $\Delta$ 100-600,  $\Delta$ 006-100,  $\Delta$ 100-214 y  $\Delta$ 308-600, nomenclatura que indica las kilobases que comprende el fragmento de ADN eliminado (Figura 4 y 5). Es importante resaltar que en todas estas mutantes están presentes los genes *panCB*. Los resultados demuestran que todas las cepas con diferentes deleciones sobre p42f presentan una disminución en el crecimiento en medio mínimo con respecto a la cepa silvestre (Figura 6), siendo la deleción  $\Delta$ 100-600 la que presenta la mayor disminución. Analizando las otras deleciones, encontramos que en la deleción  $\Delta$ 006-100 que comprende una región de

aproximadamente 94 kb eliminadas y presenta una disminución aproximada de 20% en el crecimiento (Figura 6). En la deleción  $\Delta 100-214$ , se presenta una disminución aproximada del 10% de crecimiento (Figura 6). En la deleción  $\Delta 308-600$ , eliminamos 292 kb, aproximadamente, la disminución del crecimiento es de casi el 25%. Como se menciona al principio de este párrafo, la deleción  $\Delta 100-600$  presenta una disminución de 65%, aproximadamente, siendo este resultado parecido al que obtenemos cuando una cepa curada de p42f se crece en presencia de pantotenato, o bien cuando esta misma cepa se complementa con los genes *panCB* en *trans* (ver Villaseñor *et al.*, 2011, “Figure 2b y 3”).

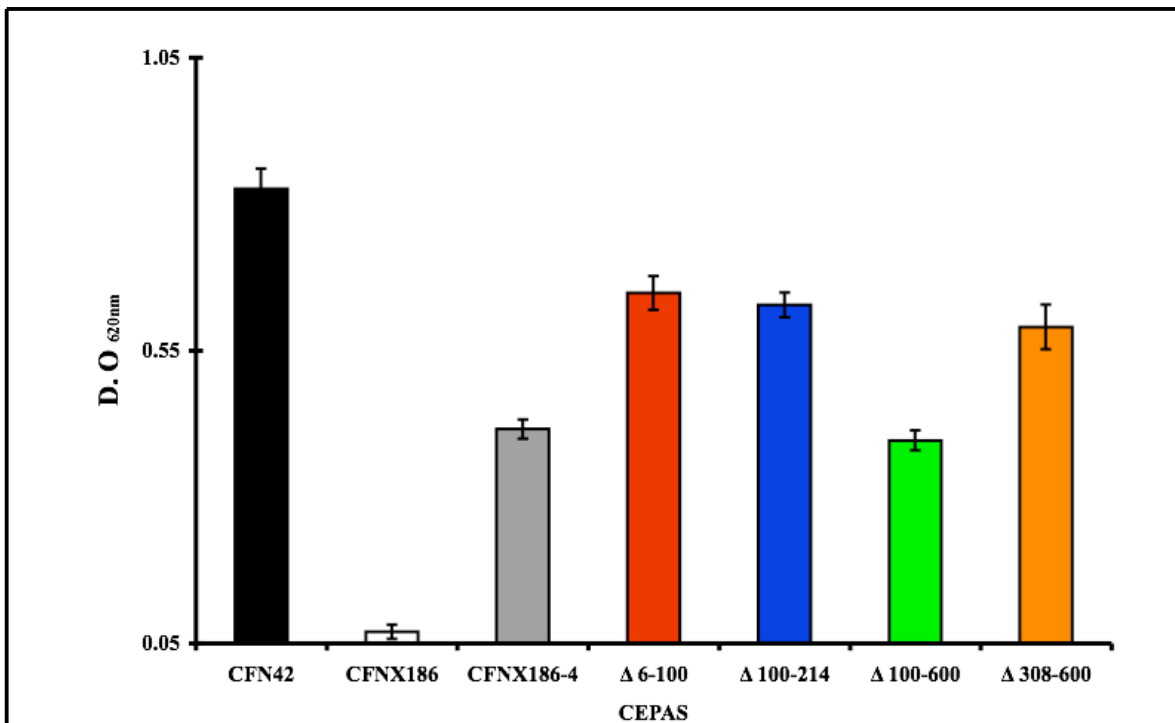
Es importante resaltar que, aunque contamos con deleciones que cubren grandes regiones de p42f, como se puede apreciar en la figura 6; aun no hemos sido capaces de eliminar la región del plásmido comprendida entre las kilobases 214 a 308, por lo que esta región se mantiene inexplorada.



**Figura 4.** Diseño de las deleciones realizadas en el plásmido p42f. Cada una de las deleciones se llevó a cabo mediante el sistema Cre/loxP. Los semicírculos de colores indican el fragmento del plásmido que fue eliminado. Los mismos colores corresponden a las barras de crecimiento de la figura 6.

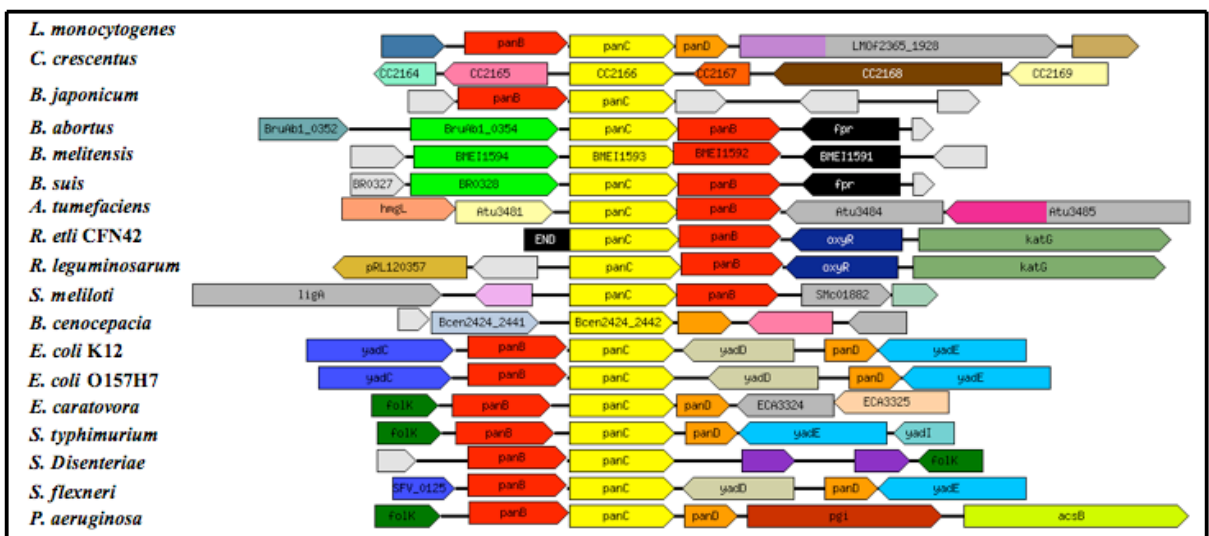


**Figura 5.** Electroforesis tipo Eckhardt de las deleciones realizadas sobre diferentes sectores del plásmido p42f. Cada una de las deleciones realizadas sobre el p42f genera un nuevo plásmido de tamaño variable que se ve reflejado en el movimiento electroforético.

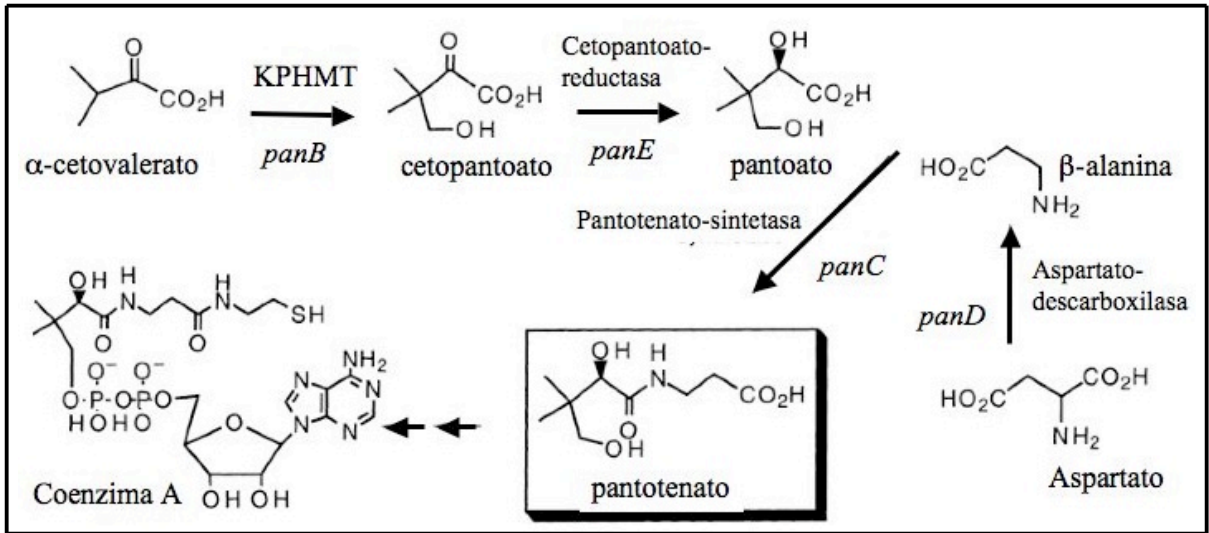


**Figura 6.** Crecimiento de distintas cepas que carecen de diferentes fragmentos del p42f medido por D. O a 620nm. Todas las cepas fueron inoculadas a una D. O. inicial de 0.05. Las gráficas representan el crecimiento después de 15 horas en MM con succinato-amoniaco sin pantotenato. Las barras muestran el promedio de al menos tres experimentos independientes. CFNX186 es la derivada curada del plásmido p42f y CFNX186-4 es la misma cepa, pero complementada con los genes *panCB*

**Caracterización de la vía de síntesis de pantotenato en los rhizobios.** La presencia de los genes *panCB* en plásmido, nos llevó a investigar dónde se encuentran codificados el resto de los genes necesarios para la vía de síntesis de esta vitamina. En *E. coli*, los genes para la síntesis de pantotenato, *panCB*, se encuentran muy cerca de otro gene que también participa en dicha vía, *panD* (Figura 7). Este último codifica para la enzima L-aspartato-1-d Descarboxilasa (ADC), que se encarga de descarboxilar aspartato para producir  $\beta$ -alanina (Figura 8). Contrario a lo que ocurre en otras bacterias, en la vecindad de los genes *panCB* de *R. etli* no está localizado *panD*. Lo más interesante es que en la secuencia completa del genoma de *R. etli* CFN42 no está anotado el gene *panD* ni la enzima ADC.



**Figura 7.** Organización genómica de los genes para la síntesis de pantotenato en diferentes bacterias. En la mayoría de estos ejemplos (*L. monocitogenes*, *E. coli*, *E. caratovora*, *S. typhimurium*, *S. flexneri* y *P. aeruginosa*) los genes *panCB* (amarillo y rojo, respectivamente) se encuentran cerca de otro gene, *panD* (anaranjado), encargado de sintetizar  $\beta$ -alanina, crucial para sintetizar pantotenato. En *R. etli*, *R. leguminosarum*, *S. meliloti*, *B. suis*, *B. abortus*, *B. melitensis*, y *B. japonicum*, solo *panCB* se encuentran cerca uno del otro y no se identifica a *panD* cerca de estos. En algunos casos (*C. crescentus* y *B. cenocepacia*) se distingue *panC* (amarillo) y *panD* (anaranjado), pero no *panB*. En el caso de *A. tumefaciens*, sabemos, por la anotación que existe un *panD*, sin embargo no esta cerca de *panCB*, aunque los tres están codificados sobre el cromosoma II lineal. La figura fue generada mediante el uso del servidor GeConT ([http://bioinfo.ibt.unam.mx/gecont\\_1/](http://bioinfo.ibt.unam.mx/gecont_1/)).



**Figura 8.** Vía de síntesis de pantotenato. La vía de síntesis de pantotenato ha sido ampliamente descrita en enterobacterias, la vía básicamente consta de dos pasos: síntesis de β-alanina, a partir de la descarboxilación de aspartato y la síntesis de pantoato y β-alanina se lleva a cabo mediante la enzima PanC (Tomado y modificado de von Delft *et al.*, 2001).

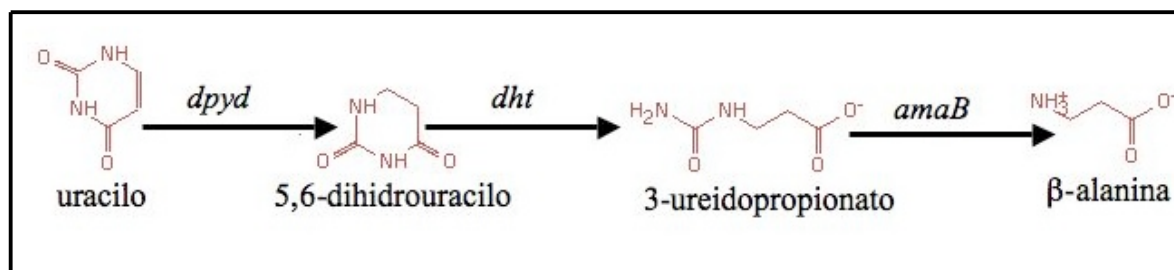
Buscamos el gene *panD* en la base de datos de rhizobios, ([www.xbase.ac.uk/rhizodb/](http://www.xbase.ac.uk/rhizodb/)) y encontramos que en *A. tumefaciens* C58, una bacteria filogenéticamente cercana a *R. etli* CFN42, sí está presente el gene *panD* que codifica para una ADC. Posteriormente y de acuerdo con los programas de búsqueda de similitud BLASTP y PsiBLAST, utilizando la ADC presente en *A. tumefaciens* como “query”, encontramos que dicha enzima está ausente en los genomas de otros rhizobios como *R. leguminosarum* 3841 y *S. meliloti* 1021 (Tabla 4). La ausencia de esta enzima se extiende a otros miembros del orden rhizobiales (Tabla 4). El mapa de vías metabólicas de KEGG señala la ausencia de la enzima ADC y sugiere que *R. etli* CFN42 podría sintetizar β-alanina a partir de la degradación de uracilo, a través de 3 pasos enzimáticos mediados por: dihidropirimidina deshidrogenasa (*dpyd*, EC:1.3.1.1), fenilhidantoinasa (*dht*, EC:3.5.2.2), y alantoato amidohidrolasa (*amaB*, EC:3.5.1). Esta vía es inusual en bacterias (todas las mutantes *panD* reportadas son auxótrofas de pantotenato), pero común en organismos eucariontes, principalmente en plantas (Zrenner *et al.*, 2006) (Figura



9). Esta aparente ausencia de ADC, nos ha llevado a proponer la hipótesis de que la síntesis de  $\beta$ -alanina en *R. etli* se lleva a cabo a partir de la degradación de uracilo.

**Tabla 4.** Enzimas involucradas en la síntesis de  $\beta$ -alanina. El producto de *panD* descarboxila aspartato y produce  $\beta$ -alanina. Los productos de *dpyd*, *dht* y *amaB* codifican enzimas que degradan uracilo obteniendo como producto final  $\beta$ -alanina.

BACTERIA/vía metabólica	<i>panD</i>	<i>dpyd</i>	<i>dht</i>	<i>amaB</i>
<i>A. tumefaciens</i> C58	+	-	+	+
<i>A. sp</i> ATCC3749	+	-	+	+
<i>A. rizhogenes</i>	+	-	+	+
<i>A. radiobacter</i> K84	-	+	+	-
<i>A. vitis</i> S4	-	+	+	-
<i>S. meliloti</i> 1021	-	+	+	+
<i>S. medicae</i>	-	-	+	-
<i>S. fredii</i> NGR234	+	+	+	-
<i>M. loti</i>	+	+	+	+
<i>M. loti</i> R74	+	-	+	-
<i>M. sp</i> BNC1	+	+	+	+
<i>B. japonicum</i>	+	-	+	+
<i>B. sp</i> ORS278	-	-	+	-
<i>R. etli</i> CFN42	-	+	+	+
<i>R. leguminosarum</i>	-	-	+	+
<i>R. leguminoarum</i> bv trifolii WSM2304	-	-	+	+
<i>R. leguminosarum</i> bv trifolii 1325	-	-	+	+
<i>R. etli</i> CIAT 652	-	-	+	-
<i>M. extorquens</i> AM1	-	-	-	-
<i>M. radiotolerans</i> JCM2831	-	-	+	-
<i>M. nodulans</i> ORS2060	-	-	+	-
<i>N. hamburgensis</i> X14	-	-	-	-
<i>B. abortus</i> S19	+	-	-	-
<i>B. melitensis</i>	+	-	+	+



**Figura 9.** Vía de síntesis de  $\beta$ -alanina a partir de la degradación de uracilo. La vía involucra tres enzimas codificadas por *dpyd* (dihidropirimidina deshidrogenasa: EC:1.3.1.1), *dht* (fenilhidantoinasa: EC:3.5.2.2), y *amaB* (alantoato amidohidrolasa: EC:3.5.1.-)

Para comprobar la hipótesis, realizamos una mutación en el gene *amaB* cuyo producto cataliza el último paso de la vía de degradación de uracilo obteniéndose como producto final

$\beta$ -alanina. La correcta interrupción del gene *amaB* se validó mediante ensayos tipo Southern. Una vez hecho esto, procedimos a evaluar su posible auxotrofia creciéndola en medio químicamente definido sin  $\beta$ -alanina, sin embargo, el crecimiento de esta cepa no disminuye, lo cual sugiere fuertemente que la vía de síntesis de  $\beta$ -alanina no ocurre por la vía de degradación de uracilo. Por medio de análisis realizados mediante búsquedas tipo BLASTP encontramos un gene parálogo de *amaB*, aunque la identidad es muy baja (32%), debemos hacer una mutación en el parálogo, así como obtener una doble mutante, para evaluar su participación en la síntesis de  $\beta$ -alanina. Nuestra segunda hipótesis propone que la síntesis de  $\beta$ -alanina podría llevarse a cabo por una proteína no ortóloga de ADC, y que no hemos sido capaces de identificar con métodos convencionales dada su lejanía con respecto a ADC *bona fide*. Para tratar de confirmar nuestra segunda hipótesis, construimos dos mutantes *panD*, por un lado, una mutante *panD* de *A. tumefaciens* y por otro lado una mutante *panD* en *E. coli* K12, las cuales serán complementadas con un banco del genoma total de *R. etli*, esperando encontrar un fragmento de ADN que codifique para la proteína no ortóloga de PanD. La otra alternativa es realizar experimentos de mutagénesis al azar en *R. etli* CFN42 para obtener una auxótrofa de  $\beta$ -alanina o de pantotenato

**Generación de las mutantes *panD* de *A. tumefaciens* C58 y *E. coli* K12, y evaluación del crecimiento en MM.** La mutación en *A. tumefaciens* se realizó interrumpiendo el gene *panD* con un “cassette” de resistencia a espectinomicina, mediante una doble recombinación. La mutación en *E. coli* se realizó mediante intercambio alélico, sustituyendo el gene *panD* por un “cassette” de resistencia a kanamicina (Datsenko y Wanner, 2000). Para verificar si son auxótrofas de  $\beta$ -alanina, ambas mutantes se crecieron en MM en

ausencia del respectivo aminoácido. La mutante *panD::ΩSp* de *A. tumefaciens* continuó siendo protótrofa de β-alanina ya que no perdió la capacidad de crecer en medio mínimo. El resultado anterior nos lleva a proponer tres hipótesis: primero, la ADC de *A. tumefaciens* no es funcional. Segundo, existe más de una vía de síntesis de β-alanina en esta bacteria. Tercero, posiblemente, la síntesis de β-alanina en *A. tumefaciens* se realiza a través de la degradación de uracilo, ya que se ha reportado recientemente que en *A. tumefaciens* es posible obtener, con fines biotecnológicos, β-alanina a partir de la degradación de uracilo, ya que la enzima *N*-carbamoil-β-alanina amidohidrolasa (NCβAA) ha sido caracterizada por Martínez-Gómez y colaboradores, donde han encontrado que *in vitro* dicha enzima hidroliza *N*-carbamoil-β-alanina produciendo CO<sub>2</sub>, NH<sub>3</sub> y β-alanina, así como otros β-aminoácidos y taurina (Martínez-Gómez *et al.*, 2009). Contrario al resultado que tuvimos con *A. tumefaciens*, la mutante *ΔpanD::km* en *E. coli* si mostró un fenotipo de auxotrofia por β-alanina muy claro, por esto, decidimos comenzar nuestros experimentos de complementación con la cepa mutante en *E. coli*.

**Complementación de la mutante *E. coli* *ΔpanD::km*.** Una vez que logramos obtener la mutante *panD* de *E. coli*, nos dimos a la tarea de realizar la complementación con un banco genómico de *R. etli* CFN42, para esto, realizamos una cruce triparental en medio rico PY entre la cepa mutante *panD* resistente a rifampicina (receptora), una cepa que porta un plásmido “helper” que lleva los genes *tra* para conjugación y la cepa donadora, en este caso una cepa de *E. coli* HB101 la cual porta segmentos del genoma de *R. etli* clonados en el cósmido pLAFR1 Tc<sup>r</sup>. Después de 24 h, la cruce se resuspendió en una solución de Sulfato de Magnesio-Tween y diluciones de la misma se platearon en medio M9 sin pantotenato,

con tetraciclina 10µg/ml y rifampicina 25µg/ml para seleccionar a las mutantes complementadas, así como en medio rico PY con rifampicina y tetraciclina como control de la eficiencia de conjugación. En un primer experimento no crecieron transconjugantes en el medio selectivo, pero sí en el medio control. Lo anterior indicó que los eventos de conjugación se llevaron a cabo de manera positiva, pero no se logró complementar la auxotrofia de la mutante *panD* de *E.coli*. La falta de complementación nos llevó a investigar si una mutante de *E. coli* puede ser complementada con genes de *R. etli*, para descartar posibles resultados negativos en nuestra complementación heteróloga, mutantes de *E. coli* en los genes *panC* y *panB*, denominadas JW129 y JW130, respectivamente, auxótrofas de pantotenato fueron complementadas con el plásmido pTV4 (genes *panC* y *panB* de *R. etli* CFN42) y pCos24 (*panC*, *panB*, *katG*, *oxyR* y 20kb más, de *R. etli* CFN42), los cuales se utilizaron en el trabajo publicado (Villaseñor et al., 2011) y del cual se hace mención previa. Las cepas mutantes complementadas (JW127/pTV4, JW129/pCos24, JW130/pTV4 y JW130/pCos24) recuperan en su totalidad la capacidad para crecer en medio M9 sin pantotenato, lo cual nos demuestra que los genes provenientes de *R. etli* sí funcionan en *E. coli*. Sin embargo, la mutante *panD* de *E. coli* no se complementó con la copia de *panD* de *A. tumefaciens* en *trans* (datos no mostrados). Cuando comparamos PanD en *E. coli* contra PanD en *A. tumefaciens*, encontramos que existe un índice de identidad de 42% y una similitud de 55%, mientras que, en el caso de PanC y PanB de *R. etli*, contra *E. coli*, encontramos un índice de identidad de 43% y 38%, respectivamente. Lo anterior pone de manifiesto que aún cuando los índices de identidad y similitud sean elevados, no significa que siempre se puede realizar una complementación heteróloga. La falta de complementación con el banco también podría deberse a la ausencia del fragmento de ADN en el cual esté codificada la supuesta ADC.

Esto, sin duda, nos plantea un nuevo reto: investigar si en *R. etli* la síntesis de  $\beta$ -alanina se realiza a través de la descarboxilación de aspartato, lo que nos lleva a incluir dentro de nuestra estrategia experimental montar el ensayo para medir la actividad de aspartato descarboxilasa a partir de un extracto crudo de proteína. El resultado de este ensayo determinará si continuamos con la complementación heteróloga o implementamos los experimentos de mutagénesis al azar para obtener una cepa de *R. etli* auxótrofa de  $\beta$ -alanina.

## DISCUSIÓN

En este trabajo encontramos, de manera experimental, que dos de las cuatro enzimas requeridas para la biosíntesis de pantotenato (Figura 8), están codificadas en plásmidos en *R. etli* y *R. leguminosarum*. Estos resultados indican que se requiere de más de un replicón para la síntesis de esta vitamina. Trabajos previos sobre la síntesis de lipopolisacárido, el principal componente estructural de la membrana externa de bacterias Gram negativas, así como la síntesis de tiamina en rhizobios demuestran la participación de enzimas codificadas en un plásmido (p42b) y el cromosoma (García de los Santos y Brom, 1997; Miranda-Ríos *et al.*, 1997). La conectividad funcional entre enzimas de los siete replicones también ha sido sugerida mediante un análisis *in silico* de todo el genoma de *R. etli* CFN42 con NEBULON, un método bioinformático que predice relaciones funcionales entre enzimas (González *et al.*, 2006; Janga *et al.*, 2005). Todos estos resultados fortalecen la hipótesis de que en las bacterias con genomas multireplicón se requiere de una cooperación funcional entre los diferentes cromosomas y plásmidos para que la célula pueda llevar a cabo algunas de sus funciones metabólicas básicas.

Por los resultados presentados en nuestro trabajo, sabemos que, además de las dos enzimas necesarias para la síntesis de pantotenato, existen otros genes que participan en el crecimiento óptimo de la bacteria en un medio definido, pues una cepa de *R. etli* que carece de todo el plásmido p42f (CFNX186), solo reestablece su crecimiento en un 50% si se complementa con la copia silvestre de los genes *panCB*, o se crece en presencia de pantotenato. La localización de dichas funciones está siendo analizada mediante la eliminación consecutiva de fragmentos del genoma que den origen a una disminución en su cinética de crecimiento. Hasta ahora, todas las deleciones muestran una disminución en el

crecimiento que va desde un 15% hasta un 65% (Figura 6). De acuerdo con nuestros datos, en todo el plásmido p42f hay funciones importantes para el crecimiento en medio mínimo que contribuyen en menor ( $\Delta 100-214$ ) o mayor grado ( $\Delta 006-100$ ,  $\Delta 308-600$ , y  $\Delta 100-600$ ) al crecimiento óptimo de *R. etli* CFN42 (Figura 6). En este momento, nos encontramos analizando cada una de las regiones eliminadas con la finalidad de encontrar posibles blancos que se relacionen con nuestro fenotipo para realizar mutaciones en ellos o bien diseñar subdelecciones que nos permitan encontrar genes importantes. Hasta ahora hemos encontrado 18 genes relacionados con el metabolismo de diferentes compuestos, como azúcares y ácidos grasos, los cuales se encuentran situados en diferentes regiones del plásmido. En el plásmido p42f se encuentran codificados 49 reguladores transcripcionales, siendo los más representados los pertenecientes a las familias LysR, AraC, GntR y TetR (González *et al.*, 2006). Hasta ahora, solo se ha descrito la importancia de *oxyR*, el regulador de la única catalasa presente en *R. etli* (Vargas *et al.*, 2003), sin embargo, en este momento, no podemos descartar la posibilidad de que el fenotipo de *R. etli* sea debido también a la importancia de otros reguladores transcripcionales presentes en el p42f que, al ser eliminados en una cepa curada de dicho plásmido, la expresión de genes dentro dentro del mismo p42f, o incluso fuera de él, se vean afectados en su expresión.

La arquitectura multireplicón y su conectividad funcional entre los diferentes replicones, debido a la presencia de enzimas que catalizan procesos metabólicos fundamentales, podrían ser parte de un proceso evolutivo dirigido hacia la formación de cromosomas adicionales, en el cual los plásmidos y los mecanismos de recombinación juegan un papel fundamental. Aunado a todo esto, p42f es un replicón en el que están conservados el contenido y el orden de los genes en comparación con otras cepas y especies de *Rhizobium*, tales como *R. leguminosarum* biovariedades *viciae* y *trifolii*, además de que el contenido de G

+ C es muy parecido al del cromosoma, 61% (González *et al.*, 2006), así como al de otros replicones importantes previamente descritos, como p42e (Landeta *et al.*, 2011). Lo anterior indica que, p42f es un replicón que se ha mantenido dentro del genero por mucho tiempo, y descarta la posibilidad de que haya sido adquirido por eventos de transferencia horizontal (González *et al.*, 2006).

La ventaja de los genomas multireplicón no está del todo clara, de acuerdo con Slater y colaboradores, el tener más de un replicón puede permitir la acumulación de genes cuando el cromosoma I no puede ya soportar mayor carga genética (Slater *et al.*, 2009). Sin embargo, parece que dentro de las *Rhizobiaceas*, diferentes especies parecen soportar la acumulación de genes de diferentes maneras, ya que en especies de *Bradyrhizobium* y *Mezorhizobium*, solo encontramos cromosomas I de tamaño considerable, con pocos, o en algunos casos, ningun plásmido asociado a dichos genomas. En contraste, cepas de *Agrobacterium* y *Rhizobium* poseen mas de un cromosoma o un número variable de replicones de gran tamaño, que van desde uno hasta seis. Dichos replicones muestran acumulación genética al parecer variable, como se ve reflejado en su tamaño, el cual varia de acuerdo a la cepa. Sin embargo, debemos considerar también la posibilidad de que el movimiento de grupos de genes esenciales hacia un replicón diferente al cromosoma I, sea simplemente un evento que observamos en un momento de la evolución de la bacteria, sin que esto implique ventajas selectivas. Movilizaciones adicionales de genes esenciales, del cromosoma I a un replicón diferente, podrían consolidar la naturaleza esencial de dicho replicón. Sin embargo, queda claro que la organización genética, incluso de los genes esenciales, es mucho más complejo y dinámico de lo que se ha pensado.

Con el número de genomas bacterianos secuenciados en aumento, la necesidad de homogenizar la nomenclatura de estos replicones con funciones esenciales, se hace presente.



Desde nuestro punto de vista, el hecho de que el plásmido p42f debiera ser considerado un cromosoma en formación o un cromido se apoya en el hallazgo de dos genes que son esenciales para el crecimiento de la bacteria, pero además la, la compleja auxotrofia de una cepa sin p42f, de acuerdo a nuestros resultados, nos hace pensar en que el número de genes esenciales podría ir en aumento.

Un dato muy importante que se deriva de nuestro trabajo es la versatilidad en la manera de sintetizar  $\beta$ -alanina que tienen los rizobios, contrario a lo establecido por el estudio de las ADC de enterobacterias. Como pudimos corroborarlo de manera experimental, una mutante *panD* de *A. tumefaciens* no presenta auxotrofia por  $\beta$ -alanina, aun cuando al parecer tiene un PanD *bona fide* y de copia única en el genoma. Este resultado pudiera deberse a la presencia de la enzima N-carbamil- $\beta$ -alanina amidohidrolase, la cual por experimentos *in vitro* sabemos que cataliza la hidrólisis de N-carbamoyl- $\beta$ -alanina (producto de la degradación de uracilo) y produce  $\beta$ -alanina (Martínez-Gómez *et al.*, 2009). Comparaciones tipo BLAST indican que ésta enzima está duplicada en *A. tumefaciens* y *R. etli* aunque entre ambos existe un 32% y 48% de identidad y similitud, respectivamente, por lo que el análisis genético-funcional requiere de dobles mutantes.

Nuestros estudios preliminares sobre la síntesis de  $\beta$ -alanina en *R. etli*, abren la posibilidad de encontrar diferencias respecto a la síntesis de este aminoácido precursor de pantotenato, entre el modelo más estudiado (enterobacterias) y rizobios. La importancia de entender dichas diferencias es el tema en el cual están enfocando nuestros esfuerzos en la actualidad y esperamos que un futuro cercano, podamos entender y explicar la síntesis de moléculas de importancia metabólica en *R. etli* CFN42.

## CONCLUSIONES.

El plásmido p42f de *R. etli* CFN42 posee elementos importantes que, en conjunto, permiten a la bacteria crecer de manera óptima en un medio químicamente definido. En este trabajo, encontramos que parte de ese grupo de genes está conformado por *panCB*, los cuales son indispensables para la síntesis de pantotenato, precursor esencial de la coenzima A.

La presencia de los genes *panCB* en plásmidos es poco común dentro del orden rhizobiales, pues está restringido a *Rhizobium etli* y *Rhizobium leguminosarum*.

De acuerdo a nuestro análisis filogenético, la inusual presencia de *panCB* en plásmidos no se debe a un evento de transferencia horizontal sino más bien a una transferencia intragenómica del cromosoma a plásmidos.

Por sí solos, los genes *panCB* no son capaces de restaurar en su totalidad el crecimiento en medio definido de una cepa que carece de todo el plásmido p42f. Lo anterior sugiere que existen otras funciones necesarias para un crecimiento óptimo.

En los genomas de *R. etli* y otros rhizobios, no está codificada la enzima ADC típica de enterobacterias.

## PERSPECTIVAS.

**Identificación de elementos importantes en el p42f.** Consideramos de gran relevancia identificar las funciones codificadas en el plásmido p42f, diferentes de *panCB* que son necesarias para el crecimiento en MM. Creemos que el esquema de deleciones consecutivas es el apropiado para ir delimitando la región involucrada, hasta llegar a un conjunto de genes cuya contribución al crecimiento pueda ahora ser analizado por mutaciones.

**Explorar si *R. etli* CFN42 realiza descarboxilación de aspartato para obtener  $\beta$ -alanina.** Se debe montar un ensayo enzimático de descarboxilación de aspartato y determinar si *R. etli* sintetiza  $\beta$ -alanina por la vía clásica pero con una ADC atípica. La ausencia de actividad enzimática indicaría que estas bacterias utilizan una vía de síntesis alternativa.

**Realizar un análisis bioinformático analizando los dominios de ADC.** Utilizando la secuencia de PanD de *A. tumefaciens*, podemos realizar una búsqueda de los ortólogos mediante BLAST. Posteriormente, tomamos las secuencias con los mejores *e-value* y mediante la utilización del programa MEME buscamos motivos en las secuencias obtenidas de BLAST. Con los motivos que se obtengan, realizamos una búsqueda, utilizando el programa MAST, sobre todo el genoma de *R. etli* CFN42, con cada uno de los motivos previamente obtenidos. Como una alternativa, podríamos realizar un análisis filogenético de las ADC, con la finalidad de evaluar la divergencia estructural de estas proteínas y que nos dé indicios sobre posibles proteínas remotamente relacionadas con ADC pero que pudieran catalizar la misma función. Este análisis filogenético podría ayudarnos a entender si la ausencia de esta enzima está restringida al orden *Rhizobiales* o es una tendencia en todo el grupo de las  $\alpha$ -proteobacterias.

**Experimentos de mutagénesis al azar.** Con la finalidad de obtener una cepa de *R. etli* auxótrofa de  $\beta$ -alanina, podríamos implementar, con la utilización de transposones, ensayos de mutagénesis al azar e identificar aquellas colonias que crezcan en presencia de  $\beta$ -alanina, pero que no crezcan si este aminoácido está ausente.

## BIBLIOGRAFÍA

- Allardet-Servent, A., S. Michaux-Charachon, E. Jumas-Bilak, L. Karayan, and M. Ramuz.** 1993. Presence of one linear and one circular chromosome in the *Agrobacterium tumefaciens* C58 genome. *J. Bacteriol.* **175**: 7869-7874
- Ayres, E. K., V. J. Thomson, G. Merino, D. Balderes, and D. H. Figurski.** 1993. Precise deletions in large bacterial genomes by vector-mediated excision (VEX). The *trfA* gene of promiscuous plasmid RK2 is essential for replication in several gram-negative hosts. *J. Mol. Biol.* **230**: 174-185
- Barbour, A. G., and C. F. Garon.** 1987. Linear plasmids of the bacterium *Borrelia burgdorferi* have covalently closed ends. *Science.* **237**: 409-411
- Bavishi, A., A. Abhishek, L. Lin, and M. Choudhary.** 2010. Complex prokaryotic genome structure: rapid evolution of chromosome II. *Genome.* **53**: 675-687
- Bendich, J. A., and K. Drlica.** 2000. Prokaryotic and eukaryotic chromosomes: what's the difference? *Bioessays.* **22**: 481-486
- Bittner, A. N., A. Foltz, and V. Oke.** 2007. Only one of five *groEL* genes is required for viability and successful symbiosis in *Sinorhizobium meliloti*. *J. Bacteriol.* **189**: 1884-1889
- Bresler, V., W. L. Montgomery, L. Fishelson, and P. E. Pollak.** 1998. Gigantism in a bacterium, *Epulopiscium fishelsoni*, correlates with complex patterns in arrangement, quantity, and segregation of DNA. *J. Bacteriol.* **180**: 5601-5611
- Brom S, Garcia-de los Santos A, Stepkowski T, Flores M, Dávila G, Romero D, Palacios R.** 1992. Different plasmids of *Rhizobium leguminosarum* bv. Phaseoli are required for optimal symbiotic performance. *J Bacteriol.* **174**:5183-5189
- Brom, S., A. García de los Santos, L. Cervantes, R. Palacios, and D. Romero.** 2000. In *Rhizobium etli* symbiotic plasmid transfer, nodulation competitiveness and cellular growth require interaction among different replicons. *Plasmid.* **44**: 34-42
- Cevallos, M. A., R. Cervantes-Rivera, and R. M. Gutiérrez-Ríos.** 2008. The *repABC* plasmid family. *Plasmid.* **60**: 19-37
- Chaconas, G., and K. Kobryn.** 2010. Structure, function and evolution of linear replicons in *Borrelia*. *Annu. Rev. Microbiol.* **64**: 185-202
- Crespi, M., E. Messens, A. B. Caplan, M. van Montagu, and J. Desomer.** 1992. Fasciations induction by the phytopathogen *Rhodococcus fascians* depends upon a linear plasmid encoding a cytokinin synthase gene. *EMBO J.* **11**: 795-804
- D' Elia, M. A., M. P. Pereira, and E. D. Brown.** 2009. Are essential genes really essential?.

*Trends in Microbiology*. **17**: 433-438

**Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U S A.* **97**: 6649-6645

**Del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas.** 1998. Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev.* **62**: 434-464

**Duigou, S., K. G. Knudsen, O. Skovgaard, E. S. Egan, A. Lobner-Olesen, and M. K. Waldor.** 2006. Independent control of replication initiation of the two *Vibrio cholerae* chromosomes by DnaA and RctB. *J. Bacteriol.* **188**: 6419-6424

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

**Egan E. S., M. A. Fogel, and M. Waldor.** 2005. Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol. Microbiol.* **56**: 1129-1138

**Egan, E. S., and M. K. Waldor.** 2003. Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell.* **114**: 521-530

**Ferdows, M. S., and A. G. Barbour.** 1989. Megabase-sized linear DNA in the bacterium *Borrelia burgdorferi*, the Lyme disease agent. *Proc. Natl. Acad. Sci. U S A.* **86**: 5679-5673

**Finan T. M., S. Weidner, K. Wong, J. Buhrmester, P. Chain, F. J. Vorhölter, I. Hernandez-Lucas, A. Becker, A. Cowie, J. Gouzy, B. Golding, and A. Pühler.** 2001. The complete sequence of the 1683 kb pSymB megaplasmid from the N<sub>2</sub>-fixing endosymbiont *Sinorhizobium meliloti*. *Proc. Natl. Acad. Sci. U S A.* **98**: 9889-9894

**Gao, J., and L. L. Chen.** 2010. Theoretical methods for identifying important functional genes in bacterial genomes. *Res. Microbiol.* **161**:1-8

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

**Gerdes, K., M. Howard, and F. Szardenings.** 2010. Pushing and pulling in prokaryotic DNA segregation. *Cell.* **141**: 927-942

**Gerdes, S., R. Edwards, M. Kubal, M. Fonstein, R. Stevens, and A. Osterman.** 2006. Essential genes on metabolic maps. *Curr. Opin. Biotechnol.* **17**: 448-456

**Ghosh, S. K., S. Hajra, A. Paek, and M. Jayaram.** 2006. Mechanisms for chromosomes and plasmid segregation. *Annu. Rev. Biochem.* **75**: 211-241

**González, V., R. I. Santamaría, P. Bustos, I. Hernández-González, A. Medrano-Soto, G. Moreno-Hagelsieb, S. C. Janga, M. A. Ramírez, V. Jiménez-Jacinto, J. Collado-Vides, and G. Dávila.** 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., and E. Martínez-Romero.** 2000. Multiresistance genes of *Rhizobium etli* CFN42. *Mol. Plant Microbe Interact.* **13**: 572-577

**Guo, X., M. Flores, P. Mavingui, S. Fuentes, G. Hernández, G. Dávila, and R. Palacios.** 2003. Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. *Genome Res.* **13**: 1810-1817

**Harrison, P. W., R. P. Lower, N. K. Kim, and J. P. Young.** 2010. Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends in Microbiology.* **18**:141–148

**Heidelberg J. F., J. A. Eisen, W. C. Nelson, R. A. Clayton, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, L. Umayam, S. R. Gill, K. E. Nelson, T. D. Read, H. Tettelin, D. Richardson, M. D. Ermolaeva, J. Vamathevan, S. Bass, H. Qin, I. Dragoi, P. Sellers, L. McDonald, T. Utterback, R. D. Fleishmann, W. C. Nierman, O. White, S. L. Salzberg, H. O. Smith, R. R. Colwell, J. J. Mekalanos, J. C. Venter, C. M. Fraser.** 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature.* **406**: 477-483

**Hu, M., X. Zhang, E. Li, and Y. Feng.** 2010. Recent advancements in toxin and antitoxin systems involved in bacterial programmed cell death. *Int. J. Of Microbiol.* **2010**: 781430

**Itaya, M., and T. Tanaka.** 1997. Experimental surgery to create subgenomes of *Bacillus subtilis* 168. *Proc. Natl. Acad. Sci. USA.* **94**: 5378-5382

**Jackowski, S.** Biosynthesis of pantothenic acid and coenzyme A. In *Escherichia coli* and *Salmonella*: cellular and molecular biology. Edited by Neidhardt, F. C., R. Curtis III, J. L. Ingraham, E. C. C. Lin, K. B. Low. Washington D. C. ASM Press. 1996: 1310-1324

**Janga, S. C., J. Collado-Vides, and G. Moreno-Hagelsieb.** 2005. Nebulon: a system for the inference of functional relationships of gene products from rearrangement of predicted operons. *Nucleic Acids Research.* **33**: 2521-2530

**Jordan, I. K., K. S. Makarova, J. L. Spouge, Y. I. Wolf, and E. V. Koonin.** 2001. Lineage-specific gene expansions in bacterial and archaeal genomes. *Genome Res.* **11**: 555-565

**Jumas-Bilak, E., S. Michaux-Charachon, G. Bourg, D. O'Callaghan, and M. Ramuz.** 1998a. Differences in chromosome number and genome rearrangements in the genus *Brucella*. *Mol. Microbiol.* **27**: 99-106

- Jumas-Bilak, E., S. Michaux-Charachon, G. Bourg, M. Ramuz, and A. Allardet-Servent.** 1998b. Unconventional genomic organization in the alpha subgroup of the Proteobacteria. *J. Bacteriol.* **180**: 2749-2755
- Kahng, L. S., and L. Shapiro.** 2003. Polar localization origins in the multipartite genome *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. *J. Bacteriol.* **185**: 3384-3391
- Kolsto, A.** 1997. Dynamic bacterial genome organization. *Mol. Microbiol.* **24**: 241-248
- Komaki, K., and H. Ishikawa.** 2000. Genomic copy number of intracellular bacteria symbionts of aphids varies in response to developmental stage and morph of their host. *Insect Biochem Mol Biol.* **30**: 253-258
- Krawiec, S., and M. Riley.** 1990. Organization of the bacterial chromosome. *Microbiol. Rev.* **54**: 502-539
- Landeta, C., A. Dávalos, M. A. Cevallos, O. Geiger, S. Brom, and D. Romero.** 2011. Plasmids with a chromosome-like role in Rhizobia. *J. Bacteriol.* **193**: 1317-1326
- Latorre, A., R. Gil, F. J. Silva, and A. Moya.** 2005. Chromosomal stasis *versus* plasmid plasticity in aphid endosymbiont *Buchnera aphidicola*. *Heredity.* **95**: 339-347
- Lewis, K.** 2010. Persister cells. *Annu. Rev. Microbiol.* **64**: 357-372
- Lin, Y. S., H. M. Kieser, D. A. Hopwood, and C. W. Chen.** 1993. The chromosomal DNA of *Streptomyces lividans* 66 is linear. *Mol. Microbiol.* **10**: 923-933
- MacLean, A. M., T. M. Finan, and M. J. Sadowsky.** 2007. Genomes of the symbiotic nitrogen-fixing bacteria of legumes. *Plant Physiol.* **144**: 615-622
- MacLellan, S. R., C. D. Sibley, and T. M. Finan.** 2004. Second chromosomes and megaplasmids in bacteria, p. 529–542. In B. E. Funnell and G. J. Phillips (ed.), *Plasmid biology*. ASM Press, Washington, DC.
- Maldonado, R., J. Jiménez, and J. Casadesús.** 1994. Changes of ploidy during the *Azotobacter vinelandii* growth cycle. *J. Bacteriol.* **176**: 3911-3919
- Mann, S., and Y. P. Chen.** 2010. Bacterial genomic G + C composition: eliciting environmental adaptation. *Genomics.* **95**: 7-15
- Marri, P. R., L. Harris, K. Houmiel, S. C. Slater, and H. Ochman.** 2008. The effect of chromosome geometry on genetic diversity. *Genetics.* **179**: 511-516
- Martínez-Gómez, A. I., S. Martínez-Rodríguez, J. Pozo-Dengra, D. Tessaro, S. Servi, J. M. Clemente- Jiménez, F. Rodríguez-Vigo, and F. J. Las Heras-Vázquez.** 2009. Potential application of *N*-Carbamoyl- $\beta$ -alanine amidohydrolase from *Agrobacterium tumefaciens* C58



for  $\beta$ -amino acid production. *Appl. Environm. Microbiol.* **75**: 514-520

**Mendell, J. E., K. D. Clements, J. H. Choat, and E. R. Angert.** 2008. Extreme polyploidy in a large bacterium. *Proc Natl Acad Sci USA.* **105**: 6730-6734

**Michaux, S., J. Paillisson, M. J. Carles-Nurit, G. Bourg, A. Allardet-Servent, and M. Ramuz.** 1993. Presence of two independent chromosomes in the *Brucella melitensis* 16M genome. *J. Bacteriol.* **175**: 701-705

**Miranda-Ríos, J., C. Morera, H. Taboada, A. Dávalos, S. Encarnación, J. Mora, and M. Soberón.** 1997. Expression of thiamin biosynthetic genes (*thiCOGE*) and production of symbiotic terminal oxidase *cbb3* in *Rhizobium etli*. *J. Bacteriol.* **179**: 6887-6893

**Moran, N. A.** 2002. Microbial minimalism: genome reduction in bacterial pathogens. *Cell.* **108**: 583-586

**Moriya, Y., M. Itoh, S. Okuda, A. Yoshizawa, and M. Kanehisa.** 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* **35**: W182-W185

**Moya, A., R. Gil, A. Latorre, J. Peretó, M. Garcillán-Barcía, and F. de la Cruz.** 2009. Towards minimal bacterial cells: evolution vs design. *FEMS Microbiol. Rev.* **33**: 225-235

**Okada, K., T. Iida, K. Kita-Tsukamoto, and T. Honda.** 2005. Vibrios commonly possess two chromosomes. *J. Bacteriol.* **187**: 752-757

**Rocha, E. P.** 2008. The organization of the bacterial genome. *Annu. Rev. Genetics.* **42**: 211-233

**Rodriguez-Quñones, F., M. Maguire, E. J. Wallington, P. S. Gould, V. Yerko, J. A. Downie, and P. A. Lund.** 2005. Two of three *groEL* homologues in *Rhizobium leguminosarum* are dispensable for normal growth. *Arch. Microbiol.* **183**: 253-265

**Rosenberg, S. M., and P. J. Hastings.** 2004. Adaptive point mutation and adaptive amplification pathways in the *Escherichia coli* Lac system: Stress responses producing genetic change. *J. Bacteriol.* **186**: 4838-4843

**Rosenblueth, M., and E. Martínez-Romero.** 2004. *Rhizobium etli* maize populations and their competitiveness for root colonization. *Arch. Microbiol.* **181**: 337-344

**Schneiker S., O. Perlova, O. Kaiser, K. Gerth, A. Alici, M. O. Altmeyer, D. Bartels, T. Bekel, S. Beyer, E. Bode, H. B. Bode, C. J. Bolten, J. V. Choudhuri, S. Doss, Y. A. Elnakady, B. Frank, L. Gaigalat, A. Goesmann, C. Groeger, F. Gross, L. Jelsbak, L. Jelsbak, J. Kalinowski, C. Kegler, T. Knauber, S. Konietzny, M. Kopp, L. Krause, D. Krug, B. Linke, T. Mahmud, R. Martinez-Arias, A. C. McHardy, M. Merai, F. Meyer, S. Mormann, J. Muñoz-Dorado, J. Perez, S. Pradella, S. Rachid, G. Raddatz, F. Rosenau, C. Rückert, F. Sasse, M. Scharfe, S. C. Schuster, G. Suen, A. Treuner-Lange, G. J. Velicer, F. J. Vorhölter, K. J. Weissman, R. D. Welch, S. C. Wenzel, D. E.**

- Whitworth, S. Wilhelm, C. Wittmann, H. Blöcker, A. Pühler, and R. Müller.** 2007. Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat Biotechnol.* **25**: 1281-1289
- Slater S. C., B. S. Goldman, B. Goodner, J. C. Setubal, S. K. Farrand, E. W. Nester, T. J. Burr, L. Banta, A. W. Dickerman, I. Paulsen, L. Otten, G. Suen, R. Welch, N. F. Almeida, F. Arnold, O. T. Burton, Z. Du, A. Ewing, E. Godsy, S. Heisel, K. L. Houmiel, J. Jhaveri, J. Lu, N. M. Miller, S. Norton, Q. Chen, W. Phoolcharoen, V. Ohlin, D. Ondrusek, N. Pride, S. L. Stricklin, J. Sun, C. Wheeler, L. Wilson, H. Zhu, and D. W. Wood.** 2009. Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. *J. Bacteriol.* **191**:2501–2511
- Sobral, B. W., R. J. Honeycutt, A. G. Atherly, and M. McClelland.** 1991. Electrophoretic separation of the three *Rhizobium meliloti* replicons. *J. Bacteriol.* **173**: 5173-5180
- Suwanto, A., and S. Kaplan.** 1989. Physical and genetic mapping of the *Rodobacter sphaeroides* 2.4.1 genome: presence of two unique circular chromosomes. *J. Bacteriol.* **171**: 5850-5859
- Suzuki, N., M. Inui, and H. Yukawa.** 2008. Random genome deletion methods applicable to prokaryotes. *Appl. Microbiol. Biotechnol.* **79**: 519-526
- Thanbichler, M., and L. Shapiro.** 2008. Getting organized, how bacterial cells move proteins and DNA. *Nature Rev.* **6**: 28-40
- Thomas, C. M.** 2004. Evolution and population genetics of bacterial plasmids. *Plasmid Biology Chapter.* **25**: 509-528
- Toro, E., and L. Shapiro.** 2010. Bacterial chromosome organization and segregation. *Cold Spring Harbor Perspect. Biol.* **2**: a000349
- Tun-Garrido, C., P. Bustos, V. González, and S. Brom.** 2003. Conjugative transfer of p42a from *Rhizobium etli* CFN42, which is required for mobilization of the symbiotic plasmid, is regulated by quorum sensing. *J. Bacteriol.* **185**:1681-1692
- Vargas, M. C., S. Encarnación, A. Dávalos, A. Reyes-Pérez, Y. Mora, A. García de los Santos, S. Brom, and J. Mora.** 2003. Only one catalase, *katG* is detectable in *Rhizobium etli*, and is encoded along with the regulator *oxyR* on a plasmid replicon. *Microbiology.* **149**: 1165-1176
- Villaseñor, T., S. Brom, A. Dávalos, L. Lozano, D. Romero, and A. García de los Santos.** 2011. Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*. *BMC Microbiology.* **11**:66
- von Delft, F., A. Lewendon, V. Dhanaraj, T. L. Blundell, C. Abell, and A. G. Smith.** 2001. The Crystal structure of *E. coli* pantothenate synthetase confirms it as a member of the cytidylyltransferase superfamily. *Structure.* **9**: 430-450

Wattam, A. R., K. P. Williams, E. E. Snyder, N. F. Jr Almeida, M. Shukla, A. W. Dickerman, O. R. Crasta, R. Kenyon, J. Lu, J. M. Shallom, H. Yoo, T. A. Ficht, R. M. Tsois, C. Munk, R. Tapia, C. S. Han, J. C. Detter, D. Bruce, T. S. Brettin, B. W. Sobral, S. M. Boyle, and J. C. Setubal. 2009. Analysis of ten *Brucella* genomes reveals evidence for horizontal gene transfer despite a preferred intracellular lifestyle. *J. Bacteriol.* **191**: 3569-3579

White O., J. A. Eisen, J. F. Heidelberg, E. K. Hickey, J. D. Peterson, R. J. Dodson, D. H. Haft, M. L. Gwinn, W. C. Nelson, D. L. Richardson, K. S. Moffat, H. Qin, L. Jiang, W. Pamphile, M. Crosby, M. Shen, J. J. Vamathevan, P. Lam, L. McDonald, T. Utterback, C. Zalewski, K. S. Makarova, L. Aravind, M. J. Daly, K. W. Minton, R. D. Fleischmann, K. A. Ketchum, K. E. Nelson, S. Salzberg, H. O. Smith, J. C. Venter, and C. M. Fraser. 1999. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science.* **286**: 1571-1577

Wood D. W., J. C. Setubal, R. Kaul, D. E. Monks, J. P. Kitajima, V. K. Okura, Y. Zhou, L. Chen, G. E. Wood, N. F. Jr. Almeida, L. Woo, Y. Chen, I. T. Paulsen, J. A. Eisen, P. D. Karp, D. Sr. Bovee, P. Chapman, J. Clendenning, G. Deatherage, W. Gillet, C. Grant, T. Kutuyavin, R. Levy, M. J. Li, E. McClelland, A. Palmieri, C. Raymond, G. Rouse, C. Saenphimmachak, Z. Wu, P. Romero, D. Gordon, S. Zhang, H. Yoo, Y. Tao, P. Biddle, M. Jung, W. Krespan, M. Perry, B. Gordon-Kamm, L. Liao, S. Kim, C. Hendrick, Z. Y. Zhao, M. Dolan, F. Chumley, S. V. Tingey, J. F. Tomb, M. P. Gordon, M. V. Olson, and E. W. Nester. 2001. The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**: 2317-2323

Xu, Q., M. Dziejman, and J. Mekalanos. 2003. Determination of the transcriptome of *Vibrio cholerae* during intrainestinal growth and midexponential phase *in Vitro*. *Proc. Natl. Acad. Sci. U S A.* **100**: 1286-1291

Xu, Y., B. Labedan, and N. Glansdorff. 2007. Surprising arginine biosynthesis: a reappraisal of the enzymology and evolution of the pathway in microorganisms. *Microbiol. Mol. Biol. Rev.* **71**: 36-47

Yamaichi, Y., M. A. Fogel, M. K. Waldor. 2007. *par* genes and the pathology of chromosome loss in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U S A.* **104**: 630-635

Yost, C. K., A. M. Rath, T. C. Noel, and M. F. Hynes. 2006. Characterization of genes involved in erythritol catabolism in *Rhizobium leguminosarum* bv. *viciae*. *Microbiology.* **152**: 2061-2074

Young J. P., L.C. Crossman, A. W. Johnston, N. R. Thomson, Z. F. Ghazoui, K. H. Hull, M. Wexler, A. R. Curson, J. D. Todd, P.S. Poole, T. H. Mauchline, A. K. East, M. A. Quail, C. Churcher, C. Arrowsmith, I. Cherevach, T. Chillingworth, K. Clarke, A. Cronin, P. Davis, A. Fraser, Z. Hance, H. Hauser, K. Jagels, S. Moule, K. Mungall, H. Norbertczak, E. Rabinowitsch, M. Sanders, M. Simmonds, S. Whitehead, and J.

**Parkhill.** 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol.* **7**: R34

**Yuan, J., Y. Yamaichi, and M. K. Waldor.** 2011. The three *Vibrio cholerae* chromosome II-encoded ParE toxins degrade chromosome I following loss of chromosome II. *J. Bacteriol.* **193**: 611-619

**Zhang, R., H. Y. Ou, and C. T. Zhang.** 2004. DEG, A Database of Essential Genes. *Nucleic Acids Research.* **32**: 271-272

**Zhang, R., and Y. Lin.** 2009. DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes. *Nucleic Acids Research.* **37**:455-458

**Zrenner R., M. Stitt, U. Sonnewald, and R. Boldt.** 2006. Pyrimidine and purine biosynthesis and degradation in plants. *Annu. Rev. Plant Biol.* **57**: 805-836

**Zuerner, R. L., J. L. Hermann, and I. Saint Girons.** 1993. Comparison of genetic maps for two *Leptospira interrogans* serovars provides evidence for two chromosomes and intraspecies heterogeneity. *J. Bacteriol.* **175**: 5445-5451