



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

“Análisis de una Genoteca construida a
partir de Hibridación Sustractiva entre
las cepas Ch24-10 y CFN 42 de
Rhizobium etli”

T E S I S
QUE PARA OBTENER EL TÍTULO DE
B I Ó L O G A
P R E S E N T A :

MARTHA GUADALUPE LÓPEZ GUERRERO



DIRECTORA:
DRA. ESPERANZA MARTÍNEZ ROMERO

2006



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

1. Datos del alumno
López
Guerrero
Martha Guadalupe
56 22 78 97
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
097114064
2. Datos del tutor
Dra
Esperanza
Martínez
Romero
3. Datos del sinodal 1
Dra
Mónica
Rosenblueth
Laguette
4. Datos del sinodal 2
Dr
Víctor Manuel
González
Zúñiga
5. Datos del sinodal 3
Dr
César
Rodríguez
Chávez
6. Datos del sinodal 4
M en C
Alfonso José
Vilchis
Peluyera
7. Datos del trabajo escrito
Análisis de una Genoteca construida a partir de
Hibridación Sustractiva entre las cepas Ch24-10 y CFN 42
de *Rhizobium etli*
96 p
2006

Este trabajo se realizó en el Centro de Ciencias Genómicas de la Universidad Nacional Autónoma de México, en el Programa de Ecología Genómica, en el periodo comprendido entre septiembre de 2004 y noviembre de 2006. Se realizó bajo la dirección de la Dra. Esperanza Martínez Romero y la Dra. Mónica Rosenblueth Laguette y contó con el apoyo de CONACYT 40997-Q.

AGRADECIMIENTOS

A la Dra. Esperanza Martínez por darme la oportunidad de ser alumna del Taller y de realizar el proyecto de tesis en el Programa que dirige. Gracias por contagiarme su entusiasmo por la ciencia, por la confianza y por motivarme a continuar con el proyecto.

A la Dra. Mónica Rosenblueth, por haber guiado mi trabajo en el laboratorio. Gracias por enseñarme y acompañarme en el proceso de la realización de la tesis. Le agradezco su amistad y todo el apoyo que me dio siempre, su confianza y su sinceridad.

A todos y cada uno de los integrantes del Programa de Ecología Genómica, por su compañerismo y su apoyo en el laboratorio.

A Marco Antonio Rogel, integrante del Programa, por la realización de la Hibridación Sustractiva y por las secuencias proporcionadas, así como por el apoyo en la realización del proyecto.

A Jorge Luis Ocampo Espíndola por su apoyo en el análisis de las secuencias.

A la M.C. Rosa Isela Santamaría Gutiérrez integrante del Programa de Genómica Evolutiva del CCG, por la realización de la comparación de las secuencias de *Rhizobium etli* Ch24-10 con las secuencias de las cepas de la misma especie.

Gracias a los sinodales, Dr. Víctor Gonzáles, Dr. César Rodríguez, M. en C. Alfonso Vilchis y Dra. Mónica Rosenblueth por el tiempo invertido en la revisión de este trabajo, por sus comentarios y correcciones que me permitieron fortalecer este trabajo.

DEDICATORIAS

A mi mamá, Guadalupe Guerrero, una mujer que es ejemplo de fortaleza, perseverancia y dedicación. Gracias por tu apoyo en todo momento, por estar conmigo siempre, por dejar que me equivocaré y por toda tu paciencia. Gracias por motivarme y ayudarme a conseguir mis metas. Por todo tu amor.

A mi Tío, Pedro Guerrero, un padre para mí, quien se ha preocupado por mí siempre. Gracias por confiar en mí y por motivarme a luchar por lo que quiero. Con tu ejemplo de dedicación y de perseverancia aprendí que uno puede lograr lo que desea. Gracias por tu apoyo incondicional y por toda la confianza que tienes en mí.

A mi abuela, Mary Mary por todo su cariño y todo su apoyo. Gracias por compartir conmigo tantas historias, por consentirme y por confiar en mí, gracias por quererme tanto siempre, gracias porque se que soy tu consentida aunque no les digas a los demás para no herir sus sentimientos. Gracias por todo tu cariño abue.

A mis tíos Oliva, Ángeles, y Alejandro por creer siempre en mí, por todas sus enseñanzas, por su cariño. Por los momentos buenos.

Gracias a mi Tío Luís Villeda, por todo el apoyo que siempre le ha dado a mi familia, a mi mama, a mis hermanas y a mí. Gracias por estar con nosotros en los momentos buenos y en los difíciles también.

A mis hermanas Laura y Fabiola. Gracias por cuidarme y por todo su amor. Ustedes me formaron y definieron gran parte de mi carácter, desde la infancia y hasta ahora. Las admiro muchísimo y las quiero.

Gracias a mi hermana Sandra, mi compañera de la infancia y mi gran amiga. Gracias herma porque me has enseñado muchas cosas en la vida, a entender que siempre tenemos algo bueno por lo que sentirnos bien y por lo que debemos estar contentos. Gracias por darme fortaleza, por escucharme siempre. Por tu amor. La vida contigo ha sido mejor, gracias por todo lo bueno que siempre me das a mí y a todos lo que tenemos la dicha de tenerte cerca.

Gracias a mis sobrinos porque ellos llenan de alegría mi vida, Lalo, Gaby, Lalito, Monste, Mónica, Diego. Ustedes son lo máximo. Los amo, aunque muchos sean más altos que yo. Siempre me inspiran y me hacen muy feliz.

Gracias a mi primo José Luís Guerrero, que además de mi primo es mi mejor amigo. Gracias por entenderme completamente, siempre. Por ver más allá de mis palabras, por entender lo que siento y por compartir. Por creer en mí y por estar siempre cerca.

Gracias a Miguel Ángel, mi cuñado, por ser mi mejor amigo, por escucharme, por aguantarme, por todos sus consejos, por preocuparte por mí, por motivarme siempre a dar más y a seguir creyendo.

A mis amigas de la prepa 5: Vanessa, Magali, Lulú y Angie con quienes viví de los mejores momentos en mi vida. Gracias por seguir siendo parte de ella. Gracias por permitirme compartir la vida con ustedes, por acompañarme en los momentos más felices y más tristes. Gracias por ser parte de mi vida. Gracias por las papas de la amistad.

Gracias Angélica por que siempre me recuerdas que la vida es hermosa y que esta tiene cosas maravillosas, gracias porque tu me recuerdas que existen personas excepcionales cerca de mí, como tú.

Gracias a mis mejores amigos de la Facultad: Ángel, Paulina, Miriam, Tere, Homero, Luis R., Cristina M. y José Luis, por todo su apoyo y por su confianza, por sus palabras de aliento siempre, por escucharme y por acompañarme. Gracias por seguir siendo mis amigos y por haber estado todo el tiempo conmigo. Gracias por estar conmigo en una de las mejores etapas de mi vida.

A Verónica, Yanet por ser mis amigas en Cuernavaca, por aceptarme como soy. Por todo su apoyo durante la realización de la tesis.

A Edith. Gracias por ser la amiga que eres, por las historias, por la Chavela jr. En muy poco tiempo te llegue a querer mucho y eres una gran amiga.

A Karla por mostrarme que siempre puedo ser mejor, por confiar en mi y apoyarme en momentos decisivos de mi vida, personal y académica. Gracias por hacerme ver el mejor lado de mí, por tu cariño, tu compañía, por mostrarme otro lado de la vida. Te quiero mucho.

“El que no posee el don de maravillarse
ni de entusiasmarse más le valdría estar muerto,
porque sus ojos están cerrados”

Albert Einstein

ÍNDICE

RESUMEN	1
1. INTRODUCCIÓN	
1.1 INTERACCIONES BENÉFICAS PLANTA-BACTERIA	2
1.1.1 LA RIZÓSFERA.....	2
1.1.2 FIJACIÓN BIOLÓGICA DE NITRÓGENO	4
1.1.2.1 SIMBIÓISIS LEGUMINOSA-RIZOBIO	6
<i>REGULACIÓN MOLECULAR</i>	8
1.1.3 RIZOBACTERIAS PROMOTORAS DEL CRECIMIENTO EN PLANTAS.....	10
<i>ASPECTOS MOLECULARES</i>	12
1.2 IMPORTANCIA DEL FRIJOL Y DEL MAÍZ	
1.2.1 EL FRIJOL	13
1.2.2 EL MAÍZ	14
<i>ASPECTOS AGRÍCOLAS</i>	15
2. ANTECEDENTES	
2.1 <i>Rhizobium etli</i>	18
3. JUSTIFICACIÓN	21
4. OBJETIVOS	
4.1 GENERALES	22
4.2 PARTICULARES	22
5. ESTRATEGIA EXPERIMENTAL	23

6. MATERIALES Y MÉTODOS

6.1 CEPAS, PLÁSMIDOS Y MEDIOS DE CULTIVO	25
6.2 PCR DE COLONIAS BACTERIANAS	26
6.3 ELECTROFORESIS / GELES DE AGAROSA	28
6.4 OBTENCIÓN DE SECUENCIAS	28
6.5 ANÁLISIS DE SECUENCIAS	28
6.6 OBTENCIÓN DE ADN TOTAL Y PLASMÍDICO	30
6.7 DIGESTION DE ADN CON ENZIMAS DE RESTRICCIÓN	30
6.8 GELES TIPO ECKHARDT	30
6.9 SOUTHERN BLOT	
6.9.1 TRANSFERENCIA DE ADN DE COLONIAS BACTERIANAS	31
6.9.2 TRANSFERENCIA DE ADN EN GEL	32
6.9.3 OBTENCIÓN DE SONDAS	32
6.9.4 HIBRIDACIÓN	33

7. RESULTADOS

7.1 SECUENCIAS CON IDENTIDAD AL GENOMA DE LA CEPA CFN42	35
7.2 SECUENCIAS SIN IDENTIDAD AL GENOMA DE LA CEPA CFN42	40
7.2.1 SECUENCIAS CON IDENTIDAD SIGNIFICATIVA A PROTEÍNAS PUTATIVAS O CONOCIDAS	42
7.2.2 LOCALIZACIÓN EN EL GENOMA	
<i>BANCO DE CÓSMIDOS</i>	48
<i>ADN TOTAL</i>	49
<i>PLÁSMIDOS</i>	50
7.2.3 COMPARACIÓN DE SECUENCIAS CON EL GENOMA DE OTRAS CEPAS DE <i>R. etli</i>	52

8. DISCUSIÓN Y CONCLUSIONES	
8.1 ANÁLISIS DE LAS SECUENCIAS CON IDENTIDAD AL GENOMA DE LA CEPA CFN42 .	54
8.2 ANÁLISIS DE LAS SECUENCIAS SIN IDENTIDAD AL GENOMA DE LA CEPA CFN42 .	55
8.2.1 LOCALIZACIÓN EN EL GENOMA	57
8.2.2. SECUENCIAS CON IDENTIDAD SIGNIFICATIVA A PROTEÍNA CON FUNCIÓN PUTATIVA O CONOCIDA / LOCALIZACIÓN EN EL GENOMA	58
<i>Secuencias relacionadas con fagos</i>	58
<i>Secuencias relacionadas con la replicación, recombinación y reparación del ADN</i>	58
<i>Reguladores transcripcionales</i>	62
<i>Transporte celular y Regulación de respuesta</i>	64
<i>Proteínas varias</i>	68
8.2.3 SECUENCIAS CON IDENTIDAD A OTRAS CEPAS DE <i>Rhizobium etli</i>	75
8.2.4 CONSIDERACIONES GENERALES	75
9. PERSPECTIVAS	76
10. ANEXO	
10.1 HIBRIDACIÓN SUSTRACTIVA.....	77
10.2 BASE DE DATOS	80
11. REFERENCIAS	80

INDICE DE FIGURAS

Figura 1. Fijación Biológica de Nitrógeno	5
Figura 2. Invasión de las raíces por <i>Rhizobium</i>	7
Figura 3. Estrategia experimental	24
Figura 4. Mapa del vector de clonación pCR 2.1-TOPO	27
Figura 5. Transferencia de ADN de colonias bacterianas	34
Figura 6. Hibridación con el banco de cósmidos	48
Figura 7. Hibridación con ADN total de las cepas CFN42 y Ch24-10	47
Figura 8. Hibridación con ADN plasmídico de las cepas CFN42 y Ch24-10	50
Figura 9. Hibridación sustractiva de ADN	79

INDICE DE TABLAS

Tabla 1. Cepas y plásmidos	25
Tabla 2. Componentes de la reacción de PCR	26
Tabla 3. Ciclos de la PCR	26
Tabla 4. Secuencias con identidad al genoma de <i>R. etli</i> CFN42	36
Tabla 5. Organización en grupos funcionales de las secuencias con identidad a proteínas putativas o conocidas	43
Tabla 6. Resultados del Southern blot de las 23 clonas seleccionadas	41
Tabla 7. Identidad de 87 secuencias de la librería de la cepa Ch24-10 con secuencias genómicas parciales de cuatro cepas de <i>R. etli</i> : Kim 5, Bra 5, 8C-3 y CIAT 894	52
Tabla 8. Secuencias de la librería genómica de <i>R. etli</i> Ch24-10 sin identidad al genoma de <i>R. etli</i> CFN42 comparadas mediante blastx	80

INDICE DE GRÁFICOS

Grafico 1. Longitud de los fragmentos clonados de la librería	40
Gráfico 2. Distribución porcentual de las 120 secuencias con respecto a la identidad que presentaron en el blastx	41
Gráfico 3. Distribución porcentual de las secuencias de identidad alta a proteínas putativas. Organización en grupos funcionales	42
Gráfico 4. Porcentaje de cobertura de las secuencias de aminoácidos de la cepa CH24-10 comparadas con las secuencias protéicas del ortólogo correspondiente	56

RESUMEN

Las bacterias pueden establecer relaciones de distintos tipos con las plantas, como patógenos o como organismos benéficos. En el segundo grupo encontramos bacterias que son capaces de colonizar el ambiente rizosférico, los tejidos de las plantas como endófitos y algunas que pueden formar nódulos fijadores de nitrógeno.

Por ejemplo, *Rhizobium etli* bv. phaseoli, simbiote específico del frijol es capaz de colonizar la rizósfera y los tejidos del maíz como endófito y promover su crecimiento. Este fenómeno al parecer ha sido favorecido la práctica agrícola conocida como milpa, en la que el frijol y el maíz son sembrados juntos.

Al estudiar la diversidad genética de las cepas de *R. etli* en campo se aisló la cepa Ch24-10, la cual presentó el genotipo dominante en suelo, rizósfera, raíz y además el único patrón en tallo de maíz. Esta cepa presenta algunas diferencias con la cepa tipo de *R. etli* CFN42, entre las que destacan, diferente perfil plasmídico, mayor competitividad para nodular frijol y mejor capacidad de fijar nitrógeno. Además, la cepa Ch24-10 promueve el crecimiento del maíz, es resistente a una fitoalexina secretada por éste e inhibe el crecimiento de hongos patógenos.

Para estudiar las diferencias genómicas entre *R. etli* CFN42 y *R. etli* Ch24-10, se generó una librería genómica o genoteca mediante hibridación sustractiva, la cual contiene el genoma exclusivo de la cepa Ch24-10 con respecto a la cepa tipo CFN42.

El objetivo de este trabajo fue secuenciar y analizar la genoteca, para conocer de manera aproximada la función de los genes únicos de la cepa Ch24-10 respecto a la cepa CFN42 que pudieran ser determinantes en la interacción con el maíz, la alta competitividad, la capacidad de promover el crecimiento y de inhibir el crecimiento de hongos patógenos. Además se propuso conocer su distribución y localización en el genoma e identificar la presencia de éstos genes en otras cepas de *R. etli*.

Se obtuvieron 189 secuencias distintas, 69 de ellas presentaron identidad al genoma de *R. etli* CFN42. Las 120 secuencias restantes (sin identidad al genoma de la cepa CFN42), tuvieron en promedio una longitud de 530 nucleótidos. Al analizar éstas 120 secuencias, 56 presentaron identidad significativa a alguna proteína anotada y 64 de ellas no.

Las 56 secuencias con identidad significativa se agruparon en dos categorías: a) con función putativa o conocida (41 secuencias), y b) hipotéticas (15 secuencias). Las clonas del primer grupo se organizaron en categorías de acuerdo a grupos funcionales por la función del gen ortólogo. Se identificaron algunas proteínas que pudieran participar en la colonización del maíz y la promoción de su crecimiento. Por ejemplo, una cisteína proteasa similar a ubiquitina, proteínas involucradas en la recepción y transmisión de señales (proteína con dominio EAL e histidina cinasa), proteínas relacionadas con la captura y el transporte de hierro (lisina-ornitina monooxigenasa y TonB) y proteínas involucradas en el quorum sensing (Regulador transcripcional de la familia RiorF y proteína similar a LuxA).

Finalmente, se seleccionaron 23 clonas para conocer su ubicación en el genoma y para localizarse en un banco de cósmidos con el fin de determinar su cercanía en el genoma, 7 clonas se localizaron en plásmido y 12 dieron señal con uno o varios cósmidos y algunas de ellas en el mismo. También se compararon 87 secuencias con el genoma de otras cepas de *R. etli*, 27 presentaron identidad significativa. El genoma incluido en la librería está distribuido en el cromosoma y en los tres plásmidos de *R. etli* Ch24-10 y existen elementos comunes con otras cepas de la especie.

1. INTRODUCCIÓN

1.1 INTERACCIONES BENÉFICAS PLANTA-BACTERIA

Las bacterias interactúan con las plantas de distintas maneras, bien como patógenos o como organismos benéficos. Dentro del tipo de interacciones benéficas encontramos la simbiosis entre rizobios y leguminosas, en la que las bacterias forman nódulos fijadores de nitrógeno en las raíces de las plantas. Otro tipo de interacción benéfica es el caso de las rizobacterias promotoras del crecimiento, PGPR's (por las siglas ingles, Plant growth promoting rhizobacteria), que pueden promover el crecimiento de las plantas mediante distintos mecanismos, colonizando el ambiente rizosférico, y también colonizando los tejidos de las plantas como endófitos.

1.1.1 LA RIZÓSFERA

El volumen de suelo alrededor de las raíces que es influenciado directamente por su actividad se ha denominado rizósfera (Hisinger *et al.*, 2005). El carbono total que es transferido al suelo por las raíces se conoce como rizodeposición y se estima que entre el 5% al 20% del carbono generado por fotosíntesis es secretado a la rizósfera, aunque se ha estimado que en condiciones de estrés nutricional en la planta, puede llegar a superar el 40% (Nguyen, 2003; Walker *et al.*, 2003; Singh *et al.*, 2004).

De manera general, los compuestos que secretan las raíces se conocen como exudados y constituyen un amplio rango de compuestos orgánicos que incluyen azúcares, polisacáridos, aminoácidos, ácidos orgánicos, péptidos, proteínas y enzimas, lisados provenientes de las células muertas de las plantas, mucílago, compuestos fenólicos y otros metabolitos secundarios (Rosenblueth *et al.*, 2001).

A la rizodeposición se le han atribuido varias funciones en el suelo como el mantenimiento del contacto suelo-raíz, la lubricación de las puntas de los pelos radiculares, protección contra la desecación, la estabilización de los microagregados del suelo, la absorción y el almacenamiento selectivo de iones (Walker *et al.*, 2003).

La rizósfera es colonizada por bacterias, hongos, protistas y nemátodos, los cuales reciben protección contra la desecación, la temperatura y el estrés por luz (Brencic & Winans, 2005). Los microorganismos en la rizósfera realizan funciones importantes como la descomposición de la materia orgánica, además mantienen la estructura del suelo y las relaciones hídricas (Singh *et al.*, 2004).

Para que se establezca una interacción planta-microorganismo es necesario que exista un reconocimiento mutuo y especificidad entre el huésped y el hospedero (Brencic & Winans, 2005). Los exudados actúan como mensajeros que permiten el reconocimiento entre los participantes de la interacción y el inicio de las interacciones biológicas y fisiológicas entre las raíces y los microorganismos (Walker *et al.*, 2003). La rizodeposición regula las comunidades bacterianas que se encuentran en la rizósfera ya que son seleccionados aquellos grupos taxonómicos y funcionales que son más competitivos para colonizarla (Singh *et al.*, 2004).

Las bacterias colonizan a las plantas a través de aberturas que se forman al emerger las raíces laterales o en la zona meristemática apical de la raíz. (rev. en Rosenblueth & Martínez-Romero, 2006). La colonización, está limitada por distintos factores tanto de la planta como de la bacteria. Entre los primeros se encuentran el estado fisiológico, el genotipo y el estado de desarrollo (Rosenblueth & Martínez-Romero, 2006). En cuanto a las bacterias, existen estrategias que les permiten colonizar de forma competitiva entre las que se encuentran: elevada tasa de crecimiento, movilidad, capacidad de catabolizar compuestos específicos de los exudados, capacidad de adaptarse a condiciones ambientales cambiantes (en el suelo), producción de sustancias que le permiten adherirse y/o penetrar la raíz, resistencia a antibióticos y la capacidad de detoxificar otras sustancias dañinas presentes en los exudados, síntesis de sideróforos y sustancias antimicrobianas (Rosenblueth *et al.*, 2001; Compant *et al.*, 2005).

Otros factores que afectan la colonización son el tipo de suelo, la concentración de nitrógeno y carbono, así como la presencia de herbicidas y fertilizantes. Se ha visto que la colonización se reduce a altas concentraciones de nitrógeno porque esto origina una disminución de la sacarosa, fuente de carbono para las bacterias (Rosenblueth & Martínez-Romero, 2006).

1.1.2 FIJACIÓN BIOLÓGICA DE NITRÓGENO

El nitrógeno es un elemento esencial para el mantenimiento de la vida, es un componente fundamental de los ácidos nucleicos, de las proteínas y de otros componentes celulares. En la atmósfera es el componente mayoritario, su forma gaseosa (N_2) comprende aproximadamente el 78% de los gases que la conforman, a pesar de esto, el nitrógeno atmosférico no puede ser utilizado directamente por la mayoría de los organismos ya que entre los dos átomos del N_2 existe un triple enlace, lo que convierte en un elemento muy estable y relativamente inerte, de esta manera, el nitrógeno disponible es generalmente escaso en los ecosistemas, lo que lo convierte en un factor limitante en el crecimiento de las plantas. Para que éstas puedan usar el nitrógeno, el N_2 tiene que ser convertido a una forma química disponible como el amonio (NH_4^+), el nitrato (NO_3^-), o el nitrógeno orgánico (por ejemplo la urea: $(NH_3)_2CO$) (Chen *et al.*, 2003; Galloway *et al.*, 1995; Jurtshuk, 1996)

La Fijación Biológica del Nitrógeno (FBN) es una de las vías por las cuales el nitrógeno se incorpora al ciclo biogeoquímico global, transformándolo en una forma asimilable por las plantas (Figura 1). Este proceso permite que el nitrógeno sea repuesto continuamente al suelo, ya que se libera constantemente a la atmósfera por el proceso denominado desnitrificación (Chen *et al.*, 2003; Ludwig & Poolé, 2003).

La FBN es una actividad que realizan exclusivamente organismos procariontes (eubacterias y arqueobacterias metanogénicas) (Elmerich 2001), bien por organismos de vida libre como por bacterias que se asocian simbióticamente con plantas. En las primeras destacan los rizobios y en las segundas las leguminosas (Lodwig & Poolé, 2003; Mylona *et al.*, 1995).

Los rizobios pertenecen al orden Rhizobiales de las α -proteobacterias. Se agrupan en los géneros *Rhizobium*, *Sinorhizobium*, *Bradyrhizobium*, *Azorhizobium* y *Mesorhizobium* (Lodwig & Poolé, 2003; Spaink, 2000; Trinchant *et al.*, 2001). Estas bacterias son capaces de sobrevivir y reproducirse en el suelo (Denison & Kiers, 2004), pero son mas conocidos por que tienen la capacidad única de inducir la formación de estructuras especializadas en las raíces de la planta denominadas nódulos fijadores de nitrógeno (Lodwig & Poolé, 2003; Spaink, 2000; Trinchant *et al.*, 2001).

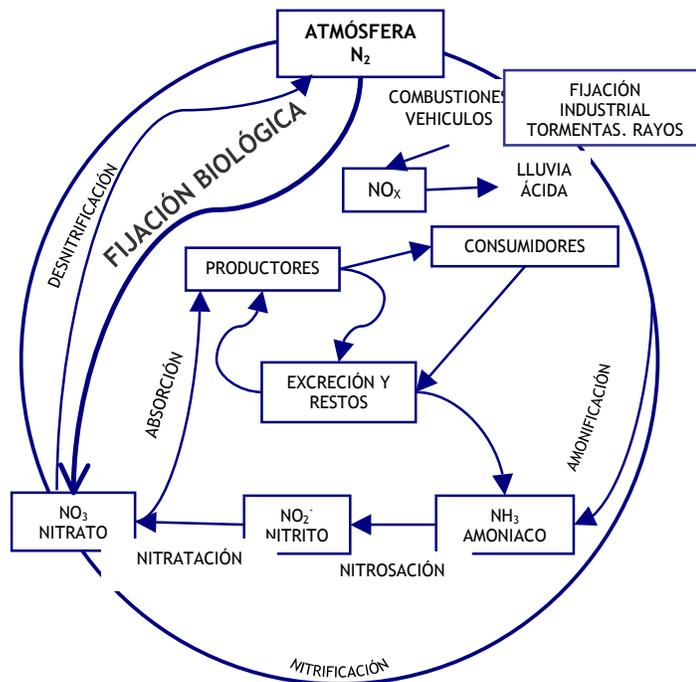


Figura 1. Ciclo Global del Nitrógeno.

La integración del N_2 al ciclo global del nitrógeno no es un proceso exclusivamente biológico, los rayos de las tormentas proporcionan aproximadamente el 10% del nitrógeno fijado en la tierra, mientras que la fertilización industrial (producción de fertilizantes químicos) provee de aproximadamente el 25% del total mundial. Sin embargo, Los procesos biológicos aportan el 60% aproximadamente del nitrógeno fijado en todo el mundo. (Chen *et al.*, 2003; Zahran, 1999)

1.1.2.1 SIMBIÓISIS LEGUMINOSA-RIZOBIO

Simbiosis puede ser definida como un aumento recíproco en la adecuación, es decir, el efecto que tiene alguno de los participantes hacia el otro que permite que aumente su posterior supervivencia o reproducción (Selosse *et al.*, 2005).

La simbiosis leguminosa-rizobio comienza con dos organismos de vida libre y termina con una íntima coexistencia celular (Long, 2001). El proceso comienza con la infección bacteriana de los pelos radiculares y la inducción del curvamiento de éstos, a continuación las bacterias son invaginadas por la pared celular de la planta, posteriormente se induce la formación del nódulo y se forma el hilo de infección que va creciendo a través de las células vegetales hasta alcanzar al primordio del nódulo. Las bacterias se dividen en el hilo de infección y permanecen inicialmente en el exterior de la célula vegetal, a continuación son liberadas por endocitosis al nódulo que ha terminado de formarse (Figura 2).

Dentro de los nódulos se forman las estructuras simbióticas fundamentales, gracias a la diferenciación de la bacteria y de las membranas celulares de la planta. Esta estructura es denominada unidad peribacteroidal o simbiosoma y está constituida por la membrana peribacteroidal (derivada de la planta), el bacteroide (bacteria diferenciada capaz de fijar nitrógeno) y el espacio peribacteroidal (Figura 2) (Lodwig & Poolé, 2003; Perret *et al.*, 2000; Schultze & Kondorosi, 1998).

Existen dos tipos estructurales de nódulos, determinados e indeterminados. Los nódulos determinados tienen un meristemo no persistente, son de forma redonda y tienen varios bacteroides contenidos en un simbiosoma, este tipo de nódulos son los que se forman en plantas como la soya y el frijol. Los nódulos indeterminados, por el contrario, tienen un meristemo persistente, son de forma cilíndrica y contienen un bacteroide por simbiosoma, este tipo lo encontramos en las plantas de alfalfa, trébol y chícharo (Lodwig & Poolé, 2003; Schultze & Kondorosi, 1998; Skorupska *et al.*, 2006).

La FBN es un proceso altamente sensible al oxígeno, por lo que los niveles de éste elemento tienen que ser regulados. En el nódulo existe una barrera de difusión a este elemento, que tiene como resultado una baja concentración de oxígeno libre alrededor del bacteroide, la leghemoglobina es la proteína que transporta el oxígeno a los bacteroides y controla los niveles de este elemento dentro del nódulo (Mylona *et al.*, 1995; Wang *et al.*, 2001, Brencic & Winans, 2005,).

Dentro del nódulo, los bacteroides utilizan a la enzima nitrogenasa para catalizar la reducción del N₂ a amonio. La reacción de reducción del nitrógeno es la siguiente (Elmerich, 2001; Trinchant *et al.*, 2001):



El amonio es el primer compuesto primario estable de la fijación biológica (Lodwig & Poolé, 2003), éste es secretado a la planta para su incorporación en los aminoácidos glutamina y asparagina (especialmente en las plantas con nódulos indeterminados), o en ureidos, derivados de las purinas (mayoritariamente en leguminosas con nódulos determinados) (Juergen & Poolé, 2006). La planta proporciona al bacteroide la fuente de carbono necesaria para la fijación del nitrógeno. El carbono suministrado es producto de los fotosintatos que son transportados por el floema hacia el nódulo en forma de sacarosa, la cual entra al nódulo y se transforma en ácidos dicarboxílicos, compuestos considerados la fuente primaria de carbono para los bacteroides *in planta* (Lodwig & Poolé, 2003).

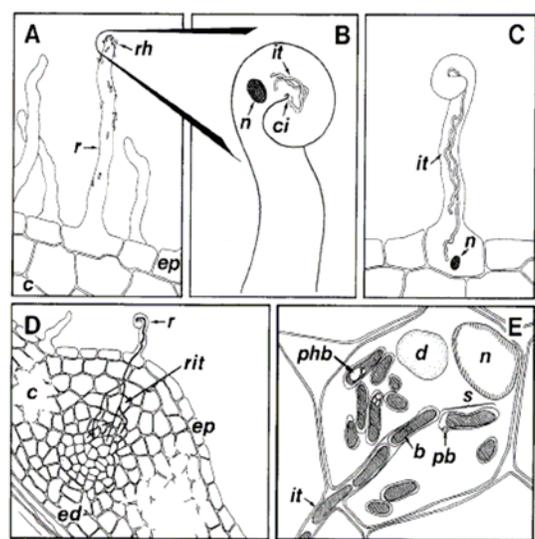


Figura 2. Invasión de las raíces por *Rhizobium*.

A. El rizobio (*rh*) coloniza la rizósfera y se une a los pelos radiculares (*r*). B. Los factores Nod inducen el curvamiento del pelo radicular permitiendo la entrada de las bacterias al centro de infección (*ci*). Los núcleos de la planta (*n*) preceden al crecimiento del hilo de infección (*it*). C. Aún acompañadas de los núcleos (*n*) un creciente hilo de infección (*it*) alcanza la base del pelo radicular. D. El hilo de infección se ramifica (*rit*) cerca del primordio del nódulo formado por la división de las células corticales. E. Los bacteroides (*b*) son liberados del hilo de infección (*it*) y forman los simbiosomas (*s*) (tom. de Perret *et al.*, 2000).

REGULACIÓN MOLECULAR

De manera general, se propone que existen diferentes tipos de señales simbióticas presentes durante la formación de los nódulos. El primer grupo lo componen los flavonoides, que son secretados por las raíces de la planta e interactúan con las proteínas NodD, activadores transcripcionales bacterianos de la familia LysR, que funcionan como sensores ambientales de las señales de la planta.

Otro grupo de señales es sintetizado cuando el complejo flavonoide-NodD activa la transcripción de los genes *nod*. El producto de estos genes, los factores Nod, provocan la deformación de los pelos radiculares, la formación del primordio del nódulo y permiten la entrada del rizobio a la raíz a través del hilo de infección. Los genes *nod* se encuentran organizados en varios operones en el genoma bacteriano (en el cromosoma o en los plásmidos simbióticos), y son regulados transcripcionalmente por las cajas *nod* que los preceden (Broughton *et al.*, 2000; Schultze & Kondorosi, 1998).

Los factores Nod están conformados por un esqueleto de tres a cinco β -1,4-*N*-acetilglucosamina con ácidos grasos unidos en el residuo del azúcar no reducido. Además tienen varias substituciones en los residuos reducidos y no reducidos de los azúcares terminales, constituyen uno de los determinantes de la especificidad del hospedero (Mylona *et al.*, 1995).

Otro grupo de compuestos importantes en el establecimiento de la infección y la invasión del nódulo son componentes celulares externos, como las proteínas del sistema de secreción tipo III y los glicanos de superficie celular (exopolisacáridos (EPS), lipopolisacáridos (LPS), antígenos K y glucanos cíclicos). Los EPS son importantes en la infección del hospedero; los LPS tienen un papel primordial en estadios tardíos de la invasión del nódulo, en la liberación de las bacterias del hilo de infección y en el desarrollo del simbiosoma, además están involucrados en la supresión de la respuesta de respuesta de la planta, al igual que los antígenos K y los glucanos cíclicos (Broughton *et al.*, 2000; Schultze & Kondorosi, 1998).

Una vez que se ha formado el nódulo, la fijación del nitrógeno es el proceso fisiológico central y está dirigido por dos grupos principales de genes: *nif* y *fix*, que son inducidos por el regulador transcripcional NifA y por el factor sigma RpoN (Mylona *et al.*, 1995).

Los genes *nif* codifican para las proteínas que constituyen a la enzima nitrogenasa (*nifHDK*), para las enzimas involucradas en la biosíntesis del cofactor de la nitrogenasa Fe-Mo (*nifENB*), para el regulador transcripcional NifA, y para proteínas que son requeridas para la actividad completa de la nitrogenasa (*nifSWX*) (Brensic & Winans, 2005).

Se propone que los genes *fixABCX* codifican para las proteínas que intervienen en una cadena de transporte de electrones, que genera energía para la nitrogenasa. La función del grupo *fixGHIS* es desconocida, mientras que el grupo *fixNOQP* codifica para el citocromo oxidasa membranal que es requerido para la respiración del rizobio en condiciones ambientales de bajo oxígeno. Los genes *fixL*, *fixJ* y *fixK* codifican para proteínas reguladoras (Brensic & Winans, 2005).

La interacción simbiótica en la FBN muestra un alto grado de especificidad, a cada estado del desarrollo del nódulo, corresponden programas genéticos específicos que son regulados recíprocamente. La colonización, crecimiento y desarrollo de las bacterias y de las células vegetales debe ser coordinado, y es regulado por la comunicación entre planta y bacteria (Schultze & Kondorosi, 1998). Un ejemplo de que la regulación entre huésped y hospedero es tan especializada se observa en la simbiosis de la alfalfa y *Sinorhizobium meliloti*, en donde los bacteroides tienen un estado de diferenciación celular similar al de las células de la planta (poliploides y de tamaño grande, respecto a las células de otros tejidos o a las bacterias no diferenciadas en bacteroides), fenómeno que es controlado por factores producidos por la planta, presentes solo en los nódulos de tipo indeterminado (Mergaert *et al.*, 2006).

1.1.3 RIZOBACTERIAS PROMOTORAS DEL CRECIMIENTO EN PLANTAS

Se ha reconocido ampliamente que las bacterias que se asocian a las plantas tienen un efecto benéfico en ellas, ya que estimulan su crecimiento y las protegen contra organismos patógenos. Este grupo de bacterias se ha denominado Bacterias promotoras del crecimiento en plantas (PGPR, del inglés, Plant growth promoting rhizobacteria). Las PGPR resultan de interés agronómico ya que permiten el aumento en el rendimiento de los cultivos y la disminución del uso de agroquímicos (que tiene como resultado la disminución de los efectos nocivos y de los gastos económicos que generan, dado el aumento constante de los costos, en particular en regiones no desarrolladas en el mundo). Además, debido a su uso potencial como control biológico, son una alternativa al uso de pesticidas, debido a la demanda actual de productos agrícolas libres de éstos agroquímicos (Kennedy *et al.*, 2004; Compant *et al.*, 2005).

El incremento en el rendimiento del cultivo es medido en parámetros agronómicos, como el aumento de los siguientes procesos: porcentaje de germinación, altura de la planta, área foliar, contenido de clorofila, biomasa, longitud de la raíz, contenido de nitrógeno (Cleyet-Marel, 2001).

Las PGPR colonizan tanto la superficie de la raíz vegetal como la rizósfera, además son capaces de llegar al interior de las raíces y establecer poblaciones endofíticas sin causarle daño a la planta. Algunas de ellas pueden pasar la barrera endodérmica y atravesar el córtex radicular hacia el sistema vascular y establecerse como endófitos en tallo, hojas y frutos. La capacidad de las bacterias de colonizar diferentes tejidos refleja su habilidad para adaptarse a diferentes nichos, de esta manera pueden formarse asociaciones íntimas entre ellas y su hospedero (Compant *et al.*, 2005).

En general, las bacterias rizosféricas y las endofíticas utilizan mecanismos similares para promover el crecimiento de las plantas y para el control de fitopatógenos. Entre los procesos involucrados en el aumento del crecimiento se encuentran la producción de fitohormonas y vitaminas, el aumento de la toma de nutrientes como fósforo y potasio (por la solubilización de estos o por un efecto en la morfología y fisiología de las raíces), producción de sideróforos que aumentan la toma de hierro (Compant *et al.*, 2005; Kennedy *et al.*, 2004; Rosenblueth & Martínez-Romero, 2006).

La supresión de patógenos ocurre mediante la inducción de la resistencia sistémica de la planta, la competencia en la colonización y la producción de sustancias antimicrobianas (Compant *et al.*, 2005; Kennedy *et al.*, 2004; Rosenblueth & Martínez-Romero, 2006).

Algunas PGPR (rizosféricas como endófitas), también son capaces de promover el crecimiento gracias a la FBN. Dentro de este grupo de bacterias se ha descrito a *Azotobacter chroococcum*, *Clostridium spp.*, *Azospirillum spp.*, *Herbaspirillum seropedicae*, *Burkholderia vietamensis* y *Acetobacter diazotrophicus* (Kennedy *et al.*, 2004).

La capacidad competitiva de las PGPR es un punto importante que debe tomarse en consideración cuando se quiere utilizar de manera eficiente en campo. Las bacterias deben sobrevivir en la rizósfera, utilizar los nutrientes exudados por la planta, proliferar, ser capaces de colonizar el sistema radicular completo de la planta y de competir efectivamente con los microorganismos nativos (Bloemberg & Lugtenber, 2001).

APECTOS MOLECULARES

Las bacterias tienen un patrón de expresión genético particular cuando interactúan con la planta, desde la colonización hasta la sobrevivencia dentro de ella y por supuesto aquellos involucrados en la promoción del crecimiento (Rosenblueth & Martínez-Romero, 2006).

Varios genes son importantes para colonizar la rizósfera y la superficie de las raíces, por ejemplo, los que están involucrados en la síntesis de los LPS, (particularmente importante el antígeno O en la adherencia y la evasión del sistema de defensa de la planta) (Lerouge & Vanderleyden, 2001). Otros genes importantes son los involucrados en la capacidad metabólica de sintetizar aminoácidos (con un efecto en la tasa de crecimiento) y aquellos involucrados en la formación del flagelo (importante en la motilidad y la quimiotaxis) (Simons *et al.*, 1996).

Recientemente, se ha identificado en *Pseudomonas putida*, un grupo diverso de genes que son expresados en la rizósfera del maíz, involucrados en funciones probables como la formación de la envoltura celular, quimiotaxis y motilidad, transporte, secreción, mecanismos de defensa, regulación, metabolismo energético, estrés, detoxificación y síntesis de proteínas (Ramos-González et al, 2005).

En el caso de los endófitos, se sabe que las bacterias expresan genes relacionados con la asimilación del nitrógeno, metabolismo del carbono, crecimiento y mecanismos de defensa (Rosenblueth & Martínez-Romero, 2006)

1.2 IMPORTANCIA DEL FRIJOL Y DEL MAÍZ

1.2.1 EL FRIJOL

El frijol (*Phaseolus vulgaris* L.) es uno de los cultivos más antiguos en América y es una de las leguminosas más consumidas mundialmente (Mora-Núñez, 1997), constituye la fuente más importante de proteína en la dieta de los países latinoamericanos y del Este de África (Broughton *et al.*, 2003). Actualmente se produce en los cinco continentes, siendo África y América los mayores consumidores (Mora-Núñez, 1997), mientras que los mayores productores del mundo son Brasil, China, India, México y EUA (Martínez-Romero, 2003).

Gracias a estudios morfológicos, a la tipificación genética de la faseolina y al estudio de isoenzimas, se ha establecido que existen dos centros de diversificación genética de esta leguminosa: Mesoamérica (México, Centroamérica y Colombia), y la región de los Andes en Sudamérica (Ecuador, Perú y Argentina) (Aguilar *et al.*, 2004).

Se ha propuesto que la domesticación del frijol ocurrió hace aproximadamente 7,000 a 10,000 años y que fue un proceso múltiple e independiente en los dos centros de origen (Graham & Ranalli, 1997; Martínez-Romero, 2003). En México, en el Valle de Tehuacán en Puebla se han encontrado restos fósiles con una antigüedad de aproximadamente 7,000 años (Mora-Núñez, 1997). En Mesoamérica, el maíz y el frijol son cultivados comúnmente en milpa, un sistema agrícola tradicional en el que se siembran juntos y en el cual, el maíz funciona como el soporte en el que el frijol se enreda. Se propone que mediante este sistema fueron domesticados juntos hace más de 4,000 años (Martínez-Romero, 2003).

El uso de la milpa proporciona ventajas agrícolas, como la reducción de riesgos al tener más de un cultivo en el campo al mismo tiempo, se aprovecha mejor la superficie usada para cultivo y puede contribuir al control de la humedad y reducir la expansión de patógenos (Graham & Ranalli, 1997), además se ha propuesto que promueve la fijación del nitrógeno en el frijol, lo que tiene efectos positivos tanto en el frijol como en el maíz (Crews & Peoples, 2004; Martínez-Romero, 2003).

1.2.2 EL MAÍZ

El maíz (*Zea mays*) pertenece a la familia de las gramíneas y junto con el trigo y el arroz, pertenece al grupo de cereales importantes en el mundo. El maíz y el trigo juntos constituyen más de la mitad de los cereales que se consumen mundialmente. El maíz representa la mayor fuente de energía, proteína y forraje en todos los continentes (Jenkins *et al.* 1998), es fuente de alimento tanto para el hombre como para distintos animales y constituye una materia prima básica de la industria de transformación (producción de almidón, aceite, proteínas, bebidas alcohólicas, edulcorantes y combustible) (FAO, 1993).

En América Latina es el cereal de consumo principal, su cultivo se realiza en ambientes muy diversos, en altitudes que van desde el nivel del mar hasta más de 3,000 m de altura, en temperaturas bajas o altas y regímenes de humedad contrastantes, en ambientes semiáridos y húmedos (CIMMYT, 2001).

Se considera que el maíz fue una de las primeras plantas cultivadas por el hombre hace entre 7 000 y 10 000 años y que México es el centro de domesticación (al igual que el frijol). La evidencia más antigua del uso del maíz como alimento para el hombre se encontró en México, en el valle de Tehuacán, Puebla, y tiene una antigüedad de 5 000 años (FAO, 1993; Berthaud, 2001; Paliwal *et al.*, 2001).

Es ampliamente aceptado que el maíz se originó a partir del teosintle (*Euchlaena mexicana*), esta teoría que se basa en evidencia fósil, en que ambos poseen 10 cromosomas que son homólogos casi en su totalidad y en el retrocruzamiento reiterado que existe en la actualidad entre las dos plantas en algunas zonas de México y Guatemala, donde el teosintle crece entre los cultivos de maíz (Paliwal, *et al.*, 2001; FAO, 1993).

En México el maíz es muy diverso, se estima que tan solo en el estado de Oaxaca existen más de de 15 variedades criollas (variedad local producida a través del tiempo mediante la selección que hacen los mismos agricultores) (Berthaud, 2001).

ASPECTOS AGRÍCOLAS

En las últimas décadas, la producción y el consumo de los granos básicos (como el maíz, el frijol, el trigo y el sorgo), en México y en Centroamérica han sido afectadas por cambios políticos, sociales y económicos, como la integración de las economías en el mundo, el final de los conflictos armados internos y la implementación de programas de ajustes estructurales en estas dos regiones (Sain & López-Pereira, 1999).

Como parte de la Reforma agraria, originada por la Revolución mexicana, se realizó el reparto de tierras expropiadas en la forma legal de ejido que establece la tierra como propiedad común. El sistema de repartición ubicó a los campesinos en las áreas más pobres, donde básicamente practican agricultura de subsistencia. Las mejores tierras, provistas de riego se repartieron principalmente entre agricultores comerciales (Fraser & Restrepo-Estrada, 1996).

Gracias a los subsidios otorgados por el gobierno, hasta algunas décadas después de la Reforma agraria, las condiciones para los agricultores fueron favorables, respecto a los precios de producción (Fraser & Restrepo-Estrada, 1996). Posteriormente, con la reducción o eliminación de los subsidios, las crisis económicas (desregulación del tipo de cambio e inflación) y a los programas de apertura comercial (Tratados de libre comercio), el sector agrícola se enfrentó a una crisis importante ya que los costos de producción se elevaron y el precio de los productos se redujo (Fraser & Restrepo-Estrada; Sain & López-Pereira, 1999).

Además de estos factores, la crisis en el campo mexicano fue mayor debido a la reducción de la inversión pública en este sector (Fraser & Restrepo-Estrada, 1996). Este conjunto de elementos generó que la producción agrícola se estancara. En el norte del país mejoró la situación agrícola a partir del uso de las innovaciones tecnológicas implementadas a partir de la Revolución verde (uso de fertilizantes y pesticidas, monocultivos en grandes extensiones agrícolas) y por la construcción de obras hidroagrícolas realizadas entre 1940 y 1950 como parte de un programa gubernamental (Vega-Valdivia & Ramírez-Romero, 2004).

En el caso específico del maíz, para finales de 1960 México no era autosuficiente en su producción y el rendimiento nacional promedio era muy bajo, de 1 tonelada/ hectárea (tn/ha). A finales de esta década y principios de 1970 se comenzó a importar y el maíz generado por los campesinos era

fundamentalmente para el autoconsumo. La producción aumentó solamente entre los productores campesinos y empresariales que manejaban riego, que utilizaban variedades de alto rendimiento y que utilizaron paquetes tecnológicos (agroquímicos como fertilizantes y plaguicidas) (Vega-Valdivia & Ramírez-Romero, 2004).

En la década de 1980 se implementaron programas orientados a fomentar la producción de maíz, sin embargo, la importación continuó siendo la fuente principal de materia prima para la industria.

Mundialmente se siembran 140 millones de hectáreas con maíz, el 60% (96 millones), se encuentra en países en desarrollo y más de la mitad de esta área se localiza en cuatro países: China (26 millones), Brasil (12 millones), México (7.5 millones) e India (6 millones) (Pingali, 2001). Del 60% que se produce, solamente el 40% es consumido ahí. El bajo rendimiento en campo en los países en desarrollo es responsable de la diferencia entre la producción y el consumo mundial (Pingali, 2001).

El promedio de rendimiento en países industrializados es mayor de 8 t/ha (toneladas / hectárea), mientras que en los países en desarrollo es tan solo de 3 t/ha. Esta diferencia de 5 t/ha se debe a que existen distintas condiciones climáticas, así como al uso de distintas tecnologías agrícolas como el uso de riego, fertilizantes y plaguicidas (Pingali, 2001).

En México, la superficie sembrada con maíz pasó de 7.6 millones de ha en 1980 a 8.3 millones en 2002. Los rendimientos globales de maíz en t/ha pasaron de 1.629 en 1988 a 2.578 t/ha en el 2002, un crecimiento de casi una tonelada en el periodo. En 1991 el rendimiento pasó la barrera de las 2 toneladas por ha. En el norte del país, específicamente en Sinaloa, se cultiva maíz de riego con semillas híbridas de alto rendimiento que tienen un rendimiento promedio de 8 toneladas por ha (Vega-Valdivia & Ramírez-Romero, 2004).

La fertilización con nitrógeno ha sido una de las herramienta más utilizadas para incrementar el rendimiento de los cultivos, especialmente en los cereales como el maíz, el trigo y el arroz (Hirel *et al.*, 2001). El maíz requiere entre 9 y 11 Kg. de N para producir 1 tn de biomasa y considerando que la mayoría de los suelos en el mundo son deficientes en nitrógeno, debe aplicarse fertilizantes nitrogenados para obtener un rendimiento óptimo en su cultivo. Generalmente, la urea es la

fuentes de nitrógeno más útiles, sin embargo, menos del 50% es utilizado por las plantas, la baja eficiencia de su uso es debida a la volatilización del NH_3 , la desnitrificación y a la lixiviación (Kennedy *et al.*, 2004)

La volatilización y la desnitrificación contaminan la atmósfera se forman gases de invernadero como N_2O , NO y NH_3 , por otro lado, mientras que la lixiviación del $\text{NO}_3^- \text{N}$ origina toxicidad en el agua subterránea. Además de los problemas ambientales que genera la fertilización con nitrógeno, el uso prolongado de urea en los sistemas agrícolas genera una disminución del contenido de materia orgánica en el suelo. A la contaminación ambiental y a la degradación del suelo en los sistemas agrícolas debe considerarse, como se mencionó en un capítulo anterior, que el costo de este tipo de insumos agrícolas tiene un precio elevado que repercute directamente en los costos de producción (Kennedy *et al.*, 2004). Por lo anterior se ha propuesto que una alternativa a la aplicación de fertilizantes nitrogenados es el uso de las Rizobacterias promotoras del crecimiento (Martínez-Romero, 2003).

2. ANTECEDENTES

2.1 *Rhizobium etli*

Rhizobium etli bv. *phaseoli* se ubica filogenéticamente en la subdivisión alfa de las Proteobacterias dentro del Orden Rhizobiales, Familia Rhizobiaceae, en el grupo de *Rhizobium/Agrobacterium*. Es una bacteria gram negativa, aeróbica, con forma bacilar. La especie fue propuesta en 1993 a partir de la reclasificación de las cepas tipo I de *Rhizobium leguminosarum* biovar *phaseoli*. (Segovia *et al.*, 1993).

El frijol es nodulado por *R. etli*, especie que forma la mayoría de los nódulos en esta leguminosa en México (Segovia *et al.*, 1993). Además de *R. etli*, en México se ha aislado de nódulos de frijol a *R. galicum* (que también se ha encontrado en Francia y Austria). El frijol en otros países es nodulado por diferentes especies de rizobios como *R. tropici* (Brasil, Colombia, Francia, Moroco, Kenya, Senegal y Gambia), *R. leguminosarum* bv. *phaseoli* (Inglaterra, Francia, España, Colombia, Brasil, Tunisia) y *R. giardinii* (Francia, España y Brasil) (Martínez-Romero, 2003)

La rotación de cultivos entre plantas leguminosas y no leguminosas promueve que los rizobios simbioses de las primeras colonicen a las segundas. Por ejemplo, en la rotación de cultivos en Egipto entre el trébol y el arroz *Rhizobium leguminosarum* bv. *trifolii* simbiote del trébol, es un endófito del arroz del cual promueve su crecimiento. Otro caso se encuentra en la rotación de cultivos entre *Sesbania rostrata* con el arroz, el trigo, la cebada y la canola, en el que *Azorhizobium caulinodans*, simbiote de *S. rostrata* fue encontrado como endófito de estos cultivos (rev. en Gutierrez-Zamora & Martínez-Romero, 2001)

En el caso de la milpa, el sembrar al maíz y al frijol juntos permite que las raíces crezcan juntas y se mezclen, de esta manera se comparte el ambiente rizosférico. Dados los antecedentes, en el grupo de trabajo se propuso saber si *R. etli* era capaz de colonizar al maíz, en el 2001 se publica el trabajo en el que se encuentra que además de ser el simbiote específico del frijol, *R. etli* es un endófito natural del maíz y es capaz de promover su crecimiento (Gutiérrez-Zamora & Martínez-Romero, 2001).

En este trabajo se obtuvieron aislados de nódulo de frijol, rizósfera y raíz de maíz. Se encontró que hay un conjunto común de tipos cromosomales de *R. etli* en el frijol y en el maíz y que existen diferentes cepas que colonizan a una misma planta. Los aislados presentaron el mismo ribotipo (obtenido por patrones de restricción del gen 16S ribosomal) que presentan las cepas de referencia de *R. etli* CFN42 y Bra5.

Algunos de estos aislados no tienen la capacidad para nodular al frijol y no se identificó la presencia del gen *nif* en su genoma, por lo que estos resultados permitieron inferir que los genes *nif* no son necesarios para colonizar al maíz.

Cuando se realizaron ensayos de colonización se observó que los aislados de maíz son más competitivos para colonizar al cereal respecto a la cepa tipo CFN42 y promueven el crecimiento del maíz entre el 10% al 100%, dependiendo de la variedad que se utilizó. No se encontró que *R. etli* fuera capaz de fijar nitrógeno midiendo este proceso indirectamente mediante el ensayo de reducción de acetileno (Gutiérrez-Zamora & Martínez-Romero, 2001).

En un trabajo posterior, se analizaron aislados de *R. etli* provenientes de suelo, nódulos del frijol, rizósfera e interior del tallo de maíz. La cepa **Ch24-10** fue presentó el único perfil plasmídico y tipo electroforético en tallo obtenido mediante MLEE (del inglés, Multilocus enzyme electrophoresis), y el más abundante en raíz y en rizósfera. Se identificó la presencia del gen *nif* en esta cepa, el cual se localizó en el plásmido b, que se definió como plásmido simbiótico (Rosenblueth & Martínez-Romero, 2004).

La existencia de tipos electroforéticos y perfiles plasmídicos dominantes en el tallo y en la raíz, como el que presenta la cepa Ch24-10, permite proponer que existen genotipos mejor adaptados para colonizar el maíz (Rosenblueth & Martínez-Romero, 2004).

La cepa Ch24-10 es más competitiva en la colonización de la raíz y de la rizósfera del cereal, comparada con algunos aislados de nódulo de frijol y rizósfera. Además, es más tolerante al 6-methoyl-2-benzoxazolinone (MBOA), un compuesto que resulta de la degradación del DIMBOA (2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3-one), una fotoalexina producida por el maíz capaz de inhibir el crecimiento de hongos y bacterias patógenas (Rosenblueth & Martínez-Romero, 2004).

Recientemente se ha encontrado que es capaz de fijar mayor cantidad de nitrógeno y es más competitiva en la nodulación del frijol que la cepa CFN42 (datos no publicados). En ensayos de laboratorio se ha observado que la cepa Ch24-10 inhibe el crecimiento de hongos en caja (datos no publicados), y cuando el frijol y el maíz se inoculan con esta cepa se inhibe el crecimiento de éstos. En estudios en campo se ha visto que mejora el rendimiento de ambos cultivos (datos no publicados).

Para conocer los determinantes genéticos que permiten a *R. etli* Ch24-10 colonizar el maíz de manera competitiva así como de promover su crecimiento, en el laboratorio se han seguido varias estrategias. Una de ellas fue obtener mutantes al azar utilizando un transposón fusionado a un gen reportero sin promotor (Tn5gusA) con el objetivo de analizar la inducción de genes con exudados de maíz. Se analizaron más de mil mutantes y se obtuvo una que se inducía más de tres veces, al secuenciar la región aledaña al transposón y compararla con la secuencia de la cepa CFN42 se encontró que la mutación se encontraba en una ATPasa y un gen aledaño a esta es una poligalacturonasa (Vilchis, 2004).

Otra estrategia fue la de mutagenizar de manera dirigida genes que resultaran de interés directo. Se mutó el gen que codifica para la enzima necesaria para catabolizar la prolina (prolina deshidrogenada) que es el aminoácido más abundante en los exudados de maíz. El gen se induce en presencia de prolina y de exudados de maíz y de frijol, y no crece en medio mínimo con prolina como única fuente de carbono. Tuvo el mismo fenotipo de colonización de la rizósfera del maíz que la cepa silvestre Ch24-10, sin embargo, sí estuvo afectada en la competencia en nodulación respecto a la cepa silvestre CFN42 (Escalante, 2006).

3. JUSTIFICACIÓN

Se sabe que el genoma cepa-específico es el que confiere características determinadas como la patogenicidad, la resistencia a antibióticos y la estructura de la superficie bacteriana (Akopyants *et al.*, 1998; Dobrindt *et al.*, 2003; Koide *et al.*, 2004).

Mediante un experimento de hibridación ADN-ADN se determinó que las cepas de *R. etli* CFN42 y Ch24-10 comparten el 85% del genoma. El genoma exclusivo de la cepa Ch24-10 con respecto a la cepa CFN42 pudiera ser el que contiene los determinantes de la interacción de *R. etli* Ch24-10 con el maíz, su alta competitividad para colonizar la rizósfera y el interior de la planta, así como la supresión de patógenos. Con el fin de aislar el genoma exclusivo de la cepa Ch24-10 se realizó en el laboratorio un experimento de hibridación sustractiva entre ambas cepas y se generó una genoteca con el ADN presente en la cepa CH24-10 y no en la cepa CFN42 (datos no publicados).

El estudio de los genes únicos en la cepa Ch24-10 permitirá generar el conocimiento sobre su genoma completo e identificar aquellos que pudieran ser importantes en la interacción con el maíz, conocimiento que pudiera extrapolarse a otros endófitos y a las PGPR.

Dada la importancia que tiene el maíz y el frijol en el mundo y especialmente en México, por ser centro de origen y diversificación de ambos, el conocimiento generado sobre el genoma de *R. etli* Ch24-10, podría permitir su aplicación en agricultura, con el fin de disminuir el uso de agroquímicos (fertilizantes nitrogenados y fungicidas) y contribuir al desarrollo una agricultura menos contaminante.

4. OBJETIVOS

4.1 GENERALES

- Secuenciar la genoteca generada por hibridación sustractiva entre las cepas de *Rhizobium etli* Ch24-10 y CFN42.
- Analizar el genoma que es diferente en la cepa Ch24-10 con respecto a la cepa CFN42, para conocer la función putativa de las secuencias obtenidas.
- Obtener una aproximación del papel de los posibles genes en la interacción *Rhizobium etli* Ch24-10 con el maíz.

4.2 PARTICULARES

- Determinar si las secuencias se localizan en plásmido.
- Determinar si las secuencias se encuentran contiguas en el genoma.
- Determinar la presencia de elementos comunes de la genoteca en otras cepas de *R. etli*.

5. ESTRATEGIA EXPERIMENTAL

A partir de la genoteca de *R. etli* Ch24-10 generada de los experimentos de hibridación sustractiva con el genoma de la cepa CFN42, se siguió la siguiente estrategia experimental para alcanzar los objetivos de este trabajo (Figura 3):

Se obtuvieron las secuencias de las clonas de la librería y el primer análisis que se realizó fue comparar las secuencias entre sí para agrupar, en su caso, aquellas que fueran iguales. Posteriormente, las secuencias diferentes se compararon con el genoma de *R. etli* CFN42 mediante un BLAST local (blastn y blastx), con el objetivo de detectar aquellas clonas que presentaran identidad al genoma de la cepa CFN42 y colocarlas en un grupo aparte.

Aquellas secuencias de la librería de *R. etli* Ch24-10 que no presentaron identidad con el genoma de la cepa CFN42 se analizaron usando el BLAST, en el servidor del NCBI, utilizando el programa blastx. Con los resultados del BLAST se generó una pequeña base de datos, organizada en dos grupos generales: 1) secuencias con identidad significativa a una proteína anotada y 2) secuencias con identidad no significativa.

Las secuencias con identidad significativa se agruparon posteriormente en dos categorías: a) secuencias con identidad a proteínas con función putativa o conocida y b) secuencias con identidad a proteínas hipotéticas. Las secuencias con identidad significativa a proteínas de función putativa o conocida se tradujeron a aminoácidos y se les realizó un blastp en el servidor del NCBI, con el fin de identificar la presencia de dominios conservados.

Algunas de las secuencias de la librería de la cepa Ch24-10 que no presentaron identidad con el genoma de la cepa CFN42 se compararon también mediante BLAST (blastn local), con las secuencias genómicas parciales de otras cepas de *R. etli*, Kim5, Bra 5, 8C3 y CIAT894 con el fin de obtener una aproximación de la presencia de elementos genéticos comunes entre ellas.

Mediante hibridación tipo Southern blot, algunas clonas de la librería de la cepa Ch24-10 (sin identidad al genoma de la cepa CFN42), se seleccionaron para conocer su ubicación en el genoma, además las clonas se hibridaron contra un banco de cósmidos de *R. etli* Ch24-10 con el objetivo de determinar la cercanía en el genoma de las secuencias.

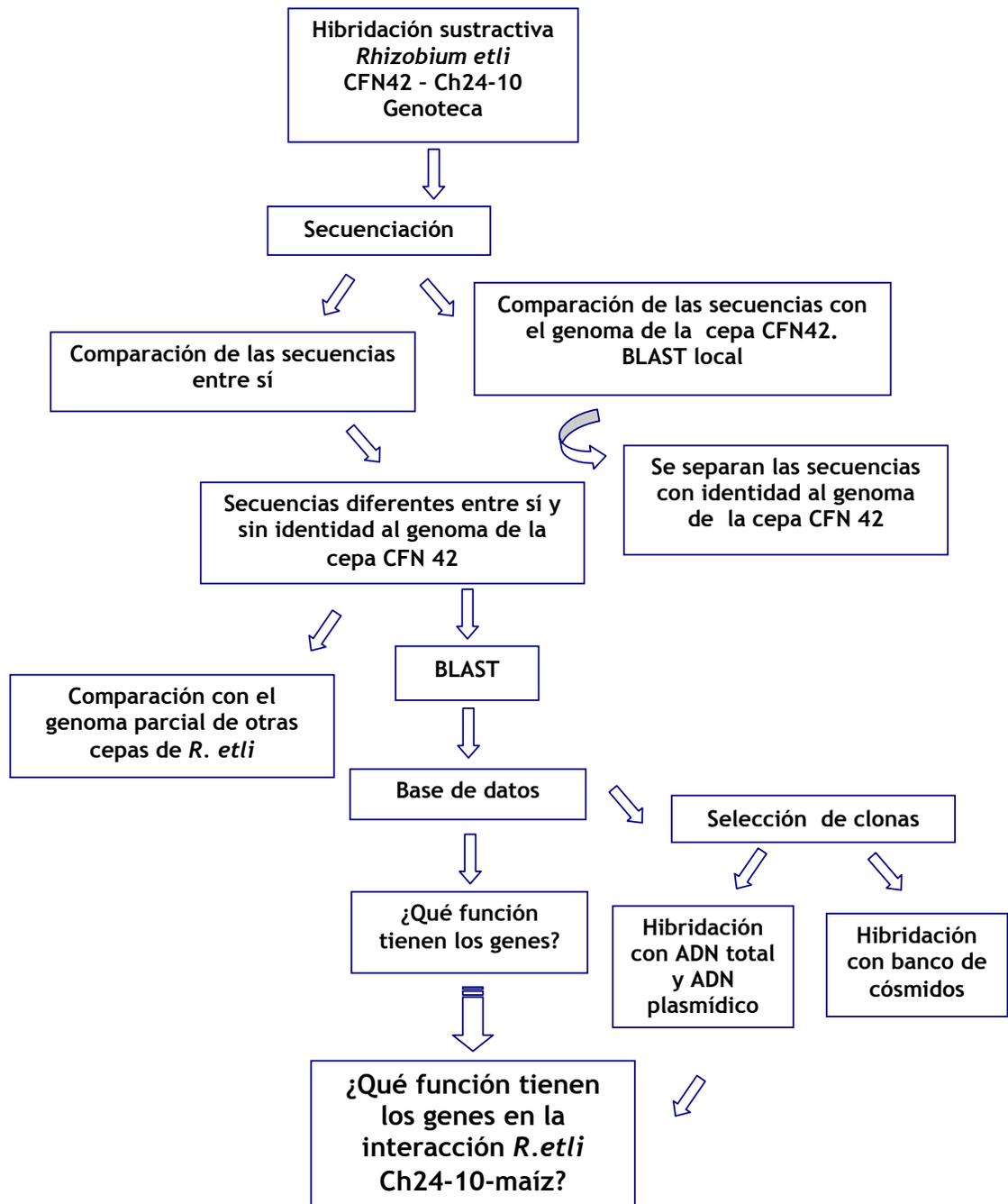


Figura 3. Estrategia experimental.

6. MATERIALES Y MÉTODOS

6.1 CEPAS, PLÁSMIDOS Y MEDIOS DE CULTIVO

Los vectores y las cepas utilizadas de *R. etli* y *Escherichia coli* se muestran en la tabla 1, donde se indican algunas de sus características.

Para crecer a *R. etli* se utilizó medio PY (5 g de peptona de caseína, 3 g de extracto de levadura, 1 g CaCl₂ · 2H₂O / L) y el antibiótico Ácido nalidíxico (Nal), 20µg/ml.

Para *E. coli* se utilizó medio LB (10 g de peptona de caseína, 5 g de extracto de levadura, 1 gr de NaCl / 1 L), para las colonias de la librería se usaron los antibióticos Km₄₅ (Kanamicina, 45 µg/ml) y Ap₁₀₀ (Ampicilina, 100 µg/ml, mientras que para las que contienen los cósmidos se utilizó Tc₁₀ (Tetraciclina 10 µg/ml).

Tabla 1. Cepas y plásmidos utilizados.

Cepas / Plásmidos	Características	Referencia o fuente
<i>R. etli</i> CFN42	Cepa tipo	Segovia <i>et al.</i> , 1993
<i>R. etli</i> Ch24-10	Tipo silvestre, aislada de tallo de maíz	Rosenblueth & Martínez-Romero, 2004
Genoteca <i>E. coli</i> Vector pCR2.1 TOPO (fig. 4)	Km ^R y Ap ^R Vector de clonación	Invitrogen Corporation
Banco de cósmidos de <i>R. etli</i> <i>E. coli</i> VC5257 Vector pLARF1	Tc ^R	Vilchis, 2004 Stratagene

6.2 PCR DE COLONIAS BACTERIANAS

Las colonias de la librería se inocularon en medio LB sólido con Km₄₅Amp₁₀₀ y se incubaron toda la noche a 37°C. La Reacción en cadena de la polimerasa (PCR, por sus siglas en inglés Polymerase chain reaction) se realizó en 50 µl de acuerdo a las especificaciones del Kit *Taq* Polymerase Recombinant (Invitrogen Corporation), con algunas modificaciones para realizar la PCR utilizando el ADN molde a partir de colonias bacterianas (Tabla 2).

En ADN templado se colocó en el tubo con los 50 µl de reacción, tomando con una punta estéril una pequeña cantidad de la colonia fresca, se agitó en un vortex y posteriormente se realizaron los ciclos de amplificación de la PCR en el termociclador GeneAmp PCR System 2700 de Applied Biosystems como se describe en la Tabla 3. La concentración y la presencia de una sola banda de la PCR se confirmaron por electroforesis en gel de agarosa. La secuencia de los oligonucleótidos (primers o cebadores) específicos para amplificar los fragmentos clonados de la librería es la que sigue (ver Figura 5):

M13 Forward (-20) 5'-GTAAAACGACGGCCAG-3', M13 Reverse 5'-CAGGAAACAGCTATGAC-3'

Tabla 2. Componentes de la reacción de PCR

Componentes de la mezcla	Volumen (µl)	Concentración final
Buffer 10X sin Mg	5	1x
MgCl ₂ 50mM	1.5	1.5 mM
Mezcla de dNTP 10mM	1	0.2 mM de cada uno
Oligos (20mM)	1 de cada uno	0.2 mM de cada uno
Taq polimerasa	0.25	1.25 unidades
Agua miliQ estéril	Se ajusta a 50 µl	
ADN	Colonia bacteriana	-

Tabla 3. Ciclos de la PCR

Ciclos	
94° C 10 minutos	Este primer ciclo permite lisar las células para liberar el ADN que será usado como templado
94° C 3 minutos 55° C 1 minuto 72° C 1 minuto 30 ciclos	Ciclos de la PCR
72° C 7 minutos	Extensión final

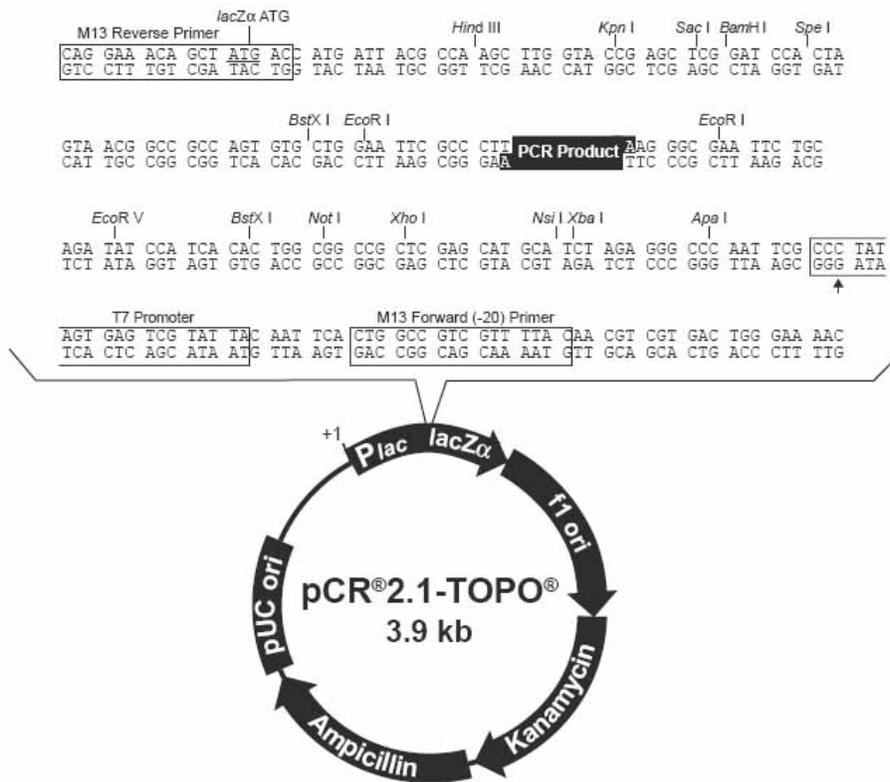


Figura 4. Mapa del vector de clonación pCR 2.1-TOPO (Invitrogen Corporation)
 En el sitio marcado como “PCR product” es donde se encuentran clonados los fragmentos de la genoteca. Se muestran también los sitios de unión de los oligonucleótidos M13 que se localizan a 300 pares de bases dirección arriba y abajo del fragmento clonado, por tanto al analizar la secuencia es necesario eliminar estos extremos.

6.3 ELECTROFORESIS / GELES DE AGAROSA

Los geles se prepararon con un 1% de agarosa Tipo II disuelta en Tris Acetatos 1X (Trizma base 48.4 g / L, 11.42 ml x L, 0.5 M EDTA, pH8). Las muestras se colocaron en el pozo del gel y se corrieron en una cámara de electroforesis (Amersham Pharmacia Biotech) a 95 volts durante 60 minutos en Tris acetatos 1X.

6.4 OBTENCIÓN DE SECUENCIAS

Una vez amplificados los fragmentos clonados y determinada su concentración en gel (superior a 50 ng/ μ l), los productos de la PCR fueron secuenciados por un servicio comercial.

6.5 ANÁLISIS DE SECUENCIAS

Las secuencias se analizaron con el Programa Chromas versión 1.62 (Technelisyum), para observar un cromatograma de ésta. Las secuencias se guardaron en el formato fasta para ser analizadas con el programa BioEdit versión 7.0.5.2 (Hall, 1999). Utilizando este programa es posible realizar modificaciones a la secuencia, lo que permitió eliminar los extremos que correspondían al vector de clonación (Figura 4), y corregir de ser necesario, algún error en la secuencia al comparar con el archivo original en Chromas.

Las secuencias se compararon entre sí mediante alineamientos en el servidor de GeneBee (Brodsky *et al.*, 1995), para determinar la existencia de secuencias iguales repetidas y continuar el análisis solamente de aquellas que fueran distintas.

Posteriormente, las secuencias de la librería de *R. etli* Ch24-10 diferentes entre sí, se compararon con el genoma de la cepa CFN42 secuenciado y anotado recientemente (González *et al.*, 2006), mediante BLAST local, utilizando los programas blastn y blastx (Altschul *et al.*, 1997), para descartar las secuencias que tuvieran un porcentaje de identidad igual o mayor al 45%.

El grupo de secuencias sin identidad al genoma de la cepa CFN42 se analizaron posteriormente con el BLAST (Basic Local Alignment Tool) (Altschul *et al.*, 1997) en el servidor del NCBI con el programa blastx.

El BLAST es un método de búsqueda de identidad entre secuencias y está basado en una sólida teoría estadística, es un método que fundamentalmente busca

secuencias homólogas (origen evolutivo común). Para establecer si dos o más secuencias (de nucleótidos o aminoácidos) son homólogas generalmente se considera la identidad o similitud entre ellas, dado que esta es una variable que se expresa numéricamente y que se relaciona con una variable probabilística (Valor de E, que expresa la probabilidad de que la identidad sea resultado del azar, mientras más cercano a cero sea este valor, el valor de E es más significativo).

Al realizar el BLAST, si los resultados muestran un porcentaje de identidad alto entre secuencias, un valor de E cercano a cero y un puntaje (score) alto (que expresa la longitud de las secuencias con identidad), existe una baja probabilidad de se hayan originado de manera independiente evolutivamente y de que la similitud entre ellas sea resultado del azar (Koonin & Galperin, 2003).

Por otra parte, el programa blastn hace una comparación entre secuencias nucleotídicas, mientras que el blastx realiza lo hace entre secuencias de aminoácidos (realizando la traducción en los seis marcos de lectura).

Al realizar el BLAST de las secuencias de la librería de la cepa Ch24-10 sin identidad al genoma de la cepa CFN42, los valores de corte que se utilizaron para definir identidad significativa en este trabajo fueron los siguientes: identidad (igual o superior al 30%), valor de E menor a 10^{-5} y score superior a 40.

Los datos que se tomaron en cuenta para construir la base de datos con la información del BLAST fueron los datos del primer “Bit” de los resultados (por considerarse el más significativo estadísticamente), son los siguientes:

- Porcentaje de identidad, valor de E y score.
- Nombre de la proteína con la que presentó identidad así como su tamaño, además del número de acceso a la información existente sobre esta secuencia.
- Especie a la que pertenece la secuencia de referencia, el marco de lectura (blastx) y la posición en las secuencias donde se presentó la similitud.

Las secuencias con identidad a proteínas de función putativa o conocida se tradujeron a aminoácidos utilizando el programa BioEdit (Hall, 1999), y se realizó

BLAST (Altschul *et al.*, 1997), en el servidor del NCBI utilizando el programa blastp, con el fin de identificar la presencia de dominios conservados.

Algunas de las secuencias de la cepa Ch24-10 sin identidad al genoma de la cepa CFN42, se compararon mediante BLAST (blastn local) (Altschul *et al.*, 1997) con la secuencias genómicas parciales de las cepas de *R. etli*, Kim5, Bra 5, 8C3 y CIAT894.

6.6 OBTENCIÓN DE ADN TOTAL Y PLASMÍDICO.

Para extraer el ADN total se utilizó el Kit GenomicPrep Cells and Tissue DNA Isolation (Amersham Biosciences). El ADN plasmídico se obtuvo de acuerdo al protocolo del High Pure Plasmid Isolation Kit (Roche Applied Bioscience). Se verificó la concentración (utilizando ADN del fago λ digerido con *Hind* III, New England Biolabs) y la integridad del ADN por electroforesis en gel de agarosa.

6.7 DIGESTIÓN DE ADN CON ENZIMAS DE RESTRICCIÓN

El ADN total de las cepas Ch24-10 y CFN42 se digirió con la enzima de restricción *EcoRI* (New England Biolabs). Se diluyó en el buffer NE para *EcoRI* (New England Biolabs) a una concentración final de 1X (50mM NaCl, 100 mM Tris-HCl, 10mM MgCl₂, 0.025 Triton X-100, pH 7.5), el ADN total y la enzima. Para digerir 1 μ g de ADN se requirieron 20 unidades de enzima. La reacción se incubó por 4 horas a 37°C y posteriormente se inactivó la enzima a 65°C por 20 minutos. Se verificó la digestión del ADN por electroforesis en gel de agarosa.

6.8 GELES TIPO ECKHARDT

Los geles tipo Eckhardt (Eckhardt, 1978 modificado por Hynes & O'Connell, 1990) permiten visualizar el perfil plasmídico de las bacterias y se realizaron de la siguiente manera: se preparó un gel de agarosa al 0.7% en Tris Boratos (TB) 1x (10.8 g Trizma base, 0.93g EDTA, 5.5g ácido bórico / L) y 1% de SDS (dodecil sulfato de sodio).

Las cepas de *R. etli* CFN42 y Ch24-10, se sembraron en PY sólido (PYNal₃₀) cuatro días antes de realizar el Eckhardt, al tercer día se inocularon en PY líquido y

se incubaron a 28°C en agitación durante toda la noche. Posteriormente se tomaron 200 µl de cada cultivo en un tubo Eppendorf (a una densidad óptica de 0.5 medida a 600 nm, medida en el espectrofotómetro DU650 de Beckman), y se colocaron en hielo. Se agregaron 0.5 ml de Sarcosyl y se centrifugó, un pulso. Se desechó el sobrenadante y se mantuvieron en hielo.

La pastilla celular se resuspendió cuidadosamente en 20 µl de solución E1 [sacarosa al 10% disuelta en TB 1X y ARNasa (10mg /ml en TB), a la solución E1 se le añadió al momento de usarse unos granitos de lisozima]. La pastilla resuspendida se cargó inmediatamente en el pozo del gel, sumergido en TB 1X, se corrió el gel en una cámara de electroforesis durante 5 minutos a 15 volts y posteriormente a 80 volts por 8 horas a -4°C.

6.9 SOUTHERN BLOT

6.9.1 TRANSFERENCIA DE ADN DE COLONIAS BACTERIANAS

Para ubicar las clonas seleccionadas de la librería en el banco de cósmidos de la cepa Ch24-10 (Vilchis, 2004) se analizaron más de 1000 colonias organizadas en juegos de 48 colonias de la siguiente manera (fig. 5):

Se marcaron con números los filtros de nylon Hybond N+ (Amersham Biosciences) y se esterilizaron. Una vez estériles se colocaron con pinzas estériles en la caja Petri con medio LB Tc₁₀ (cósmidos clonados en un el vector pLARF1). Las colonias se replicaron en los filtros a partir de colonias frescas y de dejaron crecer toda la noche.

Para transferir el ADN se sacaron los filtros de las cajas Petri y se colocaron en soluciones de transferencia, formado en cada una un puente con papel Whatmann tal como ilustra la figura 6. Los filtros se colocaron en el puente durante 7 minutos en solución desnaturalizante (1.5 M NaCl, 0.5 M NaOH), posteriormente se colocaron en otro puente 5 minutos en solución neutralizante (1.5 M NaCl, 0.5 M Tris HCl, 0.001 M EDTA) y 5 minutos en otro puente con solución neutralizante. Finalmente 5 minutos en un puente con SSC 2x (0.3 M NaCl, 0.03 M citrato de sodio). Los filtros se dejan secar y se sumergen en etanol al 100% durante 3 minutos. Una vez que están secos se fija el ADN primero con luz ultravioleta

utilizando un Spectrolinker (XL-1000 UV Crosslinker de Spectronics Corporation), y posteriormente horneando por dos horas a 80°C.

6.9.2 TRANSFERENCIA DE ADN EN GEL

El ADN total de las cepas de *R. etli* Ch24-10 y CFN42 digerido con *EcoRI* y se corrió en geles de agarosa al 1% (500 ng por pozo) durante una hora (en un pozo el DNA de la cepa Ch24-10 y en el siguiente el ADN de la cepa CFN42, y así sucesivamente).

Posteriormente el ADN de los geles se transfirió por capilaridad a una membrana de nylon Hybond-N+ (Amersham Biosciences) de la siguiente manera: con agitación, el gel de agarosa se sumergió en la solución de depurinación (11 ml HCl, 989 ml de agua destilada), entre 20 y 30 minutos, después en la solución desnaturalizante (1.5 M NaCl, 0.5 M NaOH), durante 30 minutos y finalmente en la solución neutralizante (1.5 M NaCl, 0.5 M Tris HCl, 0.001 M EDTA) de 30 a 50 minutos. Enseguida, el gel se colocó sobre un vidrio, encima la membrana de Hybond-N+ y tres capas de papel Whatmann previamente humedecidos con SSC 10X (1.5 M NaCl, 0.15 M citrato de sodio), toallas de papel, otro vidrio y un peso sobre él. Se niveló y se dejó transferir toda la noche. Al día siguiente la membrana se retiró y se dejó secar a temperatura ambiente. Una vez seca, el ADN se fijó con luz ultravioleta (crosslinker) y horneando a 80°C durante dos horas. El gel tipo Eckhardt se transfiere de la misma manera.

6.9.3 OBTENCIÓN DE SONDAS

Para conocer la ubicación y cercanía en el genoma de las clonas, se realizó PCR (en 100 µl de reacción, cómo se indica en el apartado 6.2), a las 23 clonas seleccionadas aleatoriamente. Las reacciones de PCR se analizaron por electroforesis en gel de agarosa. La banda correspondiente se purificó del gel de acuerdo a las especificaciones del High Pure PCR Product Purification Kit (Roche Applied Science).

6.9.4 HIBRIDACIÓN

Las membranas con el ADN de los cósmidos de la cepa Ch24-10, ADN total y ADN plasmídico (Eckhardt) de las cepas Ch24-10 y CFN42, se prehibridaron con el Rapid Hyb Buffer (Amersham Biosciences) a una temperatura de 65°C, las membranas con ADN plasmídico durante toda la noche y aquellas con ADN total y ADN de los cósmidos durante 4 horas. Transcurrido el tiempo de prehibridación, las sondas purificadas de gel se marcaron con α -³²PdCTP con el kit Redirpime II Random Prime Labelling System (Amersham Biosciences) y se colocaron en los tubos de hibridación con las membranas prehibridadas. La hibridación se realizó a una temperatura de 65°C, toda la noche. Posteriormente, las membranas se lavaron a 65°C dos veces con SSC 2X / SDS 0.1%, dos veces con SSS 1X / SDS 0.1% y finalmente dos veces con SSC 0.5% / SDS 0.1%, (cada lavado durante 10 minutos). Al finalizar los lavados, las membranas se enjuagaron con SSC 2X y se mantuvieron húmedas dentro de una bolsa de plástico sellada y se expusieron a una autorradiografía (Kodak) a -80°C. Posteriormente, dependiendo del marcado se reveló la autorradiografía

Para localizar los cósmidos se realizó una primera hibridación que consistió en mezclar las sondas y marcarlas juntas para identificar aquellos cósmidos que dieran una señal positiva con cualquiera de las sondas. El siguiente paso fue hibridar los cósmidos positivos (fueron 16), con cada una de las sondas por separado para identificar qué cósmido hibridó con qué sonda específicamente.

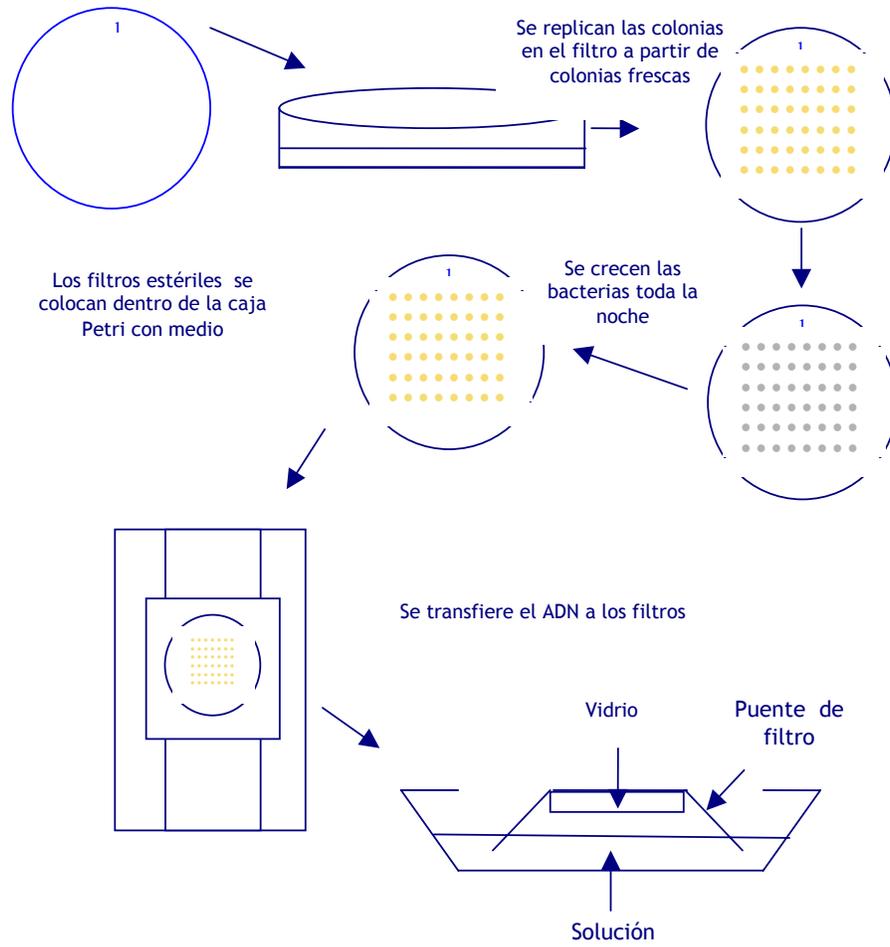


Figura 5. Transferencia de ADN de colonias bacterianas.

7. RESULTADOS

7.1 SECUENCIAS CON IDENTIDAD AL GENOMA DE LA CEPA CFN42

De acuerdo a los resultados del blastx, se obtuvieron 69 secuencias de la genoteca de la cepa Ch24-10 con identidad con genoma de la cepa CFN42 (organizado en 7 replicones organizados en 1 cromosoma y 6 plásmidos) (González *et al.*, 2006).

De este grupo de secuencias, 23 de ellas (33%), se localizan en los plásmidos: 1 en el p42a, 5 en el p42b, 2 en el p42c, 9 en el plásmido simbiótico psym (p42d), 2 en el p42e y 4 en el p42f. De las 69 secuencias, 16 presentaron identidad con el genoma de *Rhizobium leguminosarum* bv. *viciae* 384, en un porcentaje mayor a 40%, un valor de E menor de 10^{-5} y un score superior a 80. En la tabla 4A, 4B y 4C se encuentran los resultados resumidos de este análisis.

Tabla 4. Secuencias con identidad al genoma de *R. etli* CFN42.

Tabla 4A. Secuencias con identidad mayor de 75% (en más de 60 aminoácidos) y valor de E menor de 10⁻⁵.

Se muestra la identidad en el número de aminoácidos (aa) y su porcentaje (%), la longitud (nucleótidos) de la secuencia de la librería de la cepa Ch24-10, el valor de E, el score y el número de veces que se repitió cada secuencia.

SECUENCIA / LONGITUD (nn)	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
9-48 / 348	ADN polimerasa III subunidad alfa	100/105 (95%)	214	9.00E-55	C	1
12-4 / 650	Secuencia de inserción putativa, transposasa	135/155 (87%)	267	3.00E-70	psym	1
6-18 / 579	Probable regulador transcripcional, familia LuxR	73/92 (79%)	146	5.00E-34	C	2
8-39 / 818	Probable regulador transcripcional, familia AraC	231/244 (94%)	469	4.00E-131	p42f	1
9-10 / 878	Regulador transcripcional putativo	71/71 (100%)	145	3.00E-33	C	1
seq2 / 454	Proteína sensora del sistema de dos componentes, histidina cinasa*	99/106 (93%)	171	1.00E-41	C	1
5-27 / 711	Caja sensora putativa / familia de proteínas con dominio GGDEF *	68/84 (80%)	139	1.00E-31	C	1
4-4 / 255	Translocasa, exportadora de proteínas	73/73 (100%)	149	4.00E-35	C	2
9-28 / 349	Probable transportador ABC de azúcares, proteína de unión al sustrato*	91/104 (87%)	212	5.00E-54	p42b	1
11-5 / 460	Transportador ABC de azúcares, permeasa	140/140 (100%)	218	6.00E-56	psym	1
seq8 / 274	Transportador ABC de azúcares, permeasa	68/68 (100%)	131	8.00E-30	p42b	1
1-32 / 612	Probable transportador ABC de oligopéptidos, permeasa	158/161 (98%)	317	2.00E-85	p42f	1
6-2 / 503	Probable transportador ABC de azúcares, permeasa	103/105 (98%)	108	7.00E-23	p42b	1
5-24 / 396	Probable transportador ABC de azúcares, proteína de unión al sustrato	89/90 (98%)	154	2.00E-36	C	1

(Psym: p42d, *Secuencias con identidad al genoma de *Rhizobium leguminosarum* bv. *viciae* 384, C: cromosoma, p; plásmido).

Tabla 4A. Continuación...

SECUENCIA / LONGITUD (nn)	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
3-20 / 825	Probable transportador ABC de azúcares, proteína de unión al sustrato	94/108 (88%)	197	5.00E-49	p42b	1
4-21 / 730	Putativa toxina RTX tipo hemolisina, proteína de unión a calcio*	168/231 (72%)	358	2.00E-97	C	1
7-19 / 452	Hemolisina*	107/108 (99%)	218	9.00E-56	C	1
5-25 / 471	Proteína de quimiotaxis, aceptora de metilos	88/88 (100%)	177	2.00E-43	C	1
6-32 / 551	Probable succinato semialdehído dehidrogenasa [NADP+]	171/171 (100%)	292	6.00E-78	psym	1
7-6 / 839	Proteína de autoagregación (proteína de adherencia)	165/180 (91%)	332 bits	9.00E-90	C	1
seq34 / 567	Proteína transmembranal, biosíntesis de flagelo	91/92 (98%)	179	6.00E-44	C	1
8-48 / 351	Probable reductasa FMN dependiente de NAD(P)H	79/103 (76%)	164	1.00E-39	p42c	1
seq17 / 564	Endopeptidasa Clp, subunidad 3	87/89 (97%)	178	1.00E-43	C	1
seq25 / 394	Putativa L-iditol 2-dehidrogenasa	107/108 (99%)	222	4.00E-57	psym	1
seq36 / 818	Ciclohidrolasa II de GTP	166/170 (97%)	327	5.00E-88	psym	1
seq49 / 838	Ceto-ácido reductoisomerasa	207/253 (81%)	389	6.00E-107	C	1
seq53 / 555	Nitrogenasa (iron protein)	176/179 (98%)	309	4.00E-83	psym	1
3-41 / 781	Proteína hipotética RHE_CH00091	66/70 (94%)	134	3.00E-30	C	1
1-39 / 546	Proteína hipotética RHE_CH02651	99/129 (76%)	201	1.00E-50	C	2
5-14 / 363	Proteína hipotética RHE_CH01108	86/109 (78%)	184	1.00E-45	C	4
3-33 / 902	Proteína hipotética RHE_CH03202	77/97 (79%)	152	2.00E-35	C	1
5-12 / 795	Proteína hipotética RHE_CH01631	136/154 (88%)	281	2.00E-74	C	1
6-3 / 560	Proteína hipotética RHE_CH03237*	84/97 (86%)	124	3.00E-27	C	2
seq20 / 547	Proteína hipotética RHE_CH02651	99/129 (76%)	201	1.00E-50	C	1
7-1 / 832	Proteína hipotética RHE_CH01631	140/161 (86%)	288	3.00E-76	C	2

Tabla 4B. Secuencias con identidad al menor de 75% y/o menos de 60 aminoácidos y valor de E menor de 10^{-5} .

SECUENCIA / LONGITUD (nm)	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
10-1 / 631	Secuencia de inserción putativa, transposasa*	49/63 (77%)	109	7.00E-23	p42b	1
4-13 / 765	Probable regulador transcripcional, familia MerR*	38/40 (95%)	82.8	1.00E-14	C	4
seq7 / 386	Proteína reguladora específica de Nif	41/45 (91%)	83.6	2.00E-15	psym	1
seq37 / 725	Porina, proteína de membrana externa	62/63 (98%)	87	6.00E-16	p42e	1
9-25 / 291	Proteína putativa de biosíntesis de exopolisacáridos	54/64 (84%)	111	9.00E-24	C	1
4-3 / 492	Adenilato ciclasa*	74/153 (48%)	157	2.00E-37	C	1
3-37 / 516	Probable glicosiltransferasa*	38/66 (57%)	82-4	8.00E-15	C	2
5-47 / 227	Gama-carboximuconolactone decarboxilasa*	31/31 (100%)	65.5	8.00E-10	p42f	1
12-29 / 769	Reductasa putativa*	43/43 (100%)	91.7	3.00E-17	C	1
8-16 / 815	Cetol piruvato transferasa*	133/220 (60%)	263	8.00E-69	C	1
8-18 / 866	Cetol piruvato transferasa*	133/220 (60%)	263	4.00E-69	C	1
seq38 / 443	GDP-L-fucosa sintasa*	80/134 (59%)	172	4.00E-42	C	2
seq52 / 911	Fosfomanomutasa	66/104 (63%)	133	1.00E-29	C	1
9-29 / 541	Proteína hipotética RHE_CH01538	55/70 (78%)	109	5.00E-23	C	1
seq47 / 512	Proteína hipotética RHE_CH01807	50/57 (87%)	110	2.00E-23	C	1
8-27 / 561	Proteína hipotética RHE_CH02021*	43/57 (75%)	93.6	4.00E-18	C	1
seq50 / 229	Proteína hipotética RHE_CH00223	46/46 (100%)	95.5	7.00E-19	C	1
2-36 / 357	Proteína hipotética RHE_PE00109	24/26 (92%)	57	3.00E-07	p42e	7
seq13 / 279	Proteína hipotética RHE_PD00011	33/40 (82%)	67.4	2.00E-10	psym	1
seq16 / 351	Proteína hipotética RHE_CH00444	28/29 (96%)	57	3.00E-07	C	1

Tabla 4B. Continuación...

SECUENCIA / LONGITUD	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
----------------------	-----------	-----------	-------	------------	---------------------------	---------------------

(nn)						
8-11 / 286	Proteína hipotética RHE_PF00054	33/60 (55%)	67	3.00E-10	p42f	1
9-17	Proteína hipotética RHE_CH01905*	85/89 (95%)	169	4.00E-41	C	3
7-35/ 521	Proteína hipotética RHE_CH02761	33/43 (76%)	66.2	6.00E-10	C	9
5-46 / 365	Proteína hipotética RHE_PC00203*	38/41 (92%)	129	5.00E-29	p42c	2
6-1 / 576	Proteína hipotética RHE_CH01114*	45/89 (50%)	161	2.00E-38	C	1
6-14 / 413	Proteína hipotética RHE_CH01669*	43/47 (91%)	90	2.00E-18	C	1
seq40 / 402	Proteína hipotética RHE_CH01114*	45/89 (50%)	80.5	2.00E-14	C	1
9-11 / 597	Proteína hipotética RHE_PA00030*	57/126 (45%)	97.4	3.00E-19	p42a	1
J11-40 / 700	Proteína hipotética RHE_PD00172	144/220 (65%)	293	4.00E-78	psym	1

Tabla 4C. Secuencias con identidad al genoma de la cepa CFN42 con un valor de E mayor de 10^{-5} .

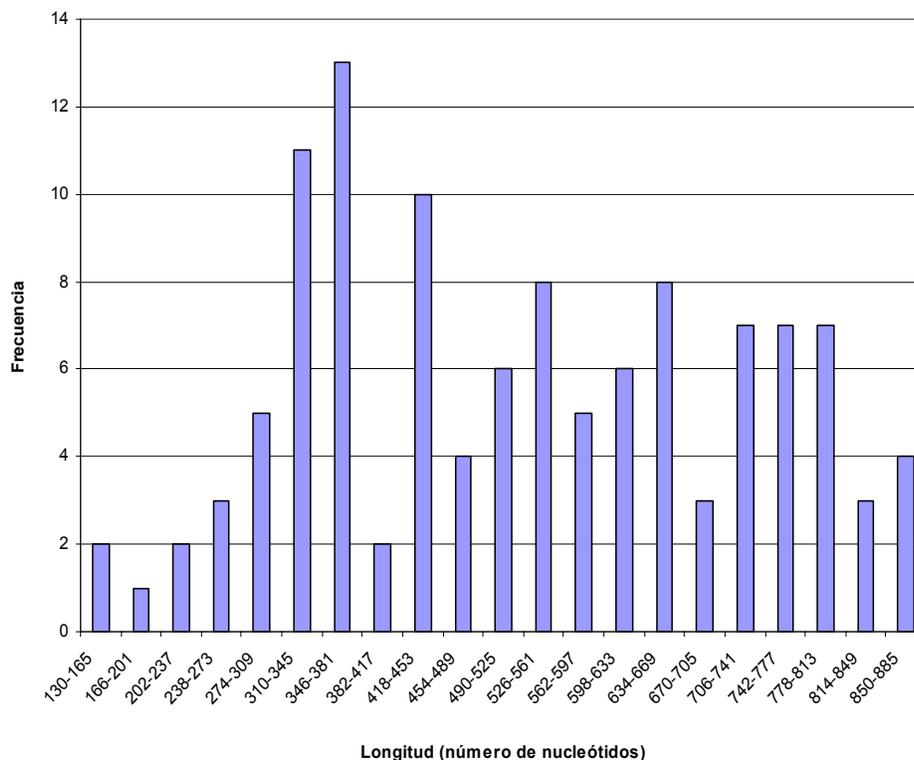
SECUENCIA / LONGITUD (nn)	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
seq30 / 548	Proteína ribosomal L33 50S	18/18 (100%)	41.2	0.02	C	1
J3-21 / 461	Proteína del sistema de dos componentes, regulador de respuesta	22/22 (100%)	47	3.00E-04	C	3
J9-40 / 633	Probable acetiltransferasa	20/23 (86%)	43.1	0.009	C	1
J5-23 / 300	Proteína hipotética RHE_CH00479	19/19 (100%)	50.4	3.00E-05	C	1
J5-2 / 414	Proteína hipotética RHE_CH03754*	49/64 (76%)	42.4	0.007	C	1

7.2 SECUENCIAS SIN IDENTIDAD AL GENOMA DE LA CEPA CFN42

Como se mencionó, de las 348 clonas que se secuenciaron 243 no presentaron identidad al genoma de la cepa CFN42. De estas 243 secuencias 120 fueron distintas entre sí, es importante notar que algunas secuencias se repetían en número variable, en algunos casos hasta 17 veces, como la secuencia 3-34, por otro lado había secuencias únicas como las secuencias seq45, 8-34 y 2-18 por mencionar algunas (Ver Anexo, en la tabla 8 se muestran las secuencias repetidas). En los casos en que se repitieron se tomó una sola secuencia para realizar los análisis posteriores.

La longitud de las 120 secuencias se encuentra en un rango que va de los 130 a los 884 nucleótidos, con un promedio de 530 (los insertos clonados se secuenciaron completos). En el gráfico 1 se muestra que los fragmentos más pequeños se ubican en el intervalo que va de 130 a 165 nucleótidos y que los más grandes tienen una longitud entre 850 y 885. La mayoría se distribuye en el rango de 300 a 800 nucleótidos.

Gráfico 1. Longitud de los fragmentos clonados de la librería.



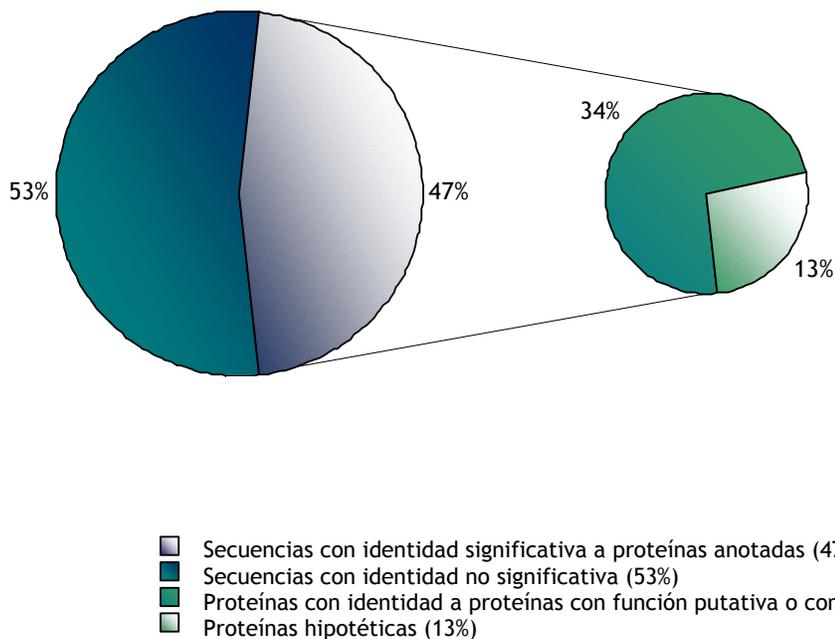
Las 120 secuencias sin identidad al genoma de la cepa CFN42 se organizaron en dos grupos generales (Gráfico 2):

- 1) 56 secuencias (47%), con identidad significativa a proteínas anotadas con una identidad igual o superior al 30%, valor de E menor a 10^{-5} y score superior a 40 (Los resultados en detalle del BLAST se muestran en el Anexo en la tabla 8)
- 2) 64 secuencias (53%), sin identidad significativa con una identidad menor al 30%, valor de E mayor de 10^{-5} y score inferior a 40) (Los datos no se muestran).

Por otro lado, las secuencias con identidad significativa se dividieron a su vez en:

- a) 41 secuencias con identidad a una proteína putativa o conocida (32% del total, 73% de este grupo).
- b) 15 secuencias con identidad a proteínas hipotéticas (13% del total, 27% de este grupo).

Gráfico 2. Distribución porcentual de las 120 secuencias con respecto a la identidad que presentaron en el blastx.



7.2.1 SECUENCIAS CON IDENTIDAD SIGNIFICATIVA A PROTEÍNAS PUTATIVAS O CONOCIDAS.

Las secuencias con identidad a proteínas putativas o conocidas se agruparon en cinco categorías de acuerdo al grupo funcional del ortólogo correspondiente (Gráfico 3). El grupo con mayor número de secuencias es el relacionado con replicación, recombinación y reparación del ADN, en segundo lugar se ubica el grupo relacionado con transporte celular y regulación de respuesta. Este grupo presentó identidad por un lado, a proteínas cuya función esta comprobada experimentalmente (TonB y gen inducible por el hospedero de *R. leguminosarum*), y por otro a proteínas putativas con función conservada (el resto).

Al realizar la traducción a aminoácidos de las secuencias nucleotídicas con identidad a proteínas putativas, algunas de ellas mostraron identidad a dominios conservados, en la mayoría de los casos correspondientes al resultado del blastx (tabla 5 y en detalle en el Anexo, tabla 8).

Gráfico 3. Distribución porcentual de las secuencias de Identidad alta a proteínas putativas. Organización en grupos funcionales.

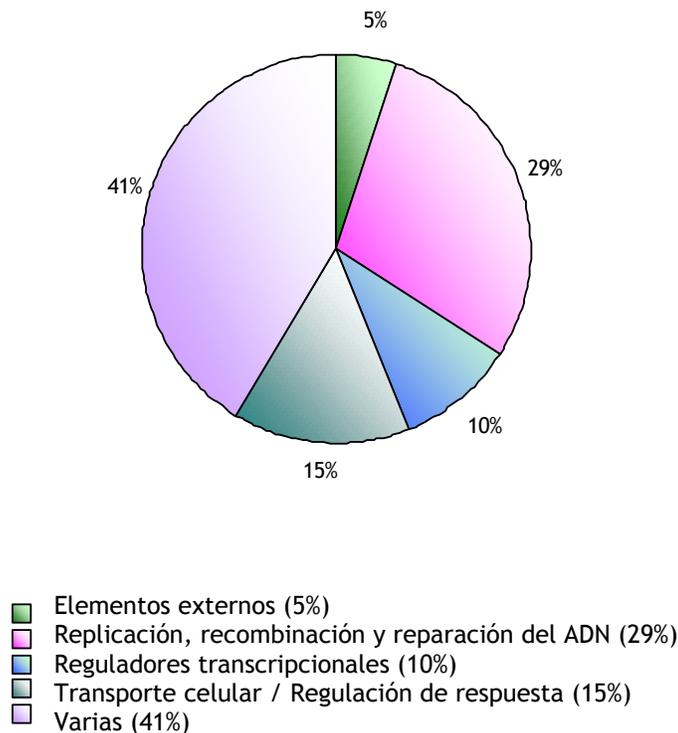


Tabla 5. Organización en grupos funcionales de las secuencias con identidad a proteínas putativas o conocidas.

Se muestra la longitud de la secuencia proteica del ortólogo en aminoácidos (aa), así como de la secuencia de la librería de la cepa Ch24-10 en nucleótidos (nn). Se muestra también la identidad en número de aa y su porcentaje (%) correspondiente y el porcentaje que cubre la secuencia de la librería traducida a aminoácidos respecto a la longitud de la proteína con la que presentó similitud. Se indica el dominio al que presentó identidad la secuencia de la librería después de su traducción a aminoácidos y la realización del blastp y en su caso el plásmido con el que hibridaron.

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD (aa) / (%)	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Relacionadas con fagos	• Proteína mayoritaria de la cápside gp7 / 227 proteína hipotética de <i>R. leguminosarum</i> bv. viciae 3841 similar a esta de fago	7-10 / 638	79/138 (57%)	93	x
	• Integrasa / 999	seq45 / 347	23/67 (34%)	12	Integrasa de fago
Replicación, recombinación y reparación del ADN	• Transposasa y derivados inactivados / 781	3-25 / 641	70/195 (35%)	28	x
	• ***Transposasa para secuencia de inserción ISRM18/ 507	6-40 / 453	143/147 (97%)	30	Core de la integrasa
	• * Transposasa para secuencia de inserción ISRM18 / 507	4-40 / 755	123/132 (93%)	50	x
	• Transposasa / 266	8-30 / 396	35/58 (60%)	50	x
	• ADN metilasa N4/N6 / 613	1-16 / 542	117/179 (65%)	29	ADN metilasa
	• ADN metilasa N4/N6 / 560	3-46 / 559	111/179 (62%)	33	ADN metilasa específica de adenina
	• ADN metilasa específica de adenina / 194	2-18 / 225	43/70 (61%)	39	de adenina
	• ADN metilasa N4/N6 / 938	9-9 / 470	45/107 (42%)	17	x
	• ADN metilasa específica de citosina C-5 / 710	7-11 / 756	38/90 (42%)	35	x
	• Helicasa caja DEAD/DEAH dependiente de ATP / 811	10-18 / 522	101/114 (88%)	21	Enzima de restricción I
• Helicasa C terminal, caja DEAD/DEAH, dependiente de ATP / 963	4-37 / 734	83/148 (56%)	25	DEAD helicasa	
• Endonucleasa de restricción (subunidad helicasa) / 999	5-4 / 516	124/177 (70%)	17	ADN o ARN helicasa	

(*pa, **pb, ***pc; X = no se encontró similitud con algún dominio conservado)

Tabla 5. Continuación...

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD (aa) / (%)	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Reguladores Transcripcionales	<ul style="list-style-type: none"> Regulador transcripcional similar a GntR / 218 Hélice –vuelta-hélice	3-35 / 307	38/98 (38%)	47	GntR
	<ul style="list-style-type: none"> * Regulador tipo LysR / 328 	4-32 / 745	127/249 (51%)	76	LysR
	<ul style="list-style-type: none"> Regulador transcripcional tipo RioFf 87 / Similar al dominio de unión al gen <i>traR</i> (activa el operón de bioluminiscencia) / 62 	6-23 / 627	34/57 (59%)	100	LuxR
	<ul style="list-style-type: none"> Hélice-vuelta-hélice tipo AraC / 315 	8-34 / 808	68/90 (75%)	85	ATT AraC
Transporte celular / Regulación de respuesta	<ul style="list-style-type: none"> Proteína sensora, traducción de señal, histidina cinasa / 384 [Regulador transcripcional con dominio PAS, tipo-AAA ATPasa]	7-31 / 750	89/150 (59%)	65	PAS
	<ul style="list-style-type: none"> Diguanilato ciclasa/Fosofodiesterasa con dominio GGDEF & EAL /912 	7-14 / 745	34/63 (53%)	27	EAL
	<ul style="list-style-type: none"> Diguanilato ciclasa-Fosofodiesterasa con dominio GGDEF & EAL / 713 [Proteínas con este dominio pueden actuar como sensores]	2-1 / 329	34/49 (69%)	15	EAL
	<ul style="list-style-type: none"> Gen inducible por el hospedero (ORF 3) / 295 [Importante para competencia en la nodulación]	7-22 / 576	123/147 (83%)	65	X
	<ul style="list-style-type: none"> TonB. Proteína de membrana, traduce energía a receptores de membrana externa. Transporte de hierro / 454 	9-20 / 336	62/106 (58%)	25	X
<ul style="list-style-type: none"> Canal iónico. Tipo transportador de potasio / 343 	Seq12 / 503	35/62 (55%)	49	X	

Tabla 5. Continuación...

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD aa / %	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Proteínas varias	<ul style="list-style-type: none"> Acetiltransferasa / 340 [Metabolismo de lípidos]	12-27 / 168	25/44 (56%)	16	X
	<ul style="list-style-type: none"> Glicosiltransferasa familia 2, involucrada en la biogénesis de la pared celular / 307 	2-34 / 735	46/152 (30%)	80	X
	<ul style="list-style-type: none"> Glicosiltransferasa (síntesis de core del lipopolisacárido) / 384 	8-17 / 862	29/95 (30%)	75	Glicosiltransferasa
	<ul style="list-style-type: none"> Glicosiltransferasa familia 8 / 243 Proteína involucrada en la biosíntesis de polisacáridos / 489 Proteína transportadora de polisacáridos / 488 [Proteína de biosíntesis de polisacáridos, generalmente son integrales de membrana. La familia incluye a RfbX parte del operón de biosíntesis del antígeno O] [Biosíntesis de disacáridos, polisacáridos]	7-43 / 315 3-47 / 448 5-38 / 628	38/103 (36%) 46/152 (30%) 29/95 (30%)	43 21 42	Glicosiltransferasa Proteína de biosíntesis de polisacáridos Proteína de biosíntesis de polisacáridos
<ul style="list-style-type: none"> Glicósidohidrolas, familia 19 / 201 [Dominio: quitinasa]	1-40 / 657	45/86 (52%)	100	X	

Tabla 5. Continuación...

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD aa / %	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Proteínas varias	<ul style="list-style-type: none"> • *** Metiltransferasa / 352 [Transferencia de metilo del S-adenosil-L-metionina (AdoMet) a cualquier átomo de nitrógeno, oxígeno o carbono. Modifica DNA, RNA, proteínas, moléculas pequeñas como el catecol con fines regulatorios]	4-20 / 432	41/54 (75%)	41	X
	<ul style="list-style-type: none"> • Proteína LuxA Putativa / 354 Cadena alfa de la luciferasa monooxigenasa [La luciferasa bacteriana es una flavina monooxigenasa que cataliza la oxidación de una cadena larga de aldehído provocando la emisión de energía en forma lumínica]	4-6 / 346	35/56 (62%)	32	Luciferasa bacteriana
	<ul style="list-style-type: none"> • Fitanoil CoA dioxigenasa (PhyH) / 259 [PhyH es una enzima peroxisomal que cataliza el primer paso en la oxidación-alfa del ácido fitánico Es una oxidorreductasa]	6-31 / 746	63/107 (58%)	98	Fitanoil coA dioxigenasa
	<ul style="list-style-type: none"> • * Lisina/ornitina monooxigenasa IucD / 540 [Biosíntesis de metabolitos secundarios, transporte y catabolismo] n6-oxigenación de lisina, y el primer paso en la síntesis de la aerobactina	3-45 / 761	68/131 (52%)	56	Lisina/ornitina monooxigenasa IncD
<ul style="list-style-type: none"> • Nitrilohidratasa / 153 [Hidratación de compuestos nitrilo. Actúa en las cadenas cortas de nitrilos alifáticos convirtiéndolos en el aminoácido correspondiente. No actúa en los aminoácidos ni en los nitrilos aromáticos]	12-22 / 137	27/31 (87%)	29	X	

Tabla 5. Continuación...

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD aa / %	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Proteínas varias	<ul style="list-style-type: none"> • Formato dehidrogenasa / 126 [Puede actuar en el metabolismo del glioxilato/dicarboxilato, metabolismo del metano y en el sistema de dos componentes. Oxidoreductasa. Actúa en el grupo aldehído o grupo oxo de los donadores, el NAD ⁺ o NADP ⁺ son los aceptores]	8-35 / 344	21/25 (84%)	90	X
	<ul style="list-style-type: none"> • ** Ulp1 proteasa, dominio catalítico C-terminal/ 1441 	12-30 / 813	54/151 (35%)	19	X
	<ul style="list-style-type: none"> • Ulp1 proteasa, dominio catalítico C-terminal / 1441 	7-21 / 594	90/176 (51%)	14	PEPTIDASA, UPI PROTEASA
	<ul style="list-style-type: none"> • Serina proteasa, familia de Tripsina / 523 	11-39 / 822	39/99 (39%)	52	TRIPSINA LIKE SERINA PROTEASA
<ul style="list-style-type: none"> • ATPasa / 392 	8-3 / 610	132/200 (66%)	52	ATPasa	

7.2.2 LOCALIZACIÓN EN EL GENOMA

BANCO DE CÓSMIDOS

Se obtuvieron los siguientes resultados de la localización de las 23 clonas seleccionadas de la genoteca de la cepa Ch24-10, con el banco de cósmidos: 12 clonas dieron señal positiva con uno, dos o cuatro cósmidos. Algunas clonas hibridaron en el mismo cósmido, por ejemplo 6-40, 4-40 y 7-21 que dieron señal positiva con el cósmido 3.

En la figura 6 se observa un ejemplo de la señal positiva de hibridación ya que no se muestran todas las imágenes de las autorradiografías. En la tabla 5, columna 1, se muestran los resultados detallados de esta hibridación.

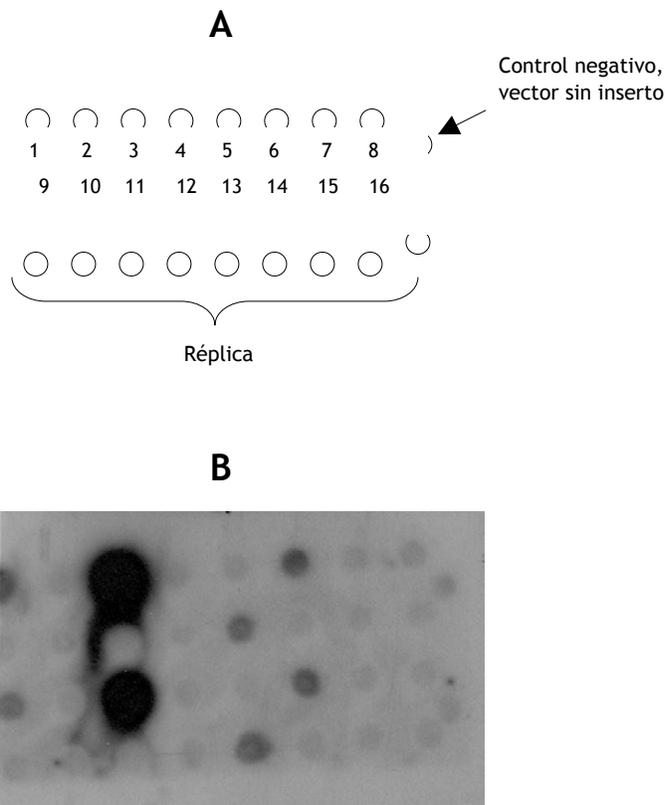


Figura 6. Hibridación con el banco de cósmidos.

A. Patrón de distribución de los cósmidos, control negativo y réplica. B. Ejemplo de hibridación positiva. Clona J4-40, señal positiva con el cósmido 3.

ADN TOTAL

Las 23 clonas seleccionadas de la librería de la cepa Ch24-10 para conocer su localización en el genoma, no hibridaron con el genoma de la cepa CFN42, incluido como control negativo, por el contrario, todas lo hicieron con el genoma de la cepa Ch24-10. En la tabla 6, columna 2 se encuentran los resultados en detalle y en la figura 7 se muestra una imagen de la autorradiografía con el resultado de la hibridación de la clona 4-32.

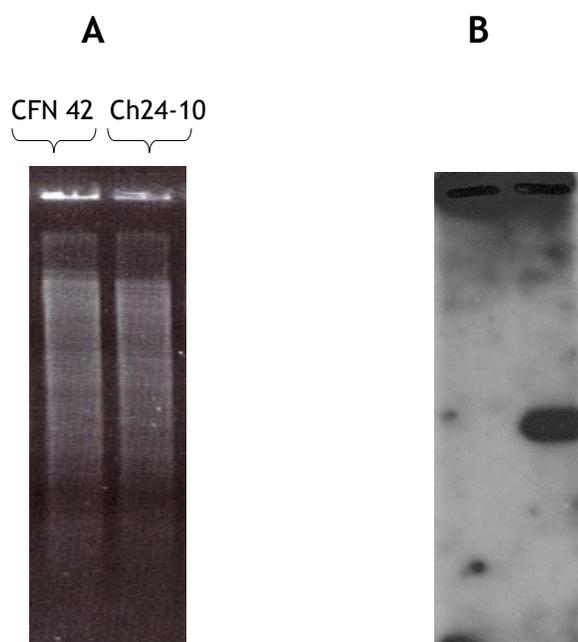


Figura 8. Hibridación con ADN total de las cepas CFN42 y Ch24-10

A. ADN total digerido con *EcoRI* B. Ejemplo de hibridación positiva con el genoma de la cepa Ch24-10 y negativa con el de la CFN42, clona 4-32

PLÁSMIDOS

De las 23 clonas que se hibridaron contra el genoma plasmídico de las cepas Ch24-10, CFN42 ninguna hibridó con los plásmidos de la cepa CFN42. Mientras que 7 clonas dieron señal positiva con algún plásmido de la cepa Ch24-10, dos clonas se ubicaron en el plásmido a (4-40 y 3-35), una en el plásmido b (7-21) y cuatro en el plásmido c (6-40, 4-32, 4-20, 4-34) (Tabla 5, columna 3). En la figura 8 se muestra una imagen con una autorradiografía con señal positiva en plásmido c de la cepa Ch24-10.

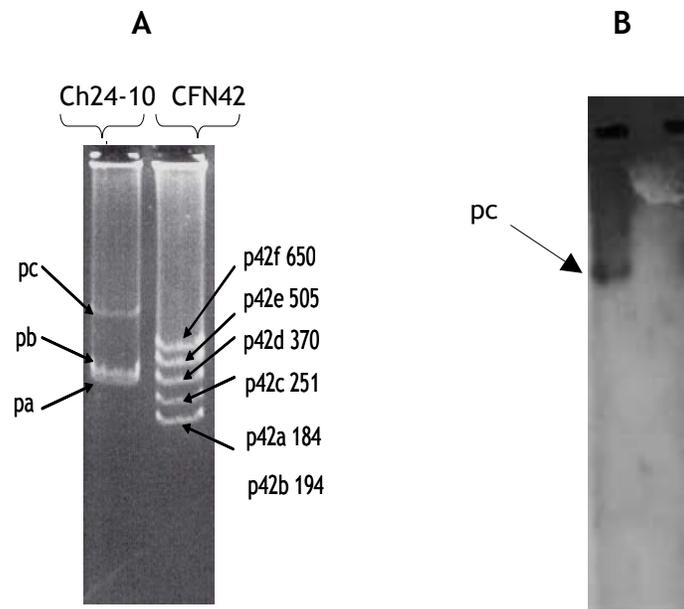


Figura 8. Hibridación con ADN plasmídico de las cepas CFN42 y Ch24-10.

A. Patrón plasmídico de *R. etli* Ch24-10 y CFN42. B Ejemplo de hibridación positiva de la clona 4-20 con el plásmido c de la cepa Ch24-10

Tabla 6. Resultados de la hibridación tipo Southern blot de ADN total y Eckhardt de las 23 clonas seleccionadas. Se muestra el número de bandas y el tamaño aproximado en Kilobases que hibridaron con el ADN total.

Secuencia	Función putativa	Columna 1 Cósmidos (número de cósmido)	Columna 2 ADN total/ No. bandas	Columna 3 Eckhardt
3-25	Transposasa	1	1 (22.2)	x
6-40	Transposasa para secuencia de inserción ISRM18	3	2 (23, 9)	pc
4-40	Transposasa para secuencia de inserción ISRM18	3	3 (9.2, 6.3)	pa
5-4	Endonucleasa	5, 11	1 (6)	X
4-37	Helicasa C, caja DEAD/DEAH	x	1 (4.2)	X
3-46	ADN Metilasa N4/N6	x	1 (9.2)	X
2-18	ADN Metilasa específica de adenina	x	1 (9.4)	X
4-32	Regulador transcripcional (Tipo LysR)	4	1 (4.1)	pc
7-22	Gen inducible por el hospedero	x	1 (3.4)	X
7-31	Proteína de respuesta. Histidina cinasa	8, 12	2 (6.4, 4)	X
6-31	Fitanoil-CoA dioxigenasa(PhyH)	x	1 (4)	X
3-45	Lisina/ ornitina monooxigenasa	x	1 (4)	pa
4-20	O-metiltransferasa	x	2 (6.4, 4)	pc
7-21	Ulp1 (familia de proteasa) Dominio catalítico C-terminal	3	3 (22, 6.3, 3.8)	pb
8-3	ATPasa	x	1 (1.6)	X
13-19	Proteína hipotética	14	2 (1.6, 1.7)	X
4-34	Proteína hipotética	x	1 (9)	pc
2-5	Proteína hipotética	x	1 (3.8)	X
4-1	Proteína hipotética	x	1 (9)	X
6-19	Proteína hipotética	x	1 (6)	X
7-30	Proteína hipotética	x	1 (6)	X
7-15	Proteína hipotética	2, 5, 9, 10	1 (2)	X
10-32	Proteína hipotética	x	1 (6)	X

X no se detectó señal en la hibridación

7.2.3 COMPARACIÓN DE SECUENCIAS CON EL GENOMA PARCIAL DE OTRAS CEPAS DE *R. etli*

Se compararon 87 clonas de la genoteca de la cepa Ch24-10 contra secuencias parciales de 4 cepas de *R. etli* aisladas de diferentes partes del mundo (Kim5 (México), Bra 5 (Brasil), 8C-3 (España) y CIAT 894(Colombia)). 30 de ellas presentaron identidad significativa a una o más cepas de *R. etli*, 13 pertenecen al grupo de identidad alta (13 con función putativa y 1 proteína hipotética) (Tabla 7A). Las 17 restantes pertenecen al grupo de secuencias con identidad baja (Tabla 7B).

Tabla 7. Identidad de 30 secuencias de la librería genómica de la cepa Ch24-10 con secuencias genómicas parciales de otras cuatro cepas de *R. etli*: Kim5, Bra 5, 8C-3 y CIAT 894
Se muestran los valores que tuvieron un puntaje (score) mayor a 100 y porcentaje de identidad mayor al 80%

7A. Identidad al genoma de distintas cepas de *R. etli* con secuencias de la librería de la cepa Ch24-10 que presentaron identidad significativa a proteínas anotadas.

Clona	Tamaño de la secuencia (nucleótidos)	Kim 5 % Identidad / Bases	Bra 5 % Identidad / Bases	8C-3 % Identidad / Bases	CIAT894 % Identidad / Bases
6-40 Integrasa	453	100 / 255	99 / 442	99 / 442	100 / 285
4-40 Integrasa	755	99 / 562	99 / 562	100 / 541	99 / 351 98 / 199
4-32 Regulador tipo LysR	745	-	96 / 386	-	89 / 748
seq46 Gen de la competitividad nodulación	576	99 / 383	98 / 437	-	-
2-34 Glicosiltransferasa (biogénesis de la pared celular)	735	-	96 / 415 95 / 105	99 / 603	-
3-47 Proteína de síntesis polisacáridos (antígeno O)	315	-	-	99 / 324	-
5-38 Proteína de síntesis de polisacáridos	628	-	-	99 / 332	-
J4-6 Lux A	346	99 / 306	-	-	-
6-31 Fitanoil dioxigenasa	746	-	96 / 446 98 / 91	-	-
3-45 Lisina / ornitina Monooxigenasa	761	99 / 710	99 / 522	99 / 572	-
7-21 UP1 proteasa	594	99 / 562 100 / 248 99 / 206	98 / 562 99 / 299 99 / 261	98 / 409	-
8-3 ATPasa	396	-	-	99 / 598	-
5-30 Proteína hipotética	877	-	96 / 132	-	-

Tabla 7B. Identidad de las secuencias de la cepa Ch24-10 que presentaron identidad no significativa con otra proteína anotada, al genoma de distintas cepas de *R. etli*.

Clona	Tamaño de la secuencia (nucleótidos)	Kim 5 % Identidad / Bases	Bra 5 % Identidad / Bases	8C-3 % Identidad / Bases	CIAT894 % Identidad / Bases
1-44	337	-	95 / 99	-	-
3-10	730	91 / 721 86 / 335 91 / 79	92 / 686 97 / 471 86 / 409	91 / 603 93 / 525 92 / 294	90 / 383 86 / 525 86 / 409
3-36	479	-	-	-	93 / 339
4-1	484	-	-	98 / 115	97 / 469
4-11	419	82 / 209	-	-	-
4-12	617	98 / 602 89 / 271 89 / 89	90 / 519 90 / 274 87 / 331	88 / 521 87 / 433 87 / 296	88 / 470 87 / 416 87 / 331
4-24	356	99 / 324	-	-	-
4-25	700	-	86 / 303	84 / 343	-
5-39	611	-	-	99 / 592	-
5-44	580	-	-	98 / 150	-
7-25	518	-	-	98 / 150	-
8-5	238	-	99 / 525	-	-
10-16	769	99 / 346	-	-	-
seq5	596	-	-	88 / 330	96 / 119
seq10	785	-	100 / 351	-	-
seq15	792	-	86 / 303	84 / 343	-
seq24	343	-	96 / 275	-	-

7. RESULTADOS

7.1 SECUENCIAS CON IDENTIDAD AL GENOMA DE LA CEPA CFN42

De acuerdo a los resultados del blastx, se obtuvieron 69 secuencias de la genoteca de la cepa Ch24-10 con identidad con genoma de la cepa CFN42 (organizado en 7 replicones organizados en 1 cromosoma y 6 plásmidos) (González *et al.*, 2006).

De este grupo de secuencias, 23 de ellas (33%), se localizan en los plásmidos: 1 en el p42a, 5 en el p42b, 2 en el p42c, 9 en el plásmido simbiótico psym (p42d), 2 en el p42e y 4 en el p42f. De las 69 secuencias, 16 presentaron identidad con el genoma de *Rhizobium leguminosarum* bv. *viciae* 384, en un porcentaje mayor a 40%, un valor de E menor de 10^{-5} y un score superior a 80. En la tabla 4A, 4B y 4C se encuentran los resultados resumidos de este análisis.

Tabla 4. Secuencias con identidad al genoma de *R. etli* CFN42.

Tabla 4A. Secuencias con identidad mayor de 75% (en más de 60 aminoácidos) y valor de E menor de 10^{-5} .

Se muestra la identidad en el número de aminoácidos (aa) y su porcentaje (%), la longitud (nucleótidos) de la secuencia de la librería de la cepa Ch24-10, el valor de E, el score y el número de veces que se repitió cada secuencia.

SECUENCIA / LONGITUD (nn)	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
9-48 / 348	ADN polimerasa III subunidad alfa	100/105 (95%)	214	9.00E-55	C	1
12-4 / 650	Secuencia de inserción putativa, transposasa	135/155 (87%)	267	3.00E-70	psym	1
6-18 / 579	Probable regulador transcripcional, familia LuxR	73/92 (79%)	146	5.00E-34	C	2
8-39 / 818	Probable regulador transcripcional, familia AraC	231/244 (94%)	469	4.00E-131	p42f	1
9-10 / 878	Regulador transcripcional putativo	71/71 (100%)	145	3.00E-33	C	1
seq2 / 454	Proteína sensora del sistema de dos componentes, histidina cinasa*	99/106 (93%)	171	1.00E-41	C	1
5-27 / 711	Caja sensora putativa / familia de proteínas con dominio GGDEF *	68/84 (80%)	139	1.00E-31	C	1
4-4 / 255	Translocasa, exportadora de proteínas	73/73 (100%)	149	4.00E-35	C	2
9-28 / 349	Probable transportador ABC de azúcares, proteína de unión al sustrato*	91/104 (87%)	212	5.00E-54	p42b	1
11-5 / 460	Transportador ABC de azúcares, permeasa	140/140 (100%)	218	6.00E-56	psym	1
seq8 / 274	Transportador ABC de azúcares, permeasa	68/68 (100%)	131	8.00E-30	p42b	1
1-32 / 612	Probable transportador ABC de oligopéptidos, permeasa	158/161 (98%)	317	2.00E-85	p42f	1
6-2 / 503	Probable transportador ABC de azúcares, permeasa	103/105 (98%)	108	7.00E-23	p42b	1
5-24 / 396	Probable transportador ABC de azúcares, proteína de unión al sustrato	89/90 (98%)	154	2.00E-36	C	1

(Psym: p42d, *Secuencias con identidad al genoma de *Rhizobium leguminosarum* bv. *viciae* 384, C: cromosoma, p; plásmido).

Tabla 4A. Continuación...

SECUENCIA / LONGITUD (nn)	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
3-20 / 825	Probable transportador ABC de azúcares, proteína de unión al sustrato	94/108 (88%)	197	5.00E-49	p42b	1
4-21 / 730	Putativa toxina RTX tipo hemolisina, proteína de unión a calcio*	168/231 (72%)	358	2.00E-97	C	1
7-19 / 452	Hemolisina*	107/108 (99%)	218	9.00E-56	C	1
5-25 / 471	Proteína de quimiotaxis, aceptora de metilos	88/88 (100%)	177	2.00E-43	C	1
6-32 / 551	Probable succinato semialdehído dehidrogenasa [NADP+]	171/171 (100%)	292	6.00E-78	psym	1
7-6 / 839	Proteína de autoagregación (proteína de adherencia)	165/180 (91%)	332 bits	9.00E-90	C	1
seq34 / 567	Proteína transmembranal, biosíntesis de flagelo	91/92 (98%)	179	6.00E-44	C	1
8-48 / 351	Probable reductasa FMN dependiente de NAD(P)H	79/103 (76%)	164	1.00E-39	p42c	1
seq17 / 564	Endopeptidasa Clp, subunidad 3	87/89 (97%)	178	1.00E-43	C	1
seq25 / 394	Putativa L-iditol 2-dehidrogenasa	107/108 (99%)	222	4.00E-57	psym	1
seq36 / 818	Ciclohidrolasa II de GTP	166/170 (97%)	327	5.00E-88	psym	1
seq49 / 838	Ceto-ácido reductoisomerasa	207/253 (81%)	389	6.00E-107	C	1
seq53 / 555	Nitrogenasa (iron protein)	176/179 (98%)	309	4.00E-83	psym	1
3-41 / 781	Proteína hipotética RHE_CH00091	66/70 (94%)	134	3.00E-30	C	1
1-39 / 546	Proteína hipotética RHE_CH02651	99/129 (76%)	201	1.00E-50	C	2
5-14 / 363	Proteína hipotética RHE_CH01108	86/109 (78%)	184	1.00E-45	C	4
3-33 / 902	Proteína hipotética RHE_CH03202	77/97 (79%)	152	2.00E-35	C	1
5-12 / 795	Proteína hipotética RHE_CH01631	136/154 (88%)	281	2.00E-74	C	1
6-3 / 560	Proteína hipotética RHE_CH03237*	84/97 (86%)	124	3.00E-27	C	2
seq20 / 547	Proteína hipotética RHE_CH02651	99/129 (76%)	201	1.00E-50	C	1
7-1 / 832	Proteína hipotética RHE_CH01631	140/161 (86%)	288	3.00E-76	C	2

Tabla 4B. Secuencias con identidad al menor de 75% y/o menos de 60 aminoácidos y valor de E menor de 10^{-5} .

SECUENCIA / LONGITUD (nm)	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
10-1 / 631	Secuencia de inserción putativa, transposasa*	49/63 (77%)	109	7.00E-23	p42b	1
4-13 / 765	Probable regulador transcripcional, familia MerR*	38/40 (95%)	82.8	1.00E-14	C	4
seq7 / 386	Proteína reguladora específica de Nif	41/45 (91%)	83.6	2.00E-15	psym	1
seq37 / 725	Porina, proteína de membrana externa	62/63 (98%)	87	6.00E-16	p42e	1
9-25 / 291	Proteína putativa de biosíntesis de exopolisacáridos	54/64 (84%)	111	9.00E-24	C	1
4-3 / 492	Adenilato ciclasa*	74/153 (48%)	157	2.00E-37	C	1
3-37 / 516	Probable glicosiltransferasa*	38/66 (57%)	82-4	8.00E-15	C	2
5-47 / 227	Gama-carboximuconolactone decarboxilasa*	31/31 (100%)	65.5	8.00E-10	p42f	1
12-29 / 769	Reductasa putativa*	43/43 (100%)	91.7	3.00E-17	C	1
8-16 / 815	Cetol piruvato transferasa*	133/220 (60%)	263	8.00E-69	C	1
8-18 / 866	Cetol piruvato transferasa*	133/220 (60%)	263	4.00E-69	C	1
seq38 / 443	GDP-L-fucosa sintasa*	80/134 (59%)	172	4.00E-42	C	2
seq52 / 911	Fosfomanomutasa	66/104 (63%)	133	1.00E-29	C	1
9-29 / 541	Proteína hipotética RHE_CH01538	55/70 (78%)	109	5.00E-23	C	1
seq47 / 512	Proteína hipotética RHE_CH01807	50/57 (87%)	110	2.00E-23	C	1
8-27 / 561	Proteína hipotética RHE_CH02021*	43/57 (75%)	93.6	4.00E-18	C	1
seq50 / 229	Proteína hipotética RHE_CH00223	46/46 (100%)	95.5	7.00E-19	C	1
2-36 / 357	Proteína hipotética RHE_PE00109	24/26 (92%)	57	3.00E-07	p42e	7
seq13 / 279	Proteína hipotética RHE_PD00011	33/40 (82%)	67.4	2.00E-10	psym	1
seq16 / 351	Proteína hipotética RHE_CH00444	28/29 (96%)	57	3.00E-07	C	1

Tabla 4B. Continuación...

SECUENCIA / LONGITUD	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
----------------------	-----------	-----------	-------	------------	---------------------------	---------------------

(nn)						
8-11 / 286	Proteína hipotética RHE_PF00054	33/60 (55%)	67	3.00E-10	p42f	1
9-17	Proteína hipotética RHE_CH01905*	85/89 (95%)	169	4.00E-41	C	3
7-35/ 521	Proteína hipotética RHE_CH02761	33/43 (76%)	66.2	6.00E-10	C	9
5-46 / 365	Proteína hipotética RHE_PC00203*	38/41 (92%)	129	5.00E-29	p42c	2
6-1 / 576	Proteína hipotética RHE_CH01114*	45/89 (50%)	161	2.00E-38	C	1
6-14 / 413	Proteína hipotética RHE_CH01669*	43/47 (91%)	90	2.00E-18	C	1
seq40 / 402	Proteína hipotética RHE_CH01114*	45/89 (50%)	80.5	2.00E-14	C	1
9-11 / 597	Proteína hipotética RHE_PA00030*	57/126 (45%)	97.4	3.00E-19	p42a	1
J11-40 / 700	Proteína hipotética RHE_PD00172	144/220 (65%)	293	4.00E-78	psym	1

Tabla 4C. Secuencias con identidad al genoma de la cepa CFN42 con un valor de E mayor de 10^{-5} .

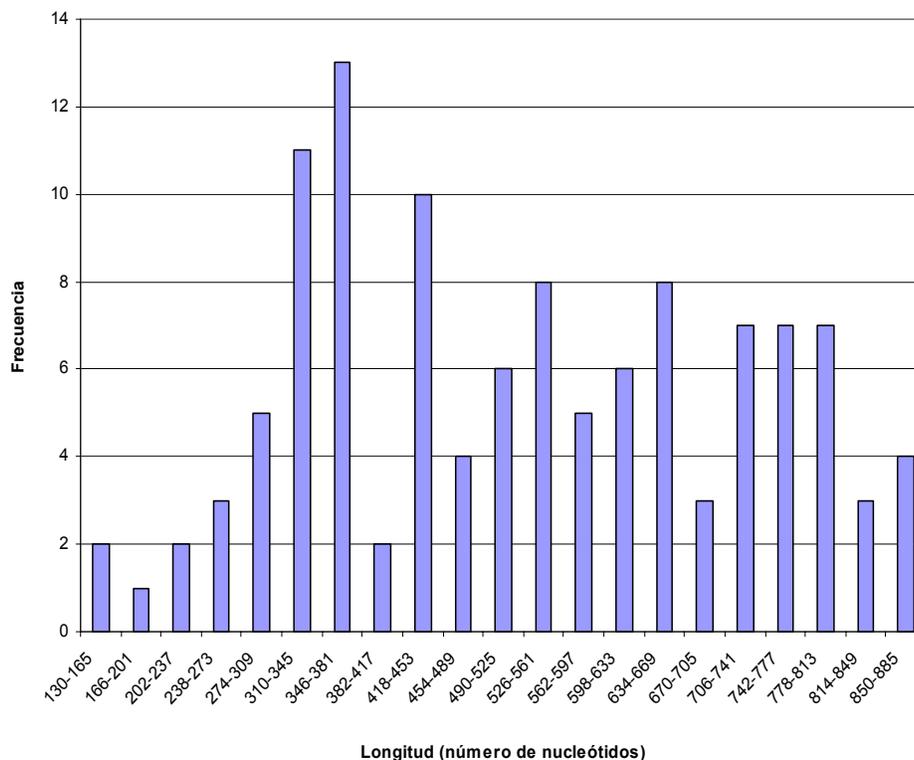
SECUENCIA / LONGITUD (nn)	ANOTACIÓN	IDENTIDAD	SCORE	VALOR DE E	LOCALIZACIÓN EN EL GENOMA	VECES QUE SE REPITE
seq30 / 548	Proteína ribosomal L33 50S	18/18 (100%)	41.2	0.02	C	1
J3-21 / 461	Proteína del sistema de dos componentes, regulador de respuesta	22/22 (100%)	47	3.00E-04	C	3
J9-40 / 633	Probable acetiltransferasa	20/23 (86%)	43.1	0.009	C	1
J5-23 / 300	Proteína hipotética RHE_CH00479	19/19 (100%)	50.4	3.00E-05	C	1
J5-2 / 414	Proteína hipotética RHE_CH03754*	49/64 (76%)	42.4	0.007	C	1

7.2 SECUENCIAS SIN IDENTIDAD AL GENOMA DE LA CEPA CFN42

Como se mencionó, de las 348 clonas que se secuenciaron 243 no presentaron identidad al genoma de la cepa CFN42. De estas 243 secuencias 120 fueron distintas entre sí, es importante notar que algunas secuencias se repetían en número variable, en algunos casos hasta 17 veces, como la secuencia 3-34, por otro lado había secuencias únicas como las secuencias seq45, 8-34 y 2-18 por mencionar algunas (Ver Anexo, en la tabla 8 se muestran las secuencias repetidas). En los casos en que se repitieron se tomó una sola secuencia para realizar los análisis posteriores.

La longitud de las 120 secuencias se encuentra en un rango que va de los 130 a los 884 nucleótidos, con un promedio de 530 (los insertos clonados se secuenciaron completos). En el gráfico 1 se muestra que los fragmentos más pequeños se ubican en el intervalo que va de 130 a 165 nucleótidos y que los más grandes tienen una longitud entre 850 y 885. La mayoría se distribuye en el rango de 300 a 800 nucleótidos.

Gráfico 1. Longitud de los fragmentos clonados de la librería.



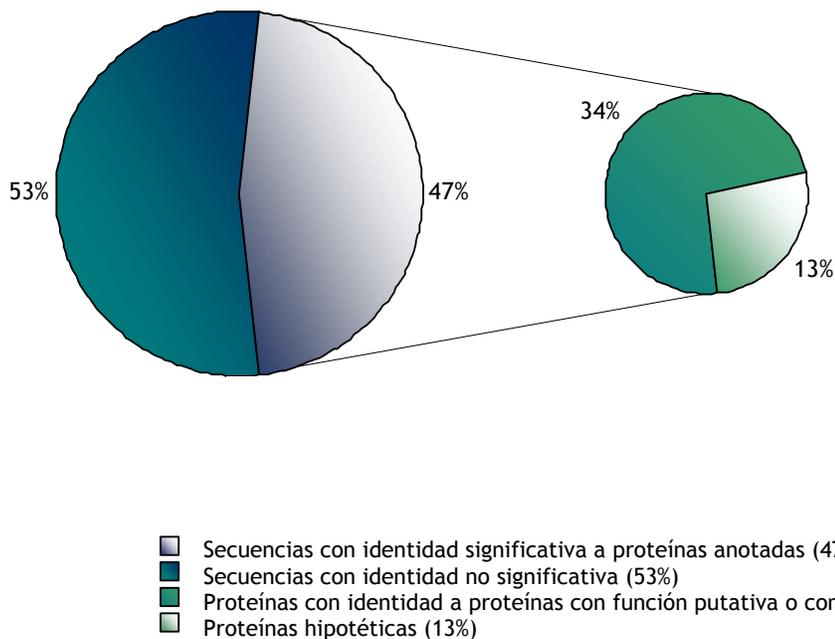
Las 120 secuencias sin identidad al genoma de la cepa CFN42 se organizaron en dos grupos generales (Gráfico 2):

- 1) 56 secuencias (47%), con identidad significativa a proteínas anotadas con una identidad igual o superior al 30%, valor de E menor a 10^{-5} y score superior a 40 (Los resultados en detalle del BLAST se muestran en el Anexo en la tabla 8)
- 2) 64 secuencias (53%), sin identidad significativa con una identidad menor al 30%, valor de E mayor de 10^{-5} y score inferior a 40) (Los datos no se muestran).

Por otro lado, las secuencias con identidad significativa se dividieron a su vez en:

- a) 41 secuencias con identidad a una proteína putativa o conocida (32% del total, 73% de este grupo).
- b) 15 secuencias con identidad a proteínas hipotéticas (13% del total, 27% de este grupo).

Gráfico 2. Distribución porcentual de las 120 secuencias con respecto a la identidad que presentaron en el blastx.



7.2.1 SECUENCIAS CON IDENTIDAD SIGNIFICATIVA A PROTEÍNAS PUTATIVAS O CONOCIDAS.

Las secuencias con identidad a proteínas putativas o conocidas se agruparon en cinco categorías de acuerdo al grupo funcional del ortólogo correspondiente (Gráfico 3). El grupo con mayor número de secuencias es el relacionado con replicación, recombinación y reparación del ADN, en segundo lugar se ubica el grupo relacionado con transporte celular y regulación de respuesta. Este grupo presentó identidad por un lado, a proteínas cuya función esta comprobada experimentalmente (TonB y gen inducible por el hospedero de *R. leguminosarum*), y por otro a proteínas putativas con función conservada (el resto).

Al realizar la traducción a aminoácidos de las secuencias nucleotídicas con identidad a proteínas putativas, algunas de ellas mostraron identidad a dominios conservados, en la mayoría de los casos correspondientes al resultado del blastx (tabla 5 y en detalle en el Anexo, tabla 8).

Gráfico 3. Distribución porcentual de las secuencias de Identidad alta a proteínas putativas. Organización en grupos funcionales.

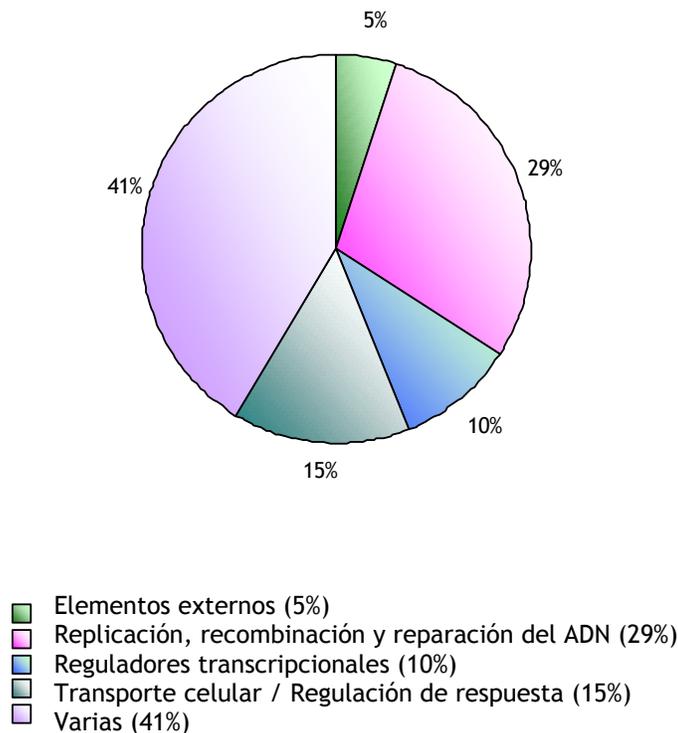


Tabla 5. Organización en grupos funcionales de las secuencias con identidad a proteínas putativas o conocidas.

Se muestra la longitud de la secuencia proteica del ortólogo en aminoácidos (aa), así como de la secuencia de la librería de la cepa Ch24-10 en nucleótidos (nn). Se muestra también la identidad en número de aa y su porcentaje (%) correspondiente y el porcentaje que cubre la secuencia de la librería traducida a aminoácidos respecto a la longitud de la proteína con la que presentó similitud. Se indica el dominio al que presentó identidad la secuencia de la librería después de su traducción a aminoácidos y la realización del blastp y en su caso el plásmido con el que hibridaron.

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD (aa) / (%)	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Relacionadas con fagos	• Proteína mayoritaria de la cápside gp7 / 227 proteína hipotética de <i>R. leguminosarum</i> bv. viciae 3841 similar a esta de fago	7-10 / 638	79/138 (57%)	93	x
	• Integrasa / 999	seq45 / 347	23/67 (34%)	12	Integrasa de fago
Replicación, recombinación y reparación del ADN	• Transposasa y derivados inactivados / 781	3-25 / 641	70/195 (35%)	28	x
	• ***Transposasa para secuencia de inserción ISRM18/ 507	6-40 / 453	143/147 (97%)	30	Core de la integrasa
	• * Transposasa para secuencia de inserción ISRM18 / 507	4-40 / 755	123/132 (93%)	50	x
	• Transposasa / 266	8-30 / 396	35/58 (60%)	50	x
	• ADN metilasa N4/N6 / 613	1-16 / 542	117/179 (65%)	29	ADN metilasa
	• ADN metilasa N4/N6 / 560	3-46 / 559	111/179 (62%)	33	ADN metilasa específica de adenina
	• ADN metilasa específica de adenina / 194	2-18 / 225	43/70 (61%)	39	de adenina
	• ADN metilasa N4/N6 / 938	9-9 / 470	45/107 (42%)	17	x
	• ADN metilasa específica de citosina C-5 / 710	7-11 / 756	38/90 (42%)	35	x
	• Helicasa caja DEAD/DEAH dependiente de ATP / 811	10-18 / 522	101/114 (88%)	21	Enzima de restricción I
• Helicasa C terminal, caja DEAD/DEAH, dependiente de ATP / 963	4-37 / 734	83/148 (56%)	25	DEAD helicasa	
• Endonucleasa de restricción (subunidad helicasa) / 999	5-4 / 516	124/177 (70%)	17	ADN o ARN helicasa	

(*pa, **pb, ***pc; X = no se encontró similitud con algún dominio conservado)

Tabla 5. Continuación...

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD (aa) / (%)	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Reguladores Transcripcionales	<ul style="list-style-type: none"> Regulador transcripcional similar a GntR / 218 Hélice –vuelta-hélice	3-35 / 307	38/98 (38%)	47	GntR
	<ul style="list-style-type: none"> * Regulador tipo LysR / 328 	4-32 / 745	127/249 (51%)	76	LysR
	<ul style="list-style-type: none"> Regulador transcripcional tipo RioFf 87 / Similar al dominio de unión al gen <i>traR</i> (activa el operón de bioluminiscencia) / 62 	6-23 / 627	34/57 (59%)	100	LuxR
	<ul style="list-style-type: none"> Hélice-vuelta-hélice tipo AraC / 315 	8-34 / 808	68/90 (75%)	85	ATT AraC
Transporte celular / Regulación de respuesta	<ul style="list-style-type: none"> Proteína sensora, traducción de señal, histidina cinasa / 384 [Regulador transcripcional con dominio PAS, tipo-AAA ATPasa]	7-31 / 750	89/150 (59%)	65	PAS
	<ul style="list-style-type: none"> Diguanilato ciclasa/Fosofodiesterasa con dominio GGDEF & EAL /912 	7-14 / 745	34/63 (53%)	27	EAL
	<ul style="list-style-type: none"> Diguanilato ciclasa-Fosofodiesterasa con dominio GGDEF & EAL / 713 [Proteínas con este dominio pueden actuar como sensores]	2-1 / 329	34/49 (69%)	15	EAL
	<ul style="list-style-type: none"> Gen inducible por el hospedero (ORF 3) / 295 [Importante para competencia en la nodulación]	7-22 / 576	123/147 (83%)	65	X
	<ul style="list-style-type: none"> TonB. Proteína de membrana, traduce energía a receptores de membrana externa. Transporte de hierro / 454 	9-20 / 336	62/106 (58%)	25	X
<ul style="list-style-type: none"> Canal iónico. Tipo transportador de potasio / 343 	Seq12 / 503	35/62 (55%)	49	X	

Tabla 5. Continuación...

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD aa / %	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Proteínas varias	<ul style="list-style-type: none"> Acetiltransferasa / 340 [Metabolismo de lípidos]	12-27 / 168	25/44 (56%)	16	X
	<ul style="list-style-type: none"> Glicosiltransferasa familia 2, involucrada en la biogénesis de la pared celular / 307 	2-34 / 735	46/152 (30%)	80	X
	<ul style="list-style-type: none"> Glicosiltransferasa (síntesis de core del lipopolisacárido) / 384 	8-17 / 862	29/95 (30%)	75	Glicosiltransferasa
	<ul style="list-style-type: none"> Glicosiltransferasa familia 8 / 243 Proteína involucrada en la biosíntesis de polisacáridos / 489 Proteína transportadora de polisacáridos / 488 [Proteína de biosíntesis de polisacáridos, generalmente son integrales de membrana. La familia incluye a RfbX parte del operón de biosíntesis del antígeno O] [Biosíntesis de disacáridos, polisacáridos]	7-43 / 315 3-47 / 448 5-38 / 628	38/103 (36%) 46/152 (30%) 29/95 (30%)	43 21 42	Glicosiltransferasa Proteína de biosíntesis de polisacáridos Proteína de biosíntesis de polisacáridos
<ul style="list-style-type: none"> Glicósidohidrolas, familia 19 / 201 [Dominio: quitinasa]	1-40 / 657	45/86 (52%)	100	X	

Tabla 5. Continuación...

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD aa / %	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Proteínas varias	<ul style="list-style-type: none"> • *** Metiltransferasa / 352 [Transferencia de metilo del S-adenosil-L-metionina (AdoMet) a cualquier átomo de nitrógeno, oxígeno o carbono. Modifica DNA, RNA, proteínas, moléculas pequeñas como el catecol con fines regulatorios]	4-20 / 432	41/54 (75%)	41	X
	<ul style="list-style-type: none"> • Proteína LuxA Putativa / 354 Cadena alfa de la luciferasa monooxigenasa [La luciferasa bacteriana es una flavina monooxigenasa que cataliza la oxidación de una cadena larga de aldehído provocando la emisión de energía en forma lumínica]	4-6 / 346	35/56 (62%)	32	Luciferasa bacteriana
	<ul style="list-style-type: none"> • Fitanoil CoA dioxigenasa (PhyH) / 259 [PhyH es una enzima peroxisomal que cataliza el primer paso en la oxidación-alfa del ácido fitánico Es una oxidorreductasa]	6-31 / 746	63/107 (58%)	98	Fitanoil coA dioxigenasa
	<ul style="list-style-type: none"> • * Lisina/ornitina monooxigenasa IucD / 540 [Biosíntesis de metabolitos secundarios, transporte y catabolismo] n6-oxigenación de lisina, y el primer paso en la síntesis de la aerobactina	3-45 / 761	68/131 (52%)	56	Lisina/ornitina monooxigenasa IncD
<ul style="list-style-type: none"> • Nitrilohidratasa / 153 [Hidratación de compuestos nitrilo. Actúa en las cadenas cortas de nitrilos alifáticos convirtiéndolos en el aminoácido correspondiente. No actúa en los aminoácidos ni en los nitrilos aromáticos]	12-22 / 137	27/31 (87%)	29	X	

Tabla 5. Continuación...

GRUPOS FUNCIONALES	FUNCIÓN / LONGITUD DE LA PROTEÍNA (aa)	SECUENCIA / LONGITUD (nn)	IDENTIDAD aa / %	PORCENTAJE DE COBERTURA	DOMINIO (blastp)
Proteínas varias	<ul style="list-style-type: none"> • Formato dehidrogenasa / 126 [Puede actuar en el metabolismo del glioxilato/dicarboxilato, metabolismo del metano y en el sistema de dos componentes. Oxidoreductasa. Actúa en el grupo aldehído o grupo oxo de los donadores, el NAD ⁺ o NADP ⁺ son los aceptores]	8-35 / 344	21/25 (84%)	90	X
	<ul style="list-style-type: none"> • ** Ulp1 proteasa, dominio catalítico C-terminal/ 1441 	12-30 / 813	54/151 (35%)	19	X
	<ul style="list-style-type: none"> • Ulp1 proteasa, dominio catalítico C-terminal / 1441 	7-21 / 594	90/176 (51%)	14	PEPTIDASA, UPI PROTEASA
	<ul style="list-style-type: none"> • Serina proteasa, familia de Tripsina / 523 	11-39 / 822	39/99 (39%)	52	TRIPSINA LIKE SERINA PROTEASA
<ul style="list-style-type: none"> • ATPasa / 392 	8-3 / 610	132/200 (66%)	52	ATPasa	

7.2.2 LOCALIZACIÓN EN EL GENOMA

BANCO DE CÓSMIDOS

Se obtuvieron los siguientes resultados de la localización de las 23 clonas seleccionadas de la genoteca de la cepa Ch24-10, con el banco de cósmidos: 12 clonas dieron señal positiva con uno, dos o cuatro cósmidos. Algunas clonas hibridaron en el mismo cósmido, por ejemplo 6-40, 4-40 y 7-21 que dieron señal positiva con el cósmido 3.

En la figura 6 se observa un ejemplo de la señal positiva de hibridación ya que no se muestran todas las imágenes de las autorradiografías. En la tabla 5, columna 1, se muestran los resultados detallados de esta hibridación.

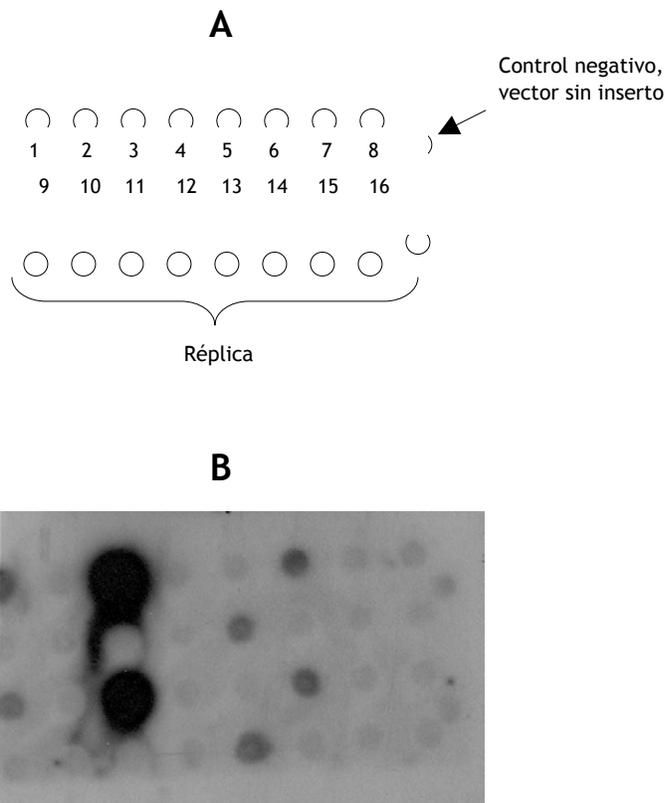


Figura 6. Hibridación con el banco de cósmidos.

A. Patrón de distribución de los cósmidos, control negativo y réplica. B. Ejemplo de hibridación positiva. Clona J4-40, señal positiva con el cósmido 3.

ADN TOTAL

Las 23 clonas seleccionadas de la librería de la cepa Ch24-10 para conocer su localización en el genoma, no hibridaron con el genoma de la cepa CFN42, incluido como control negativo, por el contrario, todas lo hicieron con el genoma de la cepa Ch24-10. En la tabla 6, columna 2 se encuentran los resultados en detalle y en la figura 7 se muestra una imagen de la autorradiografía con el resultado de la hibridación de la clona 4-32.

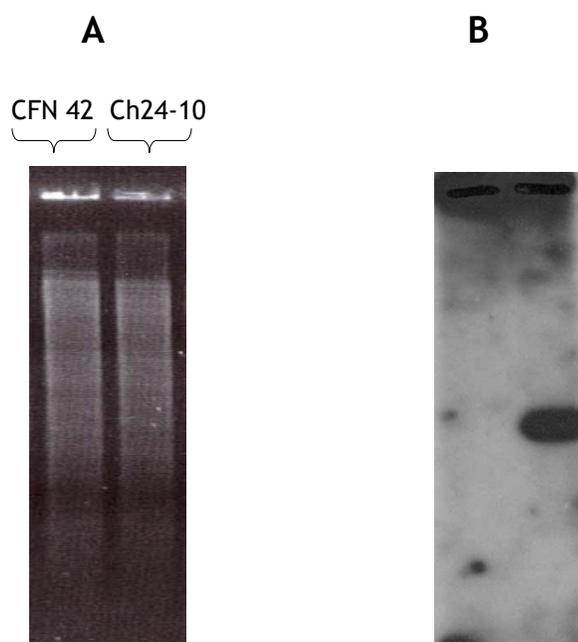


Figura 8. Hibridación con ADN total de las cepas CFN42 y Ch24-10

A. ADN total digerido con *EcoRI* B. Ejemplo de hibridación positiva con el genoma de la cepa Ch24-10 y negativa con el de la CFN42, clona 4-32

PLÁSMIDOS

De las 23 clonas que se hibridaron contra el genoma plasmídico de las cepas Ch24-10, CFN42 ninguna hibridó con los plásmidos de la cepa CFN42. Mientras que 7 clonas dieron señal positiva con algún plásmido de la cepa Ch24-10, dos clonas se ubicaron en el plásmido a (4-40 y 3-35), una en el plásmido b (7-21) y cuatro en el plásmido c (6-40, 4-32, 4-20, 4-34) (Tabla 5, columna 3). En la figura 8 se muestra una imagen con una autorradiografía con señal positiva en plásmido c de la cepa Ch24-10.

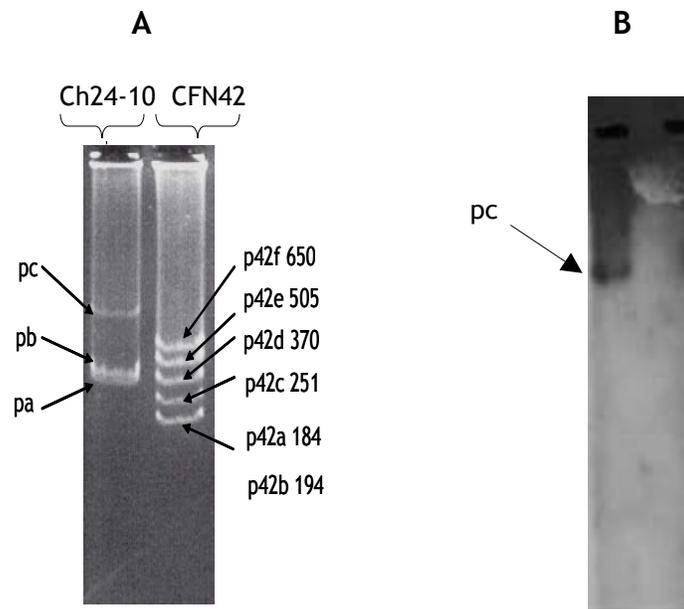


Figura 8. Hibridación con ADN plasmídico de las cepas CFN42 y Ch24-10.

A. Patrón plasmídico de *R. etli* Ch24-10 y CFN42. B Ejemplo de hibridación positiva de la clona 4-20 con el plásmido c de la cepa Ch24-10

Tabla 6. Resultados de la hibridación tipo Southern blot de ADN total y Eckhardt de las 23 clonas seleccionadas. Se muestra el número de bandas y el tamaño aproximado en Kilobases que hibridaron con el ADN total.

Secuencia	Función putativa	Columna 1 Cósmidos (número de cósmido)	Columna 2 ADN total/ No. bandas	Columna 3 Eckhardt
3-25	Transposasa	1	1 (22.2)	x
6-40	Transposasa para secuencia de inserción ISRM18	3	2 (23, 9)	pc
4-40	Transposasa para secuencia de inserción ISRM18	3	3 (9.2, 6.3)	pa
5-4	Endonucleasa	5, 11	1 (6)	X
4-37	Helicasa C, caja DEAD/DEAH	x	1 (4.2)	X
3-46	ADN Metilasa N4/N6	x	1 (9.2)	X
2-18	ADN Metilasa específica de adenina	x	1 (9.4)	X
4-32	Regulador transcripcional (Tipo LysR)	4	1 (4.1)	pc
7-22	Gen inducible por el hospedero	x	1 (3.4)	X
7-31	Proteína de respuesta. Histidina cinasa	8, 12	2 (6.4, 4)	X
6-31	Fitanoil-CoA dioxigenasa(PhyH)	x	1 (4)	X
3-45	Lisina/ ornitina monooxigenasa	x	1 (4)	pa
4-20	O-metiltransferasa	x	2 (6.4, 4)	pc
7-21	Ulp1 (familia de proteasa) Dominio catalítico C-terminal	3	3 (22, 6.3, 3.8)	pb
8-3	ATPasa	x	1 (1.6)	X
13-19	Proteína hipotética	14	2 (1.6, 1.7)	X
4-34	Proteína hipotética	x	1 (9)	pc
2-5	Proteína hipotética	x	1 (3.8)	X
4-1	Proteína hipotética	x	1 (9)	X
6-19	Proteína hipotética	x	1 (6)	X
7-30	Proteína hipotética	x	1 (6)	X
7-15	Proteína hipotética	2, 5, 9, 10	1 (2)	X
10-32	Proteína hipotética	x	1 (6)	X

X no se detectó señal en la hibridación

7.2.3 COMPARACIÓN DE SECUENCIAS CON EL GENOMA PARCIAL DE OTRAS CEPAS DE *R. etli*

Se compararon 87 clonas de la genoteca de la cepa Ch24-10 contra secuencias parciales de 4 cepas de *R. etli* aisladas de diferentes partes del mundo (Kim5 (México), Bra 5 (Brasil), 8C-3 (España) y CIAT 894(Colombia)). 30 de ellas presentaron identidad significativa a una o más cepas de *R. etli*, 13 pertenecen al grupo de identidad alta (13 con función putativa y 1 proteína hipotética) (Tabla 7A). Las 17 restantes pertenecen al grupo de secuencias con identidad baja (Tabla 7B).

Tabla 7. Identidad de 30 secuencias de la librería genómica de la cepa Ch24-10 con secuencias genómicas parciales de otras cuatro cepas de *R. etli*: Kim5, Bra 5, 8C-3 y CIAT 894
Se muestran los valores que tuvieron un puntaje (score) mayor a 100 y porcentaje de identidad mayor al 80%

7A. Identidad al genoma de distintas cepas de *R. etli* con secuencias de la librería de la cepa Ch24-10 que presentaron identidad significativa a proteínas anotadas.

Clona	Tamaño de la secuencia (nucleótidos)	Kim 5 % Identidad / Bases	Bra 5 % Identidad / Bases	8C-3 % Identidad / Bases	CIAT894 % Identidad / Bases
6-40 Integrasa	453	100 / 255	99 / 442	99 / 442	100 / 285
4-40 Integrasa	755	99 / 562	99 / 562	100 / 541	99 / 351 98 / 199
4-32 Regulador tipo LysR	745	-	96 / 386	-	89 / 748
seq46 Gen de la competitividad nodulación	576	99 / 383	98 / 437	-	-
2-34 Glicosiltransferasa (biogénesis de la pared celular)	735	-	96 / 415 95 / 105	99 / 603	-
3-47 Proteína de síntesis polisacáridos (antígeno O)	315	-	-	99 / 324	-
5-38 Proteína de síntesis de polisacáridos	628	-	-	99 / 332	-
J4-6 Lux A	346	99 / 306	-	-	-
6-31 Fitanoil dioxigenasa	746	-	96 / 446 98 / 91	-	-
3-45 Lisina / ornitina Monooxigenasa	761	99 / 710	99 / 522	99 / 572	-
7-21 UP1 proteasa	594	99 / 562 100 / 248 99 / 206	98 / 562 99 / 299 99 / 261	98 / 409	-
8-3 ATPasa	396	-	-	99 / 598	-
5-30 Proteína hipotética	877	-	96 / 132	-	-

Tabla 7B. Identidad de las secuencias de la cepa Ch24-10 que presentaron identidad no significativa con otra proteína anotada, al genoma de distintas cepas de *R. etli*.

Clona	Tamaño de la secuencia (nucleótidos)	Kim 5 % Identidad / Bases	Bra 5 % Identidad / Bases	8C-3 % Identidad / Bases	CIAT894 % Identidad / Bases
1-44	337	-	95 / 99	-	-
3-10	730	91 / 721 86 / 335 91 / 79	92 / 686 97 / 471 86 / 409	91 / 603 93 / 525 92 / 294	90 / 383 86 / 525 86 / 409
3-36	479	-	-	-	93 / 339
4-1	484	-	-	98 / 115	97 / 469
4-11	419	82 / 209	-	-	-
4-12	617	98 / 602 89 / 271 89 / 89	90 / 519 90 / 274 87 / 331	88 / 521 87 / 433 87 / 296	88 / 470 87 / 416 87 / 331
4-24	356	99 / 324	-	-	-
4-25	700	-	86 / 303	84 / 343	-
5-39	611	-	-	99 / 592	-
5-44	580	-	-	98 / 150	-
7-25	518	-	-	98 / 150	-
8-5	238	-	99 / 525	-	-
10-16	769	99 / 346	-	-	-
seq5	596	-	-	88 / 330	96 / 119
seq10	785	-	100 / 351	-	-
seq15	792	-	86 / 303	84 / 343	-
seq24	343	-	96 / 275	-	-

9. PERSPECTIVAS

Analizar la expresión de éstos genes en la rizósfera y en el interior de la raíz mediante la realización de un “macroarreglo” y mediante fusiones transcripcionales identificar el sitio exacto de su expresión.

Se obtendrá la secuencia completa de los genes putativos identificados y se analizará experimentalmente su función, realizando mutaciones dirigidas y estudiando su fenotipo en competencia en la colonización del maíz y en nodulación, promoción del crecimiento y supresión de patógenos.

Además de obtendrán secuencias de tamaño mayor a 1.5 kilobases, que por las condiciones originales de amplificación del PCR en este trabajo, no se amplificaron o se amplificaron débilmente.

10. ANEXO

10.1 HIBRIDACIÓN SUSTRACTIVA

Con el objetivo de conocer la regulación molecular en distintos procesos biológicos que presentan programas de expresión genética diferente, como la organogénesis o el crecimiento celular, se diseñó una técnica que permite aislar los genes expresados en cada condición. La hibridación sustractiva fue originalmente diseñada para organismos eucariontes y se denominó SSH (Supresión subtráctiva hybridization). Esta técnica consiste en aislar el ARN de cada condición (genes transcritos) y obtener el ADN complementario mediante una reacción de transcripción inversa. Posteriormente se amplifican selectivamente, utilizando como base el efecto supresivo de la PCR: cuando se añaden extremos largos repetidos a fragmentos de ADN se puede suprimir la amplificación de secuencias no deseadas. (Diatchenko, *et al.*, 1996)

Posteriormente, en base a la SSH se generó un método para aislar el ADN específico de una cepa, (bacterial genome subtractive hybridization). La técnica consiste en aislar el ADN de las dos cepas y digerirlos por separado con una enzima de restricción. Los fragmentos de ADN de la bacteria de interés (tester) son reducidos mediante hibridación sustractiva con los fragmentos (secuencias) que también se encuentran en la cepa de referencia (driver). Los fragmentos que restan (exclusivos de la cepa de interés) son enriquecidos por PCR mediante las secuencias específicas (tester-specific sequences) y son clonados para su análisis posterior (Akopyants *et al.*, 1998)

En esta metodología (Figura 10) el ADN tanto de la cepa de interés como de la cepa de referencia (tester y driver respectivamente) son digeridos con la endonucleasa de restricción *AluI* para generar fragmentos de una longitud media de 0.5 kb. Por separado, dos diferentes adaptadores, se unen únicamente a los extremos 3' del DNA tester (porque sus extremos 5' carecen de grupos fosfato), que son ligados con diferentes alícuotas del ADN tester. Este ADN ligado a los adaptadores es desnaturalizado y mezclado (por separado) con un exceso de ADN driver (que no tiene adaptadores) para que se apareen (primera hibridación). Posteriormente, se mezclan

las dos reacciones y se agrega más ADN driver desnaturalizado para que se una a las secuencias del ADN tester que también está presente en el genoma driver. Las hebras sencillas complementarias que quedan se alinean y las secuencias con el adaptador son copiadas sobre el extremo 3'. La PCR es utilizada ahora para obtener amplificación exponencial del ADN tester con diferentes adaptadores en los extremos ("e" en la Fig. 10). La amplificación de los fragmentos con el mismo adaptador es suprimida porque los extremos se unen entre sí por la presencia de secuencias invertidas repetidas en los adaptadores ("b" en la Fig. 10) lo que inhibe la unión de los oligonucleótidos para realizar la PCR. El ADN tester con un solo adaptador en uno de los extremos se amplificará linealmente pero no exponencialmente (Akopyants *et al.*, 1998)

En el caso de la hibridación sustractiva con la cual se construyó la genoteca del ADN exclusivo de la cepa Ch24-10 con respecto a la CFN42, se realizó con el kit Clontech PCR-Select™ cDNA Subtraction Kit que permite realizar la hibridación de genomas completos y no únicamente de ADN complementario como lo marca el nombre. La metodología del kit para realizar la hibridación sustractiva se basa en la descrita por Akopyants *et al.* (1988).

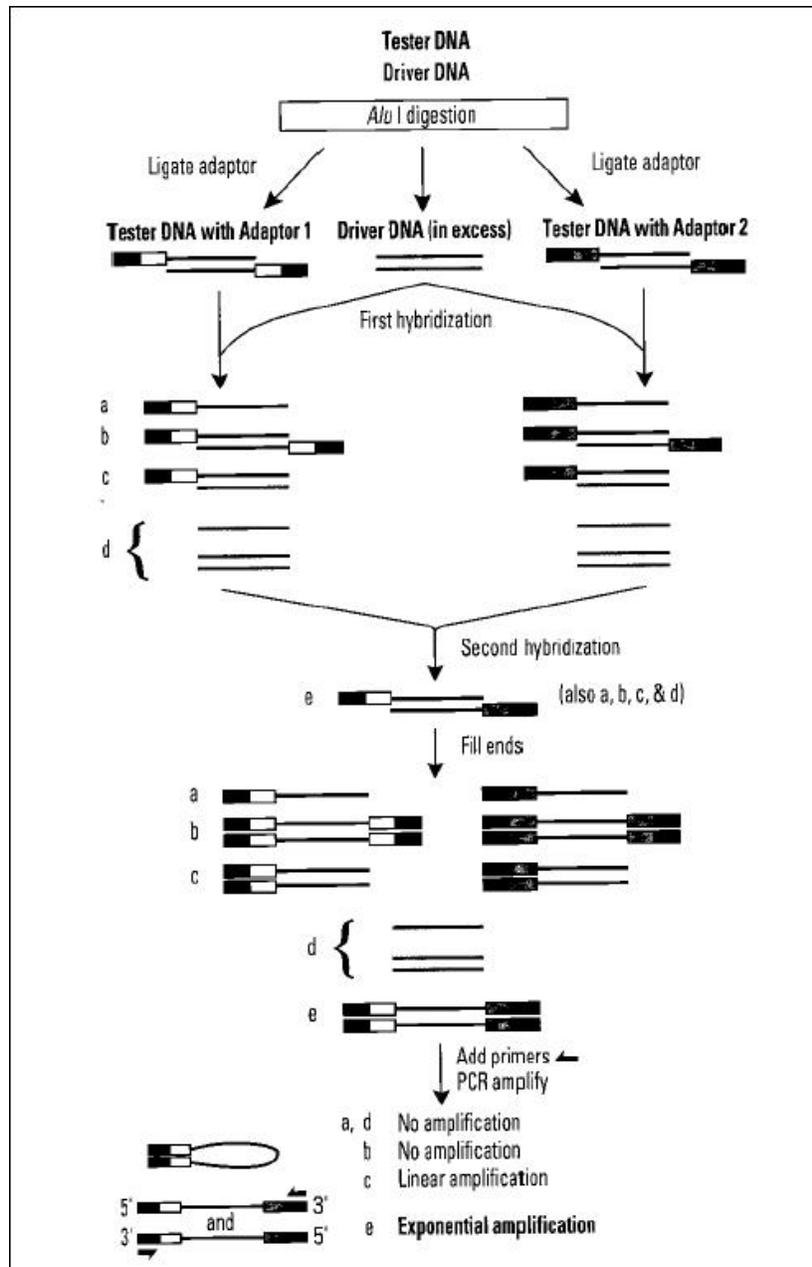


Figura 9. Hibridación sustractiva de ADN.

Las líneas continuas representan al ADN tester y driver digerido con *Alu I*. Los rectángulos representan la parte externa de los adaptadores de PCR, los que carecen del grupo fosfato en su extremo 5'. Los adaptadores 1 y 2 son idénticos en su extremo 5' y son diferentes cerca del extremo 3'. Note que después de que los extremos 3' son rellenados, las moléculas a, b y c con el adaptador 2 siguen presentes pero no se muestran (Akopyants *et al.*, 1998).

10.2 BASE DE DATOS

Tabla 8. Base de datos con los resultados del blastx de las secuencias de la librería genómica de *R. etli* Ch24-10 sin identidad al genoma de la cepa CFN42. Secuencias con identidad significativa a alguna proteína anotada de la base de datos del NCBI.

TABLA 8A. Secuencias con identidad significativa a proteínas putativas o conocidas.

CLONA	LONGITUD (NUCLEÓTIDOS)	PROTEÍNA PUTATIVA	ESPECIE / UBICACIÓN TAXONÓMICA (CLASE; ORDEN)	IDENTIDAD AMINOÁCIDOS / PORCENTAJE	SCORE	E VALUE	UBICACIÓN	FRAME	NÚMERO DE ACCESO
7-10 9-3	638	Proteína mayoritaria de la cápside gp7	<i>Streptococcus mitis phage SM1</i> -, <i>Caudovirales</i>	79/138 (57%)	153	5.00E-36	553-140 79-207	-3	gi 32441590 gb AAP81889.1
seq45	347	Proteína hipotética Rsph17025DRAFT_3854 Región: integrasa de fago	<i>Rhodobacter sphaeroides</i> <i>AATTCC 17025</i> <i>Alphaproteobacteria;</i> <i>Rhodobacterales</i>	23/67 (34%)	72	7E-12	158-358 167-233	2	gi 83370123 ref ZP_00914969.1
3-25 6-36	641	Proteína hipotética Nham_1366 Región: Transposasa y derivados inactivados	<i>Nitrobacter hamburgensis X14</i> <i>Alfaproteobacteria, Rhizobiales</i>	70/195 (35%)	115	1.00E-24	620-51 92-285	-2	gi 69928902 ref ZP_00625916.1
6-40 7-7, 9-7 11-42	453	Transposasa putativa para secuencia de inserción ISRM18	<i>Sinorhizobium meliloti</i> <i>Alfaproteobacteria, Rhizobiales</i>	143/147 (97%)	289	3.00E-77	456-16 211-357	-1	gi 15076155 emb CAC47708.1
4-40 8-2	755	Transposasa putativa para secuencia de inserción ISRM18	<i>S. meliloti</i> <i>Alfaproteobacteria, Rhizobiales</i>	123/132 (93%)	250	5E-65	434-39 1-132	-1	gi 15076155 emb CAC47708.1
J8-30 10-2	396	Transposasa y derivados inactivados	<i>Acinetobacter baumannii</i> <i>Gammaproteobacteria;</i> <i>Pseudomonadales;</i>	35/58 (60%)	66.6	2.00E-19	228-392 40-97	3	gi 90265374 emb CAJ77056.1
1-16 1-3, 2-13, seq14	542	ADN metilasa N4 / N6 putativa	<i>Corynebacterium diphtheriae</i> <i>Actinobacteria; Actinobacteridae</i>	117/179 (65%)	220	3.00E-56	548-18 234-420	-2	gi 38200728 emb CAE50429.1
3-46 2-6, 6-10, 10-12	559	ADN metilasa N4/N6	<i>Azotobacter vinelandii AVOP</i> <i>Gammaproteobacteria;</i> <i>Pseudomonadales</i>	111/179 (62%)	230	3.00E-59	560-30 32/205	-1	gi 67153570 ref ZP_00415315.1
2-18	225	COG2189: ADN metilasa específica de adenina	<i>Escherichia coli E24377A</i> <i>Gammaproteobacteria;</i> <i>Enterobacteriales</i>	43/70 (61%)	107	1.00E-22	255-16 37-106	-2	gi 75187699 ref ZP_00700966.1

TABLA 8A. Continuación...

CLONA	LONGITUD (NUCLEÓTIDOS)	PROTEÍNA PUTATIVA	ESPECIE / UBICACIÓN TAXONÓMICA (CLASE; ORDEN)	IDENTIDAD AMINOÁCIDOS / PORCENTAJE	SCORE	E VALUE	UBICACIÓN	FRAME	NÚMERO DE ACCESO
9-9	470	ADN metilasa N4 / N6	<i>Syntrophobacter fumaroxidans</i> MPOB Deltaproteobacteria	45/107 (42%)	79.3	5.00E-14	152-472 347-448	2	gi 71543696 ref ZP_00664773.1
7-11 J6-45	756	ADN metilasa C-5 citosina específica	<i>Pseudomonas syringae</i> pv. <i>Gammaproteobacteria</i> ; <i>Pseudomonadales</i>	38/90 (42%)	58.2	4.00E-07	778-545 317-404	-1	gi 66046078 ref YP_235919.1
10-18	522	Helicasa DEAD/DEAH box Dependiente de ATP	<i>Nitrobacter winogradskyi</i> Nb-255 <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	101/114 (88%)	205	7.00E-52	184-525 1-114	1	gi 74420317 gb ABA04516.1
4-37 9-38, 10-38	734	Helicasa, C-terminal: caja DEAD/DEAH, Helicasa dependiente de ATP	<i>Alkalilimnicola</i> <i>ehrlchei</i> MLHE-1 <i>Gammaproteobacteria</i> ; <i>Chromatiales</i>	83/148 (56%)	167	4.00E-40	461-21 8-154	-3	gi 78700647 ref ZP_00865103.1
5-4	516	Endonucleasa de restricción, putativa	<i>Xanthobacter autotrophicus</i> Py2 <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	124/177 (70%)	237	2.00E-61	537-19 479-655	-2	gi 89359916 ref ZP_01197736.1
3-35	307	Proteína putativa similar al regulador transcripcional <i>GntR</i>	<i>Streptomyces roseachromogenes</i> subsp. <i>Oscitans</i> <i>Bacteria</i> ; <i>Actinobacteria</i>	38/98 (38%)	53.1	4.00E-06	315-25 69-165	-3	gi 24940590 gb AAN65217.1
4-32	745	Regulador transcripcional similar al tipo <i>LysR</i>	<i>S. meliloti</i> <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	127/249 (51%)	257	3.00E-67	765-19 22-270	-3	gi 14523236 gb AAK64824.1
6-23 9-13, 9-22 12-46	627	riorf 87 . Proteína hipotética similar al dominio de unión al gen traR	<i>Agrobacterium rhizogenes</i> <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	34/57 (59%)	70.1	7.00E-11	307-137 1-57	-1	gi 10567397 db BAB16206.1
8-34 7-34	808	Hélice-vuelta-hélice tipo tipo AraC	<i>Mesorhizobium</i> sp. <i>BNC1</i> <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	68/90 (75%)	143	9.00E-33	772-503 222-311	-1	gi 68192065 gb EAN06720.1

TABLA 8A. Continuación...

CLONA	LONGITUD (NUCLEÓTIDOS)	PROTEÍNA PUTATIVA	ESPECIE / UBICACIÓN TAXONÓMICA (CLASE; ORDEN)	IDENTIDAD AMINOÁCIDOS / PORCENTAJE	SCORE	E VALUE	UBICACIÓN	FRAME	NÚMERO DE ACCESO
7-31 5-11, 12-24, seq31	750	Sensor de respuesta. Histidina cinasa. Regulador transcripcional con dominios: PAS, tipo AAA, ATPasa y de unión a ADN	<i>Mesorhizobium huakuii</i> MAFF303099 <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	89/150 (59%)	191	3.00E-47	314-763 27-176	2	gi 14022474 db BAB49083.1
7-14 7-17, 9-37, 12-44	745	Diguanilato ciclasa putativa / Fosfodiesterasa (Dominio GGDEF/ EAL)	<i>Rhodopseudomonas palustris</i> CGA009 <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	34/63 (53%)	69.3	5.00E-11	269-81 832-891	-3	gi 39651085 emb CAE29608.1
2-1	329	Diguanilato ciclasa putativa / Fosfodiesterasa (Dominio GGDEF/ EAL)	<i>B. cepacia</i> R18194 <i>Betaproteobacteria</i> ; <i>Burkholderiales</i>	34/49 (69%)	71.2	1.00E-11	332-186 596-644	-2	gi 77964774 gb ABB06155.1
7-22 seq46 seq21	576	ORF 3. Gen relacionado con la competencia en la nodulación.	<i>R. leguminosarum</i> <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	123/147 (83%)	255	4.00E-80	556-16 107-286	-3	gi 1839275 gb AAB47010.1
9-20 9-21	336	Proteína TonB Proteína de membrana, traduce energía a otros receptores de membrana / Transporte de hierro	<i>R. leguminosarum</i> <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	62/106 (58%)	52.8	5.00E-06	6-323 272-371	3	gi 13277329 emb CAC34389.1
seq12	503	Proteína transportadora de iones. Canal de potasio. Putativa	<i>Vibrio</i> sp. MED222 <i>Gammaproteobacteria</i> ; <i>Vibrionales</i>	35/62 (55%)	74.7	2.00E-12	216-31 40-101	-3	gi 86146518 ref ZP_01064841.1
12-27	168	Probable acetiltransferasa	<i>Mesorhizobium huakuii</i> MAFF303099 <i>Alphaproteobacteria</i> ; <i>Rhizobiales</i>	25/44 (56%)	55.1	1.00E-06	172-41 56-99	-1	gi 14021959 db BAB48570.1
5-38 6-7, 11-14 13-27	628	gll1781 Probable proteína de transporte de polisacáridos Proteína de membrana involucrada en exportar el antígeno-O y el ácido teicoico	<i>Gleobacter violaceus</i> <i>Cyanobacteria</i> ; <i>Gloeobacteria</i>	29/95 (30%)	57.8	3.00E-07	338-622 1-94	2	gi 37521350 ref NP_924727.1

TABLA 8A. Continuación...

CLONA	LONGITUD (NUCLEÓTIDOS)	PROTEÍNA PUTATIVA	ESPECIE / UBICACIÓN TAXONÓMICA (CLASE; ORDEN)	IDENTIDAD AMINOÁCIDOS / PORCENTAJE	SCORE	E VALUE	UBICACIÓN	FRAME	NÚMERO DE ACCESO
2-34	735	COG0463: Glicosiltransferasa familia 2, biogénesis de la pared celular	<i>Trichodesmium exythraeum</i> IMSIOI Cyanobacteria; Oscillatoriales	36/119 (30%)	73.2	1.00E-11	645-289 187-305	-2	gi 71676110 ref ZP_00673853.1
8-17 4-43	862	Putativa glicosiltransferasa de biosíntesis del "core" de lipopolisacáridos	<i>Burkholderia xenovorans</i> LB400 Betaproteobacteria; Burkholderiales;	115/258 (44%)	209	1.00E-52	777-13 1-257	-2	gi 48788658 ref ZP_00284637.1
7-43 1-26, 7-26, 10-3, seq22	315	Proteína hipotética CPS_3246 Región: Glicosiltransferasa	<i>Colwellia psychrerythraea</i> 34H Gammaproteobacteria; Alteromonadales	38/103 (36%)	74.7	1.00E-12	331-23 25-127	-2	gi 71143563 gb AAZ24036.1
3-47	315	Proteína hipotética gl13709 Región: Proteína de biosíntesis de polisacáridos	<i>Gloeobacter violaceus</i> Cyanobacteria; Gloeobacteria	46/152 (30%)	85.9 bits	5.00E-16	12-467 112-263	3	gi 35214282 db BAC91650.1
1-40	657	Proteína hipotética Región: Glicósido hidrolasa, familia 19. Dominio: Quitinasa	<i>Chlorobium phaeobacteroides</i> DSM 266 Bacteria; Chlorobi; Chlorobia	45/86 (52%)	64.7	3.00E-09	429-677 1-86	3	gi 67934818 ref ZP_00527844.1
4-20	432	Metiltransferasa-O, familia 2	<i>Ralstonia eutropha</i> JMP134 Betaproteobacteria; Burkholderiales	41/54 (75%)	88.2	3.00E-20	251-412 299-352	2	gi 72122716 gb AAZ64902.1
4-6 6-6	346	Proteína putativa LuxA Región: Monooxigenasa, sustrato: flavina	<i>Sinorhizobium</i> sp. BR816 Alphaproteobacteria; Rhizobiales	35/56 (62%)	72	8.00E-12	320-153 135-190	-3	gi 18152932 gb AAL61964.1
6-31 seq27	764	Proteína hipotética blr2133 Región: Fitanoil Co-A dioxigenasa (PhyH)	<i>B. japonicum</i> USDA 110 Alphaproteobacteria; Rhizobiales	63/107 (58%)	134	5.00E-30	339-19 1-107	-1	gi 27350387 db BAC47398.1

TABLA 8A. Continuación...

CLONA	LONGITUD (NUCLEÓTIDOS)	PROTEÍNA PUTATIVA	ESPECIE / UBICACIÓN TAXONÓMICA (CLASE; ORDEN)	IDENTIDAD AMINOÁCIDOS / PORCENTAJE	SCORE	E VALUE	UBICACIÓN	FRAME	NÚMERO DE ACCESO
12-22	137	Nitrilo hidratasa cadena alfa	<i>uncultured bacterium Dr1</i>	27/31 (87%)	56.2	4.00E-07	100-81-31	-2	gi 27657371 emb CAC83631.1
8-35	344	Formato dehidrogenasa subunidad beta Región: Coenzima F420 hidrogenasa/ dehidrogenasa, subunidad beta N terminal	<i>Methanococcus voltae</i> <i>Archaea; Euryarchaeota; Methanococci</i>	21/25 (84%)	51.6	1.00E-05	231-30595-119	3 nada	gi 14211982 gb AAK57554.1
12-30	813	Proteína hipotética blr1705 Ulp1, dominio catalítico C-terminal (ubiquitin like protease1)	<i>Bradyrhizobium japonicum USDA 110</i> <i>Alphaproteobacteria; Rhizobiales</i>	54/151 (35%)	78.6	3.00E-13	359-796506-652	2	gi 27376816 ref NP_768345.1
7-21 9-5	594	Hypothetical protein blr1705 "Ulp1 protease family, C-terminal catalytic domain"	<i>B. japonicum USDA 110</i>	90/176 (51%)	165	1E-39	59-5741214-1383	2	gi 27349958 db BAC46970.1
11-39	822	Serina proteasa putativa, familia tripsina Dominio: Serina proteasa similar a tripsina, típicamente periplásmica, contiene un C-dominio terminar PDZ	<i>Aurantimonas sp. SI85-9A1</i> <i>Alphaproteobacteria; Rhizobiales;</i>	39/99 (39%)	75.9	2.00E-12	548-841424-520	2	gi 90419802 ref ZP_01227711.1
8-3	610	ATPasa "ATPasa predicha (superfamilia AAA+) [Función general sólo predicha]	<i>Polaromonas sp. JS666</i> <i>Betaproteobacteria; Burkholderiales</i>	132/200 (66%)	232	1.00E-59	15-61479-278	3	gi 67847285 ref ZP_00502404.1
3-45 8-25, 9-2, 9-19, 9-45, 10-35	761	Probable oxigenasa Región IucD Lisina/ornitina N monooxigenasa	<i>B. japonicum USDA 110</i>	68/131 (52%)	134	4E-56	741-349181-311	-3	gi 27377242 ref NP_768771.1

TABLA 8B. Secuencias con identidad significativa a proteínas hipotéticas.

CLONA	LONGITUD (NUCLEÓTIDOS)	PROTEÍNA PUTATIVA	ESPECIE / UBICACIÓN TAXONÓMICA (CLASE, ORDEN)	IDENTIDAD AMINOÁCIDOS / PORCENTAJE	SCORE	E VALUE	UBICACIÓN	FRAME	NÚMERO DE ACCESO
2-5	880	Proteína hipotética conservada MCA2652	<i>Methylococcus capsulatus strain Bath</i>	53/72 (73%)	106	1.00E-21	544-329 1-72	-3	gi 53756933 gb AAU91224.1
4-1 5-13, 5-34, 8-14, 14-34	484	Proteína hipotética blr 4890	<i>B. japonicum</i>	120/156 (76%)	261	1.00E-68	488-21 263-418	-3	gi 27353155 db BAC50155.1
4-34	353	Proteína hipotética conservada	<i>Shewanella sp. ANA-3</i>	41/73 (56%)	83.6	3.00E-15	357-139 409-480	-3	gi 78685870 ref ZP_00850644.1
5-30	877	Proteína hipotética	<i>Xantomonas axonopodes pv. citri Str. 306</i>	34/99 (34%)	45.4	3.00E-07	727-452 60-155	-2	gi 21108486 gb AAM37099.1
6-19 9-26	594	Proteína hipotética conservada	<i>Colwellia psychrerythraea 34H</i>	83/204 (40%)	147	2.00E-34	599-9 23-219	-2	gi 71146131 gb AAZ26604.1
6-21	875	Proteína hipotética tiorf24	<i>Agrobacterium tumefaciens</i>	56/147 (38%)	85.1	4.00E-15	72-503 25-171	3	gi 6498197 db j BAA87649.1
7-30	372	Proteína hipotética alr3434	<i>Nostoc sp. 7120</i>	57/97 (58%)	125	4.00E-28	371-81 204-300	-3	gi 17132567 db j BAB75133.1
10-7 seq19	537	Proteína hipotética Nwi_1614	<i>Nitrobacter winogradskyi Nb-255</i>	42/93 (45%)	75.5	1.00E-12	542-264 33-123	-3	gi 74420676 gb ABA04875.1

TABLA 9B. Continuación...

CLONA	LONGITUD (NUCLEÓTIDOS)	PROTEÍNA PUTATIVA	ESPECIE / UBICACIÓN TAXONÓMICA (CLASE, ORDEN)	IDENTIDAD AMINOÁCIDOS / PORCENTAJE	SCORE	E VALUE	UBICACIÓN	FRAME	NÚMERO DE ACCESO
10-24	446	Proteína hipotética AGR_pAT_284p	<i>A tumefaciens str. C58</i>	32/77 (41%)	61.6	4.00E-19	449-219 13-89	-1	gi 15161963 gb AAK90576.1
10-32 1-12, 3-16	482	Proteína hipotética RNGR00302	<i>Rhizobium sp. NGR234</i>	58/97 (59%)	128	9.00E-29	309-19 8-104	-1	gi 36959003 gb AAQ87428.1
13-19	390	Proteína hipotética conservada	<i>Rhodobacter sphaeroides ATCC 17029</i>	59/94 (62%)	132	4.00E-30	296-15 445-538	-3	gi 83374719 ref ZP_00919489.1
seq5	596	Proteína hipotética MflvDRAFT_5132	<i>Mycobacterium flavescens PYR-GCK</i>	24/56 (42%)	50.8	3.00E-05	346-513 64-119	1	gi 89338010 ref ZP_01190788.1
seq35	347	Proteína hipotética GOX0407	<i>Gluconobacter oxydans 621H</i>	31/50 (62%)	51.2	7.00E-09	521-378 8-55	-3	gi 58001296 gb AAW60190.1
11-13	635	Proteína hipotética PSEEN3673	<i>Pseudomonas entomophila L48</i>	73/134 (54%)	125	1.00E-27	246-647 2-134	3	gi 104782690 ref YP_609188.1
7-15 3-43	783	Proteína hipotética Beta hélice paralela repetida	<i>Nitrobacter hamburgensis X14</i>	44/91 (48%)	83.2	1.00E-14	776-261 109-291	-2	gi 69926474 ref ZP_00623945.1

REFERENCIAS

- Amersham Biosciences.** General Electric Company (1997-2007)
[<http://www5.amershambiosciences.com>]
- Akopyants NS, A Fradkov, L Diatchenko, JE Hill, PD Siebert, SA Lukyanov, ED Sverdlov & DE Berg** (1998) PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori*. Proc. Natl. Acad. Sci. 95:13108-13113.
- Altschul SF, TL Madden, AA Schäffer, J Zhang, Z Zhang, W Miller & DJ Lipman** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.
- Atkinson EM, Palcic MM, Hindsgaul O & Long SR** (1994) Biosynthesis of *Rhizobium meliloti* lipooligosaccharide Nod factors: NodA is required for an N-acyltransferase activity. Proc. Natl. Acad. Sci. 91:8418-22
- Bellon MR, M Adato, J Becerril & D Mindek** (2005) Impact of Improved Germplasm on Poverty Alleviation: The Case of Tuxpeño-Derived Materials in Mexico. Mexico, DF. CIMMYT
- Berthaud J** (2001) La diversidad del maíz en Oaxaca, México: Preguntas sencillas sin respuestas fáciles. CIMMYT. Informe anual 2000-2001. Investigaciones mundiales para la subsistencia local.
[http://www.cimmyt.org/Research/Economics/map/impact_studies/ImpactsMaize66_97/ImpactosLA/ImpactosMaiz66_97.htm]
- Bloemberg GV & BJJ Lugtenberg** (2001) Molecular basis of plant growth promotion and biocontrol by rhizobacteria. Current opinion in Plant Biology. 4:343-350
- Brandao PFB, JP Clapp & AT Bull** (2003) Diversity of Nitrile Hydratase and Amidase Enzyme Genes in *Rhodococcus erythropolis* recovered from Geographically Distinct Habitats. Applied and Environmental Microbiology. 69:5754-5766.
- Brencic A & SC Winans** (2005) Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. Microbiology and Molecular Biology Reviews. 69:155-194.

Brodsky LI, VV Ivanov, YL Kalaydzidis, A.M Leontovich, VK Nikolaev, SI Feranchuk & VA Drachev (1995) GeneBee-NET: Internet-Based Server for Analyzing Biopolymers Structure Biochemistry (Moscow). 60:823-828

[http://www.genebee.msu.su/services/malign_reduced.html]

Brom S, L Girard, A García-de los Santos, JM Sanjuán-Pinilla, J Olivares & J Sanjuán (2002) Conservation of plasmid-encoded traits among bena-nodulating *Rhizobium* species. Applied and Environmental Microbiology. 68: 2555-2561

Broughton WJ, S Jabbouri & X Perret (2000) Keys to symbiotic harmony. Journal of Bacteriology. 182:5641-5652

Broughton, WJ, G Hernández, M Blair, S Beebe, P Gepts & J Vanderleyden (2003) Beans (*Phaseolus* spp.), model food legumes. Plant and Soil 252: 55-128

Cajens S (2003) Prophages and bacterial genomics: what have we learned so far? Molecular Microbiology. 49:277-300

Campbell JA, GJ Davies, V Bulone & B Henrissat (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochemical Journal. 326:929-942

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

CIMMYT (2001) La economía del maíz en América latina. Impactos del mejoramiento de maíz en América Latina, 1966-1997.

[http://www.cimmyt.org/Research/Economics/map/impact_studies/ImpactsMaize66_97/ImpactosLA/ImpactosMaiz66_97.htm]

Cleyet-Marel JC, M Larcher, H Bertrand, S Raptor & X Pinochet (2001) Plant Growth Enhancement by Rhizobacteria. En Morot-Gaudry JF (ed.). Nitrogen assimilation by plants. Physiological, Biochemical and Molecular Aspects. Science Publishers, Inc. India

Compant S, B Duffy, J Nowak, C Clément & EA Barka (2005) Use of Plants Growth-Promoting Bacteria for Biocontrol of Plant Diseases: Principles, Mechanisms of Action, and Future Prospects. Applied and Environmental Microbiology. 71: 4951-4959

Crews TE & MB Peoples (2004) Legume versus fertilizer sources of nitrogen: ecological tradeoffs and human needs. *Agriculture, Ecosystems and Environment*. 102:279-297.

Crosa JH & CT Walsh (2002) Genetics and Assembly Line Enzymology of Siderophore Biosynthesis in Bacteria. *Microbiology and Molecular Biology Reviews*. 66:223-249

Denison RF & ET Kiers (2004) Why are most rhizobia beneficial to their plant host, rather than parasitic? *Microbes and Infection*. 6:1235-1239

Diatchenko I, Y-FC Laut, AP Campbell, A Chenchik, F Moqadam, B Huang, S Lukyanov, K Lukyanov, N Gurskaya, ED Sverdlov & PD Siebert (1996) Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci.* 93: 6025-6030.

Dobrindt U, F Agerer, K Michaelis, A Janka, C Buchrieser, M Samuelson, C Svanborg, G Gottschalk, H Karch & J Hacker (2003) Analysis of genome plasticity in Pathogenic and Commensal *Escherichia coli* Isolates by use of DNA arrays. *Journal of Bacteriology*. 185: 1831-1840

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

Elmerich C (2001) Nitrogenase: Biochemical, Molecular and Genetic Aspects. En Morot-Gaudry JF (Ed.) Nitrogen assimilation by plants. Physiological, Biochemical and Molecular Aspects. Science Publishers, Inc. India.

Escalante V (2006) Obtención de mutantes que *Rhizobium etli* (Ch24-10) deficientes para utilizar la prolina y análisis de su fenotipo para colonizar plantas de maíz. Tesis de Licenciatura. Facultad de Ciencias Biológicas, UAEM. Programa de Ecología Genómica, Centro de Ciencias Genómicas, UNAM.

FAO, Organización de las Naciones Unidas para la Agricultura y la Alimentación (1993) El maíz en la nutrición humana. Colección Alimentación y nutrición No. 25. Serie Código FAO: 86 AGRIS: S01.

Fahraeus G (1957) The infection of clover root hair by nodule bacteria studied by a single glass technique. *J Gen Microbiol* 16:374-381

Fineran PC, L Everson, H Slater & GPC Salmond (2005) A GntR family transcriptional regulator (PigT) controls gluconate-mediated repression and defines a new, independent pathway for regulation of the tripyrrole antibiotic prodigiosin in *Serratia*. *Microbiology*. 151:3833-3845

Flores M, L Morales, A Ávila, V González, P Bustos, D García, Y Mora, X Guo, J Collado-Vides, D Piñero, G Dávila, J Mora & R Palacios (2005) Diversification of DNA sequences in the symbiotic genome of *Rhizobium etli*. *Journal of Bacteriology*. 187:7185-7192

Fraser C & S Restrepo-Estrada (1996) Comunicación para el desarrollo rural en México. En los buenos y en los malos tiempos. Departamento de Desarrollo Sostenible. Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO). [<http://www.fao.org/docrep/W3616S/w3616s00.htm>]

Galloway JN, WH Schlesinger, HI Levy, AF Michaels & JL Schnoor (1995) Nitrogen fixation: anthropogenic enhancement-environmental response. *Global Biogeochem Cycles* 9: 235-252

García-Del Portillo F, MG Pucciarelli & J Casadesús (1999) DNA adenine methylase mutants of *Salmonella typhimurium* show defects in protein secretion, cell invasion, and M cell cytotoxicity. *Proc Natl Acad Sci*. 96: 11578-11583.

Ghassan MM, Y Kfoury & U Hadi (2003) Transcription levels and Sequence Analysis of *tonB* gene in *Haemophilus parainfluenzae* versus *Haemophilus influenzae* isolates from Patients Undergoing Tonsillectomy and/or Adenoidectomy. *The Journal of Applied Research*. 3:288-394.

González V, RI Santamaría, P Bustos, I Hernández-González, A Medrano-Soto, G Moreno-Hagelsieb, S Chandra Janga, MA Ramírez, V Jiménez-Jacinto, J Collado-Vides & G Dávila (2006) The partitioned *Rhizobium etli* genome: Genetic and metabolic redundancy in seven interacting replicons. *PNAS* 103: 3834-3839.

Graham PH & P Ranalli (1997) Common bean (*Phaseolus vulgaris* L.) *Field Crops Research*. 53: 131-146.

Gormley NA, MA Watson & SE Haldford (2005) Bacterial Restriction-Modification Systems. *Encyclopedia Of Life Sciences*. John Wiley & Sons, Ltd: Chichester [<http://www.els.net>] [doi: 10.1038/npg.els.0003897]

Gutierrez-Zamora ML & E Martínez-Romero (2001) Natural endophytic association between *Rhizobium etli* and maize (*Zea mays* L.) *Journal of Biotechnology* 91:117-126

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41:95-98.

Herrero M, V De Lorenzo & JB Neilands (1998) Nucleotide Sequence of the *iucD* gene of the pColV-K30 aerobactin operon and topology of its product studied with *phoA* and *lacZ* gene fusions. *Journal Of Bacteriology.*170(1):p. 56-64

Hillerich B, J Westpheling (2006) A New GntR Family Transcriptional Regulator in *Streptomyces coelicolor* is required for Morphogenesis and Antibiotic Production and Controls Transcription of an ABC Transporter in Response to Carbon Source. *Journal of Bacteriology.* 188:7477-7487

Hynes MF & MP O'Connell (1990) Host plant effect on competition among strains of *Rhizobium leguminosarum*. *Can J Microbiology.* 36:864-869

Hirel B, P Bertin, I Quilleré, W Bourdoncle, C Attagnant, C Dellay, A. Gouy, S Cadiou, C Retailiau, M Falque & A Gallais (2001) Towards a Better understanding of the Genetic and Physiological Basis for Nitrogen use efficiency in Maize. *Plan Physiology.* 125:1258-1270

Hisinger, P, GR Gobran, PJ Gregory & WW Wenzel (2005) Rhizosphere geometry and heterogeneity arising from root mediated physical and chemical processes. *New Phytologist.* 168: 293-303

InterPro Database. European Bioinformatics Institute (2006) European Molecular Biology Laboratory [<http://www.ebi.ac.uk/interpro>]

Invitrogen Corporation (2006) [<http://www.invitrogen.com>]

Jenkins G, A Bahia Filho, K Cassman, B Keenan, D Marshall, VS Vyas, J Griffith, R Quatrano, P Sachdeva & G Gryseels (1998) Report of the Fourth External Programme and Management. Review of the Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT). Organización de las Naciones Unidas para la Agricultura y la Alimentación. [<http://www.fao.org/wairdocs/tac/x5802e/x5802e00.htm>]

Juergen P & P Poolé (2006) Metabolic changes of rhizobia in legume nodules. *TRENDS in Microbiology.*14: 161-168

Jurtshuk P (1996) Sección 1 Bacteriology. Capítulo 4. Bacterial Metabolism. En Baron S (ed) Medical Microbiology. Fourth Edition. The University of Texas Medical Branch at Galveston, USA.

Kiely, PD, JM Haynes, CH Higgins, A Franks, GL Mark, JP Morrissey & F O’Gara (2005) Exploiting new systems-based strategies to elucidate plant-bacterial interactions in the rhizosphere. *Microbial Ecology*. 3:257-66.

Kobayashi M, T Suzuki, T Fujita, M Masuda & S Shimizu (1995) Occurrence of Enzymes involved in Biosynthesis of indole-3-acetic acid from indole-3-acetonitrile in Plant-Associated Bacteria, *Agrobacterium* and *Rhizobium*. *Proc Natl Acad Sci*. 92:714-718.

Koide T, PA Zaini, LM Moreira, RZN Vencio, AY Matsukuma, AM Durham, DC Teixeira, H El-Dorry, PB Monteiro, ACR da Silva, S Verjovski-Almeida, AM da Silva & SL Gomes (2004) DNA Microarray-Based Genome Comparison of a Pathogenic and a Nonpathogenic Strain of *Xylella fastidiosa* delineates genes important for Bacterial Virulence. *Journal of Bacteriology*. 36:1229-1244

Koonin E & MY Galperin (2003) Sequence-Evolution-Function. Computational Approaches in Comparative Genomics. Kluwer Academic Publishers.

Lerouge I & J Vanderleyden (2001) O-antigen structural variation: mechanisms and possible roles in animal/plant-microbe interactions. *FEMS Microbiology Reviews*. 26:17-47

Lodwig E & PP Poolé (2003) Metabolism of *Rhizobium* Bacteroids. *Critical Reviews in Plant Sciences*. 22:37-78

Long SR (2001) Genes and Signals in the *Rhizobium*-Legume Symbiosis. *Plant Physiology*. 125:69-72

Low DA, NJ Weyand & MJ Mahan (2001) Roles of DNA Adenine Methylation in Regulating Bacterial Gene Expression and Virulence. *Infection and Immunity*. 69:7197-7204

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

Méndez-Ortiz MM & J Membrillo-Hernández (2005) Proteínas con dominios GGDEF y EAL: su función en el metabolismo bacteriano. *Revista Latinoamericana de Microbiología*. 47:130-139.

Mergaert P, T Uchiumi, B Alunni, G Evanno, A Cheron, O Catrice, AE Mausset, F Barloy-Hubler, F Galibert, A Kondorosi, E Kondorosi (2006) Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium-legume* symbiosis. *PNAS*.103: 5230-5235

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

Mora-Nuñez OA (1997) Origen e importancia del cultivo de la caraota (*Phaseolus vulgaris* L.). *Rev. Fac. Agron, Maracay*. 23:225-234.

Mossessova E & CD Lima (2000) Ulp1-SUMO Crystal Structure and Genetic Analysis Reveal Conserved Interactions and a Regulatory Element Essential for Cell Growth in Yeast. *Molecular Cell*. 5: 865-876

Mylona P, K Pawlowski & T Bisseling (1995) Symbiotic Nitrogen Fixation. *The Plant Cell*. 7:869-885

National Center for Biotechnology Information

[<http://www.ncbi.nlm.nih.gov>]

New England Biolabs [<http://www.neb.com/nebecomm>]

Nguyen C (2003) Rhizodeposition of organic C by plant: mechanisms and controls. *Agronomie* 23: 375-396.

Paliwal RL, G Granados, HR Lafitte, AD Violic & JP Marathée (2001) El Maíz en los Trópicos: Mejoramiento y Producción. Organización De Las Naciones Unidas Para la Agricultura y la Alimentación, Roma.

Perret X, C Staehelin & WJ Broughton (2000) Molecular Basis of Symbiotic Promiscuity. *Microbiology and Molecular Biology Reviews*. 64:180-201

Pingali PL (ed.). 2001. CIMMYT 1999-2000 World Maize Facts and Trends. Meeting World Maize Needs: Technological Opportunities and Priorities for the Public Sector. Mexico, D.F.: CIMMYT

Ramos-González MA, MJ Campos & J L Ramos (2005) Analysis of *Pseudomonas putida* KT2440 Gene Expression in the Maize Rhizosphere: In Vitro Expression Technology Capture and Identification of Root-Activated Promoters. Journal of Bacteriology. 187:4033-4041

Reading NC & V Sperandio (2005) Quorum sensing: the many languages of bacteria. FEMS Microbiol Lett 254:1-11

Roche [<http://www.roche.com>]

Rosenblueth M, J Martínez-Romero & E Martínez-Romero (2001) Capítulo 3. Ecología química de la rizósfera y en las simbiosis de plantas. En: Interacciones Químicas entre Organismos: Aspectos Básicos y Perspectivas de Aplicación. Anaya AL, FJ Espinosa-García, R Cruz-Ortega (eds) Instituto de ecología UNAM Plaza y Valdés. México.

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

Rosenblueth M & E Martínez-Romero (2006) Bacterial endophytes and their interactions with hosts. American Phythopatological Society. 19:827-837

Sain G & MA López-Pereira (1999) Maize Production and Agricultural Policies in Central America and Mexico. CIMMYT Economics Working Paper No. 99-02. Mexico, DF. CIMMYT

[http://www.cimmyt.org/research/Economics/map/impact_studies/ImpactsMaize66_97/ImpactosLA/pdfs/ImpactosLA_economia.pdf]

Schulz B & C Boyle (2005) The endophytic continuum. Mycol. Res. 109 (6): 661-686

Schultze M & A Kondorosi (1998) Regulation on symbiotic root nodule development. Annual Review of Genetics. 32:33-57

Segovia L, JPW Young & E Martínez-Romero (1993) Reclassification of American *Rhizobium leguminosarum* biovar phaseoli type I strains as *Rhizobium etli* sp. nov. Int J Syst Bacteriol. 43:374-377

Selosse MA, E Baudoin & P Vandenkoornhuys (2004) Symbiotic microorganisms, a key for ecological success and protection of plants. *CR Biologies*. 327:639-648

Silva C, P Vinuesa, LE Eguiarte, E Martínez-Romero & V Souza (2003) *Rhizobium etli* and *Rhizobium gallicum* nodulate common Bean (*Phaseolis vulgaris*) in a traditionally managed milpa plot in México: Populations genetics and biogeographic implications. *Applied and Environmental Microbiology*. 69:2

Singh BK, P Millard, AS Whiteley & JC Murrell (2004) Unravelling rhizosphere-microbial interactions: opportunities and limitations. *TRENDS in Microbiology*. 12:386-393

Skorupska A, M Janczarek, M Marczak, A Mazur & J Król (2006) Rhizobial exopolysaccharides: genetic control and symbiotic functions. *Microbial Cell Factories*. 5:1-7

Spaink HP (2000) Root nodulation and infection factors produced by rhizobial bacterial. *Annual Reviews of Microbiology*. 54:257-288

Stratagene [<http://www.stratagene.com>]

Technelysium Pty Ltd (2003-2007) [<http://www.technelysium.com.au>]

Trinchant JC, JJ Drevon & J Rigaud (2001) Symbiotic Nitrogen Fixation. En: Morot-Gaudry JF. Nitrogen assimilation by plants. Physiological, Biochemical and Molecular Aspects. Science Publishers, Inc. India

Tsunedomi R, H Izu, T Kawai, K Matsushita, T Ferenci & M Yamada (2003) The Activator of *gntII* Genes for Gluconate Metabolism, *gntH*, Exerts Negative Control of *gntR*-Regulated *gntI* Genes in *Escherichia coli*. *Journal of Bacteriology*. *Journal of Bacteriology*. 185: 1783-1795

Vega-Valdivia DD & P Ramírez-Romero (2004) Situación y perspectivas del maíz en México. Universidad Autónoma Chapingo.

[http://www.economia.gob.mx/pics/p/p1763/Maiz_270304.pdf]

Vilchis A (2004) Mutantes de *Rhizobium etli* que expresan β -glucoronidasa en presencia de exudados de maíz. Tesis de Licenciatura. Facultad de Ciencias Biológicas, UAEM. Programa de Ecología Genómica, Centro de Ciencias Genómicas, UNAM.

Visca P, A Ciervo & N Orsi (1994) Cloning and Nucleotide Sequence of the *pvda* gene Encoding the Pyoverdin Biosynthetic Enzyme L-Ornithine N5-Oxygenase in *Pseudomonas aeruginosa*. Journal of Bacteriology. 176:1128-1140

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

Wang ET, J Martínez-Romero & IM López-Lara (2001) Rhizobium y su destacada asociación con plantas. En: Martínez-Romero, E, J Martínez-Romero (eds). Microbios en línea. Centro de Investigación sobre Fijación del Nitrógeno. Coordinación de investigación científica. UNAM

[\[http://www.biblioweb.dgsca.unam.mx/libros/microbios\]](http://www.biblioweb.dgsca.unam.mx/libros/microbios)

Watanabe T, H Fujihara & K Furukawa (2003) Characterization of the Second LysR-Type Regulator in the Biphenyl-Catabolic Gene Cluster of *Pseudomonas pseudoalcaligenes* KF707. Journal of Bacteriology. 185:3575-3582

Wexler M , KH Yeoman, JB Stevens, NG de Luca, G Sawers & AWB Johnston (2001) The *Rhizobium leguminosarum* tonB gene is required for the uptake of siderophore and haem as sources of iron. Molecular Microbiology. 41(4):801-816

Wion D & J Casadesús (2006) N6-methyl-adenine: an epigenetic signal for DNA interactions. Nature Reviews. Microbiology. 4:183-190

Zahran HH (1999) Rhizobium-Legume symbiosis and Nitrogen Fixation under severe conditions and in an arid climate. Microbiology and Molecular Biology Reviews. Vol. 63, No. 4: 968-989