

INSTITUTO DE BIOTECNOLOGÍA
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**“CARACTERIZACIÓN FUNCIONAL DE LOS
GENES *nodT* EN *Rhizobium etli*”**

TESIS

QUE PARA OBTENER EL GRADO DE

DOCTOR EN CIENCIAS BIOQUÍMICAS

PRESENTA

M. en Bt. ARMANDO HERNÁNDEZ MENDOZA

JURADO DE EXAMEN:

PRESIDENTE:

Dr. José Luis Puente García

SECRETARIO:

MIBB Carmen Quinto Hernández

VOCALES:

Dra. Susana Brom Klanner
Dra. Cinthia Ernestina Núñez López
Dr. Lorenzo Segovia Forcella

SUPLENTES:

Dr. Guillermo Gosset Lagarda
Dr. Enrique Merino Pérez



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

A mis hijas Eriana y Rosita, y a mi hijo Armandito
A mis Padres Genaro Hernández Nava y Rosa Elena Mendoza Cruz

AGRADECIMIENTOS

En primer lugar quiero agradecer a mi tutora Carmen Quinto. Gracias por tu apoyo y confianza durante esta larga travesía. Finalmente alcanzamos llegar al puerto. Y al Dr. Federico Sánchez, su esposo, por ser un hombre que sabe escuchar y decir la palabra adecuada en el momento perfecto

A los miembros de mis comités tutorales: Dra. Brenda Valderrama, Dr. Edmundo Calva, Dr. David Romero, Dr. Victor Valdés y Dr. Alejandro Alagón. Cada uno en su momento contribuyó a que encontrara una solución para este trabajo

A los miembros de mi Jurado de Examen: Dra. Susana Brom, Dra. Cinthia Nuñez, Dr. José Luis Puente, Dr. Enrique Merino, Dr. Guillermo Gosset, Dr. Lorenzo Segovia. Sus atinados comentarios y sus correcciones contribuyeron a mejorar sustancialmente este trabajo.

A CONACYT por la beca de estudios de doctorado 90288 y a DGAPA por el apoyo económico para la impresión de tesis y para los viajes a congresos

A Ceí Abreu. Sin tu visión, trabajo, apoyo, consejos y discusiones, la publicación del artículo habría tenido otra historia. Pero encima de todo, agradezco tu amistad.

A mis compañeros de trabajo: Olivia Santana Estrada y Noreide Nava Nuñez, por el apoyo técnico, el cariño, las risas y la confianza mutua

A Aníbal Tovar Castro, Tanja Scheublin y el Dr. E. Pérez-Rueda, por contribuir a la realización de esta tesis. También a Maria Guadalupe Negrete, José Luis Ramírez y José Luis Zitlalpopocatl por el apoyo en el laboratorio y en el invernadero

A la Unidad de Computo del IBT: Ricardo Ciria, Alma Martínez Valle, Arturo Ocadiz, Abel Linares y Juan Manuel Hurtado... Muchas gracias por los "Bytes en la pBR"

A la Unidad de Docencia: Por su ayuda para agilizar los trámites: Ing. Jalil Saab (gracias por los libros), Gloria Villa y Maribel Velasco (incluyendo un beso fraterno)

A mis hermanos: Anabell, Mario Luis y Genaro Hernández Mendoza... Por el cariño fraterno, puesto a prueba siempre, y gracias a su paciencia me lo siguen otorgando

A Erika: porque a pesar de todo, la amistad ha vencido al tiempo

Por los momentos de esparcimiento, de ayuda en el trabajo y su amistad durante toda mi estancia y posteriormente en la ruta del tentempie: Edgar Dantán, Gabriel Guillén, Juan Olivares, Selene Napsuacily, Miguel Villalobos, Luis Cárdenas, Raúl Huertas, Aníbal Tovar, Franz Orellán, Georgina Ponce, Miguel Ángel Déctor y Nayeli Sánchez

Y finalmente a Caro. Por todo

ÍNDICE

ABSTRACT.....	6
RESUMEN.....	7
1.- INTRODUCCIÓN.....	8
1.1 La interacción simbiótica entre Rhizobiales y leguminosas.....	8
1.2 Mecanismos de resistencia a compuestos externos en bacterias Gram-negativas.....	10
1.3 Modelo de acción del sistema de transporte ternario RND-MFP-OMF.....	13
2.- ANTECEDENTES.....	14
2.1 <i>Rhizobium etli</i> y la secreción de factores de nodulación.....	14
2.2 Identificación y caracterización molecular de los genes <i>nodT</i> en Rhizobiaceas.....	15
3.- OBJETIVO GENERAL.....	18
3.1 OBJETIVOS PARTICULARES.....	18
4.- MATERIALES Y METODOS.....	19
4.1 Cepas bacterianas, plásmidos y condiciones de crecimiento.....	19
4.2 Secuencia nucleotídica y análisis filogenéticos.....	19
4.3 Técnicas de Genética y Biología Molecular.....	20
4.4 Aislamiento de los genes <i>nodTch</i> y <i>nodTpc</i> de <i>Rhizobium etli</i>	20
4.5 Intentos de mutagénesis de <i>nodTch</i>	21
4.6 Análisis de la estabilidad del plásmido pRK415:: <i>nodTch</i> en diferentes fondos genéticos de <i>R. etli</i>	22
4.7 Mutagénesis de los genes <i>nodTpc</i> , <i>rmiR</i> , <i>rmiA</i> y <i>rmiB</i>	22
4.8 Análisis de extensión de primer (PE), Reacción en cadena de la polimerasa (PCR) y PCR acoplada a transcriptasa reversa (RT-PCR).....	23
4.9 Pruebas de susceptibilidad a drogas en <i>R. etli</i> y mutantes derivadas.....	24
4.10 Aislamiento de fracciones enriquecidas de membrana externa de <i>R. etli</i>	24
4.11 Análisis de genomas y asignaciones de familias de proteínas.....	25
4.12 Correlación entre el tamaño del genoma y familias de proteínas.....	25
4.13 Metodología para la predicción <i>in silico</i> de los probables ligandos para las proteínas RND..	25
4.14 Número de acceso a la secuencia nucleotídica.....	27
5.- RESULTADOS Y DISCUSIÓN.....	27
5.1 Identificación de los genes <i>nodTch</i> y <i>nodTpc</i>	27
5.2 <i>nodTch</i> tiene una función esencial en <i>R. etli</i>	34
5.3 Posibles homólogos de los genes <i>nodT</i> se localizan en diferentes aislados geográficos de <i>R. etli</i> así como en <i>R. tropici</i>	38
5.4 Análisis transcripcional de los genes <i>rmiRABnodTch</i> y <i>nodTpc</i>	41
5.5 Localización subcelular de NodTch.....	44
5.6 Las Rhizobiales presentan una menor redundancia genómica de la familia de factores de membrana externa.....	46
5.7 Caracterización de las mutantes <i>rmiRAB</i> de <i>R. etli</i>	49
5.8 Análisis de los aminoácidos de las proteínas de la familia RND que interaccionan con distintos ligandos.....	50
5.9 Selección de las proteínas de la superfamilia RND.....	54
5.10 Análisis de los grupos formados por el alineamiento de los 14 residuos seleccionados.....	55
GRUPO AcrB (Amplio espectro de selección de compuestos).....	56
GRUPO TbtB (Expulsión de solventes).....	56
GRUPO CnrA (Expulsión de metales pesados).....	57
GRUPO ORF2pc (Expulsión de fluoroquinolonas).....	57
GRUPO AmeB (Sustrato desconocido).....	58

6. CONCLUSIONES.....	60
7. PERSPECTIVAS	61
8. BIBLIOGRAFIA.....	62
Apéndice 1.....	77
Apéndice 2.....	85
Apéndice 3.....	87

ABSTRACT

Rhizobium etli is a Gram-negative soil bacterium that induces nitrogen-fixing nodules on common bean roots (*Phaseolus vulgaris*). *R. etli* encodes two genes homologous to *nodT* of *R. leguminosarum*. *nodTch* is chromosomal and forms an operon, named as *rmiRABnodTch*, with new genes resembling a multidrug efflux pump of the resistance-nodulation-cell division (RND) family. *nodTch* is the last gene of this operon and can also be independently transcribed; the gene product is located in the bacterial outer membrane. Cell survival requires *nodTch* under all conditions tested. A second *nodT* gene, *nodTpc*, is encoded by plasmid *c*; *nodTpc* show a basal transcription level, but does not complement the essential function encoded by *nodTch*. NodT proteins belong to the outer membrane efflux proteins of the TolC superfamily. The number of duplications in the *tolC* gene family positively correlates with genome size in Gram-negative bacteria. Nonetheless, some alpha-proteobacteria, including *R. etli*, encode fewer outer membrane factor exporters than expected suggesting further roles in addition to detoxification.

Additionally, we mutagenized *rmiRAB*, which are cotranscribed with *nodTch*. No differences were observed in these mutants when compared to the wild-type strain neither in resistance to multiple drugs nor in the nodulation process. Analysis *in silico* of cluster generated with fourteen amino acids interacting with ligands and extracted from a multiple alignment of RND proteins, provide enough information to predict specific ligands in proteins of this family. Nonetheless, RmiB (a RND-like protein) was clustered in a group of proteins belonging to the Rhizobiaceae group which substrate remains unknown. We propose that this group of proteins could interact with specific substrates from the rhizosphere.

RESUMEN

Rhizobium etli es una bacteria Gram-negativa del suelo que induce nódulos fijadores de nitrógeno en las raíces del frijol común (). En el genoma de *R. etli* están codificados dos genes homólogos a *nodT* de *R. leguminosarum*. *nodTch* se encuentra codificado en el cromosoma y forma un operón con genes que tienen similitud con los sistemas de exclusión de diversas drogas de la familia Resistencia-Nodulación-División Celular (RND). *nodTch* es el último gen de este operón, al que hemos llamado *rmiRABnodTch*, y además se puede transcribir en forma independiente. El producto de este gen se localiza en la membrana externa. *nodTch* es requerido para la sobrevivencia celular en todas las condiciones en las que lo analizamos. El segundo gen *nodT*, *nodTpc*, está codificado en el plásmido *c. nodTpc* se transcribe con un nivel basal, y al parecer no complementa la función esencial observada en *nodTch*. NodT pertenece a la familia de proteínas TolC que son transportadores de membrana externa. El número de duplicaciones de la familia de genes *tolC* correlaciona en forma positiva con el tamaño del genoma en las bacterias Gram-negativas. Sin embargo, algunas alfa-proteobacterias donde está incluida *R. etli*, codifican en sus genomas muy pocos exportadores de membrana externa de los que se esperan, sugiriendo que el producto de estos genes podría ser más general que simplemente participar en transporte.

Además, mutagenizamos los genes *rmiRAB*, que se cotranscriben con *nodTch*. No observamos diferencias en resistencia a diversas drogas o en el proceso de nodulación, cuando se comparan estas mutantes con la cepa silvestre. Para obtener mayor información, realizamos agrupamientos de 14 residuos extraídos de un alineamiento múltiple de proteínas RND, los cuales están involucrados en la interacción con ligandos. Estos residuos nos dieron suficiente información como para poder predecir ligandos específicos para ciertos grupos de proteínas. Sin embargo, RmiB (que es una proteína tipo RND) fue agrupada con proteínas con un ligando desconocido, pero que todas pertenecen al grupo Rhizobiaceae. Proponemos que el sustrato específico de estas proteínas podría localizarse en la rizósfera.

1.- INTRODUCCIÓN

1.1 La interacción simbiótica entre Rhizobiales y leguminosas

Las bacterias del género rhizobia (*Azorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* y *Sinorhizobium*, colectivamente conocidas como *Rhizobium*) y las raíces de las plantas leguminosas establecen una relación simbiótica que se inicia con un intercambio mutuo de señales químicas que se difunden en la rizósfera (32, 42). Las raíces de la planta secretan flavonoides que sirven como quimioatrayentes y que además interactúan con las proteínas reguladoras bacterianas NodD, las cuales forman parte de la familia de reguladores bacterianos LysR. Esta interacción hace que NodD funcione como activador transcripcional de los genes de nodulación (*nod*, *nol* y *noe*), uniéndose a elementos promotores de la transcripción que son conocidos como cajas de nodulación (cajas *nod*) (31, 33, 43). Estas cajas de nodulación son secuencias reguladoras en *cis*, altamente conservadas y con una longitud de 47 pb (31, 115). Los productos de los genes de nodulación se encargan de la síntesis del esqueleto principal de moléculas de naturaleza lipoquitooligosacárida conocida como factores Nod (FNs) así como en su posterior decoración y secreción. Los FNs, en concentraciones nanomolares, son capaces de inducir una serie de respuestas en la planta, que van desde la deformación de los pelos radicales hasta la formación de primordios de nódulo (19).

Los genes de nodulación se han localizado en plásmidos de alto peso molecular conocidos como plásmidos simbióticos (pSym), aunque algunos genes *nod* han sido localizados también en cromosoma, como ocurre en la bacteria de crecimiento lento *Bradyrhizobium japonicum*. Las bacterias atrapadas en el ápice del pelo radical penetran a través de una invaginación de la membrana plasmática conocida como hilo de infección. Finalmente, el microorganismo logra alcanzar la región del nódulo en donde infecta las células. Una célula infectada del nódulo puede contener varios miles de microsimbiontes, cada uno recubierto de una membrana conocida como membrana peribacteroidal (MPB). Esta membrana se deriva de la membrana plasmática, aunque comparte propiedades con la membrana vacuolar (101, 129). Una vez en el nódulo, el microsimbionte se diferencia a bacteroide, fija el nitrógeno atmosférico en amonio el cual es transferido a la planta que a

su vez proporciona a la bacteria fuentes de carbono, nutrientes y un nicho para vivir (Fig. 1).

Las Rhizobiales, además de que establecen una interacción simbiótica con plantas leguminosas, también deben contener con diversos elementos externos que incluyen cambios de pH, estrés oxidativo, presencia de compuestos tóxicos, etc. Esta simbiosis en un principio se desarrolla en forma parecida a un proceso infeccioso por lo que la planta al inicio activa una respuesta de defensa que finalmente los microsimbiontes logran resistir (9).

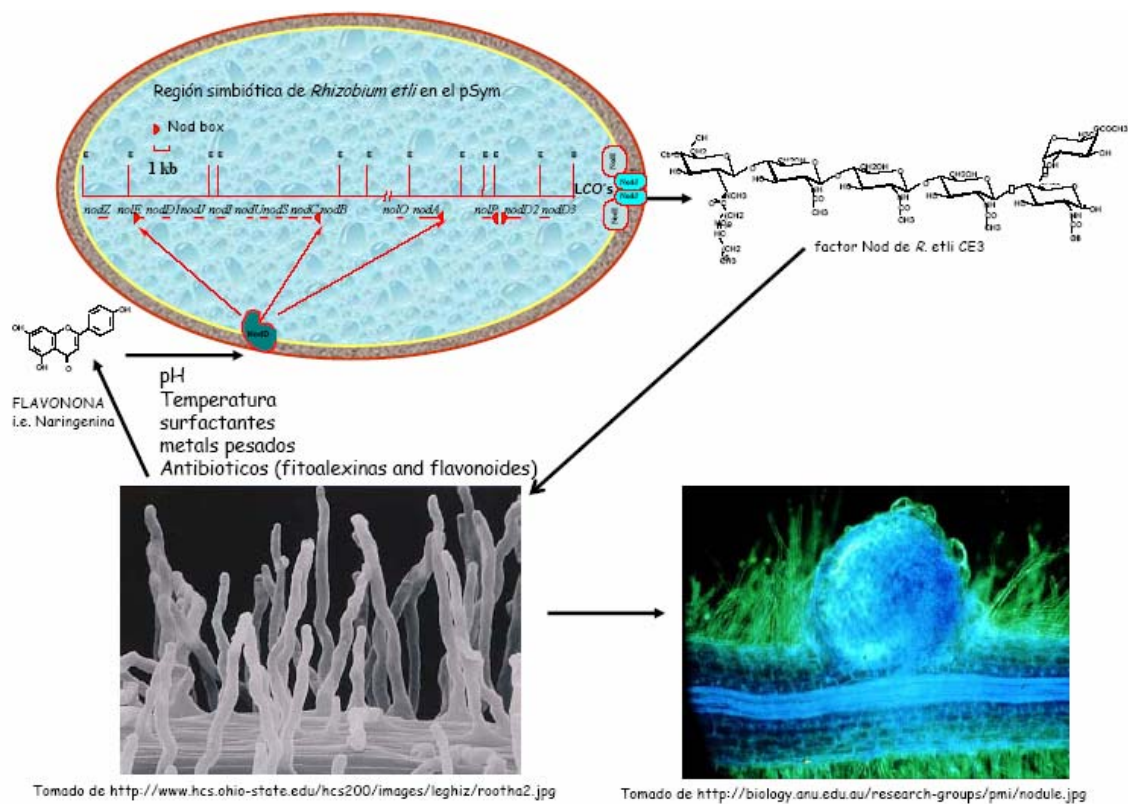


Fig. 1.- Representación esquemática de la interacción *Rhizobium*-raíces de leguminosa

1.2 Mecanismos de resistencia a compuestos externos en bacterias Gram-negativas

Existen diferentes mecanismos por los cuales las bacterias Gram-negativas contienen con la presencia de compuestos tóxicos. Anteriormente se creía que uno de los principales mecanismos de resistencia a drogas era la baja permeabilidad de la membrana externa (78). Sin embargo, actualmente prevalece la idea de que la resistencia intrínseca de las bacterias a diferentes compuestos dañinos de naturaleza heterogénea como antibióticos, detergentes, solventes o colorantes entre otros (conocidos en general como drogas), es debido a la presencia de bombas de exclusión (78). Los análisis genómicos indican que las bacterias tienen una gran cantidad de genes que codifican para probables exportadores de drogas. Estos transportadores se han clasificado en 5 familias: MFS (major facilitator superfamily), SMR (small multidrug resistance), RND (resistance/nodulation/ cell division), MATE (multidrug and toxic compound extrusion) y ABC (ATP-binding cassette) (77, 96) (Fig. 2). Las primeras 3 clases de transportadores son antiportadores H^+ /droga. MATE es un antiportador Na^+ /droga (18) y los de la familia ABC utilizan la hidrólisis de ATP como fuerza motriz. Los primeros transportadores descritos fueron los antiportadores Tetraciclina/ H^+ de la familia de transportadores MFS, los cuales están codificados en plásmidos. La familia MFS es la más grande de los exportadores de drogas, pero la mayoría de sus miembros no se expresan en condiciones normales (119). La resistencia propia de las bacterias Gram-negativas se atribuye a los exportadores tipo RND, como AcrB y MexB (59, 79), mientras que los exportadores tipo MFS son los que mayoritariamente funcionan en bacterias Gram-positivas (13, 74, 104). En las células de mamíferos, los transportadores más importantes de resistencia a drogas pertenecen a la familia ABC (26), aunque en bacterias hay muy pocos ejemplos de este tipo (50, 126). Los transportadores tipo RND usualmente funcionan como un complejo tripartita junto a una proteína de fusión a membrana (27) y un canal de membrana externa (88) y exportan drogas directamente al exterior celular, atravesando el periplasma y usando como fuerza motriz a los protones (76, 138).

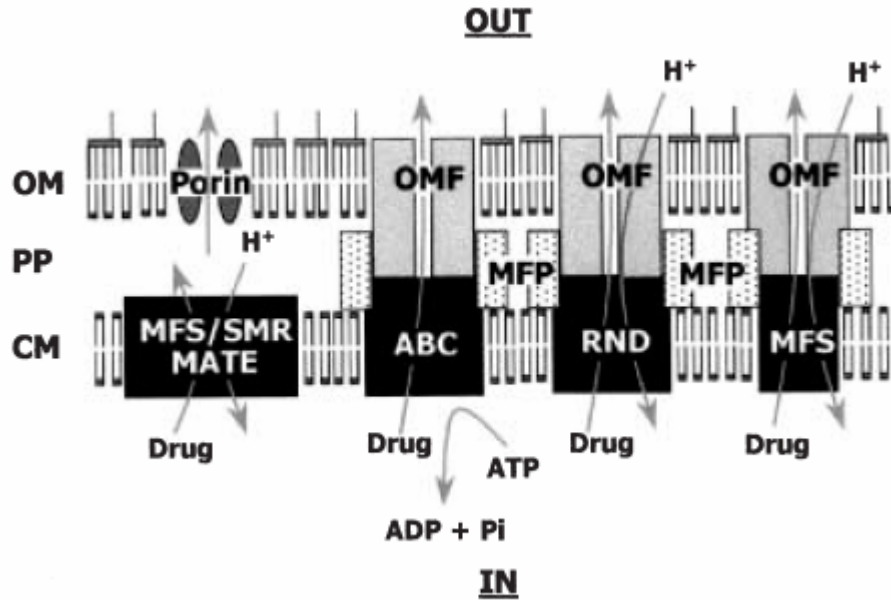


Fig. 2.- Representación esquemática de los diferentes tipos de bombas de exclusión de drogas presentes en bacterias. OM Membrana externa; PP Periplasma, CM Membrana Citoplásmica. Tomado de (91)

Se ha demostrado que proteínas transportadoras de membrana externa que pertenecen a la familia llamada factores de membrana externa (OMF: Outer membrane factor) son capaces de interactuar con la mayoría de los transportadores tipo RND y con algunos tipo ABC y MFS, como el caso de TolC de *Escherichia coli* (81) (Fig. 3). En *Pseudomonas aeruginosa*, la proteína de membrana externa OprM que es un OMF puede interactuar a su vez con varios transportadores tipo RND (61, 84, 85). Esto sugiere que las proteínas OMF pueden tener funciones adicionales a la exclusión de drogas. Se ha demostrado que TolC, además de su papel en el transporte de moléculas tan diversas como toxinas (α -hemolisina) o antibióticos (112), también participan en la segregación de cromosomas (40).

Los sustratos exportados por las proteínas tipo RND y sus complejos proteicos incluyen compuestos aniónicos, catiónicos, zwitterónicos y neutrales, aunque en su gran mayoría son de naturaleza anfífilica, que incluyen antibióticos, colorantes, detergentes, metales pesados y solventes (79). Esta familia es la que tiene un mayor rango de especificidad de sustratos (79) que la de los exportadores de mamíferos como MDR y MRP, que exportan principalmente cationes orgánicos (28) y aniones orgánicos (47),

respectivamente. Esta amplia especificidad por sustrato se debe principalmente a la estructura propia de estos transportadores.

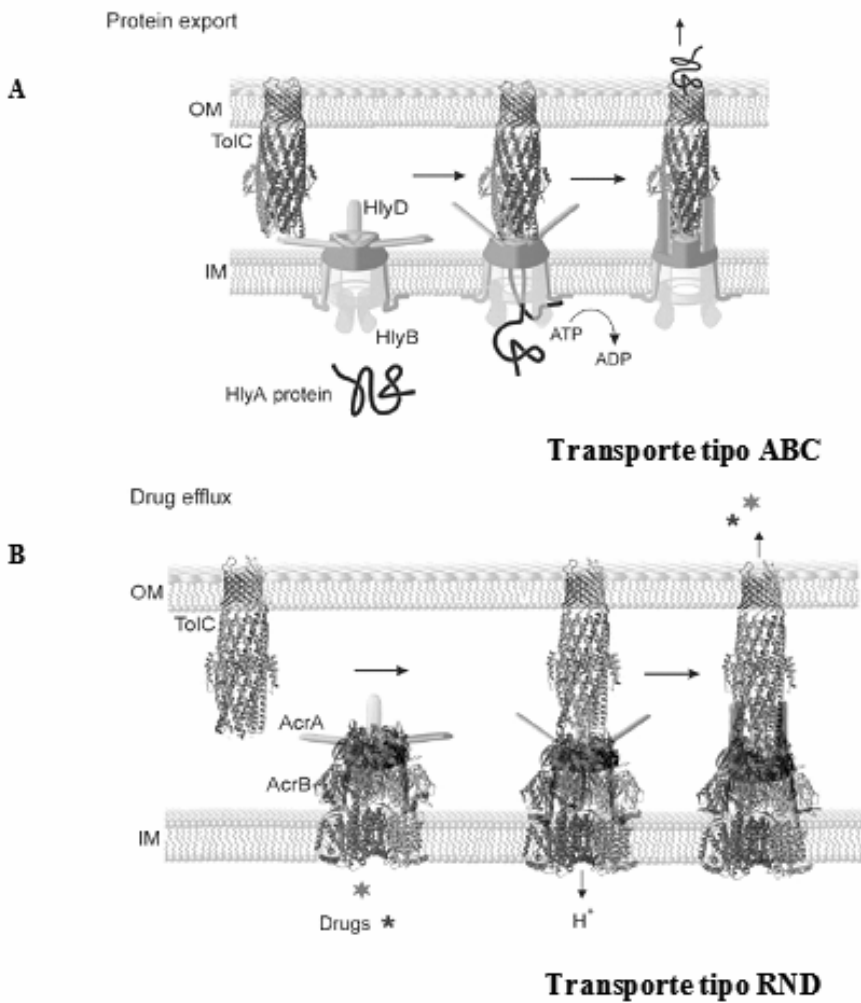


Fig. 3.- Esquema representativo de la multifuncionalidad de TolC al interactuar con diversos sistemas de transporte. En este caso se representa su función como exportador de α -hemolisina en conjunto con un sistema de transporte tipo ABC (A) y en la resistencia a drogas diversas actuando con transportadores tipo RND (B). Tomado de (51)

1.3 Modelo de acción del sistema de transporte ternario RND-MFP-OMF

La estructura tridimensional de AcrB de *E. coli*, un miembro representativo de las proteínas tipo RND, indica que estas proteínas forman homotrómeros y cada subunidad está formada por la secuencia de más de 1000 residuos de aminoácidos, con 12 segmentos transmembranales y dos grandes horquillas periplásmicas compuestas cada una de aproximadamente 300 residuos de aminoácidos lo cual genera una enorme cavidad central. La punta de este dominio periplásmico trimérico forma un embudo y su borde tiene las dimensiones apropiadas para que AcrB pueda interactuar con la punta del barril de hélices que tiene TolC. La cavidad central formada por las horquillas periplásmicas es muy amplia (35 Å de diámetro) pero el poro central parece estar esencialmente cerrado, o muy estrecho. En la unión entre subunidades, a la altura de la parte periplásmica de la bicapa lipídica interna, se forma un pequeño agujero llamado vestíbulo, que se dirige hacia la cavidad central. La estructura tridimensional cristalizada de AcrB en presencia de drogas, indica que diferentes residuos dentro de la cavidad central pueden interactuar con diferentes moléculas, lo cual sugiere que esa región tiene un papel preponderante en la selección del sustrato.

Por otro lado, se ha propuesto que los tres componentes de las bombas de exclusión de drogas tipo RND se encuentran en forma independiente una de la otra, de tal forma que la porina se encuentra unida al peptidoglicano de la membrana externa, mientras que la proteína de fusión de membrana (MFP: membrane fusion protein) se encuentra unida a la parte superior de la bomba RND pero parcialmente plegada. En presencia del sustrato, la conformación de la bomba RND se modifica permitiendo el anclaje de la porina OMF por medio de la MFP. Cuando el complejo ternario se encuentra unido, se refuerzan estas interacciones. Una vez que esto ocurre, los protones son capturados en los pares iónicos transmembranales y ocurre una conformación remota acoplada que hace que el poro central se abra. Esto es seguido por un transporte activo desde la cavidad central del embudo y posteriormente a través del OMF (Fig. 4).

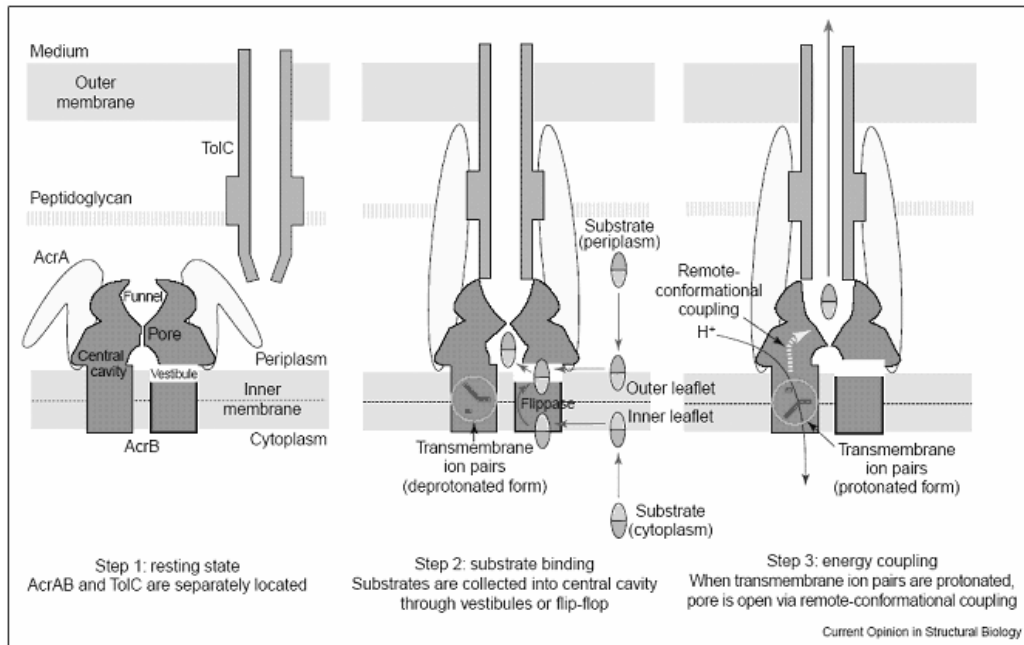


Fig. 4. Esquema representando el proceso mediante el cual las bombas de exclusión de drogas tipo RND llevan a cabo su función. Tomado de (71).

2.- ANTECEDENTES

2.1 *Rhizobium etli* y la secreción de factores de nodulación

Rhizobium etli CE3 (83, 111) es una bacteria Gram-negativa con seis plásmidos de alto peso molecular nombrados **a**, **b**, **c**, **d**, **e** y **f** (17). La mayoría de los genes de nodulación en *R. etli* se han localizado en el plásmido **d** o pSym (35). En nuestro laboratorio estamos interesados en la caracterización de genes que son de importancia para la interacción simbiótica que establece esta bacteria con la leguminosa *Phaseolus vulgaris*. Entre los trabajos reportados en este laboratorio se encuentra la descripción de la organización genómica de los genes de nodulación comunes *nodABC* de *R. etli* (127), la caracterización de los transportadores de los factores de nodulación *nodIJ* (20, 128) o la identificación del gen *nodS* (130), entre otros.

El gen *nodI* codifica para una proteína de 311 aa del tipo ABC, con una predicción en su masa molecular de 34.3 kDa. El gen *nodJ* codifica para una proteína hidrofóbica de 259 aa, con una probable masa molecular de 27.7 kDa (29). *nodIJ* han sido localizados en las especies rizobianas estudiadas genéticamente, y parecen no ser indispensables para la nodulación (29, 107, 128, 135). NodI está asociada a la membrana citoplásmica interna de *R. leguminosarum* (108) y presenta similitud significativa con KpsT, BexA y CtrD de *E. coli*, *Haemophilus influenzae* y *Neisseria meningitidis*, respectivamente. Estas proteínas participan en la secreción de polisacáridos capsulares y conservan los sitios de unión a ATP, características de la familia de proteínas transportadoras tipo ABC. Por otro lado, NodJ es una proteína hidrofóbica transmembranal sin similitud significativa con proteínas transmembranales dedicadas al transporte de polisacáridos; sin embargo, su perfil hidropático es muy similar al de las proteínas transmembranales KpsM, BexB y CtrC que participan, junto con KpsT, BexA y CtrD, respectivamente, en la secreción de compuestos (128). *nodJ* presenta en su extremo amino una secuencia similar al motivo de cierre de leucina que podría estar involucrado en la formación del multímero de NodJ o en la interacción con otras proteínas.

En nuestro laboratorio, por medio de una cinética de secreción de los FNs en *R. etli* se observó que en mutantes *nodI* o *nodJ* estos compuestos se detectan en el medio extracelular 4 horas después de la inducción de la expresión de los genes *nod* con

naringenina cuando en la cepa silvestre ocurre a tiempos tan cortos como una hora (20). Resultados similares fueron reportados en *R. leguminosarum* bv. *viciae* (116). Además, los FNs se acumulan intracelularmente en las mutantes, reflejando que *nodI* y *nodJ* tienen un papel relevante en la secreción de esos compuestos. Por otro lado, las mutantes en *nodI* y *nodJ* de *R. etli* presentan un retraso en el tiempo de aparición de nódulos, así como una reducción en el número que de éstos se forma en la raíz de *P. vulgaris*.

2.2 Identificación y caracterización molecular de los genes *nodT* en Rhizobiaceas.

El gen *nodT* fue inicialmente identificado en *R. leguminosarum* bv. *viciae* formando parte de un operón con los genes *nodMN* y en *R. leguminosarum* bv. *trifolii* con *nodABCIJ* (120). En *R. leguminosarum* bv. *phaseoli* este gen fue localizado fuera del plásmido simbiótico y no se ha encontrado que esté formando parte de un operón con genes de nodulación o que tenga elementos reguladores cercanos parecidos a cajas *nod* (99). En *R. leguminosarum* bv. *trifolii* se localizó aproximadamente a 125 nucleótidos después del gen *nodJ*, sin que se observara alguna región de regulación en esa secuencia. En *R. leguminosarum* bv. *viciae* se localizó un ORF con una similitud a nivel de aminoácidos del 60% con el producto del gen *nodT*, aunque en lugar de iniciar con el codón ATG, inicia con el codón TTG, el cual ha sido observado como codón de inicio en aproximadamente el 1% de los genes bacterianos. Tampoco en esta zona se localizó alguna caja de nodulación u otro tipo de regulador transcripcional descrito para este gen. En cambio, se observó una región que podría servir como sitio de unión a ribosomas. Mutaciones en *nodT* de *R. leguminosarum* bv. *viciae* y *R. leguminosarum* bv. *trifolii* no presentan un fenotipo claro en el proceso de nodulación y se ha propuesto que esto es probablemente debido a la presencia de copias homólogas a *nodT* presentes en las bacterias estudiadas (las cuales han sido localizadas por hibridación tipo Southern). Estas copias no han sido caracterizadas funcionalmente y podrían complementar el fenotipo de las mutaciones reportadas previamente (120). En el caso de *R. leguminosarum* bv. *viciae*, una mutación en este gen no provoca un efecto en la nodulación, pero en *R. leguminosarum* bv. *trifolii* cepa TA1, una cepa a la cual le falta en forma natural *nodT* y que nodula a una variedad de tréboles, la complementación con *nodT* le permite nodular un trébol subterráneo llamado Woogenellup, sin embargo los autores no discuten cual podría ser la causa de este efecto (56, 57). Por lo

tanto, el papel de *nodT* aun es desconocido. Hasta ahora no existen datos publicados en donde se obtengan dobles mutantes de los genes *nodT* en las diferentes Rhizobiales donde ha sido descrita la presencia de estos genes (99).

NodT tiene un péptido señal de tránsito en el extremo amino terminal, característico de las lipoproteínas de membrana externa y muestra similitud hacia el extremo carboxilo terminal con dominios conservados en las proteínas de membrana externa TolC, AprF, HlyD y CyaE (100), las cuales están involucradas en el transporte de diversas moléculas. Se propuso que NodT podría participar junto con NodI y NodJ en la excreción de FNs (100); sin embargo, los experimentos reportados por Spaink *et al* (116) no apoyan esta idea.

Los análisis de secuencia de las proteínas NodT han dado como resultado el clasificarlos dentro de la familia de proteínas OMF (outer membrane factors), la cual está formada por genes que se han localizado formando parte de operones que producen complejos proteicos encargados de la secreción de diversos compuestos como lipopolisacáridos en procariontes Gram-negativos, ácidos teicóicos que están asociados con las membranas y las paredes de bacterias Gram-positivas, polisacáridos capsulares y exopolisacáridos de bacterias Gram-positivas y negativas, y los lipooligosacáridos secretados como moléculas señales por bacterias que forman nódulos en plantas leguminosas (133).

La ubicación de NodT dentro de esta familia se realizó en base al análisis filogenético de las secuencias de aminoácidos de NodT comparándolos con los de diferentes transportadores de membrana externa (88). Los genes NodT también han sido clasificados como pertenecientes a la familia de transportadores TolC (112).

En el presente trabajo se presenta la identificación de dos copias del gen *nodT* en *R. etli* CE3. Una copia está localizada en el cromosoma y la otra en el plásmido c (38). El gen *nodT* cromosomal (*nodTch*), forma un operón con genes que presentan identidad con genes que codifican para proteínas de los sistemas de exclusión de diversas drogas tipo RND. *nodTch*, además, se transcribe en forma independiente del operón y esto hace sentido con el hecho de que este gen presenta una función esencial para la bacteria. En concordancia con los niveles de identidad que presenta con proteínas de la familia OMF, NodTch fue localizada en la fracción de membrana externa. Mutaciones en los genes ubicados río arriba de *nodTch*, a los cuales denominamos *rmiRAB* por razones que se explicarán

posteriormente, no presentan un fenotipo en nodulación o en resistencia a compuestos dañinos (antibióticos, colorantes, detergentes, etc.). Sin embargo, basados en análisis *in silico* de filogenia y de los residuos de interacción de estas proteínas con ligandos, proponemos que su función puede estar relacionada con la exclusión de compuestos propios de los procesos de interacción planta-microorganismos.

Finalmente, el gen *nodT* de plásmido c (*nodTpc*) presenta una expresión basal en las condiciones en que realizamos los experimentos. La mutación de este gen no presenta un fenotipo en la nodulación o en la resistencia a compuestos tóxicos y probablemente no complementa funcionalmente a *nodTch*.

3.- OBJETIVO GENERAL

Caracterizar funcionalmente a los genes *nodT* presentes en *Rhizobium etli*

3.1 OBJETIVOS PARTICULARES

- 1) Aislar y secuenciar los loci de *nodT* en *R. etli*.
- 2) Analizar la organización transcripcional de los genes *nodT*.
- 3) Mutagenizar los genes *nodT*, y en dado caso, aquellos genes con los que se estén co-transcribiendo, para identificar su posible función.
- 4) Hacer análisis *in silico* de los genes *nodT*, así como de los genes con los cuales puedan formar operones con el fin de sugerir sus posibles funciones.

4.- MATERIALES Y METODOS

4.1 Cepas bacterianas, plásmidos y condiciones de crecimiento

Las cepas bacterianas y los plásmidos usados en este estudio se muestran en la tabla 1. Las cepas de *R. etli* se cultivaron en medio de cultivo “Peptone Yeast” (PY) (83) a 30 °C. *E. coli* se cultivó en medio de cultivo Luria-Bertani (LB) (106) a 37 °C. Los plásmidos se mantuvieron en *E. coli* con antibióticos de selección apropiados: plásmido pBluescript II KS(+) (Stratagene), 100 µg/ml ampicilina (Amp); plásmido pRK415 (46), y sus derivados, 10 µg/ml tetraciclina (Tc); plásmido pJQ200SK (97), y sus derivados, 30 µg/ml gentamicina (Gm); 30 µg/ml kanamicina (Km) o 100 µg/ml espectinomicina (Sp). La cepa silvestre de *R. etli* y sus derivadas fueron seleccionadas usando combinaciones apropiadas de los siguientes antibióticos: ácido nalidíxico (Nal) 20 µg/ml; estreptomycin (Sm) 100 µg/ml; Sp100 µg/ml; Km 30 µg/ml; Gm 30 µg/ml o Tc 5 µg/ml.

4.2 Secuencia nucleotídica y análisis filogenéticos

La secuencia de nucleótidos del DNA plasmídico o de productos de PCR así como el diseño y generación de oligonucleótidos fue realizado en la Unidad de Síntesis y Secuenciación de DNA del Instituto de Biotecnología-U.N.A.M usando un secuenciador Applied Biosystems Inc. (ABI) modelo 373. Las secuencias de DNA fueron ensambladas y procesadas usando la paquetería de los software AbiView y BioEdit. El algoritmo BLAST y PSI-BLAST fueron usados para la búsqueda de similitudes con bases de datos de DNA y proteínas (5). Las predicciones de secuencias con características de promotores y terminadores de transcripción fueron realizadas usando el programa Neural Network Promoter Prediction (http://www.fruitfly.org/seq_tools/promoter.html) y el programa Terminador de la paquetería GCG (15). Los alineamientos de proteínas fueron generados usando el algoritmo CLUSTAL X tal y como está implementado por la paquetería del mismo nombre y usando los parámetros estándar (121). Los árboles filogenéticos fueron generados con la paquetería presentes en PHYLIP (30) y visualizados con el software TreeView (86)

4.3 Técnicas de Genética y Biología Molecular

Los procesamientos de DNA, incluyendo digestiones con endonucleasas de restricción, ligaciones, transformaciones y electroforesis en geles de agarosa fueron desarrollados como se ha descrito previamente (106). Las endonucleasas de restricción que se usaron provenían de de New England Biolabs, Inc., y la ligasa T4 DNA de Roche Diagnostics, GmbH, la transcriptasa reversa AMV de Roche Diagnostics, GmbH, y la polimerasa *Taq* Platinum de Invitrogen, Life Technologies, todas fueron usadas de acuerdo con las recomendaciones de los proveedores. La extracción de plásmidos de cepas de *E. coli* y DNA genómico de *R. etli* se realizó por medio de lisis alcalina (106), o usando los componentes del Plasmid Midi Kit (Qiagen Inc., Mississauga, Ontario, Canadá) o High Pure Plasmid Purification Kit (Boheringer Mannheim Inc, Mannheim). Los fragmentos de DNA fueron purificados de los geles de agarosa con los componentes del GeneClean Kit (Bio 101, Inc.) y se marcaron radiactivamente con el sistema *rediprime* II random prime labeling system (Amersham Pharmacia Biotech) usando las instrucciones de los proveedores. El RNA total fue aislado de cultivos en fase logarítmica de las diferentes cepas de *R. etli* con los componentes del Qiagen RNeasy Mini Kit (Qiagen) o con la metodología reportada por Chomczynski (23) y tratados con DNasa libre de RNAsas (Roche; 2 U de enzima por μg de RNA, 60 min a 37 °C). Se utilizaron 5 μg de RNA tratado con DNasa como templado para reacciones de transcriptasa reversa acoplado a PCR (RT-PCR) o de “extensión de primero” usando los reactivos para RT-PCR de Roche o de Biolabs. En todos los casos se siguieron los protocolos descritos por los fabricantes.

4.4 Aislamiento de los genes *nodTch* y *nodTpc* de *Rhizobium etli*

Los genes *nodT* de *R. etli* fueron aislados de una librería genómica de esta bacteria (44), por medio de hibridación en colonia usando protocolos estandarizados (106). Un cósmido conteniendo *nodTch* fue seleccionado y llamado pLAFRTch; este cósmido fue digerido con varias enzimas de restricción. Un fragmento *HindIII-BamHI* de 2.2 kilo bases (kb) con el gen *nodTch* completo y un fragmento de 7.8 kb *NotI* llevando la región cromosomal completa que incluye además de *nodTch*, a otros ORFs (que posteriormente se describirán), fueron aislados y subclonados en el plásmido pBluescript SK generando los plásmidos pHME22 y pHME8, respectivamente. Para clonar a *nodTch* en un vector de

expresión, un fragmento de 1.8 kb *XhoI-BamHI* derivado del pHME22 fue subclonado en el plásmido pRK415 (46) para generar el plásmido pRK415::*nodTch* (Tabla 1).

El cósmido que lleva el gen *nodTpc* fue seleccionado y llamada pLARFTpc. Un fragmento *EcoRI* de 3.7 kb fue subclonado en el pBluescript SK para generar el plásmido pHME40 (Tabla 1).

4.5 Intentos de mutagénesis de *nodTch*

Para mutagenizar *nodTch* por remoción de parte del gen en un solo paso de recombinación, un fragmento de 934 pb, obtenido por digestión del plásmido pHME22 con *FspI* fue subclonado en el vector suicida pJQ200SK (97) para generar el plásmido pJQ934. Este vector fue introducido en *E. coli* S17.1 para realizar cruza biparentales con *R. etli*, sin embargo no se obtuvieron transconjugantes. Para obtener una mutante *nodTch* letal condicionada, un fragmento de 1.8 kb *XhoI-BamHI* llevando el gen *nodTch* fue subclonado en el plásmido pBluescript SK. Un elemento interposon Ω -Km (4) fue introducido en el sitio *AatII* localizado en la región intragénica de *nodTch*. El gen interrumpido fue subclonado en el plásmido pJQ200SK y la construcción resultante, pHMTch::Km, fue introducida por cruce biparental en *R. etli*. Varias recombinantes sencillas fueron seleccionadas por resistencia a antibióticos. El plásmido pRK415::*nodTch* fue usado para transformar *E. coli* S17.1 para después introducirlo en una de las transconjugantes mencionadas por cruce biparental. Una transconjugante seleccionada conteniendo el pHMTch::Km integrado en el cromosoma y el plásmido pRK415::*nodTch* fue cultivada toda la noche. El cultivo se esparció sobre medio PY suplementado con sacarosa al 10% (w/v) para seleccionar derivadas donde se hubiera llevado a cabo un proceso de doble recombinación. Células merodiploides fueron obtenidas con una frecuencia de recombinación de 10^{-4} . De estas células que llevaban el gen *nodTch* interrumpido por un casete de resistencia a Km, y simultáneamente estaban complementadas con el plásmido pRK415::*nodTch*, se seleccionaron 5 para analizarlas por hibridación tipo Southern. 10 de ellas fueron crecidas en medio PY sin antibiótico para seleccionar posteriormente clonas que se hubieran curado del plásmido pRK415::*nodTch*.

4.6 Análisis de la estabilidad del plásmido pRK415::nodTch en diferentes fondos genéticos de *R. etli*

Para analizar la estabilidad de este plásmido en el fondo genómico de la cepa mutante en *nodTch* en comparación con la cepa silvestre, se introdujo el plásmido en la cepa silvestre, y se determinó la permanencia del marcador de resistencia a Tc del plásmido durante 10 días (aproximadamente 60 generaciones) a intervalos de 1 día (aprox. cada 6 generaciones). Cultivos en fase estacionaria fueron diluidos en medio fresco sin antibióticos de selección a una densidad óptica de 0.05 (600 nm de λ) y crecidas por 6, 12, 18, 24, 30, 36, 42, 48, 54 y 60 generaciones. De cada uno de estos cultivos se tomó un inóculo que se diluyó en forma serial y se esparció en medio sólido para cuantificar el número de células en las siguientes condiciones: a) sin antibiótico de selección, b) en la presencia de Nal para seleccionar cepas de *Rhizobium*, c) en Km para seleccionar mutantes en *nodTch* y d) en Tc para seleccionar cepas que llevaran el plásmido pRK415::nodTch. El número de colonias obtenidas en cada condición fue cuantificado y los resultados graficados como fracciones de unidad tomando como valor 1 a la población total (sin antibiótico de selección). Los resultados mostrados son representativos de 3 experimentos independientes. La presencia del plásmido pRK415::nodTch fue confirmada por hibridación tipo Southern.

4.7 Mutagénesis de los genes *nodTpc*, *rmiR*, *rmiA* y *rmiB*

Para obtener mutantes en *nodTpc*, *rmiR*, *rmiA* y *rmiB*, los fragmentos de restricción correspondientes (2.2 kb *EcoRI*-*BamHI* para *nodTpc*, 0.8 kb *EcoRI* para *rmiR*, 1.2 kb *HindIII* para *rmiA* y 2.6 kb *XhoI* para *rmiB*) obtenidos de los plásmidos pLARFTpc y pLAFRTch fueron subclonados independientemente en el plásmido pBluescript SK. Un casete de resistencia a Sp flanqueado con elementos Ω (4) fue introducido en cada uno de los genes arriba mencionados. Cada uno de los 4 genes interrumpidos fueron subclonados en forma independiente en el vector pJQ200SK. Los plásmidos resultantes pHMR::Sp, pHMA::Sp, pHMB::Sp o pHMTpc::Sp (Tabla 1) fueron introducidos por transformación en *E. coli* S17.1 y movilizados por conjugación a la cepa silvestre de *R. etli*. Las transconjugantes fueron seleccionadas en medio PY conteniendo Gm y Nal. Las transconjugantes fueron seleccionados en medio sólido PY suplementado con 10% (w/v) de sacarosa y las doble recombinantes en medio sólido en la presencia de los antibióticos

apropiados. Para confirmar la interrupción de los genes por los casetes de resistencia a antibióticos, se realizaron hibridaciones tipo Southern usando como sondas los genes *nodTpc* y *rmiRAB* así como el casete de resistencia a Sp.

4.8 Análisis de extensión de primero (PE), Reacción en cadena de la polimerasa (PCR) y PCR acoplada a transcriptasa reversa (RT-PCR).

Las RT-PCR y PCR se hicieron usando pares de primeros específicos de las regiones internas de los genes, diseñados para amplificar las regiones intergénicas de *rmiR* y *rmiA* (*rmiR*PCR: 5'GGGACGACAAGGAGACAC 3' y *rmiA*RT-PCR: 5'AGGGCCGAGGAGATGAAC 3'), *rmiA* y *rmiB* (*rmiA*PCR: 5'AAACCAGTCCCGATGAAG 3' y *rmiB*RT-PCR: 5'GAAATCGACAGCGTGAAG 3'), *rmiB* y *nodTch* (*rmiB*PCR: 5'AAGGACCGGATGGAAAGC 3' y *nodTch*RT: 5'ATTCAGGAGGAGCATGA 3') o de la región que cubre desde el gen *nodTch* y su probable región promotora (*nodTch*RT: 5'ATTCAGGAGGAGCATGA 3' y *nodTch*prPCR: 5'TTCGGAGTGGGTGTTATT3'). El DNA complementario (cDNA) fue obtenido usando transcriptasa reversa del Virus de Mieloblastosis Aviar (AMV Reverse Transcriptase) usando los protocolos del proveedor (Roche diagnostics, GmbH). Las condiciones de los ciclos de amplificación fueron las siguientes: un ciclo de calentamiento a 94 °C por 3 min; 35 ciclos a 94 °C por 30 s, 52 °C por 45 s, y 72 °C por 45 s seguido por una extensión final a 72 °C por 10 min. Los oligonucleótidos *nodTch*RT y *nodTpc*PE (5'ATTCTTGTGCGAAGCGAC3') fueron también usados como primeros en experimentos de PE. Los primeros (2 µM) fueron marcados radiactivamente en una reacción de 50 µl (Buffer de Reacción Forward 5X, 10 µl; T4 Polynucleotide Kinase (Gibco BRL) 2 µl (20 units); [γ -32P] ATP (10 µCi/µl, 3000 Ci/mmol), 5 µl y agua estéril, 28 µl). Cada uno de esos primeros fue usado en reacciones de PE separadas. Los cDNA generados fueron separados por co-electroforesis con una escalera de DNA secuencial generado a partir de una clona genómica apropiada usando los mismos oligonucleótidos en un gel desnaturalizante de poliacrilamida al 6%.

4.9 Pruebas de susceptibilidad a drogas en *R. etli* y mutantes derivadas

La susceptibilidad a antibióticos (ampicilina, carbanecilina, cloranfenicol, gentamicina, ácido nalidíxico, tetraciclina y kanamicina), detergentes (SDS, sarcosil, Triton X-100) y colorantes (bromuro de etidio, acriflavina y naranja de acridina) fue analizada en medio de cultivo PY conteniendo diferentes concentraciones de estos compuestos, realizando diluciones seriales. También se probó la susceptibilidad a compuestos fenólicos de plantas o a flavonoides tales como naringenina, catecol, ácido elágico, ácido gálico y al antibiótico vegetal catequina (+). Los datos se presentaron como MIC, que refleja la concentración de antibiótico más baja que inhibía el crecimiento celular visible después de 48 hrs de incubación a 30 °C. Adicionalmente, las susceptibilidades a drogas también fueron determinadas colocando pequeñas piezas de papel Whatmann 3MM embebidas con varias concentraciones de drogas colocadas sobre una capa de agar mezclado con una suspensión de bacterias (D.O. de 0.1 a 600 nm λ) para observar la formación de halos de inhibición.

4.10 Aislamiento de fracciones enriquecidas de membrana externa de *R. etli*

Las fracciones enriquecidas con membrana externa de *R. etli* (66, 95, 109) se obtuvieron de la siguiente manera: se creció toda la noche un cultivo bacteriano con los antibióticos de selección. El cultivo celular fue centrifugado a 10,000 g por 5 min y resuspendido en buffer PBS (pH 7.4), sonificado 2 veces por 30 s (con un ajuste de 3 según el sonicador Branson Sonifier 200) y se centrifugó a 10,000 g por 5 min para separar los restos celulares. El sobrenadante fue centrifugado a 10,000 g por 1 hr para separar las proteínas citosólicas (sobrenadante) de las proteínas de membrana (pastilla). La pastilla fue resuspendida en buffer PBS (pH 7.4) en la presencia de Triton X-100 al 2% (v/v) e incubada a 37 °C por 1 hr seguido por 1 hr de centrifugación a 10,000 g para separar la fracción insoluble (membrana externa) de la fracción soluble (membrana interna y fracción periplásmica). Las fracciones purificadas fueron corridas por electroforesis en SDS-PAGE como se ha descrito previamente (106).

4.11 Análisis de genomas y asignaciones de familias de proteínas

Los genomas eubacterianos completamente secuenciados disponibles fueron tomados del GenBank versión 155.0 (11). De los 270 originalmente tomados, el conjunto a analizar se redujo a 166 por eliminación de organismos que compartieron más del 80% de

proteínas ortólogas. Para asignar cada una de las proteínas a su respectiva familia se usaron los modelos Pfam (10) Outer-membrane Exporter Protein (OEP, PF02321), ATP binding-cassette ABC transporter (PF00005), Major Facilitator Superfamily (MFS, PF07690), y familia AcrB/AcrD/AcrF (PF00873) cuyo modelo es el más cercano para las proteínas RND. Cada proteína de las 166 bacterias no redundantes fue analizada con estos modelos usando HMMER (<http://hmmer.wustl.edu/>)

En el caso del modelo OEP, sólo se consideraron aquellas proteínas que tuvieran 2 dominios OEP como es el caso de las proteínas NodT.

4.12 Correlación entre el tamaño del genoma y familias de proteínas

Con la participación de Cei Abreu realizamos un análisis con escrituras Perl *ad hoc* usando BioPerl (117), el software gnuplot (<http://www.gnuplot.info/>) y varios módulos Perl (<http://www.cpan.org/>). El coeficiente de correlación linear Pearson (R) fue calculado con el módulo Estadístico:Descriptivo y su valor *p* fue estimado desarrollando 1,000 permutaciones de los datos del eje x, calculando la media y la desviación estándar de la nueva R, a partir de la cual los valores de z fueron obtenidos. Los valores de z fueron convertidos en valores *p* usando el módulo Estadístico:Descriptivo.

4.13 Metodología para la predicción *in silico* de los probables ligandos para las proteínas RND

La secuencia de aminoácidos de AcrB de *E. coli* fue usada como semilla para identificar miembros de la familia RND. Se realizó un PSI-BLAST usando un valor de corte de 10^{-5} contra la base de datos no redundante del CGBI. Se obtuvieron 565 proteínas de las cuales se seleccionaron sólo 44, en base a que estas habían sido caracterizadas funcionalmente con respecto a qué sustratos expulsan. RmiB y ORF2plc de *R. etli*, así como ORF2 de *R. leguminosarum* bv. *viciae* fueron incluidas también en este estudio debido a nuestro interés en dichas proteínas.

Los alineamientos múltiples de secuencia (MSA por sus siglas en inglés) de 44 proteínas seleccionadas fueron construidos con el programa Clustal X usando sus condiciones normales (121, 122), y fueron editadas manualmente para corregir problemas de introducciones de espacios y de secuencias no alineadas.

La estructura de AcrB (archivo PDB 1iwg) de *E. coli* K12 que se ha resuelto a 3.5 Å, fue utilizada para identificar, y al mismo tiempo extraer los 14 residuos que están reportados en la literatura como aquellos que interaccionan con diferentes tipos de droga: aquellos que están localizados en la cavidad central y que son: L25, K29, D99, D101, V382, A385, F386, F388, F458 y F459 (137) o que están localizados en el poro central, entre los que se encuentran: D101, V105, N109, Q112 y P116, que tienen un papel en la selección de sustratos y que por su ubicación podrían interaccionar con los compuestos que expulsan (70). Estos residuos fueron realineados y analizados por Máxima Parsimonia (MP) para generar grupos con estos “péptidos” artificiales. De esta manera, nuestra hipótesis predice que los grupos que se formen lo harán en base a la presencia de aquellos residuos que tienen mayor probabilidad de interactuar con un ligando dado.

Con la colaboración del Dr. E. Pérez-Rueda se evaluó el funcionamiento de la predicción de unión a ligando, comparando los agrupamientos realizados por Máxima Parsimonia (grupos observado) con sus correspondientes ligandos anotados. Con esto se calcularon los siguientes valores: a) verdaderos positivos (VP): grupo de proteínas con al menos un ligando en común; b) falsas positivas (FP): proteínas con ligandos diferentes al resto del grupo; c) Falsas negativas (FN): proteínas incluidas en un grupo con ligando diferente.; d) sensibilidad, $S_n = VP/(VP+FN)$, fracción de proteínas recuperadas del grupo inferido; e) valor de predicción positiva, $VPP = VP/(VP+FP)$, es la fracción de las proteínas y ligandos de un grupo inferido en donde concuerda la proteína y el ligando; y f) Precisión, $P = (S_n + VPP)/2$, que es el promedio de VPP y de la S_n . No se utilizó la clásica definición de especificidad $S_p = VN/(VN+FP)$ porque el criterio de evaluación se basó en la tasa de verdaderas negativas (VN), definidas como aquellas proteínas cuyo ligando no ha sido descrito experimentalmente. El número de ligandos es menor que el número de grupos (16) y entonces el porcentaje de VN siempre será cercano a 1, lo cual podría influir en la tendencia.

4.14 Número de acceso a la secuencia nucleotídica

Las secuencias nucleotídicas reportadas en esta tesis se enviaron a la base de datos GenBank bajo el número de acceso AY786417 para *rmiRABnodTch* y DQ115788 para *nodTpc*.

4.14 Número de acceso a la secuencia nucleotídica

Las secuencias nucleotídicas reportadas en esta tesis se enviaron a la base de datos GenBank bajo el número de acceso AY786417 para *rmiRABnodTch* y DQ115788 para *nodTpc*.

5.- RESULTADOS Y DISCUSIÓN

5.1 Identificación de los genes *nodTch* y *nodTpc*.

Para aislar el gen homólogo a *nodT* ubicado en el cromosoma de *R. etli*, se identificó y purificó un fragmento *NruI* de 7.8 kb del cósmido pLAFRTch (38). Este fragmento fue subclonado en el plásmido pSK y se generó el plásmido pHME8. La secuencia completa de este fragmento revela la presencia de cuatro marcos de lectura abiertos (ORF, por sus siglas en inglés) (Fig. 5A), cada uno precedido por un sitio de unión a ribosomas o región Shine-Dalgarno. El ORF1 codifica para una proteína de 182 aminoácidos (aa) y una masa molecular que se predice de 20.490 kDa. La proteína codificada por este ORF tiene una similitud significativa con las proteínas de la familia de reguladores transcripcionales tipo TetR, en particular con AmeR (48%) de *A. tumefaciens*, AcrR (25%) de *E. coli*, MexR (21%) de *P. aeruginosa* y TtgR (26%) de *Pseudomonas putida* (Fig. 5 C). El ORF2 codifica para una proteína de 403 aa y se predice que tendría un peso molecular de 42.312 kDa. Presenta una similitud significativa con las lipoproteínas de fusión de membrana que pertenecen a los sistemas de exclusión de drogas. Las similitudes más significativas se observan con AmeA (60%) de *A. tumefaciens*; AcrA (41%) y AcrE (36%) de *E. coli*, MexA (39%) y MexC (33%) de *P. aeruginosa*, TtgA (40%), ArpA (37%) y TtgD (34%) de *P. putida*; y con SmeD (40%) y con SmeA (35%) de *Stenotrophomonas maltophilia* (Fig. 5 C). El ORF3 se predice que tiene 1,036 aa y una masa molecular de 110.635 kDa. La similitud más significativa que presenta la proteína codificada por este gen se observa con bombas de exclusión de drogas diversas, específicamente con aquellas que pertenecen a los transportadores de la superfamilia RND, especialmente con AmeB (78%) de *A. tumefaciens*; AcrB (56%) y AcrD (55%) de *E. coli*; MexB (55%) y MexD (47%) de *P. aeruginosa*; TtgB (55%) ArpB (55%) y TtgE (50%) de *P. putida*, y con SmeE (51%) y SmeB (50%) de *S.*

maltophilia (Fig. 5 C). Finalmente, el ORF4 corresponde al gen que originalmente identificamos como *nodT* de *R. etli* (38). Se predice que la proteína codificada por este ORF tiene 483 aa y un peso molecular de 50.822 kDa. Como se esperaba, las similitudes más altas se presentaron con las proteínas NodT reportadas en las diferentes cepas de *Rhizobium*: NodT de *R. leguminosarum* bv. *phaseoli* (89%), NodT de *R. leguminosarum* bv. *trifolii* (66%) y NodT de *R. leguminosarum* bv. *viciae* (64%). En *A. tumefaciens*, que es un patógeno de plantas relacionado cercanamente a *Rhizobium*, se ha descrito AmeC que presenta también una similitud alta con el ORF4 (61%). También se observa similitud con proteínas o lipoproteínas de membrana externa como TolC (21%) de *E. coli*, OprJ (32%), OprN (30%) y OprM (29%) de *P. aeruginosa*; TtgC (29%), TtgF (29%), y ArpC (29%) de *P. putida*, así como SmeF (28%) y SmeC (25%) de *S. maltophilia* que tienen una función en el transporte de compuestos (Fig. 5 C). El análisis de la secuencia nucleotídica en las regiones intergénicas de los ORF1, 2, 3 y 4 no revelaron la presencia de regiones terminadoras de la transcripción, excepto en la región ubicada río abajo del ORF4, lo cual sugiere que estos ORF's podrían estar formando un operón. Basados en los niveles de similitud encontrados para cada uno de los ORF de *R. etli* antes descritos, decidimos llamarlos *rmiR*, *rmiA*, *rmiB* (*Rhizobium* multidrug identify) y *nodTch* (*nodT* chromosome)

Previamente, se describió en nuestro laboratorio (38) un fragmento *EcoRI* de 4 kb aislado del cósmido pLAFRTpc que lleva una región del plásmido c de *R. etli* de aprox. 20 kb que hibridó con una sonda *nodT* heteróloga. Dentro de este fragmento se localizan tres ORF's orientados en la misma dirección. ORF1 y 2 presentan similitud con CpxR (40%) y CpxA (33%), respectivamente (Fig. 5B). CpxAR forman un sistema de transducción de señales de dos componentes el cual responde a choque térmico u osmótico en *E. coli* (25, 67, 68) (Fig. 5B). El ORF3 codifica para una proteína de 483 aa y 50.822 kDa. La proteína codificada por este ORF presenta similitud significativa con NodT de *R. leguminosarum* bv. *phaseoli* (61%), *R. leguminosarum* bv. *trifolii* (55%), *R. leguminosarum* bv. *viciae* (55%) y con AmeC de *A. tumefaciens* (56%). También se observó similitud con TolC (21%) de *E. coli*, OprJ (34%), OprN (30%) y OprM (30%) de *P. aeruginosa*; TtgC (28%), TtgF (28%) y ArpC (28%) de *P. putida*, así como con SmeF

(26%) y SmeC (26%) de *S. maltophilia* (Fig. 5 C). Basados en el análisis de la secuencia, el ORF3 ubicado en el plásmido c fue nombrado *nodTpc* (*nodT* de plásmido c).

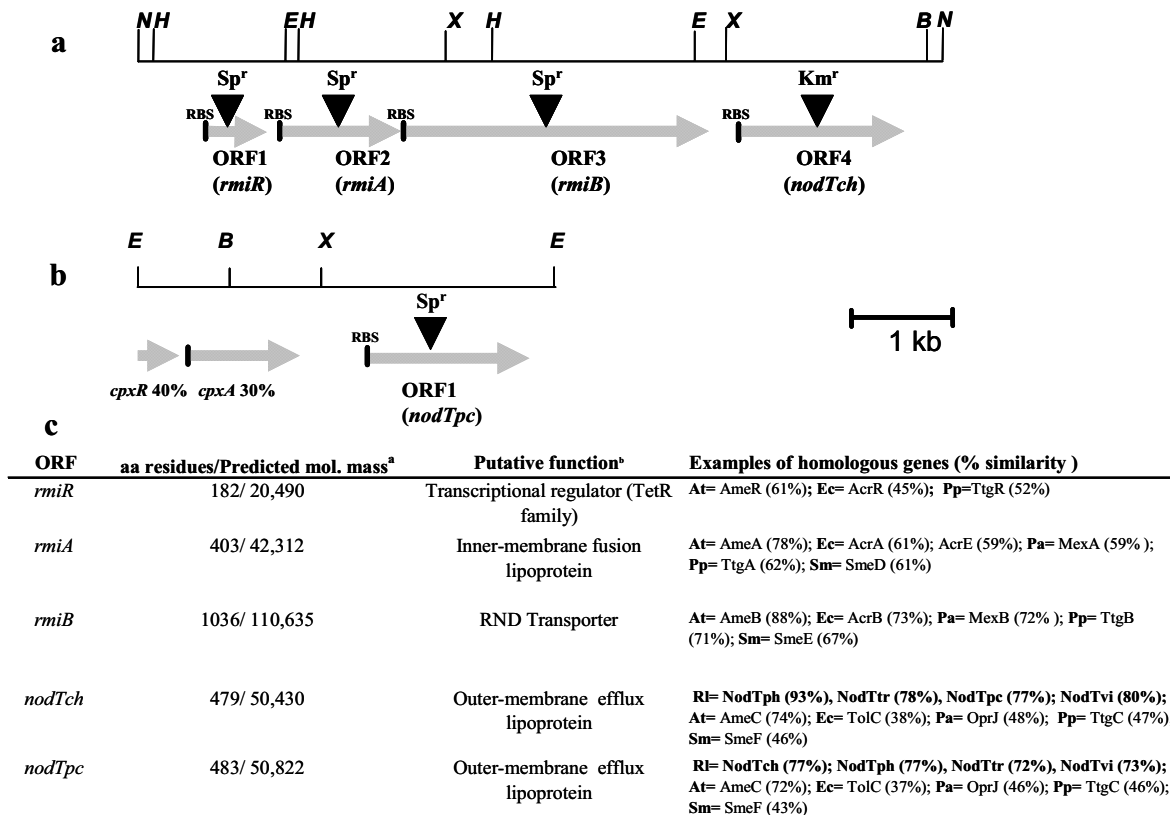


Fig. 5.- Organización genética y mapas de restricción de las regiones *rmiRABnodTch* y *nodTpc* de *Rhizobium etli*. (a) Mapa de restricción de un fragmento de DNA de 7,820 pb de una región cromosomal de *R. etli* conteniendo los ORF's indicados. (b) Mapa de restricción de fragmento de DNA de 3,703 pb correspondiente al plásmido c de *R. etli* que lleva una copia de *nodT* (*nodTpc*). Las flechas indican predicción de marcos de lectura abierta. Las cabezas de flecha indican sitios de interrupción por casetes de resistencia a Sp o Km en las mutantes *rmi* y *nodTch* de *R. etli*. RBS= Probables sitios de unión a ribosoma (c) Descripción molecular de posibles proteínas codificadas por *rmiRABnodTch* y *nodTpc*. Las abreviaturas de los microorganismos son: At, *Agrobacterium tumefaciens*; Pa, *Pseudomonas aeruginosa*; Pp, *Pseudomonas putida*; Ec, *Escherichia coli*; Sm, *Stenotrophomonas maltophilia*; Rl *Rhizobium leguminosarum* biovar ph: *phaseoli*, tr: *trifolii*, vi: *viciae*. Los números de acceso a GeneBank se indican: AmeR AmeABC (AF232237), AmrR AmrAB (AF072887), TtgR (AF238479) TtgABC (AF031417), TtgDEF (AFY19106), ArpABC (AF183959), SmeRSABC (AF173226), SmeDEF (AJ252200), MexRAB-OprM (L11616); MexCD-OprJ (U57969); AcrAB (AE000152); NodTtrifolii (P24145); NodTviciae (P15727); NodTphaseoli (PD466172). ^a Masa Molecular expresada en Daltons. ^b Como se predice con PSI-BLAST. B: *Bam*HI, E: *Eco*RI, H: *Hind*III, N: *Not*I X: *Xho*I

Con la finalidad de explorar las características de los homólogos a NodT, realizamos un alineamiento múltiple entre estas proteínas (Fig. 6A). Sin el uso de algoritmos especializados, a lo largo de la secuencia se observan varios aminoácidos conservados. En el extremo amino terminal se observaba la presencia de un péptido señal de tránsito característico de las proteínas de membrana externa. También se observan 3 regiones muy conservadas a lo largo de toda la secuencia (residuos 140-160; 282-302 y 344-364) (Fig. 6A). Previamente se había reportado que estos dominios se encuentran conservados en proteínas NodT así como en proteínas transportadoras de membrana externa, como TolC, AprF, HlyD y CyaE (100) que son transportadores de péptidos pequeños (Fig. 6A). Los niveles de identidad, usando la matriz BLOSUM62, son relativamente altos con un promedio de 68% de identidad (Fig. 6B). Las proteínas con mayor identidad son NodTch de *R. etli*, NodT de *R. leguminosarum* bv. *phaseoli* y ORF4 de *R. leguminosarum* bv. *viciae* con un promedio de 90% de identidad (Fig. 6B). Estas proteínas están codificadas por genes ubicados en cromosomas. Las proteínas NodT codificadas en los plásmidos de *R. etli* (*nodTpc*), *R. leguminosarum* bv. *viciae* y *R. leguminosarum* bv. *trifolii* presentan un promedio de identidad del 61% entre ellas (Fig. 6B). La proteína más divergentes es AmeC que está localizada en *A. tumefaciens* y la cual presentó un promedio de 56% de identidad con los NodT (Fig. 6B). Con este resultado concluimos que los genes *nodT* ubicados en cromosoma están más conservados que los de plásmido, los cuales pueden estar sometidas con mayor facilidad a una transferencia horizontal. AmeC presenta una mayor similitud con los NodT codificados en genes cromosómicos (58% de identidad) que con los codificados por genes ubicados en plásmidos (53% de identidad) (Fig. 6B), lo cual sugiere que *ameC* podría ser el ortólogo de *nodTch*.

Para analizar la presencia de dominios en NodT y compararlos con otras proteínas de la familia TolC, procedimos a realizar una búsqueda de dominios con el programa MEME/MAST disponible en la red (<http://meme.sdsc.edu/meme/intro.html>). Usando 43 secuencias de proteínas de esta familia, incluyendo a los NodT, requerimos al programa que hiciera una búsqueda de 10 dominios como máximo. El valor combinado de p (probabilidad de que un dominio se localice al azar) más bajo fue de $2.69e-12$ (una lipasa de *Helicobacter pylori*) y el mayor fue de $3.96e-149$ (CusC de *E. coli*). En el grupo de proteínas NodT se localizaron 10 dominios, mientras que las proteínas que se encargan del transporte de diversas drogas como MtrE, OprM, AdeC, OpmD, OprN y OpcM, entre otras, presentan 9 dominios (Fig. 7). Solo SrpC, una proteína transportadora de membrana externa *P. putida* que está involucrada con la resistencia a solventes y antibióticos (49), presenta los 10 dominios, lo cual sugiere que esta proteína puede tener una función similar a NodT o que NodT pudiera estar involucrada con el transporte de diversa drogas. Sin embargo, se requieren experimentos adicionales para corroborar esta hipótesis. La búsqueda de este dominio usando la matriz de valores de posición específica generada con MEME, en la base de datos SwissProt o en la de péptidos no redundantes ligados a MAST, dio como resultado a las proteínas NodT, AmeC, SrpC y otras hipotéticas (datos no mostrados). Con esto concluimos que NodT se diferencia del resto de las proteínas de la familia TolC por la presencia de un solo dominio (dominio 7) localizado en la región media, el cual pudiera tener un papel importante en la nodulación o en algún proceso común de las bacterias del grupo Rhizobiaceae.

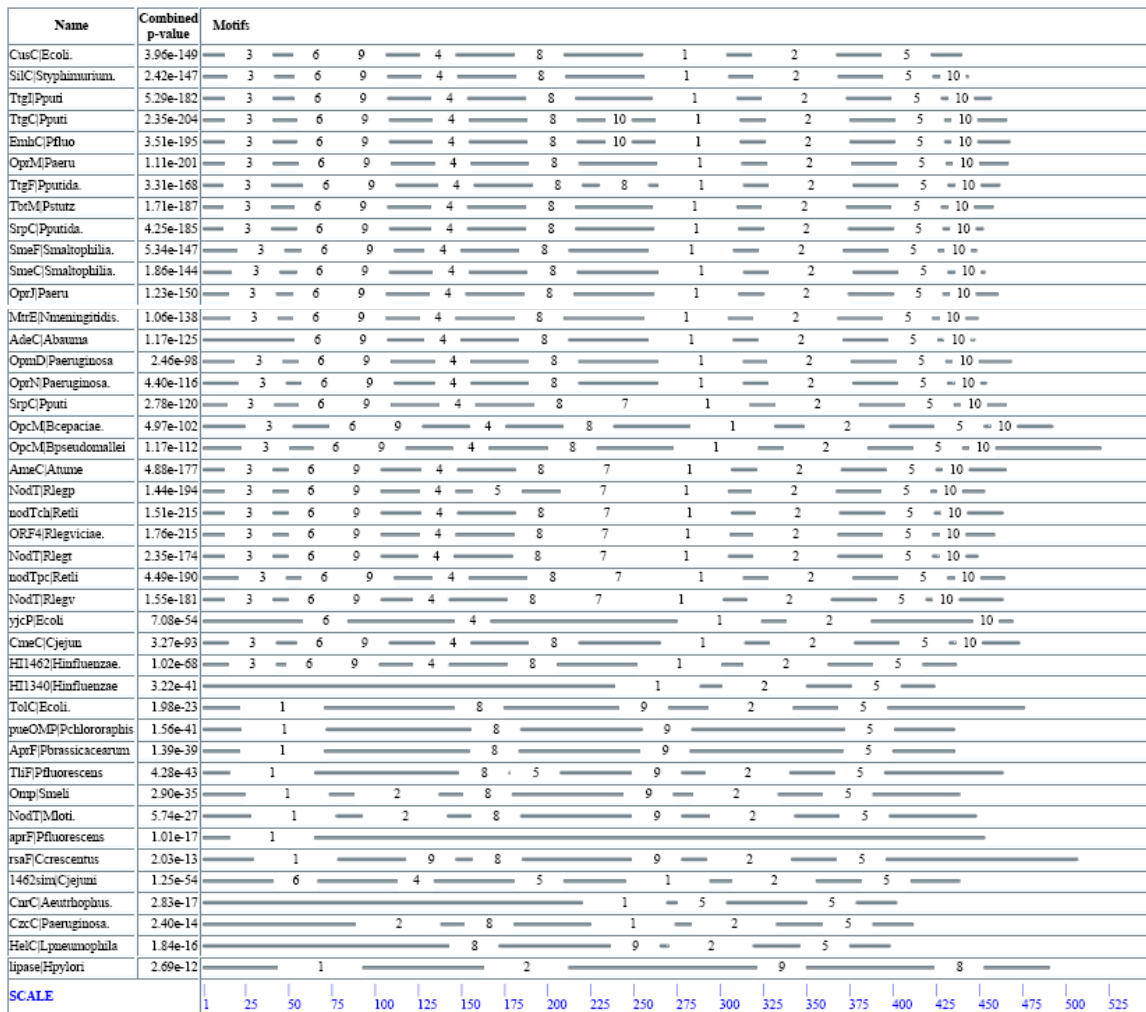


Fig. 7.- Presencia de dominios conservados en diferentes proteínas de la familia TolC. Basados en la información disponible hasta el momento de la familia NodT/TolC, se aislaron 43 secuencias de esta familia, incluyendo los genes que nosotros reportamos y homólogos localizados en las bases de datos de especies pertenecientes al grupo Rhizobiaceae. Estas secuencias fueron usadas para una búsqueda de dominios (regiones muy conservadas) con el programa MEME/MAST disponible en línea solicitando que encontrara como máximo 10 dominios. El nombre de la proteína se muestra tal y como se conoce. Abreviaturas: Abauma (*Acinetobacter baumannii*); Aeutrophus (*Alcaligenes eutrophus*); Atume (*Agrobacterium tumefaciens*); Bcepaciae (*Burkholderia cepacia*); Bpseudomallei (*Burkholderia pseudomallei*); Ccrescentus (*Caulobacter crescentus*); Cjejuni (*Campylobacter jejuni*); Ecoli (*Escherichia coli*); Hinfluenzae (*Haemophilus influenzae*); Hpylori (*Helicobacter pylori*); Lpneumophila (*Legionella pneumophila*); Mloti (*Mesorhizobium loti*); Nmeningitidis (*Neisseria meningitidis*); Paeru (*Pseudomonas aeruginosa*); Pbrassicacearum (*Pseudomonas brassicacearum*); Pchlororaphis (*Pseudomonas chlororaphis*); Pfluorescens (*Pseudomonas fluorescens*); Pputi (*Pseudomonas putida*); Pstutz (*Pseudomonas stutzeri*); Retli (*Rhizobium etli*); Rlegp (*R. leguminosarum* bv. *phaseoli*); Rlegt (*Rhizobium leguminosarum* bv. *trifolii*); Rlegviciae (*Rhizobium leguminosarum* bv. *viciae*); Smeli (*Sinorhizobium meliloti*); Smaltophilia (*Stenotrophomonas maltophilia*) y Styphimurium (*Salmonella typhimurium*).

5.2 *nodTch* tiene una función esencial en *R. etli*

Con el fin de caracterizar funcionalmente a los genes *nodT* en *R. etli*, se procedió a mutagenizar ambos genes. Para obtener una mutante con interrupción del gen *nodTch* en un evento de recombinación sencilla, un fragmento *FspI* de 934 pb que lleva una región interna de *nodTch* se clonó en el vector pJQ200SK (el cual es resistente a Gentamicina [Gm] y lleva el gen *sacB* que se induce en presencia de sacarosa y cuyo producto es tóxico en bacterias Gram-negativas), generando el plásmido pJQ934 (Material y Métodos). Este vector fue introducido por transformación en *E. coli* S17-1 para posteriormente transferirlo a *R. etli* por conjugación biparental. La selección de las posibles recombinantes fue hecha con Gm. Este experimento fue repetido al menos 6 veces y en ningún caso se obtuvieron colonias resistentes a Gm, lo cual sugiere fuertemente que *nodTch* tiene un papel en la sobrevivencia en *R. etli*. Posteriormente intentamos interrumpir el gen *nodTch* con un enfoque similar, ahora con un casete de resistencia a Km. Para esto, el fragmento de 2.2 kb *EcoRI-BamHI* que lleva el gen *nodTch* completo se interrumpió con un casete de resistencia a Km y se subclonó el fragmento completo en el vector pJQ200SK para generar el plásmido pHMETch::Km (Material y Métodos). Esta construcción se introdujo por transformación en *E. coli* S17.1 y posteriormente se transfirió a *R. etli* por conjugación biparental. Se obtuvieron recombinantes sencillas (Gm^r , Km^r , Nal^r) con una frecuencia de 10^{-4} . Para iniciar el proceso de seleccionar a una doble recombinante, una sola colonia fue crecida toda la noche en medio de cultivo líquido sin presión selectiva. Posteriormente se hicieron diluciones seriales y se tomaron 100 μ l para esparcirlos en placas de medio con agar en presencia de 5 ó 10% de sacarosa con o sin antibiótico de selección (Gm). En todos los casos, la frecuencia de colonias resistentes a sacarosa y Km observadas fue de 10^{-5} . Al analizar el crecimiento de estas colonias (aproximadamente 300) con marcadores de resistencia, el 100% de ellas resultaron ser Gm^r , indicando que al menos la parte del vector pJQ200SK que confiere la resistencia a Gm aún se mantenía en esas colonias. Al analizar el DNA de estas colonias por medio de una hibridación tipo Southern, se observó que el gen *nodTch* permanecía intacto, lo cual refuerza la idea de que este gen tiene un papel esencial en las condiciones en que realizamos nuestros ensayos.

La única forma en que fue posible interrumpir a *nodTch* con un casete de resistencia a un antibiótico (en este caso Km), fue cuando las células llevaban *in trans* una copia de *nodTch* intacta. Para generar esta mutante condicional, un fragmento de 1.8 kb *Bam*HI-*Hind*III que lleva el gen *nodTch* fue clonado bajo el control del promotor *plac* en el plásmido pRK415, generando el plásmido pRK415::*nodTch*. Este plásmido se usó para transformar *E. coli* S-17.1 y, posteriormente, introducirlo por conjugación biparental en una cepa de *R. etli* que previamente adquirió por conjugación y recombinación sencilla el plásmido pHMETch::Km. Después de la complementación con el plásmido pRK415::*nodTch*, la nueva cepa se creció en medio líquido y se esparció en medio sólido con sacarosa (5%) para seleccionar colonias doble recombinantes. En este caso, la frecuencia de recombinación observada fue de 10^{-4} . Los análisis del DNA genómico de 18 presuntas colonias doble recombinantes indicaron que el 100% de los eventos ocurrieron en el gen *nodTch* ubicado en el cromosoma y no en la copia que complementa *in trans* (datos no mostrados). Este resultado muestra al mismo tiempo que es poco probable que la letalidad observada al tratar de interrumpir a *nodTch* se deba a un efecto polar sobre genes localizados río abajo de *nodTch*, ya que con sólo introducir este gen en un vector de expresión es posible revertir el efecto de letalidad observada en una mutación en *nodTch*.

En el siguiente paso intentamos obtener una cepa curada del plásmido que complementa la mutación de *nodTch*. Para esto, cinco colonias fueron crecidas por separado en medio líquido sin antibiótico de selección por varios ciclos de generación y la población fue esparcida en medio sólido con o sin antibióticos de selección. De cada población se eligieron 100 colonias para crecerlas en medio sólido buscando colonias que fueran Tc sensibles, indicador de que habían perdido el plásmido pRK415::*nodTch*. Sin embargo, ninguna colonia perdió el plásmido que complementa a la mutación en *nodTch* (datos no mostrados). También se analizó la estabilidad de este plásmido introduciéndolo en la cepa silvestre de *R. etli* para compararla con la cepa mutante *nodTch*::Km complementada con el plásmido. Ambas cepas fueron crecidas en medio líquido sin antibiótico de selección por varias generaciones y cada 6 ciclos de generación se tomó una alícuota del cultivo para seleccionar a la población en medios sólidos con antibióticos que permitieran observar qué fracción de la población había perdido el plásmido

pRK415::*nodTch*. El resultado que se observó es que el plásmido pRK415::*nodTch* se va perdiendo en la cepa silvestre conforme va pasando el tiempo de generación en que crece la bacteria sin antibiótico de selección (Fig. 8A). Sin embargo, en la cepa que lleva interrumpido el gen *nodTch*, el plásmido pRK415::*nodTch* se mantuvo en el 100% de la población (Fig. 8B). Este experimento se corroboró seleccionando 18 colonias al final del experimento y extrayendo su DNA genómico para analizar la presencia del plásmido por medio de hibridación tipo Southern.

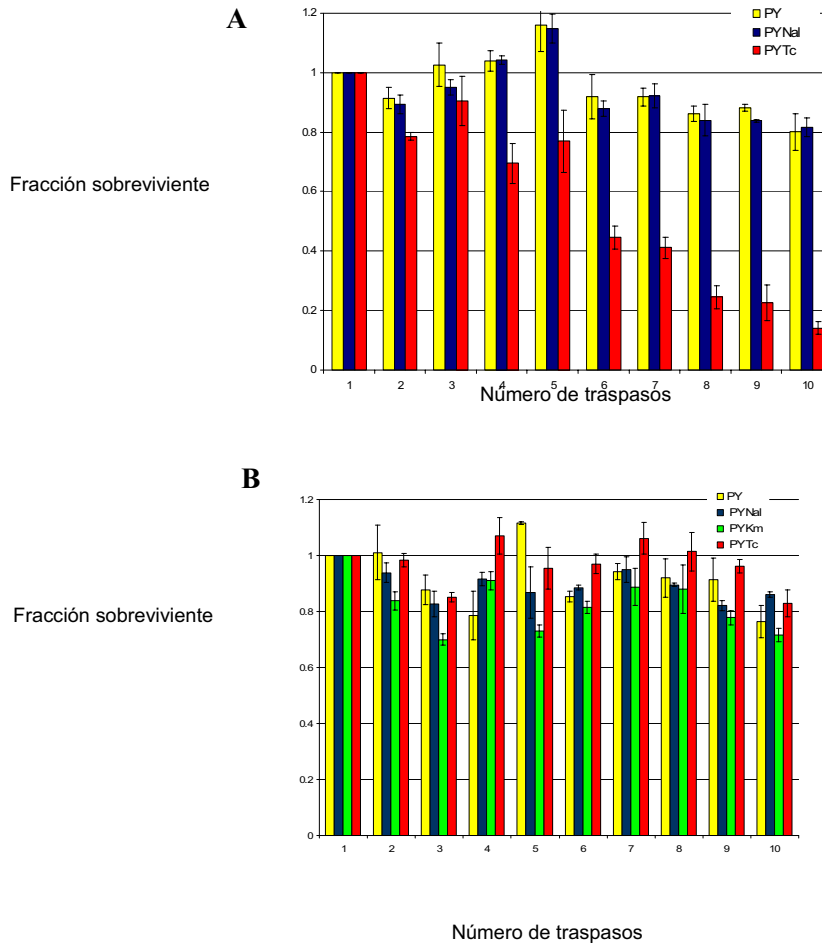


Fig. 8. – Análisis de la estabilidad del plásmido pRK415::*nodTch* en la cepa silvestre de *R. etli* y en la mutante *nodTch*::Km complementada con pRK415::*nodTch* en *R. etli*. Se muestra mediante barras la población bacteriana resistente a diferentes antibióticos crecidas en medio líquido sin presión selectiva (antibiótico Tc) durante 10 días. Se observa la herencia del plásmido pRK415::*nodTch* en la cepa silvestre de *R. etli* (A) y en la cepa de *R. etli* mutada en *nodTch*::Km previamente complementada con el mismo plásmido (B). PY.- Medio de crecimiento sin antibióticos para seleccionar la población total; o con Nal (Ácido Nalidíxico) para seleccionar la población de *R. etli*; Km (Kanamicina), o Tc (Tetraciclina).

La incapacidad de obtener una mutación en el gen *nodTch* a menos que estuviera presente un gen *nodTch* silvestre y el no poder curar el plásmido pRK415::*nodTch* en una cepa mutada en *nodTch* a pesar de ser inestable en la cepa silvestre, nos permite proponer que *nodTch* es un gen esencial para *R. etli*. Debo señalar que en el transcurso de este trabajo no realizamos una complementación con el gen *nodTpc* clonado en el vector pRK415. Sin embargo, dado que nuestros datos indicaban que *nodTpc* se transcribía en forma basal, consideramos que no era necesario llevar a cabo este experimento.

En contraste a *nodTch*, *nodTpc* fue mutagenizado fácilmente por la interrupción del gen con un casete de resistencia a Sp usando una metodología similar a la descrita arriba, sin necesidad de complementar previamente con el gen *nodTpc in trans*. En base a estos datos, concluimos que *nodTpc* no es un gen esencial para *R. etli*. El fenotipo de la mutante fue analizado durante el proceso de nodulación, y determinando si era más sensible a la presencia de diversos compuestos dañinos tales como antibióticos, detergentes, colorantes, etc. No se observaron diferencias entre la mutante y la cepa silvestre en ninguno de los ensayos llevados a cabo. Esto sugiere que sin importar el grado de similitud observado entre *nodTpc* y *nodTch*, *nodTpc* no complementa la función esencial codificada por *nodTch* pero al mismo tiempo no podemos excluir la posibilidad de que *nodTch* pueda complementar algunas de las funciones de *nodTpc*. Ante la incapacidad de obtener una doble mutante *nodTch/nodTpc*, no se puede definir si ambos genes son funcionalmente homólogos o no.

Por otro lado, se analizaron las velocidades de crecimiento tanto de la mutante *nodTch* complementada como la mutante en *nodTpc*, comparándolas contra la cepa silvestre. Las curvas de crecimiento no mostraron diferencias cuando se midió el crecimiento de estas cepas en medio PY o en medio mínimo (datos no mostrados). En estos ensayos se midió tanto la densidad óptica con un espectrofotómetro como los miligramos de proteína/mililitro.

Dentro de las proteínas analizadas con MEME/MAST anteriormente, que puedan tener una función esencial están HI1462 de *Haemophilus influenzae* que tiene un papel en la resistencia a diversas drogas según reportes recientes (123), y que por muchos años fue considerado un gen esencial (3). Otro caso es OpmD, ya que mutaciones en *opmD* de *P. aeruginosa* provocan problemas de crecimiento (1, 2). En ambos casos, estas proteínas

presentan la misma disposición de dominios que los NodT. Sin embargo, también les falta el dominio 7 que describimos anteriormente, lo que sugiere que éste dominio no es el responsable del papel esencial observado en NodTch.

Está demostrado que la duplicación de genes es la fuente primaria de nuevos genes. Los genes duplicados que son establemente preservados en los genomas usualmente tienen funciones divergentes (37). Un escenario probable ante los resultados arriba descritos, es que una vez que ocurrió la duplicación del gen ancestral *nodT*, los genes duplicados sean más susceptibles a adquirir mutaciones y dividir sus actividades en comparación con el gen ancestral. Posteriormente, uno de los genes retiene la función ancestral mientras que el otro adquiere nuevas funciones. Esta hipótesis tendría que ser explorada posteriormente en forma experimental.

5.3 Posibles homólogos de los genes *nodT* se localizan en diferentes aislados geográficos de *R. etli* así como en *R. tropici*

Con la finalidad de ver la conservación de *nodT* en *R. etli* se procedió a realizar experimentos tipo Southern usando como sonda *nodTch* o *nodTpc* sobre DNA genómico digerido con *EcoRI* de diferentes aislados geográficos de *R. etli*, así como de *R. tropici*, otro simbiote de *P. vulgaris* (65). Como control positivo se utilizó además DNA genómico de *R. leguminosarum* bv. *viciae*, una de las bacterias de donde originalmente se aislaron los genes *nodT* (120). Como se observa en la Fig. 9A, la hibridación con el gen *nodTch* dio una fuerte señal en el DNA genómico de *R. etli* CE3 a la altura de un fragmento *EcoRI* de 7.7 kb que corresponde al tamaño de la región de *nodTch* y una señal más tenue en un fragmento *EcoRI* de 4 kb que corresponde al tamaño de la región de *nodTpc*. En *R. leguminosarum* bv. *viciae* se observan tres señales que corresponden a una de las copias completas y a otra copia que se fragmenta en dos debido a un sitio de restricción *EcoRI* interno. En *R. tropici* es posible observar un único fragmento de aproximadamente 6 kb, sugiriendo la presencia de un único homólogo a *nodT* en este organismo. Con respecto a los diferentes aislados geográficos de *R. etli* se observan generalmente 2 fragmentos con diferentes intensidades y distintos tamaños con respecto a *R. etli* CE3. El único caso en el que se observó un fragmento de 7.7 kb fue en *R. etli* VIKING aunque en esta cepa no se observa ningún fragmento adicional. Existe una señal

conservada de aproximadamente 3 kb en *R. etli* Nitragin, KIM5, F14, CFNI y Brasil aunque en estos aislados geográficos también se observa una señal de aproximadamente 6 kb excepto en *R. etli* Brasil y CFNI donde sólo se observa la señal de 3 kb. Esto podría indicar que una de las copias se ha perdido en estos aislados geográficos o se ha ganado en otros y dado nuestros resultados es posible especular que ésta copia podría ser la ubicada en el plásmido c. Resultados similares se observaron cuando se usó como sonda a *nodTpc*, aunque por razones obvias en *R. etli* CE3 la señal a la altura de 4 kb es más intensa que la de 7.7 kb (Fig. 9B). Con esta sonda el patrón de bandas que se observó es similar al que se observó con *nodTch*, aunque las señales más débiles detectadas con *nodTch* eran difíciles de observar cuando se usó como sonda *nodTpc*, probablemente debido a que este gen presenta la identidad más baja entre todos los genes *nodT* (59%, Fig. 5 C). Sin embargo, en *R. etli* TAL812 se observa una señal muy intensa de 4 kb y en vez de una de 7.7 kb, se observa una señal de alrededor de 2.9 kb. Cuando se usa como sonda *nodTch*, la intensidad de las señales se invierte. Esto sugiere que en *R. etli* TAL812 *nodTpc* está conservado y que el locus *nodTch* de 7.7 kb adquirió un fragmento *EcoRI*. Con los resultados combinados de la presencia de las señales de hibridación y sus intensidades, concluimos que es muy posible que *nodTch* se encuentre conservado en los diferentes aislados geográficos de *R. etli*, aunque existen polimorfismos de estos loci dependiendo de la cepa. También se identificó un homólogo a *nodT* en *R. tropici* y dos homólogos en *R. leguminosarum* bv. *viciae* como se esperaba. Además, *nodTpc* solo está conservado en *R. etli* CE3 y en *R. etli* TAL812. Para corroborar esto se requiere aislar y secuenciar los genes de los diferentes aislados geográficos para analizar directamente su secuencia. El que solo se conserve *nodTpc* en un aislado geográfico además de *R. etli* CE3, sugiere que este gen es dispensable del genoma bacteriano. Debido a su localización en plásmido, la posibilidad de que esta copia pueda haberse ganado o perdido por transferencia horizontal en cualquiera de las cepas es muy alta. Por otro lado, el bajo número de copias de *nodT* en los diferentes aislados geográficos sugiere que en los organismos en donde sólo se detecta una copia, este gen también podría tener una función esencial tal y como observamos con la cepa *R. etli* CE3.

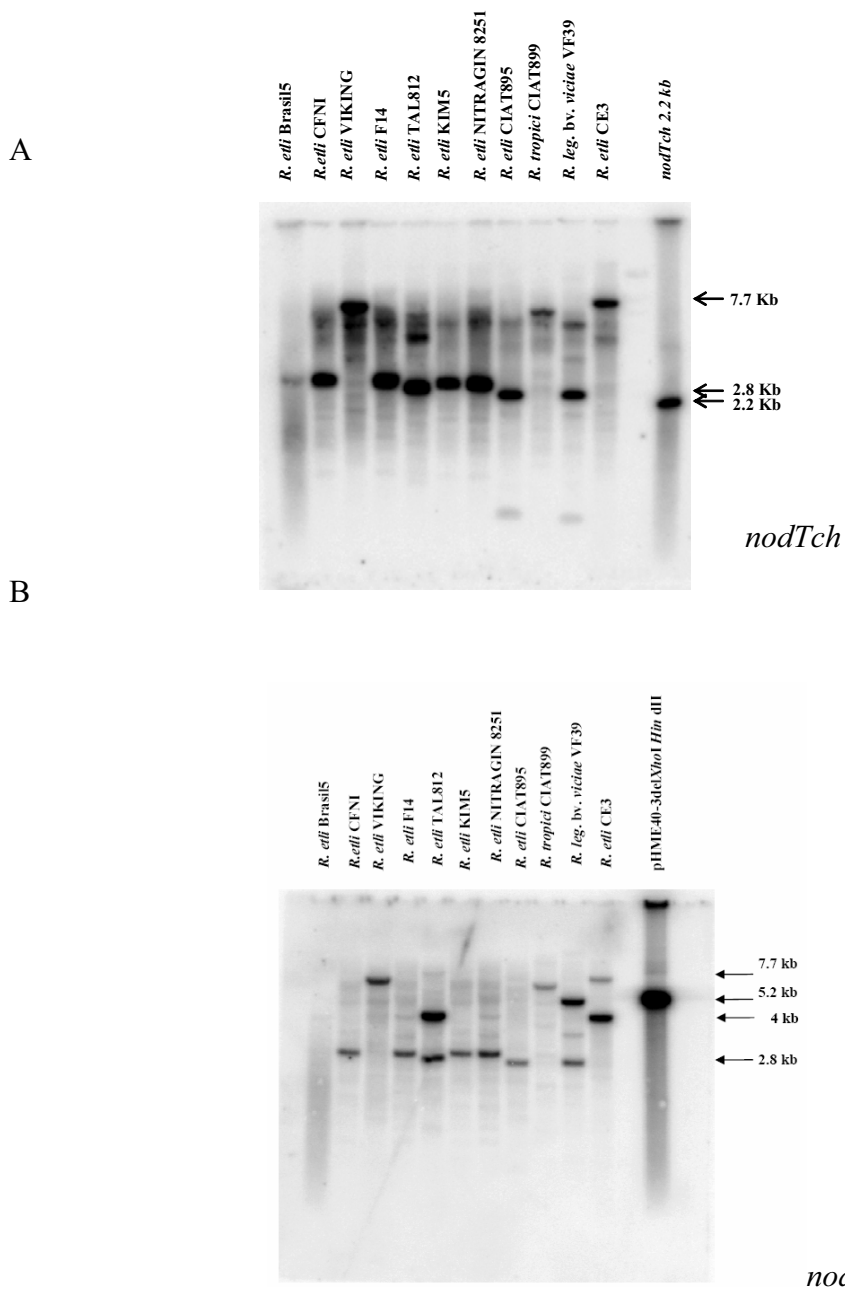


Fig. 9.- Genes homólogos a *nodT* se encuentran conservados en diferentes aislados de *R. etli* y en *R. tropici*

Hibridación tipo Southern sobre DNA genómico de diferentes aislados geográficos de *Rhizobium etli* digerido con *EcoRI*, así como de otras especies de *Rhizobium* como se indica. Se usaron como sondas fragmentos radiomarcados con ^{32}P de 1.8 kb que lleva el gen *nodTch* (A) o de 4 kb que lleva el gen *nodTpc* (B). Los números a la derecha indican el peso molecular en kilobases.

5.4 Análisis transcripcional de los genes *rmiRABnodTch* y *nodTpc*

Debido a que no se encontraron estructuras propias de los terminadores de la transcripción tipo Rho-independiente (113) usando el programa Terminator (15, 16), ni se observaron secuencias con características de los terminadores Rho-dependiente dentro de las secuencias intergénicas entre *rmiR*, *rmiA*, *rmiB* y *nodTch*, se sugiere que podrían formar un operón. Sin embargo, río abajo de *nodTch* se observa una secuencia con características de un terminador de la transcripción tipo Rho-independiente (GAATACGAACCCCGCGATCGCCTCGCGGGTTTTTTC), lo que sugiere que es el último gen del posible operón.

Para definir la organización transcripcional de *rmiRAB* y *nodTch*, realizamos RT-PCRs de las regiones intergénicas entre *rmiR* y *rmiA*, *rmiA* y *rmiB*, así como *rmiB* y *nodTch*, usando oligonucleótidos específicos diseñados de esas regiones. En cada uno de los casos se obtuvieron productos de RT-PCRs que correspondían al tamaño esperado (Fig. 10). Esto nos indica que *rmiR*, *rmiAB* y *nodTch* forman un operón. Es interesante observar que el análisis de la secuencia proteica sugiere que RmiR podría ser el regulador de este operón. El hecho de que RmiR forme parte del mismo operón con *rmiABnodTch* puede ser un caso raro pero no único. En general los genes codificando para proteínas reguladoras no están agrupados en el mismo operón con los genes estructurales, sin embargo en *B. japonicum* se reportó que *ragABrpoH3ragCD* forma un solo operón que se transcribe desde el promotor ubicado río arriba de *ragA* (53), y se observó que RagAB son miembros de un sistema regulador de dos componentes, en tanto que RpoH3 es un factor transcripcional tipo sigma32 y RagCD son proteínas con similitud a las familias de transportadores RND y MFP. Además, VarR de *Streptomyces virginiae* muestra homología con los represores de la familia TetR y regula la transcripción de *varS*, que codifica para el transportador específico de virginamicina S. En este caso, *varR* se cotranscribe con *varS* que está ubicado río arriba, formado un transcrito de 2.4 kb (73).

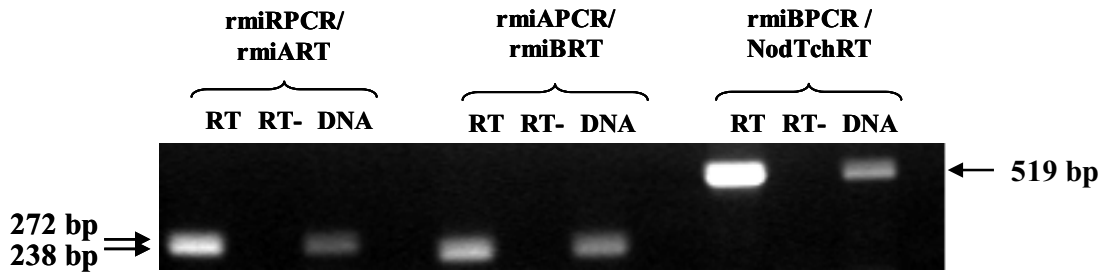


Fig. 10.- Análisis de RT-PCR indican que *rmiR*, *rmiA*, *rmiB* y *nodTch* forman un operón.

Productos de RT-PCR obtenidos con los pares de primeros *rmiR*PCR/*rmiA*RT, *rmiA*PCR/*rmiB*RT y *rmiB*PCR/*nodTch*RT indican la cotranscripción de *rmiR* y *rmiA*, *rmiA* y *rmiB* así como *rmiB* y *nodTch*, respectivamente. Los cDNAs usados como templados fueron obtenidos con reacciones de transcriptasa reversa (RT) usando RNA total de *R. etli* CE3. También se muestran las reacciones control sin RT (RT-) o llevando como templado el plásmido pHME8 con la región completa *rmiRABCnodTch* (DNA). Los números indican los tamaños esperados/observados de los fragmentos obtenidos por PCR en pb.

Previamente obtuvimos mutaciones de los genes *rmiRAB* interrumpiéndolos por separado con un casete de resistencia a Km con elementos omega en cada uno de sus extremos (ver abajo) sin que hubiera un efecto polar sobre la expresión de *nodTch*, ya que estas mutaciones no fueron letales. Esto sugirió que, además de estar formando un operón con los genes *rmiRAB*, *nodTch* podría transcribirse en forma independiente. Para probar esta hipótesis, se hizo un RT-PCR con RNA extraído de la cepa silvestre y de la cepa mutada en *rmiB*, usando como cebadores oligonucleótidos para amplificar la región final de *rmiB* y el principio de *nodTch* o sólo una región correspondiente a *nodTch*. Al usar RNA de la cepa silvestre se puede observar la amplificación de productos de 519 pb y de 174 pb correspondientes a la región ubicada entre *rmiB* y *nodTch* o solamente a la región de *nodTch*, respectivamente, pero en la cepa mutada en *rmiB* se observa un solo producto que corresponde a la región de *nodTch* (Fig. 11a). Esto demuestra que, además de cotranscribirse con *rmiRAB*, *nodTch* puede ser expresado a partir de una región promotora ubicada entre *rmiB* y *nodTch*. Los mismos resultados se observaron con las mutantes en *rmiR* y *rmiA*.

Con el objetivo de corroborar que *nodTch* se puede transcribir en forma independiente de todo el operón, se realizaron ensayos de extensión del cebador. Los resultados obtenidos muestran la presencia de un sitio de inicio de la transcripción para *nodTch* ubicado a 106 pb del codón de inicio propuesto para este gen (Fig. 11b).

Mediante el uso del programa de predicción de secuencias promotoras NNPP se identificó un probable promotor tipo σ^{70} ubicado río arriba del sitio de inicio de la transcripción localizado. Estos resultados indican que *nodTch* tiene una expresión basal independiente del operón completo. Estos datos son congruentes con el hecho de que *nodTch* tuviese una función esencial. El hecho de que *nodTch* pueda cotranscribirse con el operón completo *rmiRABnodTch* sugiere que NodTch podría estar interactuando en la bacteria junto con RmiA y RmiB como un complejo tripartito tipo RND. Sin embargo, el que además pueda transcribirse en forma independiente, abre la posibilidad de que por sí solo tenga una función como proteína de membrana externa sin interactuar con otras proteínas o formando complejos con otras proteínas que no sean tipo RmiAB, tal y como lo hacen TolC de *E. coli* y OprM de *P. aeruginosa* (52, 131, 138).

Por otro lado, se realizaron análisis de RT-PCR con cebadores que amplifican la región intergénica entre el final del ORF2 y el principio de *nodTpc* en el plásmido c. Los datos indican que *nodTpc* no se transcribe junto con los ORF's localizados río arriba, y que muy probablemente esté formando una unidad transcripcional independiente (datos no mostrados). Los análisis de extensión del cebador mostraron que este gen presenta un sitio de inicio de la transcripción localizado 28 bp río arriba del codón de inicio GUG predicho para este gen y que tiene una expresión basal en las condiciones de crecimiento usadas en este trabajo (Fig. 11c). Estos datos indican indirectamente que la expresión de *nodTpc* no complementa la función esencial que se observa para NodTch. Otra posibilidad para explicar la falta de complementación entre *nodTpc* y *nodTch* es que la proteína codificada por este mensajero se esté traduciendo a bajos niveles y no exista la suficiente cantidad de NodTpc para complementar la función esencial que se observa cuando falta NodTch.

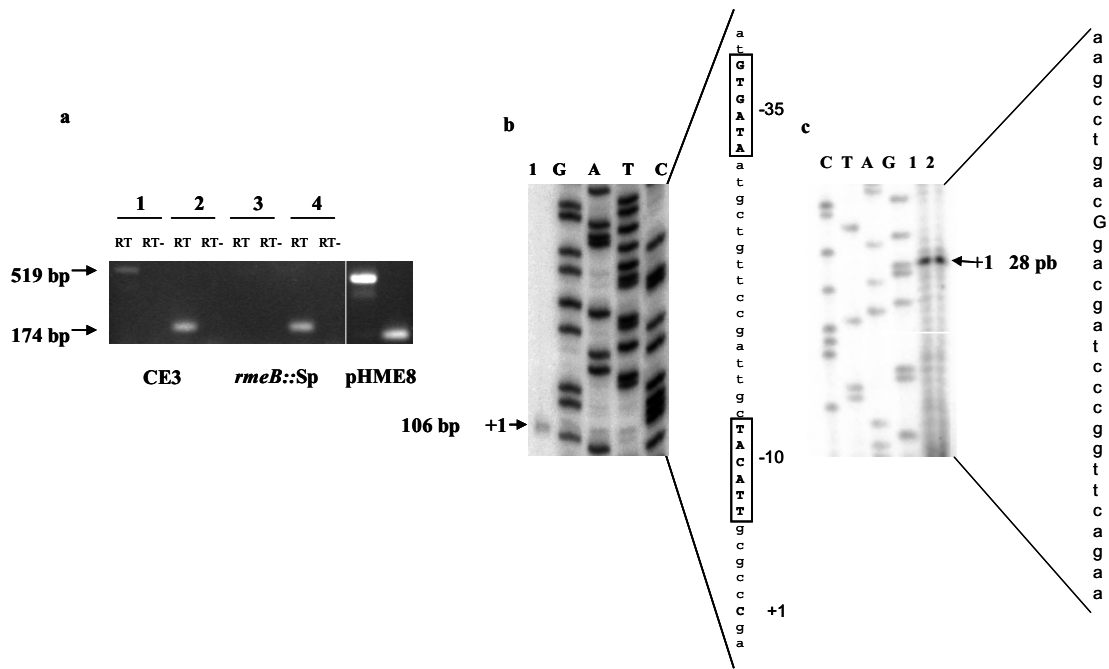


Fig. 11. Los genes *nodTch* y *nodTpc* se transcriben en forma monocistrónica.

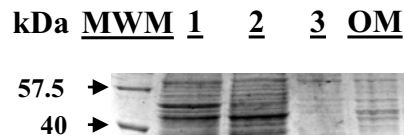
(a) Productos de RT-PCR amplificados usando pares de primeros dentro de la región intergénica entre *rmiB* y *nodT* (1 y 3: *rmiBRT/nodTRT*) y la región río abajo de la posible zona promotora de *nodTch* (2 y 4: *nodTchprPCR/ nodTchRT*). El cDNA fue generado a partir de una reacción de RT usando como templado RNA total de *R. etli* CE3 y una mutante *rmiB::Sp*. También se llevaron a cabo reacciones similares sin RT (RT-), o con el plásmido pHME8 (que lleva la región *rmiRABnodTch*). (b) Análisis de reacciones de extensión del primer (PE) para ubicar la posición del sitio de inicio de la transcripción de *nodTch* a partir de RNA total extraído de *R. etli* CE3 (línea 1). (c) Análisis de reacciones de PE para ubicar la posición del sitio de inicio de la transcripción de *nodTpc* a partir de RNA total extraído de *R. etli* CE3 (línea 1). Líneas G, A, T, y C, corresponden a DNA marcado con ^{32}P de una escalera secuencial de DNA generada con una reacción de PCR con los primeros usados para las reacciones de PE (PE_{nodTch} y PE_{nodTpc}) con DNA del plásmido pHME22 como templado. La secuencia río arriba de ambos genes *nodT* se muestra a la derecha de las figuras B y C. El sitio de inicio de la transcripción se indica con un "+1". Se muestran las propuestas cajas -10 y -35.

5.5 Localización subcelular de NodTch

El análisis que se realizó de la secuencia de aminoácidos indica que las proteínas NodT podrían ser lipoproteínas y por sus características estar localizadas en la membrana externa. Dada la alta identidad de NodTch de *R. etli* con NodT de *R. leguminosarum* bv. *viciae*, decidimos utilizar un anticuerpo generado contra esta última proteína (una

donación del Dr. A. Downie) con el objetivo de identificar a NodTch en *R. etli*. Como antecedente se ha reportado que este anticuerpo reconoce a NodT de *R. leguminosarum* bv. *phaseoli* (99), por lo tanto decidimos usarlo para inmunodetectar a NodT en *R. etli*. Sólo una proteína de 50 kDa con un pI de 6.1, que corresponde a las características predichas para NodTch fue inmunodetectada en extractos totales de *R. etli* de la cepa silvestre y en la cepa CFNX184 que está curada del plásmido c (datos no mostrados) lo que sugiere que el anticuerpo no reconoce a NodTpc o bien que NodTpc tiene un pI fuera del rango de pH que usamos para los geles bidimensionales o que la proteína se expresa a niveles tan bajos que no es inmunodetectada. NodTch también fue detectada en fracciones enriquecidas de proteínas de membrana externa (66, 95, 109) y en la fracción correspondiente a extractos totales de esta bacteria, pero no en las fracciones correspondientes al citosol y a la membrana interna, apoyando la idea de que NodTch es una proteína de membrana externa (Fig. 12a y b). El mismo anticuerpo fue capaz de reconocer a NodTch en cada uno de los extractos de proteínas totales extraídos de las mutantes en *rmiR*, *rmiA* y *rmiB* (Fig. 12c), confirmando que *nodTch* es capaz de transcribirse en forma independiente del operón. Con este anticuerpo también se corroboró que la cepa mutante en *nodTch* complementada con el plásmido pRK415::*nodTch* produce esta proteína a niveles bajos pero suficientes para permitir la viabilidad de *R. etli* y que este plásmido también expresa NodTch en *E. coli* (Fig. 12c).

A



B



C

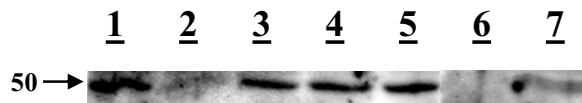


Fig. 12. NodTch de *R. etli* se localiza en la fracción enriquecida de la membrana externa y se expresa en las mutantes nulas *rmiR*::Sp, *rmiA*::Sp and *rmiB*::Sp y en la mutante letal condicionada *nodT*::Km.

(A) Muestras de diferentes fracciones enriquecidas de *R. etli* CE3 fueron analizadas en SDS-PAGE y las proteínas se detectaron por azul brillante de Coomassie. (línea 1), extracto total; (línea 2), fracción soluble; (línea 3) fracción soluble en Tritón X-100; (OM), fracción insoluble en Tritón X-100 [considerada como la fracción enriquecida de proteínas de membrana externa] (66, 95, 109). (B) Análisis por western blot de muestras obtenidas y transferidas como se describe en el panel A usando como anticuerpo anti-NodT generado en conejo (99). (C) Análisis por western blot de extractos de proteínas totales de *R. etli* CE3 (línea 1), CE3/*nodTch*::Km/pRK415::*nodTch* (línea 2); *rmiA*::sp (línea 3); *rmiB*::Sp (línea 4); *rmiR*::Sp (línea 5) usando como anticuerpo, anti-NodT. Se cargaron cantidades equivalentes de proteína total en cada pozo del gel. También se muestran las fracciones de proteína total de *E. coli* llevando el plásmido pRK415::*nodTch* sin inducir (línea 6) o inducido (línea 7) con 10 mM IPTG. MWM: patrones de marcadores de peso molecular para proteínas.

5.6 Las Rhizobiales presentan una menor redundancia genómica de la familia de factores de membrana externa.

Las proteínas NodT son miembros de la familia TolC (112), que pertenecen a la familia de proteínas OMF (88) y han sido clasificadas como miembros de la familia de proteínas de exclusión de membrana externa (Outer membrana Efflux Protein, OEP) por Pfam (10). Con el objetivo de identificar cuántas proteínas OEP existen en *R. etli*, se hizo un análisis de la presencia de esta familia en genomas bacterianos completos, el cual demostró que existe una correlación positiva entre el tamaño del genoma y el número de

genes con dominios OEP en bacterias Gram-negativas (Fig. 13a). En patógenos de plantas y de animales como *Pseudomonas* y *Burkholderia* se observa un gran número de OEP's, lo cual concuerda con su elevada resistencia a compuestos dañinos (110). Sin embargo, las Rhizobiales parecen codificar menos OEPs de los esperados dado el tamaño de sus genomas. Ya que las proteínas OEP pueden interactuar física y funcionalmente con uno o varios miembros de las familias de transportadores ABC, RND, o Major Facilitator Superfamily (MFS) (88, 96, 133, 139), decidimos realizar un análisis del número de proteínas con estos dominios en el conjunto de genomas bacterianos completamente secuenciados que recabamos y descritos en Materiales y Métodos. Los resultados indican que el número de estos transportadores también correlacionan con el tamaño de los genomas (Fig. 13b-c). En el caso de las Rhizobiales, se observa que hay un mayor número de transportadores tipo ABC que en otros organismos y este resultado se acentúa cuando se compara el número total de proteínas con dominios ABC, RND o MFS contra el número de proteínas con dominios OEP (Fig. 13d). Estos resultados podrían implicar que los transportadores ABC, RND, y MFS dan la adaptabilidad de exclusión de drogas necesaria para estos organismos en combinación con una extremadamente versátil (y así esencial) proteína OEP. En apoyo de esta hipótesis, las proteínas OEP han mostrado ser multifuncionales, como se ha observado previamente en *E. coli*, donde la mayoría de los transportadores RND, además de algunos exportadores tipo ABC y MFS, cooperan con TolC (81). Alternativamente, podría ser ventajoso para las Rhizobiales no depender de proteínas de membrana externa para el transporte, dado que en su fase de vida en simbiosis, una vez que penetra la bacteria a las células infectadas en el nódulo, la membrana externa del bacteroide interactúa físicamente con la membrana peribacteroidal (14) y las funciones en el intercambio de metabolitos es compartida por ambas fracciones membranales, aunque se ha observado que la mayoría de las funciones recaen en la membrana peribacteroidal proveniente de la planta (125).

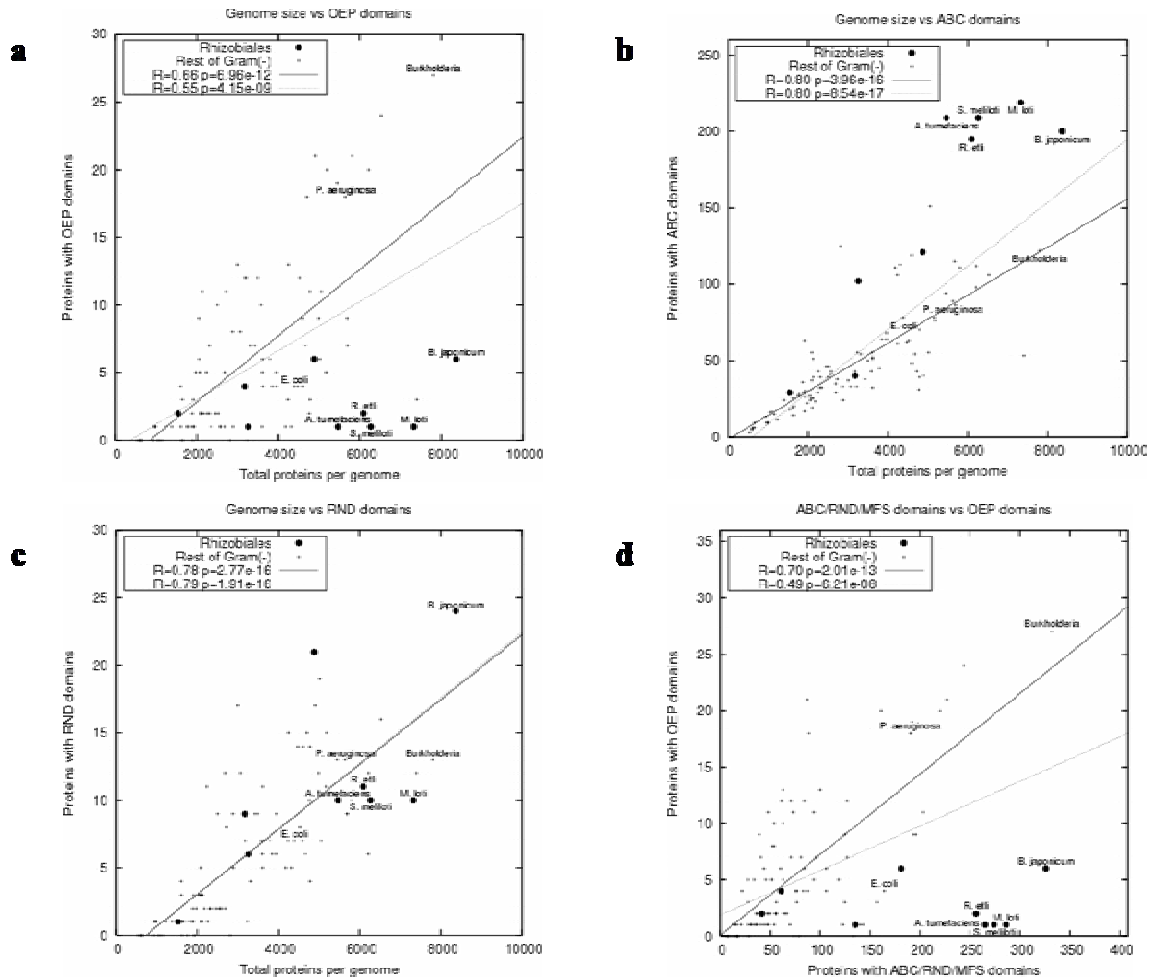


Fig. 13. Análisis del número de proteínas con dominios OEP, ABC, RND Y MFS en bacterias Gram-negativas. El tamaño de genoma correlaciona con el número de duplicaciones de los transportadores en bacterias Gram-negativas excepto para los miembros de la familia OEP en algunas alfa-proteobacterias. El número de proteínas totales codificadas por cada genoma, correlaciona significativamente con el número de proteínas conteniendo dominios OEP (a), ABC (b), y RND (c). Muchas Rhizobiales parecen tener pocas proteínas con dominios OEP (a) y más con dominios ABC (b) de las esperadas dado el tamaño del genoma. Este efecto se amplifica cuando se compara el número de proteínas con dominios ABC, RND o MFS contra aquellas con dominios OEP (d). La línea continua representa la línea de mínimos cuadrados de los datos de organismos Gram-negativos, excluyendo Rhizobiales, mientras que las líneas punteadas los incluyen. R es la correlación lineal de Pearson y p representa la probabilidad de obtener una buena correlación al azar.

En resumen, presentamos evidencia experimental que muestra la existencia de dos homólogos a *nodT* en *R. etli*, ambos localizados fuera del plásmido simbiótico: *nodTch* y *nodTpc*. *nodTch* fue localizado en el cromosoma de *R. etli* y presenta una función esencial para la bacteria en las condiciones de laboratorio analizadas. Además, este gen está formando parte de un nuevo operón que tiene identidad con las bombas para la exclusión de drogas tipo RND, aunque *nodTch* puede transcribirse en forma

independiente al operón completo. El hecho de que NodTch esté localizada en la membrana externa y de que pertenece a la superfamilia de proteínas de membrana externa TolC, nos permite sugerir que NodTch tiene un papel en la estabilidad de la membrana externa o en procesos más generales como la segregación de cromosomas en lugar de, o además de participar en el proceso de exclusión de drogas. Con respecto al gen localizado en el plásmido *c*, *nodTpc*, presenta una transcripción basal en nuestras condiciones de crecimiento, aunque no detectamos NodTpc con el anticuerpo que reconoce a NodTch. Mutaciones en este gen no presentaron ningún fenotipo en el proceso de nodulación o en el proceso de exclusión de drogas. Para definir el papel que estos dos genes tienen en *R. etli* es necesario obtener una doble mutante viable *nodTpc/nodTch*. Por otro lado, dado que nunca observamos cambios en los niveles basales de transcripción de estos genes, se vuelve importante clonar ambos genes en vectores que se repliquen y puedan expresen proteínas en *R. etli* para observar los fenotipos de una inducción o represión transcripcional de estos genes.

5.7 Caracterización de las mutantes *rmiRAB* de *R. etli*

Para analizar la función de los genes *rmiRAB* en *R. etli*, cada uno de ellos fue interrumpido por un interposon Ω -Sp (4). Para investigar el posible papel de RmiR, RmiA y RmiB en la exclusión de diversas drogas, se ensayaron las sensibilidades de estas mutantes a diferentes compuestos tóxicos (incluyendo detergentes, antibióticos, solventes y compuestos fenólicos de plantas). Los resultados obtenidos no mostraron diferencias significativas con respecto a la cepa silvestre. Tampoco observamos un efecto significativo de estas mutantes en el proceso de nodulación dado que el número de nódulos por planta, así como la morfología y las cinéticas de nodulación fueron similares en estas tres mutantes en comparación con la cepa silvestre.

En *A. tumefaciens*, existen dos sistemas de exclusión de drogas del tipo RND, que están codificados por los genes *ifeAB* (87) y por *ameRameABC* (89). En este último sistema, AmeC muestra similitud a NodT. Mutaciones en *ameC* incrementan la sensibilidad a detergentes y antibióticos, pero las mutaciones en *ameA* y *ameB* no presentan un fenotipo en este proceso (89). En forma similar, mutaciones en *rmiA* y *rmiB* tampoco tienen un fenotipo en la resistencia a drogas. Una posibilidad, es que este tipo de

proteínas esté expulsando algún compuesto muy específico de la interacción planta microorganismo que nosotros no incluimos en nuestro análisis o que no está presente en nuestros ensayos de nodulación llevados a cabo en maceta, pero que si estén presentes en el suelo. Otra posibilidad, es que exista redundancia genética. Por ejemplo, en bacterias Gram-negativas tales como *E. coli*, *P. aeruginosa* o *P. putida* (78, 92, 98) existen varios sistemas de exclusión. Es posible que en *R. etli*, otros sistemas de exclusión tipo RND o bombas de exclusión de drogas diversas puedan compensar la falta de expresión de *rmiRAB*. En *R. etli* se ha reportado la presencia de un sistema de exclusión de drogas llamado *rmrA* y *rmrB* (34). Mutaciones en estos genes en comparación con la cepa silvestre aumentan la sensibilidad a fitoalexinas, flavonoides y ácido salicílico (34). Es posible que el fenotipo de las mutantes *rmiAB* sea enmascarado por sistemas similares a *rmrAB*.

5.8 Análisis de los aminoácidos de las proteínas de la familia RND que interaccionan con distintos ligandos.

Dado que no fue posible identificar un fenotipo de los genes ubicados río arriba de *nodTch* a pesar de que se lograron obtener mutantes en ellos, se procedió a un análisis *in silico* que nos permitiera poder predecir el tipo de ligandos que interactúan con estas proteínas. Para lograr este objetivo se tomaron en cuenta varios aspectos que se describen a continuación. La ausencia de un fenotipo para RmiB puede deberse a que exista redundancia genética y que otro tipo de sistema de expulsión de drogas compense la ausencia de RmiB aunque también es posible que no hayamos seleccionado el compuesto adecuado para observar algún efecto. El análisis de la secuencia proteica de RmiB sugiere que pertenece a la familia de transportadores tipo RND. Para poder investigar qué tanta redundancia presentan estas proteínas en las bacterias, realizamos un análisis exhaustivo de secuencias en los genomas bacterianos con la ayuda del programa GeConT (21) y localizamos un promedio de seis sistemas de exclusión de drogas por genoma analizado, lo que indica que existe redundancia en estos sistemas estudiados. Por ejemplo, tanto en *P. aeruginosa* como en *P. putida* se han localizado doce copias de estos sistemas de exclusión de drogas, de los cuales sólo se han caracterizado funcionalmente siete en cada organismo (92, 94, 98), mientras que en *E. coli* K12, se ha propuesto la presencia de siete

transportadores basados en el análisis de la secuencia y la mayoría han sido completamente caracterizados (79). En el caso de *A. tumefaciens* se han identificado ocho probables bombas de exclusión del tipo RND, pero sólo dos de ellas han sido analizadas experimentalmente (87, 89). Los homólogos de AcrB de *E. coli* están organizados en su mayoría en operones con genes de la familia de proteínas de fusión a membrana (MFP, por sus siglas en inglés) (27), y ocasionalmente con genes que están codificados por proteínas o lipoproteínas de membrana externa de la familia de los factores de membrana externa (OMF, por sus siglas en inglés) (88) (Fig. 14) .

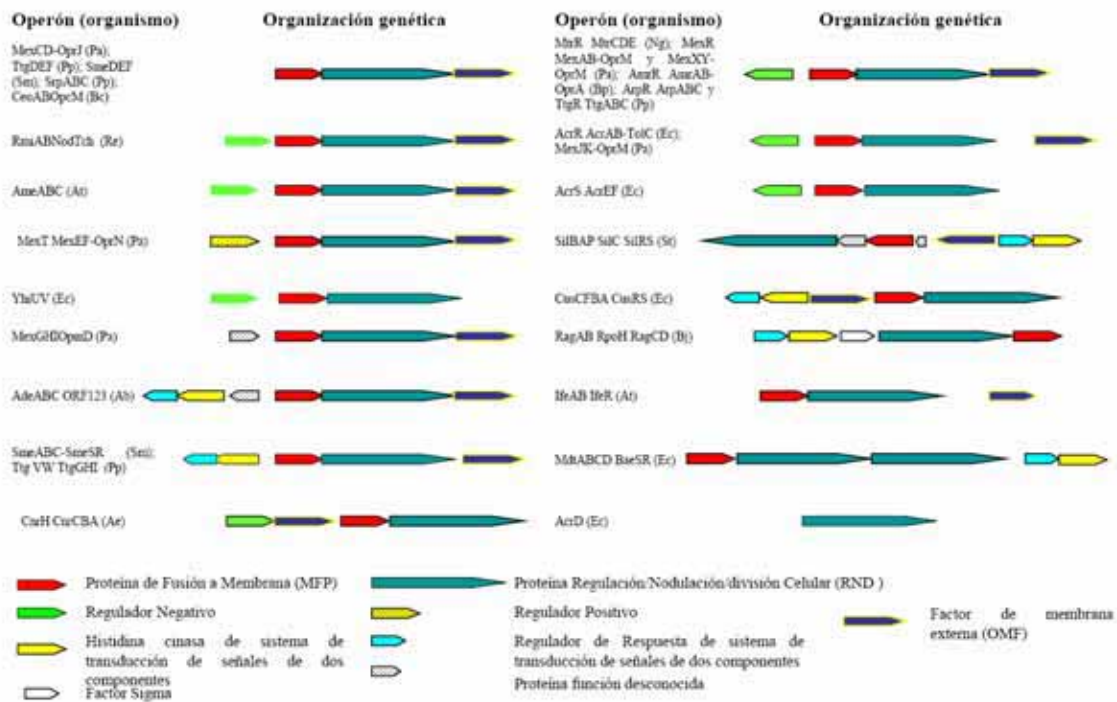


Fig. 14. Organización genética de los sistemas de transporte tipo RND.

Se muestran los reguladores y los componentes estructurales de los sistemas de exclusión de drogas más representativos de bacterias Gram-negativas. Las proteínas pertenecientes a la misma familia se representan con la misma flecha. Aunque el gen TolC no forma parte de ninguna de los sistemas de exclusión de múltiples drogas, la porina TolC se asocia a los sistemas de bombeo AcrAB, AcrEF, EmrKY, YhiUV y MdtABC de *E. coli*. De la misma forma, la porina OprM codificada por el operón mexABoprM de *P. aeruginosa* también se asocia con los sistemas de bombeo MexXY y MexJK. Los microorganismos se abrevian como sigue: Ab: *Acinetobacter baumannii*; Ae: *Alcaligenes eutrophus*; At: *Agrobacterium tumefaciens*; Bc: *Burkholderia cepacia*; Bj: *Bradyrhizobium japonicum*; Bp: *Burkholderia pseudomallei*; Ec: *Escherichia coli*; Ng: *Neisseria gonorrhoeae*; Pa *Pseudomonas aeruginosa*; Pp: *Pseudomonas putida*; Re: *Rhizobium etli*; Sm: *Stenotrophomonas maltophilia*; St: *Salmonella typhimurim*.

Esto sugiere que en *R. etli* podría haber más sistemas tipo RND. Aplicando el programa GeconT al genoma de *R. etli*, identificamos 11 sistemas tipo RND, 9 en cromosoma, 1 en el plásmido **c** y otro en el plásmido **f** (Fig. 15). Por otro lado, los análisis filogenéticos de esta familia indican que el patrón de agrupamiento correlaciona con especificidad por el sustrato expulsado. Por ejemplo, algunas proteínas que pertenecen a la familia 1 exportan metales pesados, mientras que las de la familia 2 expulsan drogas diversas (124). Se realizó un análisis filogenético usando estas proteínas de *R. etli* y se comparó con los sistemas de transporte tipo RND. Como resultado de este análisis RmiB se agrupó junto a proteínas que no tienen una función previamente descrita.



Fig. 15. Redundancia genética de los transportadores RND en *R. etli*. Se observa la presencia de 11 bombas de exclusión de drogas tipo RND (rojo) acompañadas de genes que codifican para proteínas tipo MFP (amarillo). Sólo en 2 operones se observó la presencia de genes de la familia OMF (azul) que corresponden a los dos únicos genes *nodT* reportados en este trabajo.

Existen reportes que indican que las proteínas homólogas comparten una función biológica común, pero diferente especificidad hacia los sustratos, ligandos, efectores, DNA, proteínas u otras moléculas con las cuales interactúan. Esas interacciones deben ser altamente específicas, y la identificación de residuos que participan en esa especificidad podría ser útil para determinar su preferencia por cierto tipo de compuestos (39, 48, 55, 58, 62, 80, 82, 118). Por otro lado, existe una correlación entre los aminoácidos que determinan la especificidad por un ligando y la posición que éstos ocupan dentro de un alineamiento múltiple de secuencias, donde la distribución de aminoácidos también

puede estar asociada con un agrupamiento por especificidad. En base a esto, se han diseñado una serie de algoritmos para resolver el problema de la predicción de ligandos, como el método de Traza evolutiva que explota la información acerca de las estructuras de las proteínas y de los sitios funcionales (62), mientras que otros métodos usan sólo la secuencia de proteínas (12).

En esta parte del trabajo, intentamos predecir cuál es el ligando más probable para las proteínas del tipo RND y ver si de esa manera se podría asignar una función para RmiB. Si nuestra hipótesis es correcta, es posible suponer que un ancestro genérico del dominio de la cavidad central de las proteínas RND ha divergido para expulsar diferente tipo de sustratos y que los residuos donde se lleva a cabo la unión a ligandos se han conservado en la evolución de cada grupo de proteínas con un sustrato común. Nuestra predicción se basa en la extracción de 14 residuos, a partir de un alineamiento múltiple, involucrados en interacciones con ligandos en una de las proteínas del tipo RND. Estos residuos están localizados en el poro central y en la cavidad central que se forma en el homotrímero funcional de la proteína AcrB de *E. coli* (Fig. 16) (70, 137). Además de que su estructura cristalina ya ha sido resuelta, AcrB tiene la ventaja de que exhibe una amplia variedad de especificidad de sustratos: antibióticos, desinfectantes, colorantes, detergentes y solventes (79).

Los residuos se alinearon manualmente y se reagruparon usando el algoritmo de máxima parsimonia implementado en el paquete Phylip. Los grupos de proteínas presentaban al menos un sustrato en común. Nuestros resultados podrían ser utilizados para predecir con qué ligandos van a interactuar aquellas proteínas que aún no han sido funcionalmente caracterizadas o predecir cuál es el ligando más probable para aquellas proteínas codificadas por ORFs en los genomas anotados y que aún no han sido estudiadas.

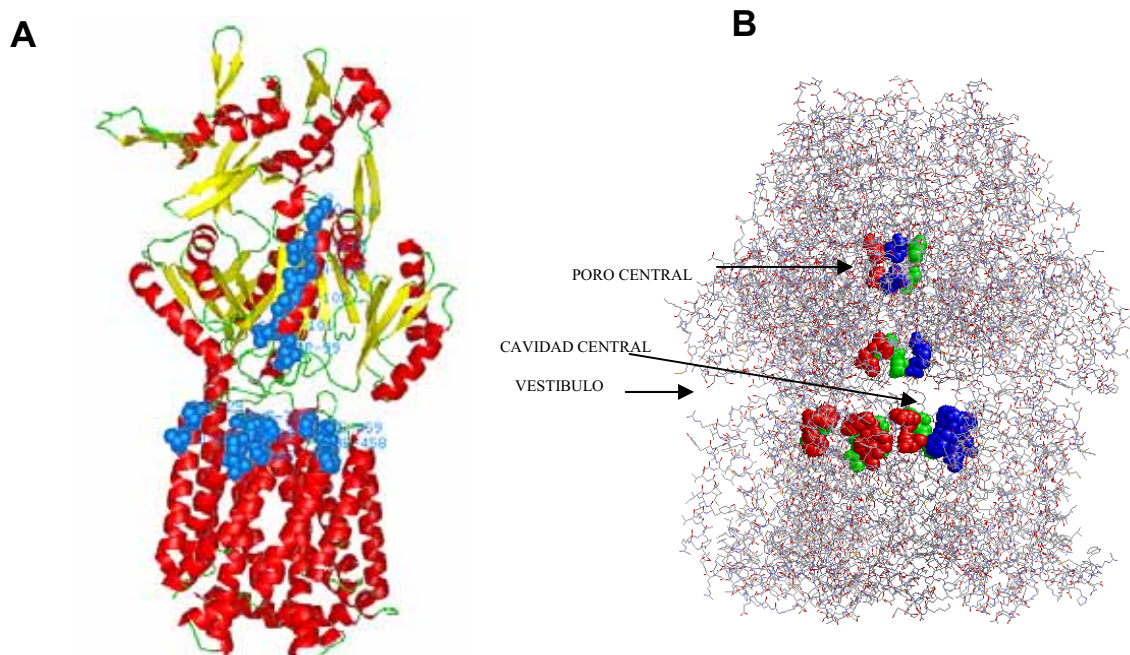


Fig. 16.- Mapeo de los residuos seleccionados en la estructura de AcrB

A) Diagrama de cinta generado con PyMol de la bomba de exclusión de diversas drogas AcrB de *E. coli* (reproducido a partir del archivo 1IWG.PDB) que muestra los residuos seleccionados para este estudio: L (25), K (29), D (99), D (101), V (382), A (385), F (386), F (388), F (458) y F (459) que son los residuos que se han reportado que pueden interactuar con ligandos (137) y V(105), N (109), Q (112), así como P (116), son los residuos que forman parte del poro central extramembranal de AcrB, el cual tiene un papel importante en el transporte de drogas. B) Diagrama de residuos con PyMol de dos subunidades de AcrB para indicar la localización de los residuos seleccionados dentro de la región de la cavidad central. La tercera subunidad fue eliminada para tener una mayor representación gráfica.

5.9 Selección de las proteínas de la superfamilia RND

Para seleccionar a miembros de la familia RND usando a AcrB como semilla, se realizó una búsqueda por Psi-BLAST en la base de datos del NCBI. Los homólogos de AcrB están ampliamente distribuidos en bacterias, y confieren una tolerancia intrínseca a diferentes drogas en los microorganismos. Con esta metodología se obtuvieron 516 proteínas, pero sólo se seleccionaron 43, en base a que han sido previamente caracterizadas funcionalmente y se conoce la especificidad o la preferencia que cada una de ellas tiene por cierto tipo de sustrato. Además, seleccionamos 4 proteínas que no están completamente caracterizadas pero que son de nuestro interés, por ser proteínas

ubicadas en el grupo de bacterias de Rhizobia. Con estas secuencias, realizamos un alineamiento múltiple (Apéndice 1). Las proteínas seleccionadas presentan un nivel de identidad del 40% en promedio entre ellas y una longitud de 1190 aa en promedio. Dentro de los compuestos que pueden ser expulsados por algunas de estas proteínas se encuentran hidrocarburos, antibióticos, detergentes y colorantes. Algunas proteínas pueden expulsar un solo tipo de compuestos y otras tienen un espectro más amplio de especificidad.

5.10 Análisis de los grupos formados por el alineamiento de los 14 residuos seleccionados

Los residuos que se encuentran localizados en la región de la cavidad central de AcrB, que es donde se propone que se acumulan las drogas que serán expulsadas, se encuentran distribuidos a lo largo de toda la secuencia, de tal forma que no hay una secuencia en particular asociada con una función de unión. A partir de un alineamiento de múltiples secuencias (MSA) donde se incluyeron las cuarenta y cuatro proteínas tipo RND seleccionadas previamente, se escogieron catorce columnas que contenían información sobre la unión a ligandos. Uno de los resultados más interesantes que obtuvimos fue que estas catorce columnas, extraídas del alineamiento de alrededor de mil doscientas, contienen la suficiente información para predecir cuál es el ligando más probable asociado a una proteína en particular. También realizamos un análisis generando logotipos de secuencias con el programa WebLogo (22) para computar el contenido de la información de las 14 columnas con residuos separándolas en los grupos que observamos y podemos observar la conservación de los residuos seleccionados una vez aislados del alineamiento total múltiple (Apéndice 2). Como un control del análisis, también se realizaron ensayos seleccionando 14 columnas en posiciones independientes localizadas al azar para reconstruir los grupos observados previamente, sin embargo no se observaron los resultados obtenidos previamente, lo cual descarta que nuestras observaciones previas hayan sido producidas al azar y confirman que existe una correlación entre los residuos que interaccionan con ciertos ligandos y su localización en el sitio de unión. Basados en esta metodología, nosotros identificamos al menos quince grupos de proteínas en los

cuales la mayoría de sus miembros comparten un sustrato común, de los cuales describiremos a continuación algunos de ellos:

GRUPO AcrB (Amplio espectro de selección de compuestos).

El grupo que incluye a AcrB de *E. coli* y *Salmonella enterica* serovar Typhimurium y MexB de *P. aeruginosa*, exhibe una especificidad amplia por compuestos tales como colorantes, detergentes y antibióticos (54, 59, 72) (Fig. 17). Como se predijo, en este grupo se conserva un alto número de residuos con respecto a la proteína seleccionada como semilla para el Psi-Blast (AcrB de *E. coli*) y sólo un residuo, K29 es variable (Tabla 2). Algunas de las proteínas que se localizan en este grupo (AcrB de *E. coli* y MexB de *P. aeruginosa*) son capaces de expulsar, además de antibióticos o detergentes, compuestos como hidrocarburos aromáticos (HA) con un bajo coeficiente de partición (e. gr. benceno) (8, 60). También se ha observado que AcrB de *S. enterica* serovar Typhimurium y MexB podrían tener un papel importante en el proceso de infección celular (41, 54). Nuestro análisis nos permite predecir que es muy probable que AcrB de *S. enterica* serovar Typhimurium esté involucrada en la resistencia a cierto tipo de solventes.

GRUPO TbtB (Expulsión de solventes).

Las proteínas TbtB, TtgH, SrpB, SepB, y TtgE se localizan en un mismo grupo. La mayoría de estas proteínas están involucradas en reducir las concentraciones intracelulares de HA (90, 102, 103, 132). Sin embargo, TbtB de *P. stutzeri* es una proteína que se encarga del transporte de Tributiltina (TBT) (45), que proviene de compuestos de organotina (Fig. 17). La TBT se usa como fungicida y alguicida y se utiliza en algunas pinturas para evitar las adherencias de algas y otros organismos marinos al casco de las embarcaciones. De este grupo, se conservan once residuos con respecto al grupo de alta especificidad de ArcB, pero se observan cambios en los residuos V382L, A385Y, y F388L/I (Tabla 2). Esto sugiere que esos cambios son fundamentales en la selectividad del ligando, mientras que el resto podría reconocer sustratos similares para expulsarlos. Nuevamente, se pone de manifiesto la capacidad de predicción de nuestro método, porque varias proteínas que exhiben un patrón de conservación de

residuos se agrupan en un mismo clado y presentan la misma preferencia por un sustrato. Esto, nos permite proponer que TbtB tiene un papel en la expulsión de HA, hecho que no ha sido explorado aún.

GRUPO CnrA (Expulsión de metales pesados)

En este grupo de proteínas se encuentran CnrA, CzcA, CusA, SilA, y MexK. La mayoría de estas proteínas se unen a ligandos similares, que en este caso son metales pesados (63, 69, 75), a excepción de MexK que está involucrada en la resistencia a tetraciclina y a una droga anfifílica biocida de amplio espectro que es el Triclosán (24)(Fig. 17). Haciendo el agrupamiento de acuerdo con la selección de los catorce residuos de interacción con ligandos, encontramos que ningún residuo está universalmente conservado en este grupo (Tabla 2). Una pregunta interesante es ¿por qué MexK, que no ha sido reportada que tenga un papel en la resistencia a metales pesados, se incluye en el grupo de proteínas que transporta metales pesados? es probable que dichas proteínas tengan una estructura con características fisicoquímicas en la cavidad central que permita la interacción con moléculas con tamaños o propiedades muy similares.

GRUPO ORF2pc (Expulsión de fluoroquinolonas)

Las proteínas ORF2pc, AdeB y MexD es otro ejemplo de cómo podemos usar este método para hacer una predicción. AdeB y MexD son capaces de expulsar principalmente fluoroquinolonas (36, 64, 93) (Fig. 17). En este grupo, la comparación de las secuencias generadas con los catorce residuos indica que son cinco residuos los que se conservan (Tabla 2). Esas posiciones pueden estar jugando un papel importante para seleccionar a las fluoroquinolonas sobre el resto de compuestos. El gen que codifica para la proteína ORF2pc fue identificada en nuestro laboratorio sólo por análisis de la secuencia. Hasta la fecha no existe evidencia experimental sobre su participación en la expulsión de un compuesto en particular. En base a este análisis, es posible proponer que esta proteína puede interactuar con compuestos como fluoroquinolonas o con características fisicoquímicas y/o estructurales similares a este tipo de compuestos.

GRUPO AmeB (Sustrato desconocido)

A este clado pertenecen RagC, RmiB, AmeB y ORF2 de *R. leguminosarum* bv. *viciae* (38, 53, 89). Todas estas proteínas están localizadas en bacterias de la familia Rhizobiaceae y ninguna se ha caracterizado funcionalmente (Fig. 17). RagC ha sido probada en la resistencia a metales pesados y a compuestos orgánicos, así como en el proceso de simbiótico que se establece entre *Bradyrhizobium* y leguminosas sin observar un efecto significativo (53). AmeB de *A. tumefaciens* y RmiB de *R. etli* no presentan un efecto claro en la resistencia a varios compuestos probados (53). La proteína RagC es la que exhibe mayor diversidad con respecto al resto de las proteínas de este grupo. Sin embargo, las proteínas que han conservado más residuos entre ellas (RmiB, AmeB y ORF2) presentan más cambios con respecto al grupo de AcrB. Uno de esos cambios es el de la prolina 116 por una serina (Tabla 2). Este cambio ocurre en el poro central cuya apertura podría estar alterada por el tamaño del nuevo residuo. Es interesante que todas las proteínas incluidas en este grupo sean del grupo de las Rhizobiales y es posible especular que los residuos seleccionados están muy conservados en este grupo debido a que comparten un sustrato común presente en las interacciones planta-microorganismo, lo cual explicaría el por qué ninguno de los compuesto que se han probado hasta ahora presenta un efecto sobre estas proteínas y hace sentido con el hecho de que tampoco fuimos capaces de encontrar un compuesto que tuviera algún efecto en las mutantes *rmiAB*.

Finalmente, se ha propuesto teóricamente que NolG es una proteína que exporta lipoquitooligosacáridos (FNs Rhizobiales) (105) y por lo mismo se dice que esta superfamilia de transportadores tiene un papel en el proceso de nodulación. Sin embargo, basados en nuestro análisis, proponemos que su papel principal es la expulsión de antibióticos del tipo de la fluoroquinolonas más que en el proceso de la simbiosis *Rhizobium*-leguminosa.

En resumen, en las 44 proteínas RND analizadas identificamos catorce residuos que pueden interactuar con diversos sustratos y que extrajimos de un alineamiento múltiple que comprendió alrededor de 1300 columnas. Estos residuos son lo suficientemente informativos para predecir cuál es su ligando más probable para un grupo en particular. Basados en esta metodología, se definieron diversos grupos con una

especificidad por un cierto ligando. Con estos resultados obtenidos *in silico*, proponemos que los sustratos expulsados por las proteínas de la familia RND son seleccionados dependiendo de los residuos que se localizan en su cavidad central y a lo largo del poro central. Esto no excluye la posibilidad de que los residuos que se localizan en la entrada de los vestíbulos, o los residuos que forman parte del poro que se proyecta hacia el citoplasma en la parte baja de la estructura, tengan un papel en la selección de los compuestos que van a llegar hasta la región de la cavidad central, tal y como ha sido propuesto con anterioridad (136). A pesar de la utilidad de esta herramienta, el papel de RmiB de *R. etli* permanece sin determinarse.

Para terminar parte de los resultados descritos en esta tesis fueron publicados en el artículo “**Diminished Redundancy of Outer Membrane Factor Proteins in Rhizobiales: a *nodT* Homolog is Essential for Free Living *Rhizobium etli***” y que actualmente se encuentra en prensa en Journal of Molecular Microbiology and Biotechnology Vol. 11, No. 6, 2006 (Apéndice 3).

6. CONCLUSIONES

- 1.- En el cromosoma de *R. etli* CFN42 se identificó un nuevo operón al que hemos denominado *rmiRABnodTch*. En el plásmido c de este mismo organismo se identificó otro homólogo de *nodT* al que se le dio el nombre de *nodTpc*.
- 2.- El gen *nodTch* se comporta como un gen esencial en las condiciones de crecimiento que utilizamos para cultivar a *R. etli* en el laboratorio.
- 3.- Además de cotranscribirse como parte del operón, *nodTch* también se transcribe independientemente del operón a partir de un promotor cuyo inicio de la transcripción se localiza 106 pb río arriba del codón de inicio.
- 4.- La familia de proteínas OMF presenta una correlación positiva en bacterias Gram negativas, ya que a mayor tamaño del genoma aumenta el número de copias presentes. Esta correlación no se mantiene en Rhizobiales, por que el número de proteínas OMF codificadas en estos genomas es sensiblemente menor.
- 5.- NodTch se localiza en la fracción de la membrana externa en concordancia con las predicciones hechas *in silico*.
- 6.- Las mutaciones obtenidas en los genes *rmiR*, *rmiA* y *rmiB*, no presentan un fenotipo claro en las exclusiones de las diversas drogas analizadas (antibióticos, detergentes, colorantes y solventes) o en el proceso de nodulación.
- 7.- *nodTpc* tiene un posible inicio de la transcripción ubicado 26 pb río arriba del codón de inicio y no se cotranscribe con otros genes, lo que sugiere que se transcribe como una unidad transcripcional independiente.
- 8.- La mutación en *nodTpc* no tiene un fenotipo alterado en nodulación, en resistencia a estrés calórico o en presencia de etanol.
- 9.- En diversos aislados geográficos de *R. etli* y en *R. tropici* se identificaron posibles homólogos a los genes *nodTch* y *nodTpc*
- 10.- Existe una correlación entre la filogenia de las proteínas transportadoras tipo RND y el sustrato que expulsan, la correlación se mantiene si sólo se toman en cuenta los residuos ubicados en el interior de la cavidad central y del poro central que se forma en el homotrímero funcional RND y que están involucrados en la interacción con los ligandos que expulsan. Gracias a esto, es posible definir grupos de proteínas RND que tienen preferencia por sustratos similares.

7. PERSPECTIVAS

- 1.- Generar una mutante letal condicionada para *nodTch*.
- 2.- Si es posible generar una mutante en *nodTch*, intentar obtener una doble mutante *nodTch/nodTpc*
- 3.- Generar una construcción que pueda regular la expresión del gen *nodTch* en un vector regulable en *R. etli* para poder controlar la cantidad de transcrito y al mismo tiempo de proteína. La sobreproducción o disminución de la concentración intracelular de NodTch podría sugerir un fenotipo.
- 4.- Sobre expresar y purificar las proteínas NodT para acoplarlas a membranas lipídicas con el fin de determinar si su papel pudiera ser el de un canal, tal y como se ha observado en el caso de OprM y TolC (7, 134).
- 5.- Con respecto a RmiRAB, ampliar la gama de compuestos a analizar para intentar identificar nuevos ligandos que interactúen con estas proteínas enfocándose a la búsqueda de nuevos compuestos que posiblemente estén presentes en la rizósfera, ya que las proteínas de la familia RND con las que RmiB tiene mayor relación se encuentran presentes en el grupo de las Rhizobiales.
- 6.- Purificar y cristalizar la proteína RmiB para posteriormente llevar a cabo una metodología de tamizado computacional en masa con diferentes compuestos químicos para identificar con cuáles pudiera estar interactuando usando programas especiales como BioCoRE (<http://www.ks.uiuc.edu/Research/biocore/>) o DrugSite (6).
- 7.- La presencia de 11 sistemas tipo RND hace también necesario la generación de mutantes múltiples que permitan determinar si existe complementación funcional entre los sistemas localizados en *R. etli*.
- 8.- Tratamientos directos de células intactas de *R. etli* en presencia de formaldehído podrían en teoría lograr un entrecruzamiento estable de NodTch con las proteínas con las que pudiera estar interactuando. Posteriormente se pueden purificar estos complejos con anticuerpos mediante ensayos de “pulldowns”.
- 9.- Para definir si RmiR es un activador transcripcional se pueden hacer ensayos con fusiones transcripcionales de las posibles regiones reguladoras a genes reporteros, o RT-PCR comparando los análisis que se obtengan de *R. etli* silvestre y la mutante en *rmiR*.

Posteriormente se puede purificar la proteína RmiR y hacer análisis de retardo en gel para identificar si actúa como una proteína con propiedades de unión a DNA.

Posteriormente se puede purificar la proteína RmiR y hacer análisis de retardo en gel para identificar si actúa como una proteína con propiedades de unión a DNA.

8. BIBLIOGRAFIA

1. **Aendekerk, S., S. P. Diggle, Z. Song, N. Hoiby, P. Cornelis, P. Williams, and M. Camara.** 2005. The MexGHI-OpmD multidrug efflux pump controls growth, antibiotic susceptibility and virulence in *Pseudomonas aeruginosa* via 4-quinolone-dependent cell-to-cell communication. *Microbiology* **151**:1113-25.
2. **Aendekerk, S., B. Ghysels, P. Cornelis, and C. Baysse.** 2002. Characterization of a new efflux pump, MexGHI-OpmD, from *Pseudomonas aeruginosa* that confers resistance to vanadium. *Microbiology* **148**:2371-81.
3. **Akerley, B. J., E. J. Rubin, V. L. Novick, K. Amaya, N. Judson, and J. J. Mekalanos.** 2002. A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. *Proc Natl Acad Sci U S A* **99**:966-71.
4. **Alexeyev, M. F., I. N. Shokolenko, and T. P. Croughan.** 1995. Improved antibiotic-resistance gene cassettes and omega elements for *Escherichia coli* vector construction and in vitro deletion/insertion mutagenesis. *Gene* **160**:63-7.
5. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**:3389-402.
6. **An, J., M. Totrov, and R. Abagyan.** 2004. Comprehensive identification of "druggable" protein ligand binding sites. *Genome Inform Ser* **15**:31-41.
7. **Andersen, C., C. Hughes, and V. Koronakis.** 2002. Electrophysiological behavior of the TolC channel-tunnel in planar lipid bilayers. *J Membr Biol* **185**:83-92.
8. **Aono, R.** 1998. Improvement of organic solvent tolerance level of *Escherichia coli* by overexpression of stress-responsive genes. *Extremophiles* **2**:239-48.
9. **Baron, C., and P. C. Zambryski.** 1995. The plant response in pathogenesis, symbiosis, and wounding: variations on a common theme? *Annu Rev Genet* **29**:107-29.
10. **Bateman, A., L. Coin, R. Durbin, R. D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E. L. Sonnhammer, D. J. Studholme, C. Yeats, and S. R. Eddy.** 2004. The Pfam protein families database. *Nucleic Acids Res* **32**:D138-41.
11. **Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler.** 2005. GenBank. *Nucleic Acids Res* **33**:D34-8.
12. **Berezin, C., F. Glaser, J. Rosenberg, I. Paz, T. Pupko, P. Fariselli, R. Casadio, and N. Ben-Tal.** 2004. ConSeq: the identification of functionally and structurally important residues in protein sequences. *Bioinformatics* **20**:1322-4.
13. **Bolhuis, H., G. Poelarends, H. W. van Veen, B. Poolman, A. J. Driessen, and W. N. Konings.** 1995. The Lactococcal *lmrP* gene encodes a proton motive force-dependent drug transporter. *J Biol Chem* **270**:26092-8.

14. **Bradley, D. J., G. W. Butcher, G. Galfre, E. A. Wood, and N. J. Brewin.** 1986. Physical association between the peribacteroid membrane and lipopolysaccharide from the bacteroid outer membrane in *Rhizobium*-infected pea root nodule cells. *J Cell Sci* **85**:47-61.
15. **Brendel, V.** 1985. Mapping of transcription terminators of bacteriophages phi X174 and G4 by sequence analysis. *J Virol* **53**:340-2.
16. **Brendel, V., and E. N. Trifonov.** 1984. A computer algorithm for testing potential prokaryotic terminators. *Nucleic Acids Res* **12**:4411-27.
17. **Brom, S., A. Garcia de los Santos, T. Stepkowsky, M. Flores, G. Davila, D. Romero, and R. Palacios.** 1992. Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. *J Bacteriol* **174**:5183-9.
18. **Brown, M. H., I. T. Paulsen, and R. A. Skurray.** 1999. The multidrug efflux protein NorM is a prototype of a new family of transporters. *Mol Microbiol* **31**:394-5.
19. **Cardenas, L., J. Dominguez, C. Quinto, I. M. Lopez-Lara, B. J. Lugtenberg, H. P. Spaink, G. J. Rademaker, J. Haverkamp, and J. E. Thomas-Oates.** 1995. Isolation, chemical structures and biological activity of the lipo-chitin oligosaccharide nodulation signals from *Rhizobium etli*. *Plant Mol Biol* **29**:453-64.
20. **Cardenas, L., J. Dominguez, O. Santana, and C. Quinto.** 1996. The role of the *nodI* and *nodJ* genes in the transport of Nod metabolites in *Rhizobium etli*. *Gene* **173**:183-7.
21. **Ciria, R., C. Abreu-Goodger, E. Morett, and E. Merino.** 2004. GeConT: gene context analysis. *Bioinformatics* **20**:2307-8.
22. **Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner.** 2004. WebLogo: a sequence logo generator. *Genome Res* **14**:1188-90.
23. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. *Anal Biochem.* **162**:156-9.
24. **Chuanchuen, R., C. T. Narasaki, and H. P. Schweizer.** 2002. The MexJK efflux pump of *Pseudomonas aeruginosa* requires OprM for antibiotic efflux but not for efflux of triclosan. *J Bacteriol* **184**:5036-44.
25. **Danese, P. N., W. B. Snyder, C. L. Cosma, L. J. Davis, and T. J. Silhavy.** 1995. The Cpx two-component signal transduction pathway of *Escherichia coli* regulates transcription of the gene specifying the stress-inducible periplasmic protease, DegP. *Genes Dev* **9**:387-98.
26. **Dean, M., A. Rzhetsky, and R. Allikmets.** 2001. The human ATP-binding cassette (ABC) transporter superfamily. *Genome Res* **11**:1156-66.
27. **Dinh, T., I. T. Paulsen, and M. H. Saier, Jr.** 1994. A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of gram-negative bacteria. *J Bacteriol* **176**:3825-31.
28. **Dutt, A., L. A. Heath, and J. A. Nelson.** 1994. P-glycoprotein and organic cation secretion by the mammalian kidney. *J Pharmacol Exp Ther* **269**:1254-60.

29. **Evans, I. J., and J. A. Downie.** 1986. The *nodI* gene product of *Rhizobium leguminosarum* is closely related to ATP-binding bacterial transport proteins; nucleotide sequence analysis of the *nodI* and *nodJ* genes. *Gene* **43**:95-101.
30. **Felsenstein, J.** 1988. Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet* **22**:521-65.
31. **Fisher, R. F., and S. R. Long.** 1993. Interactions of NodD at the *nod* Box: NodD binds to two distinct sites on the same face of the helix and induces a bend in the DNA. *J Mol Biol* **233**:336-48.
32. **Fisher, R. F., and S. R. Long.** 1992. *Rhizobium*--plant signal exchange. *Nature* **357**:655-60.
33. **Goethals, K., M. Van Montagu, and M. Holsters.** 1992. Conserved motifs in a divergent *nod* box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by LysR-type proteins. *Proc Natl Acad Sci U S A* **89**:1646-50.
34. **Gonzalez-Pasayo, R., and E. Martinez-Romero.** 2000. Multiresistance genes of *Rhizobium etli* CFN42. *Mol Plant Microbe Interact* **13**:572-7.
35. **Gonzalez, V., P. Bustos, M. A. Ramirez-Romero, A. Medrano-Soto, H. Salgado, I. Hernandez-Gonzalez, J. C. Hernandez-Celis, V. Quintero, G. Moreno-Hagelsieb, L. Girard, O. Rodriguez, M. Flores, M. A. Cevallos, J. Collado-Vides, D. Romero, and G. Davila.** 2003. The mosaic structure of the symbiotic plasmid of *Rhizobium etli* CFN42 and its relation to other symbiotic genome compartments. *Genome Biol* **4**:R36.
36. **Gotoh, N., H. Tsujimoto, M. Tsuda, K. Okamoto, A. Nomura, T. Wada, M. Nakahashi, and T. Nishino.** 1998. Characterization of the MexC-MexD-OprJ multidrug efflux system in Delta *mexA-mexB-oprM* mutants of *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* **42**:1938-43.
37. **He, X., and J. Zhang.** 2005. Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics* **169**:1157-64.
38. **Hernández-Mendoza, A.** 1999. Aislamiento y caracterización de los genes *nodT* de *Rhizobium etli* cepa CE3. Msc. U.N.A.M., Cuernavaca, Morelos, México.
39. **Hiipakka, M., K. Poikonen, and K. Saksela.** 1999. SH3 domains with high affinity and engineered ligand specificities targeted to HIV-1 Nef. *J Mol Biol* **293**:1097-106.
40. **Hiraga, S., H. Niki, T. Ogura, C. Ichinose, H. Mori, B. Ezaki, and A. Jaffe.** 1989. Chromosome partitioning in *Escherichia coli*: novel mutants producing anucleate cells. *J Bacteriol* **171**:1496-505.
41. **Hirakata, Y., R. Srikumar, K. Poole, N. Gotoh, T. Suematsu, S. Kohno, S. Kamihira, R. E. Hancock, and D. P. Speert.** 2002. Multidrug efflux systems play an important role in the invasiveness of *Pseudomonas aeruginosa*. *J Exp Med* **196**:109-18.
42. **Hirsch, A. M., M. R. Lum, and J. A. Downie.** 2001. What makes the rhizobia-legume symbiosis so special? *Plant Physiol* **127**:1484-92.
43. **Hong, G. F., J. E. Burn, and A. W. Johnston.** 1987. Evidence that DNA involved in the expression of nodulation (*nod*) genes in *Rhizobium* binds to the product of the regulatory gene *nodD*. *Nucleic Acids Res* **15**:9677-90.

44. **Huerta-Zepeda, A., L. Ortuno, G. Du Pont, S. Duran, A. Lloret, H. Merchant-Larios, and J. Calderon.** 1997. Isolation and characterization of *Rhizobium etli* mutants altered in degradation of asparagine. *J Bacteriol* **179**:2068-72.
45. **Jude, F., C. Arpin, C. Brachet-Castang, M. Capdepuy, P. Caumette, and C. Quentin.** 2004. TbtABM, a multidrug efflux pump associated with tributyltin resistance in *Pseudomonas stutzeri*. *FEMS Microbiol Lett* **232**:7-14.
46. **Keen, N. T., S. Tamaki, D. Kobayashi, and D. Trollinger.** 1988. Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene* **70**:191-7.
47. **Keppler, D., Y. Cui, J. Konig, I. Leier, and A. Nies.** 1999. Export pumps for anionic conjugates encoded by MRP genes. *Adv Enzyme Regul* **39**:237-46.
48. **Kessels, H. W., A. C. Ward, and T. N. Schumacher.** 2002. Specificity and affinity motifs for Grb2 SH2-ligand interactions. *Proc Natl Acad Sci U S A* **99**:8524-9.
49. **Kieboom, J., and J. de Bont.** 2001. Identification and molecular characterization of an efflux system involved in *Pseudomonas putida* S12 multidrug resistance. *Microbiology* **147**:43-51.
50. **Kobayashi, N., K. Nishino, and A. Yamaguchi.** 2001. Novel macrolide-specific ABC-type efflux transporter in *Escherichia coli*. *J Bacteriol* **183**:5639-44.
51. **Koronakis, V.** 2003. TolC--the bacterial exit duct for proteins and drugs. *FEBS Lett* **555**:66-71.
52. **Koronakis, V., J. Li, E. Koronakis, and K. Stauffer.** 1997. Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Mol Microbiol* **23**:617-26.
53. **Krummenacher, P., and F. Narberhaus.** 2000. Two genes encoding a putative multidrug efflux pump of the RND/MFP family are cotranscribed with an *rpoH* gene in *Bradyrhizobium japonicum*. *Gene* **241**:247-54.
54. **Lacroix, F. J., A. Cloeckaert, O. Grepinet, C. Pinault, M. Y. Popoff, H. Waxin, and P. Pardon.** 1996. Salmonella typhimurium *acrB*-like gene: identification and role in resistance to biliary salts and detergents and in murine infection. *FEMS Microbiol Lett* **135**:161-7.
55. **Lee, J. H., K. S. Voo, and D. G. Skalnik.** 2001. Identification and characterization of the DNA binding domain of CpG-binding protein. *J Biol Chem* **276**:44669-76.
56. **Lewis-Henderson, W. R., and M. A. Djordjevic.** 1991. A cultivar-specific interaction between *Rhizobium leguminosarum* bv. *trifolii* and subterranean clover is controlled by *nodM*, other bacterial cultivar specificity genes, and a single recessive host gene. *J Bacteriol* **173**:2791-9.
57. **Lewis-Henderson, W. R., and M. A. Djordjevic.** 1991. *nodT*, a positively-acting cultivar specificity determinant controlling nodulation of *Trifolium subterraneum* by *Rhizobium leguminosarum* biovar *trifolii*. *Plant Mol Biol* **16**:515-26.
58. **Li, S. S.** 2005. Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem J* **390**:641-53.

59. **Li, X. Z., H. Nikaido, and K. Poole.** 1995. Role of *mexA-mexB-oprM* in antibiotic efflux in *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* **39**:1948-53.
60. **Li, X. Z., L. Zhang, and K. Poole.** 1998. Role of the multidrug efflux systems of *Pseudomonas aeruginosa* in organic solvent tolerance. *J Bacteriol* **180**:2987-91.
61. **Li, Y., T. Mima, Y. Komori, Y. Morita, T. Kuroda, T. Mizushima, and T. Tsuchiya.** 2003. A new member of the tripartite multidrug efflux pumps, MexVW-OprM, in *Pseudomonas aeruginosa*. *J Antimicrob Chemother* **52**:572-5.
62. **Lichtarge, O., H. R. Bourne, and F. E. Cohen.** 1996. An evolutionary trace method defines binding surfaces common to protein families. *J Mol Biol* **257**:342-58.
63. **Liesegang, H., K. Lemke, R. A. Siddiqui, and H. G. Schlegel.** 1993. Characterization of the inducible nickel and cobalt resistance determinant *cnr* from pMOL28 of *Alcaligenes eutrophus* CH34. *J Bacteriol* **175**:767-78.
64. **Magnet, S., P. Courvalin, and T. Lambert.** 2001. Resistance-nodulation-cell division-type efflux pump involved in aminoglycoside resistance in *Acinetobacter baumannii* strain BM4454. *Antimicrob Agents Chemother* **45**:3375-80.
65. **Martinez-Romero, E., L. Segovia, F. M. Mercante, A. A. Franco, P. Graham, and M. A. Pardo.** 1991. *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. trees. *Int J Syst Bacteriol* **41**:417-26.
66. **Matsuyama, S., K. Inokuchi, and S. Mizushima.** 1984. Promoter exchange between *ompF* and *ompC*, genes for osmoregulated major outer membrane proteins of *Escherichia coli* K-12. *J Bacteriol* **158**:1041-7.
67. **Missiakas, D., and S. Raina.** 1998. The extracytoplasmic function sigma factors: role and regulation. *Mol Microbiol* **28**:1059-66.
68. **Missiakas, D., and S. Raina.** 1997. Signal transduction pathways in response to protein misfolding in the extracytoplasmic compartments of *E. coli*: role of two new phosphoprotein phosphatases PrpA and PrpB. *Embo J* **16**:1670-85.
69. **Munson, G. P., D. L. Lam, F. W. Outten, and T. V. O'Halloran.** 2000. Identification of a copper-responsive two-component system on the chromosome of *Escherichia coli* K-12. *J Bacteriol* **182**:5864-71.
70. **Murakami, S., N. Tamura, A. Saito, T. Hirata, and A. Yamaguchi.** 2004. Extramembrane central pore of multidrug exporter AcrB in *Escherichia coli* plays an important role in drug transport. *J Biol Chem* **279**:3743-8.
71. **Murakami, S., and A. Yamaguchi.** 2003. Multidrug-exporting secondary transporters. *Curr Opin Struct Biol* **13**:443-52.
72. **Nakamura, H., N. Hachiya, and T. Tojo.** 1978. Second acriflavine sensitivity mutation, *acrB*, in *Escherichia coli* K-12. *J Bacteriol* **134**:1184-7.
73. **Namwat, W., C. K. Lee, H. Kinoshita, Y. Yamada, and T. Nihira.** 2001. Identification of the *varR* gene as a transcriptional regulator of virginiamycin S resistance in *Streptomyces virginiae*. *J Bacteriol* **183**:2025-31.
74. **Neyfakh, A. A., C. M. Borsch, and G. W. Kaatz.** 1993. Fluoroquinolone resistance protein NorA of *Staphylococcus aureus* is a multidrug efflux transporter. *Antimicrob Agents Chemother* **37**:128-9.

75. **Nies, D. H., A. Nies, L. Chu, and S. Silver.** 1989. Expression and nucleotide sequence of a plasmid-determined divalent cation efflux system from *Alcaligenes eutrophus*. Proc Natl Acad Sci U S A **86**:7351-5.
76. **Nikaido, H.** 2000. How do exported proteins and antibiotics bypass the periplasm in Gram-negative bacterial cells? Trends Microbiol **8**:481-3.
77. **Nikaido, H.** 1996. Multidrug efflux pumps of gram-negative bacteria. J Bacteriol **178**:5853-9.
78. **Nikaido, H.** 2001. Preventing drug access to targets: cell surface permeability barriers and active efflux in bacteria. Semin Cell Dev Biol **12**:215-23.
79. **Nikaido, H., and H. I. Zgurskaya.** 2001. AcrAB and related multidrug efflux pumps of *Escherichia coli*. J Mol Microbiol Biotechnol **3**:215-8.
80. **Nilsson, M. T., and M. Widersten.** 2004. Repertoire selection of variant single-chain Cro: toward directed DNA-binding specificity of helix-turn-helix proteins. Biochemistry **43**:12038-47.
81. **Nishino, K., J. Yamada, H. Hirakawa, T. Hirata, and A. Yamaguchi.** 2003. Roles of TolC-dependent multidrug transporters of *Escherichia coli* in resistance to beta-lactams. Antimicrob Agents Chemother **47**:3030-3.
82. **Nissink, J. W., and R. Taylor.** 2004. Combined use of physicochemical data and small-molecule crystallographic contact propensities to predict interactions in protein binding sites. Org Biomol Chem **2**:3238-49.
83. **Noel, K. D., A. Sanchez, L. Fernandez, J. Leemans, and M. A. Cevallos.** 1984. *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. J Bacteriol **158**:148-55.
84. **Okamoto, K., N. Gotoh, and T. Nishino.** 2002. Alterations of susceptibility of *Pseudomonas aeruginosa* by overproduction of multidrug efflux systems, MexAB-OprM, MexCD-OprJ, and MexXY/OprM to carbapenems: substrate specificities of the efflux systems. J Infect Chemother **8**:371-3.
85. **Okamoto, K., N. Gotoh, and T. Nishino.** 2002. Extrusion of penem antibiotics by multicomponent efflux systems MexAB-OprM, MexCD-OprJ, and MexXY-OprM of *Pseudomonas aeruginosa*. Antimicrob Agents Chemother **46**:2696-9.
86. **Page, R. D.** 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci **12**:357-8.
87. **Palumbo, J. D., C. I. Kado, and D. A. Phillips.** 1998. An isoflavonoid-inducible efflux pump in *Agrobacterium tumefaciens* is involved in competitive colonization of roots. J Bacteriol **180**:3107-13.
88. **Paulsen, I. T., J. H. Park, P. S. Choi, and M. H. Saier, Jr.** 1997. A family of gram-negative bacterial outer membrane factors that function in the export of proteins, carbohydrates, drugs and heavy metals from gram-negative bacteria. FEMS Microbiol Lett **156**:1-8.
89. **Peng, W. T., and E. W. Nester.** 2001. Characterization of a putative RND-type efflux system in *Agrobacterium tumefaciens*. Gene **270**:245-52.
90. **Phoenix, P., A. Keane, A. Patel, H. Bergeron, S. Ghoshal, and P. C. Lau.** 2003. Characterization of a new solvent-responsive gene locus in *Pseudomonas putida* F1 and its functionalization as a versatile biosensor. Environ Microbiol **5**:1309-27.

91. **Poole, K.** 2004. Efflux-mediated multiresistance in Gram-negative bacteria. *Clin Microbiol Infect* **10**:12-26.
92. **Poole, K.** 2001. Multidrug efflux pumps and antimicrobial resistance in *Pseudomonas aeruginosa* and related organisms. *J Mol Microbiol Biotechnol* **3**:255-64.
93. **Poole, K., N. Gotoh, H. Tsujimoto, Q. Zhao, A. Wada, T. Yamasaki, S. Neshat, J. Yamagishi, X. Z. Li, and T. Nishino.** 1996. Overexpression of the *mexC-mexD-oprJ* efflux operon in *nfxB*-type multidrug-resistant strains of *Pseudomonas aeruginosa*. *Mol Microbiol* **21**:713-24.
94. **Poole, K., and R. Srikumar.** 2001. Multidrug efflux in *Pseudomonas aeruginosa*: components, mechanisms and clinical significance. *Curr Top Med Chem* **1**:59-71.
95. **Puente, J. L., D. Juarez, M. Bobadilla, C. F. Arias, and E. Calva.** 1995. The *Salmonella ompC* gene: structure and use as a carrier for heterologous sequences. *Gene* **156**:1-9.
96. **Putman, M., H. W. van Veen, and W. N. Konings.** 2000. Molecular properties of bacterial multidrug transporters. *Microbiol Mol Biol Rev* **64**:672-93.
97. **Quandt, J., and M. F. Hynes.** 1993. Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. *Gene* **127**:15-21.
98. **Ramos, J. L., E. Duque, M. T. Gallegos, P. Godoy, M. I. Ramos-Gonzalez, A. Rojas, W. Teran, and A. Segura.** 2002. Mechanisms of solvent tolerance in gram-negative bacteria. *Annu Rev Microbiol* **56**:743-68.
99. **Rivilla, R., and J. A. Downie.** 1994. Identification of a *Rhizobium leguminosarum* gene homologous to *nodT* but located outside the symbiotic plasmid. *Gene* **144**:87-91.
100. **Rivilla, R., J. M. Sutton, and J. A. Downie.** 1995. *Rhizobium leguminosarum* NodT is related to a family of outer-membrane transport proteins that includes TolC, PrtF, CyaE and AprF. *Gene* **161**:27-31.
101. **Robertson, J. G., B. Wells, N. J. Brewin, E. Wood, C. D. Knight, and J. A. Downie.** 1985. The legume-*Rhizobium* symbiosis: a cell surface interaction. *J Cell Sci Suppl* **2**:317-31.
102. **Rojas, A., E. Duque, G. Mosqueda, G. Golden, A. Hurtado, J. L. Ramos, and A. Segura.** 2001. Three efflux pumps are required to provide efficient tolerance to toluene in *Pseudomonas putida* DOT-T1E. *J Bacteriol* **183**:3967-73.
103. **Rojas, A., A. Segura, M. E. Guazzaroni, W. Teran, A. Hurtado, M. T. Gallegos, and J. L. Ramos.** 2003. In vivo and in vitro evidence that TtgV is the specific regulator of the TtgGHI multidrug and solvent efflux pump of *Pseudomonas putida*. *J Bacteriol* **185**:4755-63.
104. **Rouch, D. A., D. S. Cram, D. DiBerardino, T. G. Littlejohn, and R. A. Skurray.** 1990. Efflux-mediated antiseptic resistance gene *qacA* from *Staphylococcus aureus*: common ancestry with tetracycline- and sugar-transport proteins. *Mol Microbiol* **4**:2051-62.
105. **Saier, M. H., Jr., R. Tam, A. Reizer, and J. Reizer.** 1994. Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport. *Mol Microbiol* **11**:841-7.

106. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. *Molecular Cloning, A laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, New York.
107. **Scott, D. B., C. A. Young, J. M. Collins-Emerson, E. A. Terzaghi, E. S. Rockman, P. E. Lewis, and C. E. Pankhurst.** 1996. Novel and complex chromosomal arrangement of *Rhizobium loti* nodulation genes. *Mol Plant Microbe Interact* **9**:187-97.
108. **Schlaman, H. R., R. J. Okker, and B. J. Lugtenberg.** 1990. Subcellular localization of the *Rhizobium leguminosarum nodI* gene product. *J Bacteriol* **172**:5486-9.
109. **Schnaitman, C. A.** 1971. Solubilization of the cytoplasmic membrane of *Escherichia coli* by Triton X-100. *J Bacteriol* **108**:545-52.
110. **Schweizer, H. P.** 2003. Efflux as a mechanism of resistance to antimicrobials in *Pseudomonas aeruginosa* and related bacteria: unanswered questions. *Genet Mol Res* **2**:48-62.
111. **Segovia, L., J. P. Young, and E. Martinez-Romero.** 1993. Reclassification of American *Rhizobium leguminosarum* biovar *phaseoli* type I strains as *Rhizobium etli* sp. nov. *Int J Syst Bacteriol* **43**:374-7.
112. **Sharff, A., C. Fanutti, J. Shi, C. Calladine, and B. Luisi.** 2001. The role of the TolC family in protein transport and multidrug efflux. From stereochemical certainty to mechanistic hypothesis. *Eur J Biochem* **268**:5011-26.
113. **Sharp, J. A., and T. Platt.** 1984. Rho-dependent termination and concomitant NTPase activity requires a specific, intact RNA region. *J Biol Chem* **259**:2268-73.
114. **Simon, R., U. Prifer, and A. Pühler.** 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-positive bacteria. *Biotechnology* **1**:784-791.
115. **Sousa, C., J. L. Folch, P. Boloix, M. Megias, N. Nava, and C. Quinto.** 1993. A *Rhizobium tropici* DNA region carrying the amino-terminal half of a nodD gene and a nod-box-like sequence confers host-range extension. *Mol Microbiol* **9**:1157-68.
116. **Spaink, H. P., A. H. Wijfjes, and B. J. Lugtenberg.** 1995. *Rhizobium* NodI and NodJ proteins play a role in the efficiency of secretion of lipochitin oligosaccharides. *J Bacteriol* **177**:6276-81.
117. **Stajich, J. E., D. Block, K. Boulez, S. E. Brenner, S. A. Chervitz, C. Dagdigian, G. Fuellen, J. G. Gilbert, I. Korf, H. Lapp, H. Lehvaslaiho, C. Matsalla, C. J. Mungall, B. I. Osborne, M. R. Pockock, P. Schattner, M. Senger, L. D. Stein, E. Stupka, M. D. Wilkinson, and E. Birney.** 2002. The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* **12**:1611-8.
118. **Stevens, S. Y., and G. D. Glick.** 1999. Evidence for sequence-specific recognition of DNA by anti-single-stranded DNA autoantibodies. *Biochemistry* **38**:560-8.
119. **Sulavik, M. C., C. Houseweart, C. Cramer, N. Jiwani, N. Murgolo, J. Greene, B. DiDomenico, K. J. Shaw, G. H. Miller, R. Hare, and G. Shimer.** 2001. Antibiotic susceptibility profiles of *Escherichia coli* strains lacking multidrug efflux pump genes. *Antimicrob Agents Chemother* **45**:1126-36.

120. **Surin, B. P., J. M. Watson, W. D. Hamilton, A. Economou, and J. A. Downie.** 1990. Molecular characterization of the nodulation gene, *nodT*, from two biovars of *Rhizobium leguminosarum*. *Mol Microbiol* **4**:245-52.
121. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**:4876-82.
122. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**:4673-80.
123. **Treped, C. M., and J. E. Mott.** 2004. Identification of the *Haemophilus influenzae tolC* gene by susceptibility profiles of insertionally inactivated efflux pump mutants. *Antimicrob Agents Chemother* **48**:1416-8.
124. **Tseng, T. T., K. S. Gratwick, J. Kollman, D. Park, D. H. Nies, A. Goffeau, and M. H. Saier, Jr.** 1999. The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J Mol Microbiol Biotechnol* **1**:107-25.
125. **Udvardi, M. K., and D. A. Day.** 1997. Metabolite Transport across Symbiotic Membranes of Legume Nodules. *Annu Rev Plant Physiol Plant Mol Biol* **48**:493-523.
126. **van Veen, H. W., R. Callaghan, L. Soceneantu, A. Sardini, W. N. Konings, and C. F. Higgins.** 1998. A bacterial antibiotic-resistance gene that complements the human multidrug-resistance P-glycoprotein gene. *Nature* **391**:291-5.
127. **Vazquez, M., A. Davalos, A. de las Penas, F. Sanchez, and C. Quinto.** 1991. Novel organization of the common nodulation genes in *Rhizobium leguminosarum* bv. *phaseoli* strains. *J Bacteriol* **173**:1250-8.
128. **Vazquez, M., O. Santana, and C. Quinto.** 1993. The NodI and NodJ proteins from *Rhizobium* and *Bradyrhizobium* strains are similar to capsular polysaccharide secretion proteins from gram-negative bacteria. *Mol Microbiol* **8**:369-77.
129. **Verma, D. P., and Z. Hong.** 1996. Biogenesis of the peribacteroid membrane in root nodules. *Trends Microbiol* **4**:364-8.
130. **Villalobos, M. A., N. Nava, M. Vazquez, and C. Quinto.** 1994. Nucleotide sequence of the *Rhizobium etli nodS* gene. *Gene* **150**:201-2.
131. **Wandersman, C., and P. Delepelaire.** 1990. TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *Proc Natl Acad Sci U S A* **87**:4776-80.
132. **Wery, J., B. Hidayat, J. Kieboom, and J. A. de Bont.** 2001. An insertion sequence prepares *Pseudomonas putida* S12 for severe solvent stress. *J Biol Chem* **276**:5700-6.
133. **Yen, M. R., C. R. Peabody, S. M. Partovi, Y. Zhai, Y. H. Tseng, and M. H. Saier.** 2002. Protein-translocating outer membrane porins of Gram-negative bacteria. *Biochim Biophys Acta* **1562**:6-31.

134. **Yoshihara, E., H. Maseda, and K. Saito.** 2002. The outer membrane component of the multidrug efflux pump from *Pseudomonas aeruginosa* may be a gated channel. *Eur J Biochem* **269**:4738-45.
135. **Young, C., J. M. Collins-Emerson, E. A. Terzaghi, and D. B. Scott.** 1990. Nucleotide sequence of *Rhizobium loti nodI*. *Nucleic Acids Res* **18**:6691.
136. **Yu, E. W., J. R. Aires, and H. Nikaido.** 2003. AcrB multidrug efflux pump of *Escherichia coli*: composite substrate-binding cavity of exceptional flexibility generates its extremely wide substrate specificity. *J Bacteriol* **185**:5657-64.
137. **Yu, E. W., G. McDermott, H. I. Zgurskaya, H. Nikaido, and D. E. Koshland, Jr.** 2003. Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump. *Science* **300**:976-80.
138. **Zgurskaya, H. I., and H. Nikaido.** 1999. Bypassing the periplasm: reconstitution of the AcrAB multidrug efflux pump of *Escherichia coli*. *Proc Natl Acad Sci U S A* **96**:7190-5.
139. **Zgurskaya, H. I., and H. Nikaido.** 2000. Cross-linked complex between oligomeric periplasmic lipoprotein AcrA and the inner-membrane-associated multidrug efflux pump AcrB from *Escherichia coli*. *J Bacteriol* **182**:4264-7.

Tabla 1**Cepas*****R. etli***

CE3	Cepa resistente natural a Sp Derivada de CFN42	(83)
CE3/pRK415:: <i>nodTch</i>	CE3 complementada con plásmido pRK415:: <i>nodTch</i>	Este estudio
CE3 <i>nodTch</i> ::Km/ pRK415:: <i>nodTch</i>	CE3 llevando un casete de resistencia a Km interrumpiendo a <i>nodTch</i> y complementada <i>in trans</i> con plásmido pRK415:: <i>nodTch</i>	Este estudio
CE3 <i>rmiR</i> ::Sp	CE3 llevando un casete de resistencia a Sp interrumpiendo a <i>rmiR</i>	Este estudio
CE3 <i>rmiA</i> ::sp	CE3 llevando un casete de resistencia a Sp interrumpiendo a <i>rmiA</i>	Este estudio
CE3 <i>rmiB</i> ::Sp	CE3 llevando un casete de resistencia a Sp interrumpiendo a <i>rmiB</i>	Este estudio
CFNX184	<i>R. leguminosarum</i> bv. <i>phaseoli</i> CFN42 curada de pc	(17)

E. coli

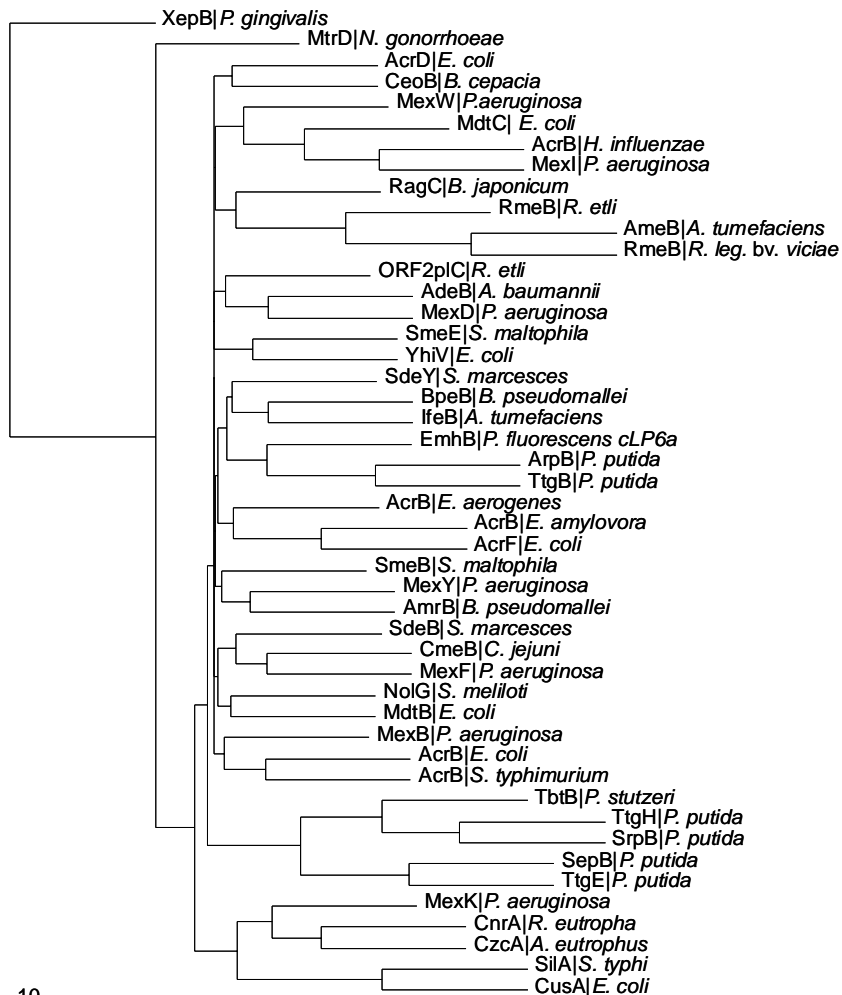
XL1Blue	<i>recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lacI, [F proAB, lacIqZDM15, Tn10(Tc^r)]</i>	Stratagene
S17.1	<i>thi, pro, hsdR (rm+), recA</i> (derivada de <i>E. coli</i> 294); lleva el plasmido RP4 (amp; tet::Mu; km::Tn7) integrado dentro del cromosoma Tc ^s ; Km ^s ; Ap ^f ; Tp ^f ; tra+	(114)

Plásmidos

pLAFRTch	pLAFR1 llevando un fragmento de DNA de 20 kb de <i>R. etli</i> con la región <i>nodTch</i>	Este estudio
pLAFRTpc	pLAFR1 llevando un fragmento de DNA de 20 kb de <i>R. etli</i> con la región <i>nodTpc</i>	Este estudio
pBluescriptSK	Vector Fagémido de clonación, 2.96 kb, MCS, Ap ^r	Stratagene
pHME8	pBluescriptSK llevando la región <i>rmiRABnodTch</i> en un fragmento de 7.8 kb <i>NotI</i>	Este estudio
pHME22	pBluescriptSK llevando la región <i>nodTch</i> en un fragmento de 2.2 kb <i>BamHI-HindIII</i>	Este estudio
pHME40	pBluescriptSK llevando a <i>nodTpc</i> en un fragmento de 3.7 kb <i>EcoRI</i>	Este estudio
pRK415	Vector de clonación de amplio rango; <i>plac</i> , MCS, Tc ^r	(46)
pRK415:: <i>nodTch</i>	Derivado de pRK415 llevando a <i>nodTch</i> en un fragmento de 1.8 kb <i>XhoI-BamHI</i>	Este estudio
pJQ200SK	Vector suicida, <i>sacB</i> Gm ^r	(97)
pHMTch::Km	Derivado de pJQ200SK con <i>nodTch</i> interrumpido con un casete de resistencia a Km	Este estudio
pHMTpc::Sp	Derivado del pJQ200SK llevando a <i>nodTpc</i> interrumpido con un casete de Sp	Este estudio
pHMA::Sp	Derivado del pJQ200SK llevando a <i>rmiA</i> interrumpido con un casete de Sp	Este estudio
pHMB::Sp	Derivado del pJQ200SK llevando a <i>rmiB</i> interrumpido con un casete de Sp	Este estudio
pHMR::Sp	Derivado del pJQ200SK llevando a <i>rmiR</i> interrumpido con un casete de Sp	Este estudio
pJQ934	Derivado del pJQ200SK llevando un fragmento interno de <i>nodTch</i> de 934 bp <i>Fsp I</i>	Este estudio

Cluster\ Position	L25 CC	K29 CC	D99 CC	D101 CC/CP	V105 CP	N109 CP	Q112 CP	P116 CP	V382 CC	A385 CC	F386 CC	F388 CC	F458 CC	F459 CC
AcrD			X	X	X (Q)	X				X	X	X (Y)	X	
MexW						X (A)	X (N)				X			
AmeB	X						X						X	
ORF2pIC	X		X			X						X	X	
SmeE*	X			X	X	X	X	X				X	X	X (M)
SdeY				X	X	X	X (S)		X (I)			X	X	
EmhB	X	X	X (N)	X	X	X	X (N)	X	X (I)	X		X	X	X
AcrF	X		X	X	X	X	X	X	X (I)		X		X	X
SmeB				X	X	X							X	X
SdeB	X (I)		X	X		X					X (L)	X	X	
NoIG					X (Q)						X (L)	X	X	X (M)
AcrB	X		X	X	X	X	X	X	X	X	X	X	X	X
TbtB	X	X		X	X	X	X	X	X (L)	X (Y)	X		X	X
CnrA														

Tabla 2. Conservación de aminoácidos en los grupos identificados por Máxima Parsimonia. Los residuos universales fueron definidos como aquellos 100 % idénticos en cada grupo. Columnas: en la primera se describe un nombre representativo para cada grupo. De la segunda a la quinceava se colocan los residuos y su conservación por posición. Por ejemplo, L25, Q112 y F458 están conservados en todas las proteínas que se incluyen en el grupo AmeB. Letras entre paréntesis son residuos conservados en un grupo en particular pero que difiere con respecto al consenso general del grupo AcrB que es el más conservado. CC.- Cavidad Central. CP.- Poro Central



PROTEIN	SUBSTRATE	ORGANISM	ACCESSION no.
XepB	EB, PU, RF	<i>Porphyromonas gingivalis</i>	AAQ65732
MtrD	EB, BL, FA, TX	<i>Neisseria gonorrhoeae</i>	U60099
AcrD	AG, BS, FU, NO	<i>Escherichia coli</i>	P24177
CeoB	FQ, CM, TM (all operon)	<i>Burkholderia cepacia</i>	AAB58161
MexW	CM, FQ, TC, EB, AC, BL	<i>Pseudomonas aeruginosa</i>	AAG07763
MdtC	NO, BS	<i>Escherichia coli</i>	BAC06609
AcrB	NO, ML, CV, EB	<i>Haemophilus influenzae</i>	AAC22555
MexI	V, BL, AG. A mutation is fully resistant to Tc	<i>Pseudomonas aeruginosa</i>	AAG07594
RagC	UK	<i>Bradyrhizobium japonicum</i>	CAA12353
rmeB	UK	<i>Rhizobium etli</i>	in process
AmeB	UK	<i>Agrobacterium tumefaciens</i>	AAG09746
ORF2	UK	<i>Rhizobium leguminosarum</i> bv. <i>viciae</i>	rhiz600e08.q1n
ORF2p C	UK	<i>Rhizobium etli</i>	
AdeB	FQ, TC, CM, AG, ML, BL, EB, TM	<i>Acinetobacter baumannii</i>	AAL14440
MexD	FQ, TC, CM, AG, CP, TR	<i>Pseudomonas aeruginosa</i>	NP_253288
SmeE	ML, EB, FQ, TC, CM	<i>Stenotrophomonas maltophilia</i>	CAC14595
YhiV	ML, EB, BS, R6G, TPP, SDS,	<i>Escherichia coli</i>	AAC76539
SdeY	ML, AC, EB, TC, FQ, TR, BKC, RH6G, H33342	<i>Serratia marcescens</i>	AB104882
BpeB	ML, AG	<i>Burkholderia pseudomallei</i>	AAQ94110
lfeB	COU (IF)	<i>Agrobacterium tumefaciens</i>	AAC25691
EmhB	CM, FQ, PAH	<i>Pseudomonas fluorescens</i> cLP6a	AAQ92181
ArpB	CM, BL, TC, ML, AG, NO	<i>Pseudomonas putida</i>	AAF73832
TtgB	CM, BL, FQ, TC, AH	<i>Pseudomonas putida</i>	AAC38671
AcrB	SDS, AC, CM, FQ, MC, TC, NO	<i>Enterobacter aerogenes</i>	CAC35725
AcrB	SDS, TC, FQ, CV, MB, EB, PLANT COMPOUNDS.	<i>Erwinia amylovora</i>	AAQ21216
AcrF	BS, AC, NO, EB, BL, CV, FA, TX, R6G, TPP, DOX	<i>Escherichia coli</i>	NP_417732
SmeB	AmeC mutant hypersensitive	<i>Stenotrophomonas maltophilia</i>	AF173226
MexY	AG, ML, AC, EB, FQ, TC	<i>Pseudomonas aeruginosa</i>	AB015853
AmrB	AG, ML	<i>Burkholderia pseudomallei</i>	AAC27754
SdeB (MexF)	FQ	<i>Serratia marcescens</i>	AAO18187
CmeB	FQ, BL, ML, RF, TC, EB, SDS, BS DC, CHA, CHE, TAU	<i>Campylobacter jejuni</i>	AAL74245
MexF	FQ, BL, CM, TM	<i>Pseudomonas aeruginosa</i>	NP_251184
NolG	UK	<i>Sinorhizobium melliloti</i>	NP_435726
MdtB	UK	<i>Escherichia coli</i>	BAC06608
MexB	AC, BL, EB, CM, CV, ML, NO, SDS, AH, SF, TC, TM, TR, AG	<i>Pseudomonas aeruginosa</i>	AAG03815
AcrB/AcrE	AC, BL, EB, CM, CV, ML, NO, SDS, AH, RF, TX, FA, BS	<i>Escherichia coli</i>	P31224
AcrB	AC, BS, FQ, TC	<i>Salmonella typhimurium</i>	NP_459470
TbtB	AH, TBT	<i>Pseudomonas stutzeri</i>	AAD15664
TtgH	AH	<i>Pseudomonas putida</i>	AAK69564
SrpB	AH	<i>Pseudomonas putida</i>	AAD12176
SepB	AH	<i>Pseudomonas putida</i>	AAO49779
TtgE	AH	<i>Pseudomonas putida</i>	CAB72258
MexK	EM, TC, TR	<i>Pseudomonas aeruginosa</i>	NP_252366
CnrA	HM	<i>Ralstonia eutropha</i>	G47056
CzcaA	HM	<i>Alcaligenes eutrophus</i>	P94177
SilA	HM	<i>Salmonella typhi</i>	AAD11749
CusA	HM	<i>Escherichia coli</i>	P38054

Fig. 17.- Análisis de la familia RND basados en los residuos de interacción a ligando y especificidad de sustrato.

A) Árbol filogenético generado con el programa Phylip usando los residuos correspondientes a la interacción a ligando reportados en AcrB

B) Tabla con la información relevante de las bombas de exclusión de drogas seleccionadas y sus sustratos expulsados. La nomenclatura es como sigue: AH: Hidrocarburos aromáticos, AC: Acriflavina, AG: Aminoglicosidos, BKC: Cloruro de Benzalconio, BL: β -Lactamas, BS: Sales biliares, CM: Cloramfenicol, CP: Cefalosporina, CV: Cristal violeta, COU: Comestrol, CHA: Ácido Cólico, CHE: Ácido quenodeoxicholico, DC: Deoxilcloato, EB: Bromuro de etidio, EM: Eritromicina, FA: Ácidos grasos FQ: Fluoroquinolonas, H33342: Hoechst 33342, HM: Metales pesados, ML: Macrolidos, MB: Azul de metileno, NO: Novobiocina, PU: Puromicina, RH6G: Rodamina 6G, RF: Rifampicina, SDS: Sulfato dodecil de sodio, TAU: Ácido Taurocolico, TBT: Tributiltina, TC: Tetraciclina, TM: Trimetropina, TR: Triclosan, TX: Triton X-100, TPP: Bromuro de tetrafenilfosfonio, UK: Desconocido, V: Vanadio

MexI Paeru LGHRWLTGGL ALLVCISLPL LYSPMKRELA PTEDQAALVT AIKAPQHANL DYVELPARKL DVVYTSIPET VSTWIING-- ----T----- --DGPAASFG GINLAAWEKR ERDASAIQSE LQGKVGDEG ----SSIFAF QLAALPGST-

Table with columns for taxonomic labels (e.g., MexI, MexW), protein names, and amino acid sequence alignments. The alignment includes position markers (1060-1190) and hyphenated amino acid codes.

Apéndice 2

Representación gráfica de los residuos de los grupos de interacción con ligandos específicos con logotipos de secuencias generados con el programa WebLogo. En algunos casos se reunieron 2 o más grupos de la fig. 17 si compartían ligandos similares. Se eliminaron aquellas proteínas que formaban grupos huérfanos. Para comparar los residuos y los grupos, ver fig. 17

>ModPfam
 CRDITANDVMTL
 >ModEcoE
 LINDVNSALGFL
 M
 >ModPfam
 RDTTANLAFLM
 >ModEcoE
 QKDAANSLFFLM



>ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF



>ModEcoE
 LTDVNDVWYTF
 >ModEcoE
 LTDVNDVWYTF
 >ModEcoE
 LTDVNDVWYTF
 >ModEcoE
 LTDVNDVWYTF
 >ModEcoE
 LTDVNDVWYTF



>ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF



>ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF



>ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF



>ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF



>ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF
 >ModEcoE
 LKDDVNDVWFFF



Apéndice 3

Artículo publicado en la revista *Journal of Molecular Microbiology and Biotechnology*, en el Volumen 11, No. 6, 2006, actualmente en prensa.

Diminished Redundancy of Outer Membrane Factor Proteins in Rhizobiales: A *nodT* Homolog Is Essential for Free-Living *Rhizobium etli*

Armando Hernández-Mendoza Noreide Nava Olivia Santana
Cei Abreu-Goodger Anibal Tovar Carmen Quinto

Departamento de Biología Molecular de Plantas, Instituto de Biotecnología,
Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

© S. Karger AG, Basel

**PROOF Copy
for personal
use only**

ANY DISTRIBUTION OF THIS
ARTICLE WITHOUT WRITTEN
CONSENT FROM S. KARGER
AG, BASEL IS A VIOLATION
OF THE COPYRIGHT.

Key Words

Rhizobium etli • Essential genes • NodT • RND efflux systems

Abstract

Rhizobium etli is a gram-negative soil bacterium that induces nitrogen-fixing nodules on common bean roots (*Phaseolus vulgaris*). *R. etli* encodes two genes homologous to *nodT* of *Rhizobium leguminosarum*. *nodTch* is chromosomal and forms an operon with new genes resembling a multi-drug efflux pump of the resistance-nodulation-cell division (RND) family. *nodTch* is the last gene of this operon and can also be independently transcribed; the gene product is located in the bacterial outer membrane. Cell survival requires *nodTch* under all conditions tested. A second *nodT* gene, *nodTpc*, is encoded by plasmid *c*; it is constitutively transcribed but does not complement the essential function encoded by *nodTch*. NodT proteins belong to the outer membrane efflux proteins of the TolC superfamily. The number of duplications in the *tolC* gene family positively correlates with genome size in gram-negative bacteria. Nonetheless, some α -proteobacteria, including *R. etli*, encode fewer outer membrane factor exporters than expected suggesting further roles in addition to detoxification.

Copyright © 2006 S. Karger AG, Basel

Introduction

The symbiotic association between soil bacteria of the genera *Azorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium*, and *Sinorhizobium* (collectively referred to as rhizobia) and legume plants in nitrogen-fixing nodules permits host plants to grow without added nitrogen fertilizer. Both legume and non-legume roots exude compounds such as flavonoids that induce the expression of some rhizobial nodulation genes (*nod*, *nol* and *noe*) which are essential for nodulation, and these products help to determine host range. Rhizobia are Gram-negative soil free-living α -proteobacteria that carry most of the nodulation genes either on large plasmids [Broughton et al., 2000] or on chromosomal segments termed symbiosis islands [Sullivan and Ronson, 1998]. These genes specify the synthesis of lipochitooligosaccharide signal molecules called Nod factors [Denarie et al., 1996] that play a pivotal role in the initial stages of symbiosis.

Rhizobium etli is the predominant symbiont found in nodules of *Phaseolus vulgaris* in Mesoamerica. The 6,530,228-bp genome sequence of *R. etli* has recently been reported [Gonzalez et al., 2006]. Most of the genes necessary for cell growth are encoded in the chromosome, whereas only a few essential genes or complete metabolic pathways are encoded in some of the six large plasmids,

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2006 S. Karger AG, Basel
1464-1801/06/0000-0000\$23.50/0

Accessible online at:
www.karger.com/mmb

Carmen Quinto
Departamento de Biología Molecular de Plantas
Instituto de Biotecnología, Universidad Nacional Autónoma de México
Cuernavaca, Morelos 62271 (México)
Tel. +52 777 329 1642, Fax + 52 777 313 6600, E-Mail quinto@ibt.unam.mx

named p42a to p42f. The majority of the nodulation genes are encoded on symbiotic plasmid p42d [Gonzalez et al., 2003].

Several *R. etli* *nod* genes including *nodABCS* and *nodIJ* have previously been identified and characterized in our laboratory [Cardenas et al., 1996; Vazquez et al., 1991; Villalobos et al., 1994]. The isolation and analysis of two *R. etli* *nodT* homologs is described in this work. In *Rhizobium leguminosarum* bv. *trifolii* *nodT* forms an operon with *nodABCII*, and with *nodMN* in *R. leg.* bv. *viciae* [Surin et al., 1990]. The function of *nodT* is not fully understood. It was reported that inactivation *nodT* in *R. leg.* bv. *trifolii* had no effect on the nodulation of *Vicia* spp., but caused a delay in the nodulation of *Trifolium subterraneum* [Canter Cremers et al., 1989]. In contrast, Surin et al. [1990] observed no effect of *nodT* mutation either in *R. leg.* bv. *viciae* or *R. leg.* bv. *trifolii* on any of the plants tested, including *T. subterraneum*. In a third study [Lewis-Henderson and Djordjevic, 1991], it was reported that the cloned *nodT* gene from *R. leg.* bv. *trifolii* ANU843 allowed the *R. leg.* bv. *trifolii* strain TAI (a strain naturally lacking *nodT*) to nodulate one specific cultivar (Woo-genellup) of *T. subterraneum*. In *R. leg.* bv. *phaseoli*, a gene homologous to *nodT* was identified outside the symbiotic plasmid, but mutations in this gene were not obtained after several attempts [Rivilla and Downie, 1994]. It has also been established that *nodT* is not involved in the synthesis or secretion of known Nod factors [Spaink et al., 1995]. Because their amino acid sequences resemble outer membrane proteins involved in secretion, including TolC, PrtF, CyaE and AprF [Rivilla et al., 1995], NodT proteins have been classified as members of the TolC family, which belong to the outer membrane factor family [Paulsen et al., 1997].

The present work shows that *R. etli* encodes for two paralogs of *nodT* located outside the symbiotic plasmid, one on the chromosome and the other on p42c. The chromosomal copy (*nodTch*) is located downstream of three contiguous open reading frames (ORFs), the predicted protein products of which show high similarity to resistance-nodulation-cell division (RND) efflux systems in Gram-negative bacteria. The second *nodT* homolog, present on plasmid c (*nodTpc*), was constitutively transcribed. *nodTch* is essential for viability under all conditions tested, yet *nodTpc* did not complement this vital function. *R. etli*, as well as some other α -proteobacteria, appears to code for less TolC-like paralogs than would be expected given its genome size suggesting additional roles for *nodTch*.

Results and Discussion

Identification of *nodT* Homologs in *R. etli*

To confirm that *nodT* is encoded by *R. etli* CE3, total genomic DNA digested with *EcoRI* was analyzed by DNA blot hybridization using *R. leg.* bv. *phaseoli* *nodT* as probe. Two hybridization fragments of 7.7 and 3.7 kb were visualized (data not shown). Subsequent DNA blot analysis of Eckhardt gels containing *R. etli* wild-type and plasmid-cured derivatives [Brom et al., 1992] demonstrated that the 7.7 kb *EcoRI* hybridization fragment was located on the bacterial chromosome and the 3.7 kb *EcoRI* fragment on plasmid c (data not shown). Therefore *R. etli* CE3 strain contains two *nodT* genes. These results were further confirmed by the analysis of the recently available complete genome of *R. etli* [Gonzalez et al., 2006].

To isolate and functionally characterize these *nodT* genes, a *R. etli* genomic library [Huerta-Zepeda et al., 1997] was screened using the *R. leg.* bv. *phaseoli* *nodT* probe. Six colonies carrying a 7.7 kb *EcoRI* fragment and four colonies carrying a 3.7 kb *EcoRI* fragment were recovered. Colonies pLARFTch (carrying the 7.7 kb *EcoRI* fragment) and pLARFTpc (carrying the 3.7 kb *EcoRI* fragment) were selected for further analysis. A 7.82-kb *Not I* DNA fragment isolated from pLARFTch (fig. 1a) was entirely sequenced. This sequence contained four contiguous ORFs, each preceded by a potential ribosome binding site (fig. 1a). The final ORF (ORF4) was located 152-bp downstream of ORF3 (fig. 1a) and encodes a protein with similarity to NodT proteins from *R. leg.* bvs. *trifolii* (78% similarity), *viciae* (80%), and *phaseoli* (93%). It also shows high similarity (74%) to the AmeC protein from *Agrobacterium tumefaciens*, a plant pathogen also member of the Rhizobiaceae family (fig. 1c).

The 3.7 kb *EcoRI* fragment isolated from pLARFTpc contained a single ORF (fig. 1b) with the most probable start codon being a GTG preceded by a good potential ribosome binding site. This ORF showed high similarity to the *nodT* genes from *R. leg.* bvs. *trifolii* (72%), *viciae* (73%), and *phaseoli* (77%). It also presented 77% similarity to ORF4 from *R. etli* and 72% to AmeC from *A. tumefaciens* (fig. 1c).

The predicted translation products of both genes contain N-terminal transit signal peptides of 17 and 25 aa, respectively (data not shown). This transit signal peptide is typical of outer membrane lipoproteins that are acylated following cleavage of the signal peptide [Inouye, 1979]. Based on sequence similarity to reported NodT proteins, these *R. etli* ORFs were named chromosomal (*nodTch*) and plasmid c *nodT* (*nodTpc*). The products of

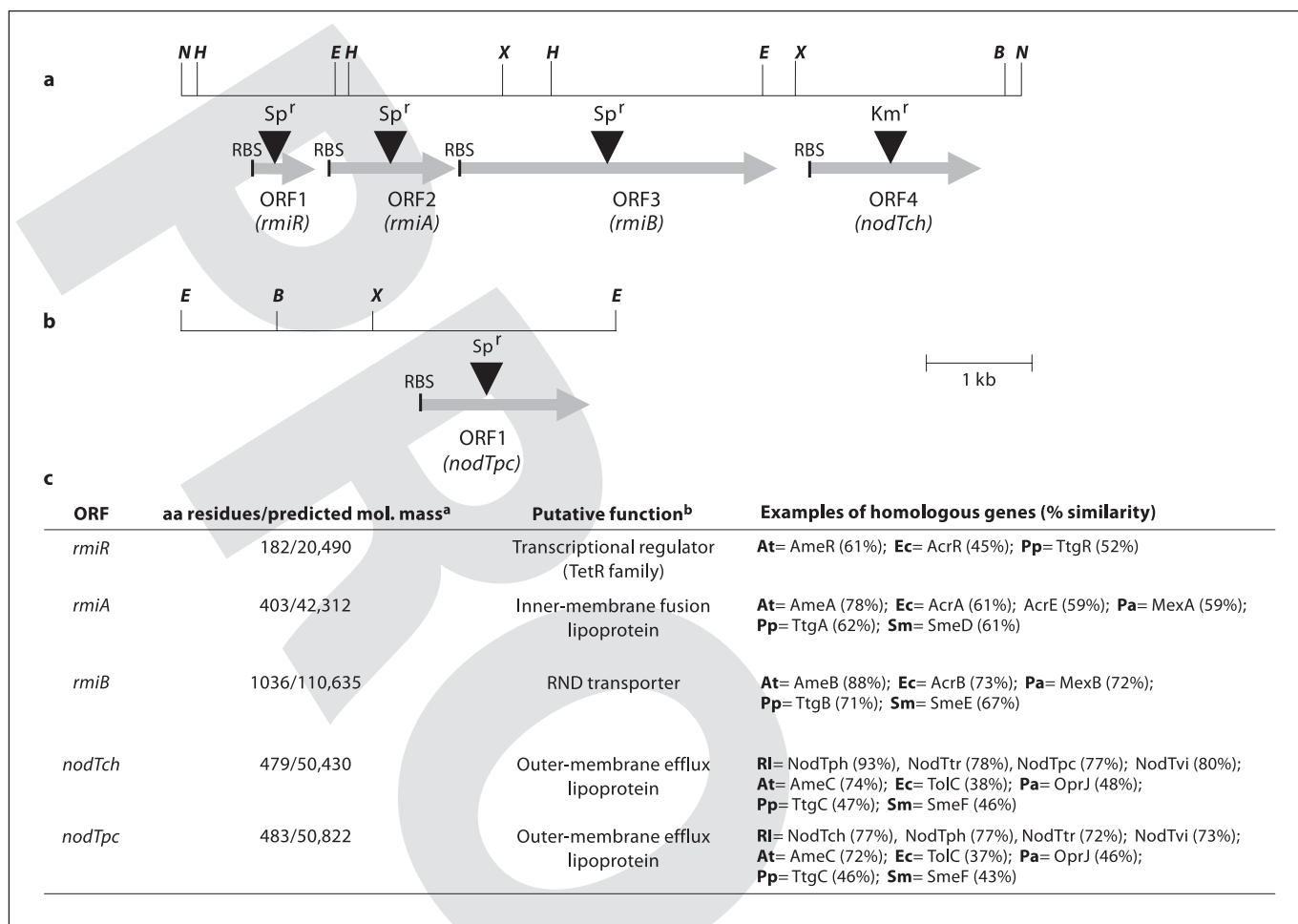


Fig. 1. Genetic organization of the *Rhizobium etli* *rmiRABnodTch* region and the *nodTpc* restriction map. **a** Restriction map of a 7,820 bp chromosomal *R. etli* DNA fragment. **b** Restriction map of a 3,703 bp *R. etli* plasmid c DNA fragment. Solid arrows indicate predicted open reading frames. RBS = Putative ribosome binding sites. Arrowheads indicate sites of insertion of spectinomycin or kanamycin resistance cassettes in *R. etli* *rmi* and *nodTch* mutants. **c** Molecular characterization of proteins encoded by *rmiRABnodT* and *nodTpc*. Microorganisms are abbreviated and printed in boldface. At = *Agrobacterium tumefaciens*; Pa = *Pseudomonas aeruginosa*; Pp = *Pseudomonas putida*; Ec = *Escherichia coli*;

Sm = *Stenotrophomonas maltophilia*; Rl = *Rhizobium leguminosarum* biovar, ph = *phaseoli*, tr = *trifolii*, vi = *viciae*. GeneBank accession numbers: AmeR AmeABC (AF232237); AmrR AmrAB (AF072887); TtgR (AF238479); TtgABC (AF031417); TtgDEF (AFY19106); ArpABC (AF183959); SmeRSABC (AF173226); SmeDEF (AJ252200); MexAB-OprM (L11616); MexCD-OprJ (U57969); AcrAB (AE000152); NodTtrifolii (P24145); NodTviciae (P15727); NodTphaseoli (PD466172). ^aMolecular mass expressed in Daltons. ^bAs predicted by PSI-BLAST. B = BamHI; E = EcoRI; H = HindIII; N = NotI, X = XhoI.

both genes also show significant similarity to other outer membrane efflux lipoproteins from Gram-negative bacteria (fig. 1c).

Chromosomal nodT Encodes an Essential Function in *R. etli*

To define the functional role of NodTch in *R. etli*, several attempts were made to obtain a *nodTch* mutant. A 934-bp *FspI* internal fragment of *nodTch* was cloned into

the suicide vector pJQ200SK to generate pJQ934 (table 1). This plasmid can be mobilized but fails to replicate in *R. etli*. pJQ200SK carries *gtmR* (gentamicin resistance) and the *sacB* gene from *Bacillus subtilis* that codes for levan sucrase; *sacB* expression confers lethality in Gram-negative strains grown in sucrose. pJQ934 was transferred by conjugation to the *R. etli* wild-type strain to disrupt the *nodTch* gene, but we were unable to obtain transformants on rich or chemically defined media (data not shown). To

Table 1. Strains and plasmids

Strains		
<i>R. etli</i>		
CE3	A natural streptomycin-resistant derivative of CFN42	[Noel et al. 1984]
CE3/pRK415:: <i>nodTch</i>	CE3 complemented with plasmid pRK415:: <i>nodTch</i>	This study
CE3 <i>nodTch</i> ::Km/ pRK415:: <i>nodTch</i>	CE3 harboring a kanamycin cassette inserted in <i>nodTch</i> complemented in <i>trans</i> with plasmid pRK415:: <i>nodTch</i>	This study
CE3 <i>rmiR</i> ::Sp	CE3 carrying a spectinomycin cassette inserted in <i>rmiR</i>	This study
CE3 <i>rmiA</i> ::sp	CE3 carrying a spectinomycin cassette inserted in <i>rmiA</i>	This study
CE3 <i>rmiB</i> ::Sp	CE3 carrying a spectinomycin cassette inserted in <i>rmiB</i>	This study
CFNX184	<i>R. leguminosarum</i> bv. <i>phaseoli</i> CFN42 cured of pc	[Brom et al., 1992]
<i>E. coli</i>		
XL1Blue	<i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>hsdR17</i> , <i>supE44</i> , <i>relA1</i> , <i>lacI</i> , [F <i>proAB</i> , <i>lacIqZDM15</i> , Tn10(Tc ^r)]	Stratagene
S17.1	<i>thi</i> , <i>pro</i> , <i>hsdR</i> (rm+), <i>recA</i> (derived from <i>E. coli</i> 294); carries plasmid RP4 derivative (amp; tet::Mu; km::Tn7) integrated into chromosome Tc ^s ; Km ^s ; Ap ^r ; Tp ^r ; tra+	[Simon et al., 1983]
Plasmids		
pLARFTch	pLARF1 carrying a ca. 20-kb <i>R. etli</i> DNA fragment containing the <i>nodTch</i> region	This study
pLARFTpc	pLARF1 carrying a ca. 20-kb <i>R. etli</i> DNA fragment containing the <i>nodTpc</i> region	This study
pBluescriptSK	Phagemid cloning vector, 2.96 kb, MCS, Ap ^r	Stratagene
pHME8	pBluescriptSK carrying <i>rmiRABnodTch</i> region on a 7.8-kb <i>NotI</i> fragment	This study
pHME22	pBluescriptSK carrying <i>nodTch</i> on a 2.2-kb <i>BamHI-HindIII</i> fragment	This study
pHME40	pBluescriptSK carrying <i>nodTpc</i> on a 3.7 kb <i>EcoRI</i> fragment	This study
pRK415	Broad-host-range cloning vector; <i>plac</i> , MCS, Tc ^r	[Keen et al., 1988]
pRK415:: <i>nodTch</i>	pRK415 derivative carrying <i>nodTch</i> on a 1.8-kb <i>XhoI-BamHI</i> fragment	This study
pJQ200SK	<i>sacB</i> suicide vector Gm ^r	[Quandt and Hynes, 1993]
pHMTch::Km	pJQ200SK derivative carrying <i>nodTch</i> gene interrupted with a kanamycin cassette	This study
pHMTpc::Sp	pJQ200SK derivative carrying <i>nodTpc</i> gene interrupted with a spectinomycin cassette	This study
pHMA::Sp	pJQ200SK derivative carrying <i>rmiA</i> gene interrupted with a spectinomycin cassette	This study
pHMB::Sp	pJQ200SK derivative carrying <i>rmiB</i> gene interrupted with a spectinomycin cassette	This study
pHMR::Sp	pJQ200SK derivative carrying <i>rmiR</i> gene interrupted with a spectinomycin cassette	This study
pJQ934	pJQ200SK derivative carrying a 934-bp <i>Fsp I</i> internal fragment of <i>nodTch</i>	This study

disrupt *nodTch* by insertion of a Km resistance cassette, plasmid pHMTch::Km was constructed (table 1) and introduced into *R. etli* by biparental mating. Single recombinants (Km^r, Gm^r, Nal^r) were obtained with a 10⁻⁴ recombination frequency. A single recombinant colony was grown overnight in liquid medium and plated on solid medium in the presence of 5% sucrose with or without Km. In both cases, the frequency of the Suc^c Km^r colonies observed was 10⁻⁵. All of these populations were Gm^r, indicating that parts of the pJQ200SK vector are retained in these cells. The genomic DNA of these cells analyzed by DNA blot hybridization showed that *nodTch* did not contain an inserted cassette (data not shown). Inactivation of the *nodTch* gene by insertion of a Km antibiotic cassette was possible only when the cells carried a func-

tional *nodTch* gene in *trans* (see Experimental Procedures). The frequency of the double recombination events observed by this method was 10⁻⁴. Southern blot analysis from 18 presumptive double recombinants indicated that all insertions occurred in the *nodTch* gene and not in the complementing copy. These results also indicate that polar effects caused by the Km cassette insertion on genes located downstream of *nodTch* are unlikely. To obtain strains cured of the complementing plasmid, a double recombinant *nodTch*::Km complemented with pHMTch::Km was grown in liquid medium without antibiotic selection for several cell division cycles and then plated on agar medium with or without antibiotic selection. No bacteria cured from the vector carrying *nodTch* were obtained, even though the same plasmid was unstable in the

wild-type strain (data not shown). Hybridization analyses using *nodTch* as a probe on genomic DNA isolated from 18 colonies at the end of the experiment confirmed this result (data not shown). By all these criteria (inability to obtain a *nodTch* mutant in the absence of a complementing wild-type copy and the stability of pRK415::*nodTch* in the mutant strain but not in the wild-type), we conclude that *nodTch* in *R. etli* is required for cell survival under the conditions tested. A similar result was reported in *R. leg. bv. phaseoli* ORF471, where a *nodT* homolog was impossible to knock out [Rivilla and Downie, 1994]. In addition, *AmeC* from *A. tumefaciens* seems to behave similarly, since an *ameC* mutant was obtained only by complementation with *ameC* in *trans* and subsequently cured of the complementing plasmid. This *ameC* mutant shows a hypersensitive drug resistance phenotype. Since the pWT201 plasmid (carrying the *A. tumefaciens* *ameC* gene) and the *A. tumefaciens* At13006 strain (an *ameC*::*Km^r* mutant) [Peng and Nester, 2001] are no longer available [Nester, personal communication], it was not possible to conduct a complementation test in *R. etli*. Nevertheless, these results strongly support the idea that members of the NodT/TolC family may have alternative functions, not only involved in the transport of noxious compounds.

In contrast to *nodTch*, *nodTpc* was easily mutated by the insertion of a Sp resistance cassette using methodology similar to that described above. The mutant phenotype was analyzed during nodulation and for sensitivity to diverse noxious compounds, as described in Experimental Procedures. No changes were observed in either assay when compared to the wild-type strain. These results suggest that despite their protein identity, *nodTpc* does not complement the essential function encoded by *nodTch*. Gene duplication is the primary source of new genes. Duplicate genes that are stably preserved in genomes usually have divergent functions [He and Zhang, 2005]. A likely scenario is that once the ancestral *nodT* gene duplication occurred, the duplicate genes were more susceptible to acquire mutations, which split their activities compared to the single ancestral gene. Subsequently, one daughter gene retained the ancestral function while the other acquired new functions. This hypothesis should be further explored experimentally.

Diminished Redundancy of Outer Membrane Factor Proteins in Rhizobiales

NodT proteins are members of the TolC family [Sharff et al., 2001], belonging to the outer membrane factor family [Paulsen et al., 1997]. As classified by Pfam [Bateman

et al., 2004], these proteins are members of the outer membrane efflux protein (OEP) family. An analysis of the presence of this family in complete bacterial genomes showed a positive correlation between genome size and the number of genes with OEP domains in Gram-negative bacteria (fig. 2a), nevertheless there are some exceptions. Animal and plant pathogens such as *Burkholderia* and *Pseudomonas* contain an extremely large number of OEPs, which agrees with their elevated resistance to noxious compounds [Schweizer, 2003]. In contrast, Rhizobiales appear to encode quite fewer OEPs than those expected given their genome size. Since OEP proteins can physically and functionally interact with one or several members of the ATP-binding cassette (ABC), resistance/nodulation/cell division (RND), or major facilitator superfamily (MFS) class proteins [Paulsen et al., 1997; Putman et al., 2000; Yen et al. 2002; Zgurskaya and Nikaido, 2000], we decided also to analyze the number of proteins with these domains in our set of fully sequenced bacteria. The results indicate that the number of these transporters also correlates quite well with genome size (fig. 2b, c). Rhizobiales appear to be slightly more enriched in ABC transporters than other organisms and this result is accentuated when we compare the total number of proteins with ABC, RND or MFS domains against the number of proteins with OEP domains (fig. 2d). These results could imply that ABC, RND, and MFS transporters provide the efflux adaptability needed by these organisms in combination with an extremely versatile (and thus essential) OEP. In support of this hypothesis, OEPs are known to be multifunctional, as shown previously in *Escherichia coli*, where all RND drug exporters, in addition to some ABC and MFS drug exporters, cooperate with TolC [Nishino et al., 2003]. Alternatively, it might be advantageous for Rhizobiales not to depend on outer membrane proteins for export, given their symbiotic phase, during which the outer membrane's importance is clearly reduced in favor of the peribacteroid membrane [Udvardi and Day, 1997]. If this were the case, these organisms would only keep the most essential OEPs.

nodTch Is Located Immediately Downstream of Genes Homologous to Members of the RND Efflux Transporter Family in Gram-Negative Bacteria

To gain additional insights into the functional role of *nodTch*, we analyzed sequences flanking this gene. Sequences upstream of *nodTch* encode ORF1, 2 and 3 oriented in the same direction as *nodTch* (fig. 1a). These three ORFs were analyzed using PSI-BLAST [Altschul et al., 1997]. ORF1 encodes a protein with high similarity to

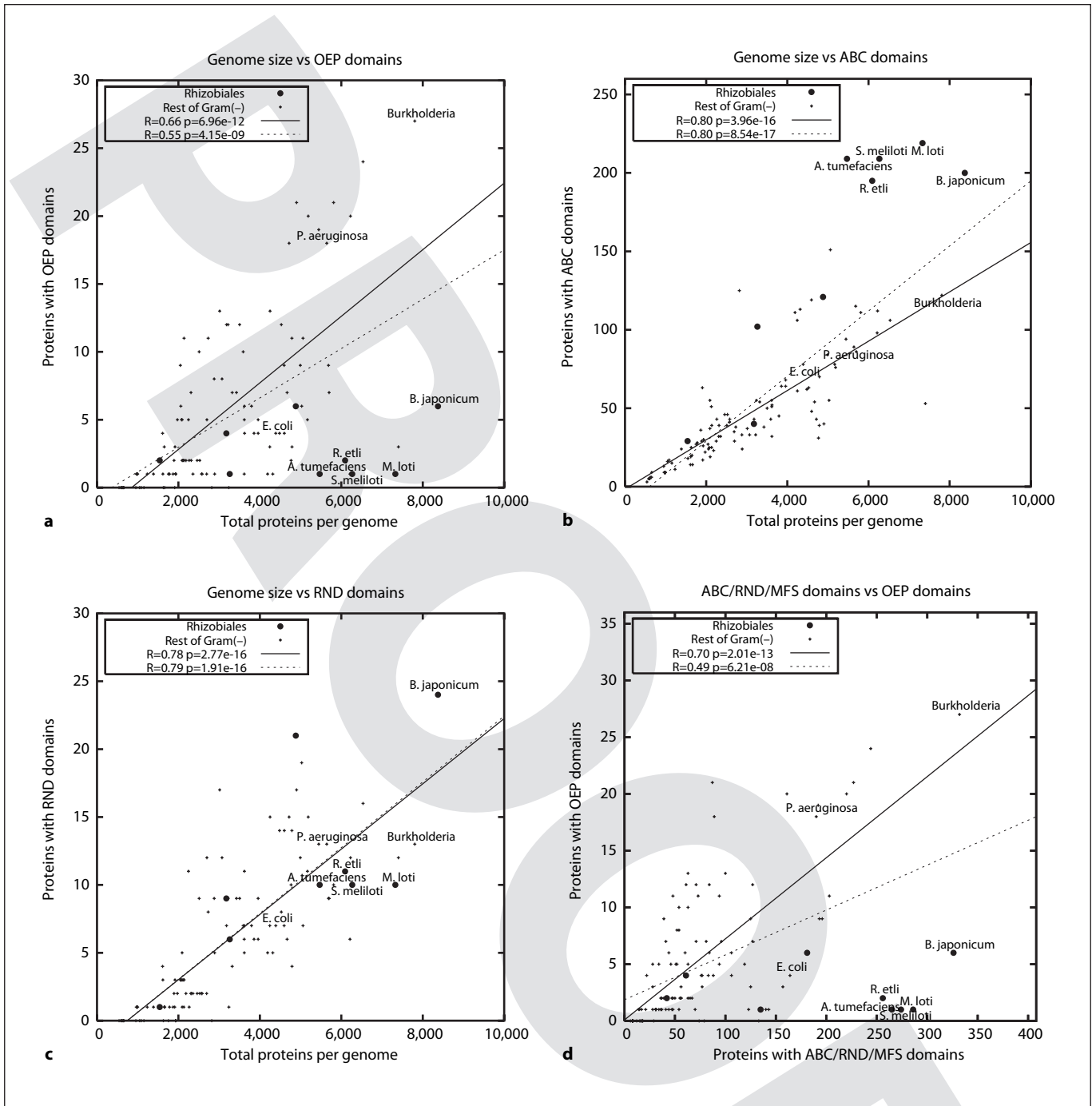


Fig. 2. Analysis of the number of proteins with OEP, ABC, RND and MFS domains in Gram-negative bacteria. Genome size correlates with the number of duplicated transporters in Gram-negative bacteria, except for members of OEP family in some α -proteobacteria. The number of total proteins encoded in each genome (genome size) significantly correlates with the number of proteins containing OEP (a), ABC (b), and RND (c) domains. Many Rhizobiales appear to have fewer proteins with OEP (a) and more

with ABC (b) domains than is expected given their genome size. This effect is amplified when comparing the number of proteins with ABC, RND or MFS domains against those with OEP domains (d). The continuous line represents the least squares fit of the data from Gram-negative organisms, excluding Rhizobiales, whilst the dotted line includes them. R is the Pearson linear correlation coefficient and p represents the probability of obtaining as good a correlation by chance.

repressor proteins of the TetR family (fig. 1c) showing the highest similarity (61%) with AmeR, which has been reported as a repressor of a putative RND-type efflux system in *A. tumefaciens*. The most conserved domain in this ORF corresponds to the DNA-binding region (data not shown). Upstream of ORF1, a putative promoter sequence was predicted (tcggcTTGGCAggttttcgctgaatagcgcTAGCATcggcaAAAGGGGTAT) with the neural network promoter prediction (NNPP) program [Reese, 2001]. ORF2 begins 154 bp downstream of ORF1. The predicted product of ORF2 exhibited high similarity to proteins of the membrane fusion protein family (fig. 1c). The best match (78% similarity) was with *A. tumefaciens* AmeA. ORF3 is 81 bp downstream of ORF2 (fig. 1b) and encodes a protein with significant similarity to members of the RND efflux transporter family; the highest match (88% similarity) was observed with AmeB of *A. tumefaciens* (fig. 1c). On the basis of the significant sequence similarity with RND-type efflux systems from Gram-negative bacteria (fig. 1c), we named ORF1, ORF2, and ORF3 as *rmiR*, *rmiA*, and *rmiB* (*Rhizobium* multidrug efflux identity), respectively.

Characterization of *R. etli* *rmiRAB* Mutants

To examine their functions, mutations were introduced into each of the *rmiRAB* genes by inserting an Ω -Sp interposon [Alexeyev et al., 1995]. To assay roles of RmiR, RmiA and RmiB in drug efflux, the strains with individual gene inactivation were tested for drug susceptibility (see Experimental Procedures). No differences were observed in these mutants when compared to the wild-type strain (data not shown). RmiRAB proteins have no detectable contribution to nodulation, because the number of nodules per plant, nodule morphology, and kinetics of nodulation were similar in all mutants and wild-type strains (data not shown).

In *A. tumefaciens*, two putative RND-type efflux systems encoded by *ifeAB* [Palumbo et al., 1998], and *ameRameABC* [Peng and Nester, 2001] have been reported, in which AmeC shows similarity to NodT and other outer membrane proteins. Mutations in *ameC* increase sensitivity to both detergents and antibiotics [Peng and Nester, 2001], while disruptions of *ameAB* do not affect sensitivity. Similarly, *R. etli* *rmiA* and *rmiB* lack a demonstrable role in drug efflux, perhaps reflecting gene redundancy. For example, in Gram-negative bacteria including *E. coli*, *P. aeruginosa* and *P. putida* [Nikaido and Zgurskaya, 2001; Poole, 2001; Ramos et al., 2002], several RND efflux systems are functional. It might be that in *R. etli*, other RND-like efflux systems or a multidrug efflux

pump different from the RND family compensate for the lack of *rmiRAB* expression. In *R. etli*, an efflux system comprised of *rmrA* and *rmrB* shows enhanced sensitivity to plant metabolites (phytoalexins, flavonoids, and salicylic acid) compared with the wild-type strain [Gonzalez-Pasayo and Martinez-Romero, 2000]. It is possible that the phenotype of *rmiAB* mutants is complemented by *rmrAB* or another efflux pump. The analysis of the *R. etli* complete genome using the GeConT program [Ciria et al., 2004] also indicates the presence of 11 paralogous RND systems. Most of them are located on the chromosome, but one is located on plasmid p42c close to *nodTpc* in the opposite orientation, and another on plasmid p42f (data not shown). The presence of several RND paralogs in *R. etli* could explain the absence of a phenotype for *rmiA* and *rmiB* mutants, because any one of them could mask the phenotype of these mutants.

nodTch Forms an Operon with *rmiRAB* and Is Also Transcribed Independently of the *rmiRABnodTch* Operon

The *nodT* genes were originally found forming part of operons with nodulation genes in *R. leg. bv. trifolii* and *viciae* [Surin et al., 1990]. In *R. leg. bv. phaseoli*, *nodT* is located outside the symbiotic plasmid and its transcriptional organization has not been defined. Sequence analysis of the *R. etli* *rmiRAB* and *nodTch* suggests that these genes could be part of a single operon. No obvious intrinsic terminator structures or Rho-dependent terminator sequences [Sharp and Platt, 1984] within the intergenic regions of *rmiR*, *rmiA*, *rmiB* and *nodTch* were predicted using the Terminator program [Brendel, 1985; Brendel and Trifonov, 1984]. Based on this analysis we propose that *rmiR*, *rmiA*, *rmiB* and *nodTch* comprise an operon. A potential Rho-independent terminator sequence (GAATACGAACCCCGCGATCGCCTCGCGGGTTT-TTTC) was found 19-bp downstream *nodTch*, suggesting that this is the final gene of the operon.

To define the transcriptional organization of *nodTch* and *rmiRAB* genes, RT-PCR products from the intergenic regions of *rmiR*, *rmiA*, *rmiB* and *nodTch* were generated using primers described in Experimental Procedures. RT-PCR products of the sizes expected for an operon were obtained (fig. 3), even though sequence analysis indicates that RmiR could be a regulator of the operon.

Interestingly, *rmiRAB* mutants were obtained by Ω -Sp interposon insertion despite the essential role observed for *nodTch*, implying that *nodTch* is also independently transcribed. To prove this, RT-PCR analysis was per-

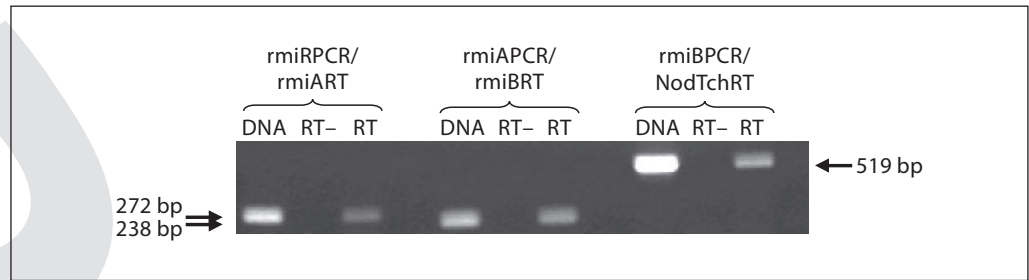


Fig. 3. RT-PCR analysis indicates that *rmiR*, *rmiA*, *rmiB* and *nodTch* form an operon. RT-PCR products obtained with *rmiR*-PCR/*rmiART*, *rmiA*-PCR/*rmiBRT* and *rmiB*-PCR/*rmiBRT* primer pairs indicate the cotranscription of *rmiR* and *rmiA*, *rmiA* and *rmiB* as well as *rmiB* and *nodTch*, respectively. Template cDNAs

were generated by Reverse Transcriptase (RT) reactions using *R. etli* CE3 total RNA. Control reactions without RT (RT-) or with pHME8 plasmid DNA carrying the complete *rmiRABCnodTch* region (DNA) are shown. Numbers indicate the observed PCR fragment sizes (basepairs).

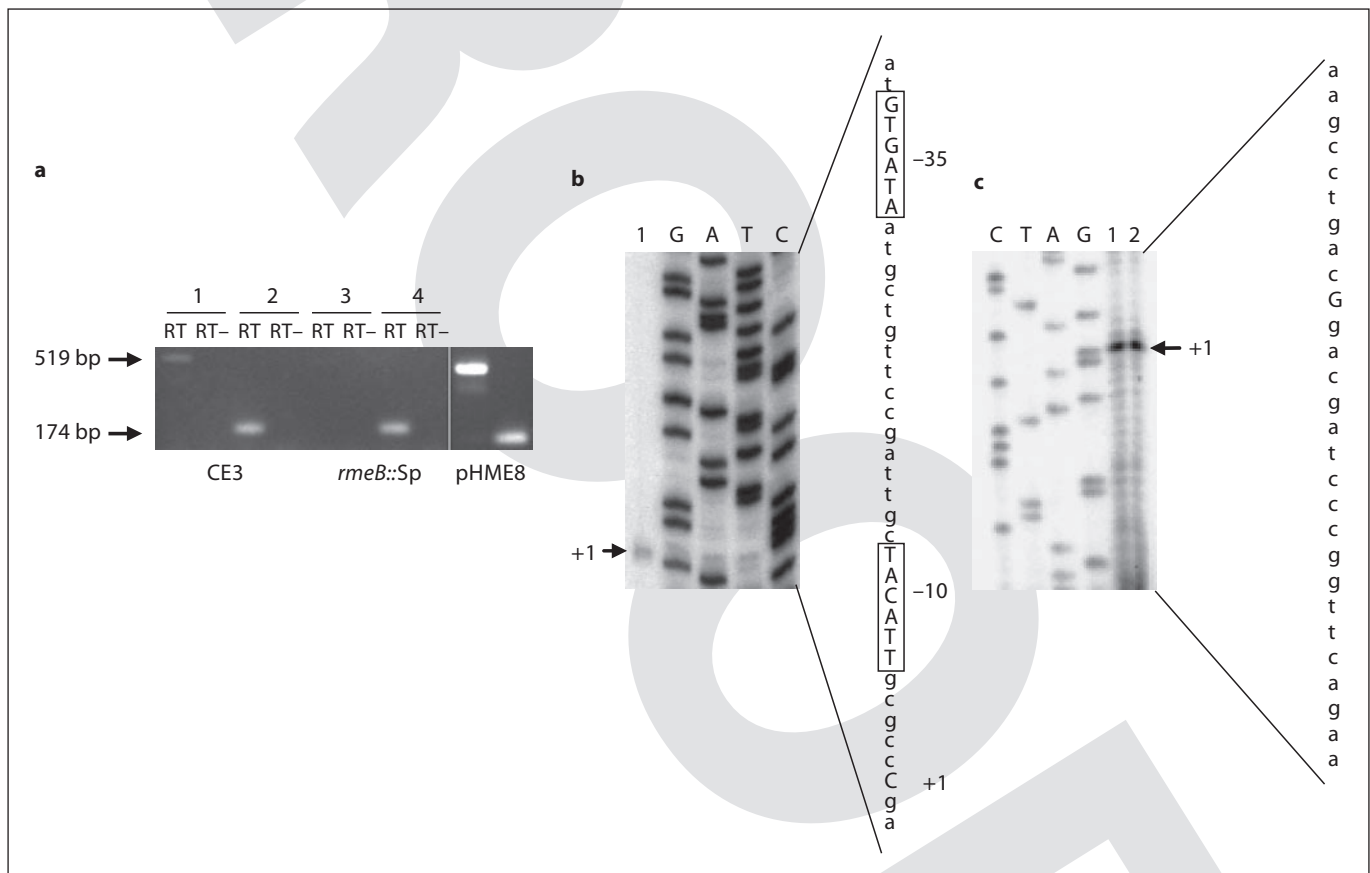


Fig. 4. *nodTch* and *nodTpc* are independently transcribed. **a** RT-PCR products amplified using pairs of primers within the intergenic *rmiB* and *nodT* region (lanes 1 and 3: *rmiBRT*/*nodTchRT*) and downstream of the putative *nodTch* promoter zone (lanes 2 and 4: *nodTchprPCR*/*nodTchRT*). cDNA was generated from a reverse transcriptase (RT) reaction using total RNA from *R. etli* CE3 and *rmiB*::*Sp* mutant as template. Similar reactions without RT (RT-), or with plasmid pHME8 (carrying *rmiRABCnodTch*) were used as controls. **b** Primer extension analysis of the *nodTch* transcription-

al start position (arrow) analyzed in total RNA isolated from *R. etli* CE3 (lane 1). **c** Primer extension analysis of the *nodTpc* transcriptional start position (arrow) using total RNA isolated from *R. etli* CE3 (lane 1). Lanes G, A, T, and C are a ³²P-labeled DNA sequencing ladder generated with the primers used for primer extension reactions (PE*nodTch* and PE*nodTpc*) and plasmid pHME22 DNA as template. Sequence upstream of *nodT* is shown on the right. The start site of *nodT* transcription is indicated by +1. The proposed -10 and -35 motifs are boxed.

formed as described in Experimental Procedures. A 519-bp PCR product containing both *rmiB* and *nodTch* regions was obtained using total RNA of the wild-type strain (fig. 4a, lane 1), but not from an *rmiB* mutant (fig. 4a, lane 3). A 174-bp PCR product corresponding to *nodTch* was observed in both strains (fig. 4a, lanes 2 and 4). These results indicate that *nodTch* is co-transcribed with *rmiB*, but also expressed from a putative promoter located between *rmiB* and *nodTch*. As expected, the same results were obtained using *rmiR* and *rmiA* mutants (data not shown).

Documentation of independent transcription was further confirmed by primer extension analysis, in which a transcription initiation site was mapped at 106 bp upstream of the *nodTch* AUG codon (fig. 4b). Using the NNPP program, a putative region with identity to σ^{70} promoters was identified (fig. 4b). Collectively, these results suggest that *nodTch* is expressed constitutively and independently of the entire operon.

RT-PCR analysis indicated that *nodTpc* is transcribed independently of other ORFs located upstream (data not shown). Primer extension analysis mapped a transcription start site located 28 bp upstream of the predicted *nodTpc* GUG start codon (fig. 4c) that is utilized constitutively under our growth conditions. Although *nodTpc* is constitutively transcribed under the conditions tested, it does not complement the essential function of *nodTch*, therefore these two genes are more likely to be involved in different cellular functions as discussed above.

Subcellular Localization of NodTch

NodTch is encoded in an operon that resembles a multidrug efflux pump. In Gram-negative bacteria, these pumps are tripartite complexes comprising a proton antiporter protein, an inner membrane adaptor protein, and an outer membrane protein. Extensive sequence analysis suggests that NodT is an outer membrane lipoprotein of the TolC superfamily. To determine NodTch localization, *R. etli* subcellular fractions (fig. 5a) were prepared and analyzed by Western blotting. A polyclonal antibody raised to *R. leg. bv. viciae* NodT (provided by Dr. A. Downie) that also recognized *R. leg. bv. phaseoli* NodT [Rivilla and Downie, 1994] was used to immunodetect NodTch in *R. etli* in protein extracts. To confirm specificity, NodTch was immunodetected with this antibody when induced with IPTG in *E. coli* harboring pRK415::*nodTch* (fig. 5c, lane 7). A protein of the expected size (50 kDa) was immunodetected in the fraction enriched for outer membrane proteins [Matsuyama et al., 1984; Puente et al., 1995; Schnaitman, 1971] and in total protein

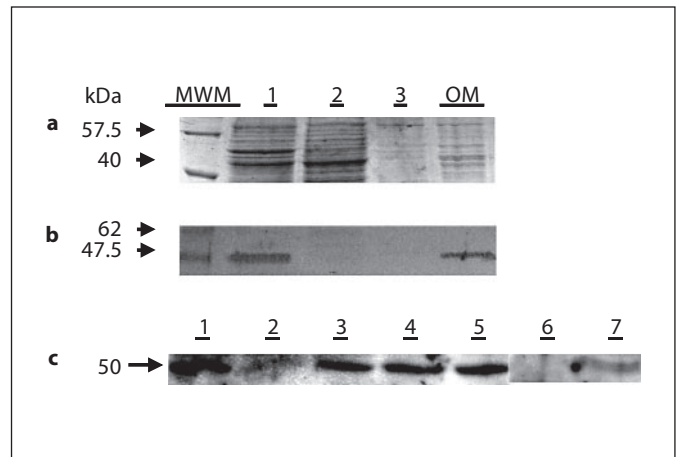


Fig. 5. **a** *R. etli* NodT is located in the outer membrane enriched fraction and is expressed in *rmiR*::Sp, *rmiA*::Sp, and *rmiB*::Sp null mutants and in the CE3/*nodT*::Km conditional null mutant. Samples from different *R. etli* CE3 fractions were analyzed on SDS-PAGE gels and stained by Coomassie brilliant blue: lane 1 = total extract; lane 2 = soluble fraction; lane 3 = Triton X-100 soluble fraction; OM = Triton X-100 insoluble fraction considered to be an outer membrane protein enriched fraction (15, 22, 27). **b** Western blot analysis of samples obtained as in **a** reacted with a rabbit anti-NodT antibody (19). **c** Western blot analysis of total protein extracts from *R. etli* CE3 (lane 1), CE3/*nodTch*::Km/pRK415::*nodTch* (lane 2); *rmiA*::sp (lane 3); *rmiB*::Sp (lane 4); *rmiR*::Sp (lane 5) using anti-NodT antibody. Equivalent amounts of total protein were loaded per lane. *E. coli* total protein fraction from a strain carrying pRK415::*nodTch* uninduced (lane 6) or induced with 10 mM IPTG (lane 7). M: Protein molecular weight marker standards. Numbers with arrows indicate the calculated molecular mass of the proteins detected.

extracts (fig. 5b, lanes 1 and OM, respectively). These data support the idea that NodTch is an outer membrane protein. The same antibody recognized NodTch in total protein extracts from each *rmiRAB* mutant, corroborating that NodTch is individually produced (fig. 5c, lanes 3–5). NodT was synthesized in an *R. etli nodTch* conditional mutant complemented with pRK415::*nodTch* (carrying *nodTch*; fig. 5c, lane 2). The antibody recognizes NodTch exclusively, since only one protein of 50 kDa, pI 6.1, was observed after two-dimensional SDS-PAGE of total protein extracts from *R. etli* wild-type or CFNX184 [Brom et al., 1992], where plasmid c is absent (data not shown). It is possible that the GUG start codon proposed for *nodTpc* might affect the translational rate of NodTpc, affecting the level of the protein. The NodTpc concentration could thus be insufficient to complement the absence of NodTch.

In summary, we have strong experimental evidence showing that there are two *nodT* homologs in *R. etli* located outside the symbiotic plasmid, *nodTch* and *nodTpc*. *nodTch* is an essential gene in *R. etli* and this gene forms part of a new operon, *rmiRABnodTch*, although *nodTch* can also be independently transcribed. NodTch was found exclusively in outer membrane enriched fractions. We propose that NodTch may play a role in outer membrane stability or in chromosome segregation instead of (or in addition to) drug efflux, since the number of OEP members present in the *R. etli* genome is exclusively limited to *nodTch* and *nodTpc*; however additional experimental evidence is needed to explore these possibilities. *nodTpc* appears to be constitutively transcribed under the growth conditions tested, although the protein was not readily detected. Mutations in this gene do not show a phenotype in nodulation or in detoxification. A viable *nodTch* mutant and a viable *nodTch/nodTpc* double mutant are required to define the roles of both genes in *R. etli*, although it is challenging to select a conditional *nodTch* allele, given its essential nature.

Experimental and Computational Procedures

Bacterial Strains, Plasmids and Growth Conditions

The bacterial strains and plasmids used in this study are shown in table 1. *R. etli* strains were grown in peptone yeast (PY) broth [Noel et al., 1984] at 30°C. *E. coli* were grown in Luria-Bertani broth [Sambrook et al., 1989] at 37°C. Plasmids were maintained in *E. coli* with appropriate antibiotic selection: plasmid pBluescript II KS(+) (Stratagene), 100 µg/ml ampicillin (Amp); plasmid pRK415 [Keen et al., 1988] and derivatives, 10 µg/ml tetracycline (Tc); plasmid pJQ200SK [Quandt and Hynes, 1993] 30 µg/ml gentamicin (Gm); pHMTch::Km (this work) 30 µg/ml Gm and 30 µg/ml kanamycin (Km); pHMA::Sp, pHMB::Sp, and pHMR::Sp 30 µg/ml Gm and 100 µg/ml spectinomycin (Sp). *R. etli* wild-type and derivatives were selected using appropriate combinations of the following antibiotics: 20 µg/ml nalidixic acid (Nal); 100 µg/ml streptomycin (Sm), 100 µg/ml Sp; 30 µg/ml Km, 30 µg/ml Gm or 5 µg/ml Tc.

Nucleotide Sequencing, Homology and Phylogenetic Analysis

Nucleotide sequencing of plasmid DNA or PCR products was performed on an Applied Biosystems Inc. model 373 sequencer at the Instituto de Biotecnología-UNAM Sequencing Facility using universal or custom primers. DNA sequences were assembled and processed using the AbiView (Tim Carver, MRC Rosalind Franklin Centre for Genomics Research Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SB, UK) and BioEdit (Tom Hall, Ibis Therapeutics) software packages. The BLAST algorithm was used to search for similarity with DNA and protein databases [Altschul et al., 1997].

Promoter and transcription terminator motif predictions were carried out using the Neural Network Promoter Prediction soft-

ware (http://www.fruitfly.org/seq_tools/promoter.html) and the Terminator program of the GCG package [Brendel, 1985; Brendel and Trifonov, 1984], respectively.

Genomes and Protein Family Assignations

Available wholly sequenced eubacterial genomes were taken from GenBank [Benson et al., 2005]. From the original 270, a reduced set of 166 were obtained by eliminating organisms that shared more than 80% of orthologous proteins. In order to find out which proteins belonged to different families, Pfam models were used [Bateman et al., 2004], consisting of the OEP (PF02321), the ABC transporter (PF00005), the MFS (PF07690), and the AcrB/AcrD/AcrF family (PF00873) which is the closest available model for RND proteins. Every protein from the 166 non-redundant bacteria was analyzed with these models using HMMER (<http://hmmer.wustl.edu/>). In the case of the OEP model, only proteins containing two OEP domains were considered.

Correlations between Genome Size and Particular Protein Families

The analysis was performed with ad hoc Perl scripts, making use of BioPerl [Stajich et al., 2002], gnuplot software (<http://www.gnuplot.info/>) and several Perl modules (<http://www.cpan.org/>). The Pearson linear correlation coefficient (R) was calculated with the Statistics::Descriptive module and its p value was estimated by performing 1,000 permutations of the x axis data, calculating the mean and standard deviation of the new R, from which z scores were obtained. The z scores were converted into p values using the Statistics::Distributions module.

Genetic and Molecular Techniques

Basic DNA procedures, including restriction endonuclease digestions, ligations, transformations, and agarose gel electrophoresis were performed as described [Sambrook et al., 1989]. Restriction endonucleases (New England Biolabs Inc.), T4 DNA ligase (Roche Diagnostics GmbH), AMV Reverse Transcriptase (Roche Diagnostics GmbH), and Platinum *Taq* polymerase (Invitrogen, Life Technologies) were acquired commercially and used according to the supplier's recommendations. The alkaline lysis method [Sambrook et al., 1989], a Plasmid Midi kit (Qiagen Inc.) or High Pure Plasmid Purification kit (Boehringer Mannheim Inc.) were used to isolate *R. etli* genomic DNA and plasmids from *E. coli* strains. DNA fragments were purified from agarose gels using the GeneClean kit (Bio 101, Inc.) and radiolabeled using the *rediprime* II random prime labeling system (Amersham Pharmacia Biotech) according to the manufacturer's instructions. Total RNA was isolated from log-phase cultures of *R. etli* using the RNeasy Mini Kit (Qiagen) or the method of Chomczynski and Sacchi [1987] and treated with RNase-free DNase (Roche; 2 U enzyme/µg RNA for 60 min at 37°C). A 5-µg sample of DNase-treated RNA was used as a template for RT-PCR or primer extension with the Roche or Biolab RT-PCR kit according to the protocol supplied by the manufacturer.

R. etli nodT homologous genes were isolated from an *R. etli* genomic library [Huerta-Zepeda et al., 1997]. Colony hybridization was performed using the *R. leg. bv. phaseoli nodT* gene as a probe labeled with ³²P, using standard protocols [Sambrook et al., 1989]. A cosmid clone carrying *nodTch* was selected and named pLARFTch. This cosmid was digested with several restriction enzymes. A 2.2-kb *HindIII-BamHI* fragment carrying the complete

nodTch sequence and a 7.8-kb *NotI* fragment carrying the entire chromosomal region, including *rmiRAB* and *nodTch* genes were isolated and subcloned into pBluescript SK to construct plasmids pHME22 and pSK8, respectively. For cloning *nodTch* in an expression vector, a 1.8-kb *XhoI*-*Bam*HI fragment derived from pHME22 was subcloned into pRK415 to yield pRK415::*nodTch*.

A cosmid clone carrying *nodTpc* was selected and named pLARFTpc. A 3,703-bp *Eco*RI fragment was subcloned into pBluescript SK to construct pHME40.

To obtain a *nodTch* disruption in a single recombination step, a 934-bp *Fsp*I fragment obtained from pHME22 was subcloned into suicide vector pJQ200SK, generating the pJQ934 vector. This vector was introduced into *E. coli* strain S17.1. Attempts to introduce this construct into *R. etli* were performed by bacterial mating, without success.

To obtain a conditional lethal *nodTch* mutant, a 1.8-kb *XhoI*-*Bam*HI fragment carrying *nodTch* was subcloned into pBluescript SK. An Ω -Km interposon element [Alexeyev et al., 1995] was introduced into an *Aat*II site located in the intragenic region of *nodTch*. The interrupted gene was subcloned into pJQ200SK, and the resulting plasmid (pHMTch::Km) was introduced into *R. etli* by biparental mating. Single recombinant strains were selected using the antibiotic resistance marker. Plasmid pRK415::*nodTch* in *E. coli* S17.1 was then introduced into one transconjugant. A transconjugant carrying pRK415::*nodTch* was grown overnight and plated on solid PY media supplemented with 10% (w/v) sucrose to induce a double recombination process. Merodiploid cells were obtained at a recombination frequency of 10^{-4} . Plasmid pRK415::*nodTch* was also introduced into *R. etli* wild-type. The stability of pRK415::*nodTch* was estimated in *R. etli* wild-type and in the conditional lethal *nodTch* mutant by measuring plasmid inheritance over a 10-day period (approx. 60 generations). To do this, stationary-phase cultures were diluted in fresh medium without any antibiotic selection to an optical density of 0.05 at 620 nm. An aliquot of this culture was removed every six generations. An inoculum of each of these time samples was serially diluted and plated onto solid medium in the following conditions: (i) without antibiotic selection; (ii) in the presence of Nal to select *Rhizobium* strains; (iii) in Km to select double recombinants, and (iv) in Tc to isolate strains carrying the complementing plasmid. The number of colonies obtained under each condition was counted. The presence of the plasmid during these growth experiments was also confirmed by Southern hybridization using total bacterial DNA samples.

To obtain *nodTpc* as well as *rmiR* and *rmiA* and *rmiB* mutants, appropriate restriction fragments (2.2 kb *Eco*RI-*Bam*HI for *nodTpc*, 0.8 kb *Eco*RI for *rmiR*, 1.2 kb *Hind*III for *rmiA* and 2.6 kb *Xho*I for *rmiB*) from plasmid pLARFTch were subcloned into pBluescript SK. An Sp resistance cassette flanked with Ω elements [Alexeyev et al., 1995] was introduced independently into the *nodTpc*, *rmiR*, *rmiA* or *rmiB* genes. The four interrupted genes were then subcloned into the suicide vector pJQ200SK. The resultant plasmids pHMR::Sp, pHMA::Sp or pHMB::Sp, pHMTpc::Sp (table 1) were introduced into *E. coli* S17.1 by transformation and mobilized into the wild-type *R. etli* strain via conjugation [Simon et al., 1983]. Transconjugants were selected on PY agar containing Gm and Nal for counterselection. Transconjugants were then plated onto PY agar containing 10% (w/v) sucrose. Double recombinants were selected in the presence of appropriate antibiotics. To confirm the insertion of antibiotic resistant cassettes, South-

ern hybridizations were carried out using *nodTpc* and *rmiRAB* genes or the Sp resistance cassette as probes to total bacterial DNA [Sambrook et al., 1989].

PCR, RT-PCR and Primer Extension Analysis

PCRs were performed using specific primer pairs designed to amplify internal regions of *rmiR* and *rmiA* (*rmiR*PCR 5' GGGAC-GACAAGGAGACAC 3' and *rmiA*PCR 5' AGGGCCGAG-GAGATGAAC 3'),

rmiA and *rmiB* (*rmiA*PCR 5'AAACCAGTCCCGATGAAG 3' and *rmiB*PCR 5' GAAATCGACAGCGTGAAG 3'), *rmiB* and *nodTch* (*rmiB*PCR: 5' AAGGACCGGATGGAAAGC 3' and *nodTch*RT: 5' ATTTTCAGGAGGCATGA 3') or from *nodTch* and the putative promoter region of *nodTch* (*nodTch*RT: 5' ATTTTCAGGAGGCATGA 3' and *nodTch*prPCR: 5' TTCGGAGTGGGT-GTTATT3'). cDNA was obtained using AMV Reverse Transcriptase (Roche Diagnostics GmbH) according to the recommendations of the supplier. Amplification was performed using the following conditions: one cycle of heating at 94°C for 3 min; 35 cycles at 94°C for 30 s, 52°C for 45 s, and 72°C for 45 s; followed by a final extension time at 72°C for 10 min. *nodTch*RT and *nodTpc*PE (5' ATTCTTGTGCGAAGCGAC3'), complementary oligonucleotides, were also used in primer extension experiments. Briefly, primers (2 μ M) were radiolabeled in a 50- μ l reaction containing 5 \times Forward Reaction Buffer, 10 μ l; T4 Polynucleotide Kinase (Gibco BRL) 2 μ l (20 units); [γ - 32 P]ATP (10 μ Ci/ μ l, 3,000 Ci/mmol), 5 μ l; and sterile water, 28 μ l. These primers were used in separate RT reactions. The generated cDNA was analyzed by co-electrophoresis in a 6% denaturing polyacrylamide gel with a DNA sequencing ladder generated from an appropriate genomic clone using the same oligonucleotides described above.

Susceptibility Tests

Susceptibility to antibiotics (ampicillin, carbanecillin, cloramphenicol, gentamicin, nalidixic acid, tetracycline and kanamycin), detergents (SDS, sarkosyl, Triton X-100), and dyes (ethidium bromide, acriflavine and acridine orange) was analyzed in PY broth containing different concentrations of these compounds. Susceptibilities to plant phenolic compounds or phenylpropanoids, such as naringenin, catechol, ellagic acid, gallic acid, and the plant antibiotic catechin (+) were also tested. Data were recorded as MICs, which reflects the lowest concentrations of antibiotic inhibiting visible cell growth after 48 h of incubation at 30°C. Additionally, drug susceptibilities were also checked by placing small pieces of 3 mm Whatmann paper embedded with several drug concentrations onto PY agar spread with tester bacterial suspensions to observe the formation of inhibition halos. In some cases, susceptibility tests were performed in liquid media using serial dilutions to measure culture growth in the presence of varying drug concentrations.

Isolation of an Outer Membrane Fraction from *R. etli*

Fractions enriched for *R. etli* outer membrane were obtained as previously described [Matsuyama et al., 1984; Puente et al., 1995; Schnaitman, 1971]. Briefly, a cell culture was grown overnight in the presence of appropriate antibiotics. Cell were collected by centrifuging at (10,000 g for 5 min) and then resuspended in PBS buffer (pH 7.4), sonicated twice for 30 s at setting 3 (Branson Sonifier 200) and centrifuged to separate cell debris. Cleared supernatants were then centrifuged at 10,000 g for 1 h to

separate the cytosolic (supernatant) and membrane proteins (pellet). The pellets were resuspended in PBS buffer (pH 7.4) in the presence of Triton X-100 2% (v/v) and incubated at 37°C for 1 h followed by centrifugation (10,000 g for 1 h) to separate the insoluble (outer membrane) and soluble (inner membrane and periplasmic) fractions. Purified fractions were evaluated by electrophoresis on SDS-PAGE gels as described [Sambrook et al., 1989].

Nucleotide Sequence Accession Number

The nucleotide sequences reported in this paper have been submitted to the GenBank DNA database under Accession No. AY786417 for *rmiRABnodTch* and DQ115788 for *nodTpc*.

References

- Alexeyev MF, Shokolenko IN, Croughan TP: Improved antibiotic-resistance gene cassettes and omega elements for *Escherichia coli* vector construction and in vitro deletion/insertion mutagenesis. *Gene* 1995;160:63–67.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–3402.
- Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR: The Pfam protein families database. *Nucleic Acids Res* 2004;32:D138–141.
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL: GenBank. *Nucleic Acids Res* 2005;33:D34–38.
- Brendel V: Mapping of transcription terminators of bacteriophages phi X174 and G4 by sequence analysis. *J Virol* 1985;53:340–342.
- Brendel V, Trifonov EN: A computer algorithm for testing potential prokaryotic terminators. *Nucleic Acids Res* 1984;12:4411–4427.
- Brom S, Garcia de los Santos A, Stepkowsky T, Flores M, Davila G, Romero D, Palacios R: Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. *J Bacteriol* 1992;174:5183–5189.
- Broughton WJ, Jabbouri S, Perret X: Keys to symbiotic harmony. *J Bacteriol* 2000;182:5641–5652.
- Canter Cremers H, Spaink HP, Wijffjes AH, Pees E, Wijffelman CA, Okker RJ, Lugtenberg BJ: Additional nodulation genes on the Sym plasmid of *Rhizobium leguminosarum biovar viciae*. *Plant Mol Biol* 1989;13:163–174.
- Cardenas L, Dominguez J, Santana O, Quinto C: The role of the *nodI* and *nodJ* genes in the transport of Nod metabolites in *Rhizobium etli*. *Gene* 1996;173:183–187.
- Ciria R, Abreu-Goodger C, Morett E, Merino E: GeConT: gene context analysis. *Bioinformatics* 2004;20:2307–2308.
- Chomczynski P, Sacchi N: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987;162:156–159.
- Denarie J, Debelle F, Prome JC: Rhizobium lipochitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. *Annu Rev Biochem* 1996;65:503–535.
- Gonzalez V, Bustos P, Ramirez-Romero MA, Medrano-Soto A, Salgado H, Hernandez-Gonzalez I, Hernandez-Celis JC, Quintero V, Moreno-Hagelsieb G, Girard L, Rodriguez O, Flores M, Cevallos MA, Collado-Vides J, Romero D, Davila G: The mosaic structure of the symbiotic plasmid of *Rhizobium etli* CFN42 and its relation to other symbiotic genome compartments. *Genome Biol* 2003;4:R36.
- Gonzalez V, Santamaria RI, Bustos P, Hernandez-Gonzalez I, Medrano-Soto A, Moreno-Hagelsieb G, Janga SC, Ramirez MA, Jimenez-Jacinto V, Collado-Vides J, Davila G: The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc Natl Acad Sci USA* 2006;103:3834–3839.
- Gonzalez-Pasayo R, Martinez-Romero E: Multi-resistance genes of *Rhizobium etli* CFN42. *Mol Plant Microbe Interact* 2000;13:572–577.
- He X, Zhang J: Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics* 2005;169:1157–1164.
- Huerta-Zepeda A, Ortuno L, Du Pont G, Duran S, Lloret A, Merchant-Larios H, Calderon J: Isolation and characterization of *Rhizobium etli* mutants altered in degradation of asparagine. *J Bacteriol* 1997;179:2068–2072.
- Inouye M: Lipoprotein of the outer membrane of *Escherichia coli*. *Biomembranes* 1979;10:141–208.
- Keen NT, Tamaki S, Kobayashi D, Trollinger D: Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene* 1988;70:191–197.
- Koronakis V, Li J, Koronakis E, Stauffer K: Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Mol Microbiol* 1997;23:617–626.
- Lewis-Henderson WR, Djordjevic MA: nodT, a positively-acting cultivar specificity determinant controlling nodulation of *Trifolium subterraneum* by *Rhizobium leguminosarum biovar trifolii*. *Plant Mol Biol* 1991;16:515–526.
- Matsuyama S, Inokuchi K, Mizushima S: Promoter exchange between *ompF* and *ompC*, genes for osmoregulated major outer membrane proteins of *Escherichia coli* K-12. *J Bacteriol* 1984;158:1041–1047.
- Nikaido H, Zgurskaya HI: AcrAB and related multidrug efflux pumps of *Escherichia coli*. *J Mol Microbiol Biotechnol* 2001;3:215–218.
- Nishino K, Yamada J, Hirakawa H, Hirata T, Yamaguchi A: Roles of TolC-dependent multidrug transporters of *Escherichia coli* in resistance to beta-lactams. *Antimicrob Agents Chemother* 2003;47:3030–3033.
- Noel KD, Sanchez A, Fernandez L, Leemans J, Cevallos MA: *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. *J Bacteriol* 1984;158:148–155.
- Palumbo JD, Kado CI, Phillips DA: An isoflavonoid-inducible efflux pump in *Agrobacterium tumefaciens* is involved in competitive colonization of roots. *J Bacteriol* 1998;180:3107–3113.
- Paulsen IT, Park JH, Choi PS, Saier MH, Jr: A family of gram-negative bacterial outer membrane factors that function in the export of proteins, carbohydrates, drugs and heavy metals from gram-negative bacteria. *FEMS Microbiol Lett* 1997;156:1–8.
- Peng WT, Nester EW: Characterization of a putative RND-type efflux system in *Agrobacterium tumefaciens*. *Gene* 2001;270:245–252.
- Poole K: Multidrug efflux pumps and antimicrobial resistance in *Pseudomonas aeruginosa* and related organisms. *J Mol Microbiol Biotechnol* 2001;3:255–264.

Acknowledgements

A. Hernández-Mendoza was supported by CONACyT grant 90288. We thank A. Downie for providing polyclonal anti-NodT antibody and R. Grande for helpful assistance. R. Ciria, A. Martínez, J.M. Hurtado and A. Ocadiz provided computational assistance. We are indebted to Drs. H. Peralta, G. Gosset, and E. Merino for providing critical comments on the manuscript. A. Hernández-Mendoza gratefully acknowledges Virginia Walbot for critically reading the manuscript and for helpful discussion.

- Puente JL, Juarez D, Bobadilla M, Arias CF, Calva E: The *Salmonella ompC* gene: structure and use as a carrier for heterologous sequences. *Gene* 1995;156:1–9.
- Putman M, van Veen HW, Konings WN: Molecular properties of bacterial multidrug transporters. *Microbiol Mol Biol Rev* 2000;64:672–693.
- Quandt J, Hynes MF: Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. *Gene* 1993;127:15–21.
- Ramos JL, Duque E, Gallegos MT, Godoy P, Ramos-Gonzalez MI, Rojas A, Teran W, Segura A: Mechanisms of solvent tolerance in gram-negative bacteria. *Annu Rev Microbiol* 2002;56:743–768.
- Reese MG: Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput Chem* 2001;26:51–56.
- Rivilla R, Downie JA: Identification of a *Rhizobium leguminosarum* gene homologous to nodT but located outside the symbiotic plasmid. *Gene* 1994;144:87–91.
- Rivilla R, Sutton JM, Downie JA: *Rhizobium leguminosarum* NodT is related to a family of outer-membrane transport proteins that includes TolC, PrtF, CyaE and AprF. *Gene* 1995;161:27–31.
- Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual*, ed 2. New York, Cold Spring Harbor Laboratory Press, 1989.
- Schnaitman CA: Solubilization of the cytoplasmic membrane of *Escherichia coli* by Triton X-100. *J Bacteriol* 1971;108:545–552.
- Schweizer HP: Efflux as a mechanism of resistance to antimicrobials in *Pseudomonas aeruginosa* and related bacteria: unanswered questions. *Genet Mol Res* 2003;2:48–62.
- Sharff A, Fanutti C, Shi J, Calladine C, Luisi B: The role of the TolC family in protein transport and multidrug efflux. From stereochemical certainty to mechanistic hypothesis. *Eur J Biochem* 2001;268:5011–5026.
- Sharp JA, Platt T: Rho-dependent termination and concomitant NTPase activity requires a specific, intact RNA region. *J Biol Chem* 1984;259:2268–2273.
- Simon R, Prifer U, Pühler A: A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-positive bacteria. *Biotechnology* 1983;1:784–791.
- Spaink HP, Wijffjes AH, Lugtenberg BJ: *Rhizobium* NodI and NodJ proteins play a role in the efficiency of secretion of lipochitin oligosaccharides. *J Bacteriol* 1995;177:6276–6281.
- Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigian C, Fuellen G, Gilbert JG, Korf I, Lapp H, Lehvaslaiho H, Matsalla C, Mungall CJ, Osborne BI, Pocock MR, Schattner P, Senger M, Stein LD, Stupka E, Wilkinson MD, Birney E: The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* 2002;12:1611–1618.
- Sullivan JT, Ronson CW: Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proc Natl Acad Sci USA* 1998;95:5145–5149.
- Surin BP, Watson JM, Hamilton WD, Economou A, Downie JA: Molecular characterization of the nodulation gene, nodT, from two biovars of *Rhizobium leguminosarum*. *Mol Microbiol* 1990;4:245–252.
- Udvardi MK, Day DA: Metabolite transport across symbiotic membranes of legume nodules. *Annu Rev Plant Physiol Plant Mol Biol* 1997;48:493–523.
- Vazquez M, Davalos A, de las Penas A, Sanchez F, Quinto C: Novel organization of the common nodulation genes in *Rhizobium leguminosarum* bv. *phaseoli* strains. *J Bacteriol* 1991;173:1250–1258.
- Villalobos MA, Nava N, Vazquez M, Quinto C: Nucleotide sequence of the *Rhizobium etli nodS* gene. *Gene* 1994;150:201–202.
- Wandersman C, Deleplaire P: TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *Proc Natl Acad Sci USA* 1990;87:4776–4780.
- Yen MR, Peabody CR, Partovi SM, Zhai Y, Tseng YH, Saier MH: Protein-translocating outer membrane porins of Gram-negative bacteria. *Biochim Biophys Acta* 2002;1562:6–31.
- Zgurskaya HI, Nikaido H: Bypassing the periplasm: reconstitution of the AcrAB multidrug efflux pump of *Escherichia coli*. *Proc Natl Acad Sci USA* 1999;96:7190–7195.
- Zgurskaya HI, Nikaido H: Cross-linked complex between oligomeric periplasmic lipoprotein AcrA and the inner-membrane-associated multidrug efflux pump AcrB from *Escherichia coli*. *J Bacteriol* 2000;182:4264–4267.