UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
DOCTORADO EN CIENCIAS BIOMÉDICAS
CENTRO DE CIENCIAS GENÓMICAS
PROGRAMA DE INGENIERÍA GENÓMICA

IDENTIFICACIÓN Y CARACTERIZACIÓN DE LOS MECANISMOS DE HOMEOSTASIS DE COBRE EN LA BACTERIA *Rhizobium etli* CFN42

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:
ANTONIO GONZÁLEZ SÁNCHEZ

Director de tesis:
DR. ALEJANDRO GARCÍA DE LOS SANTOS
Centro de Ciencias Genómicas

Miembros del comité tutor:
DRA. SUSANA BROM KLANNER
Centro de Ciencias Genómicas

DR. DANIEL GENARO SEGURA GONZÁLEZ
Instituto de Biotecnología

Cuernavaca, Morelos. Junio 2018

El presente trabajo se desarrolló en el Centro de Ciencias Genómicas de la UNAM, en el Programa de Ingeniería Genómica, bajo la dirección del Dr. Alejandro García de los Santos y forma parte de los estudios del Programa de Doctorado en Ciencias Biomédicas

Durante la realización de este trabajo se contó con el apoyo de manutención del Consejo Nacional de Ciencia y Tecnología, del Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica-UNAM IN209815 y del Programa de Apoyo para estudios de Posgrado.

AGRADECIMIENTOS

A Dios por la vida otorgada y por brindarme la oportunidad de terminar un objetivo más en mi vida personal.

A mis Padres por su apoyo incondicional, enseñándome a nunca rendirme por más difícil que parezca la situación, siendo un gran ejemplo de admiración para mí.

A mis hermanos por el apoyo y cariño.

A Mariana por su enorme apoyo y cariño que cada día hemos compartido juntos, quien me enseñó a pensar en grande y a luchar por los sueños.

A mi jefe Alejandro, por abrirme la puerta del laboratorio y darme la oportunidad de realizar mis estudios de doctorado bajo su tutoría.

A los integrantes del laboratorio con quien viví grandes momentos

A David, Susy y Pablo por sus observaciones y puntos de vista del proyecto, a lo largo de estos 5 años que he estado en el laboratorio.

A CONACYT, por la beca otorgada, así como al Programa de Doctorado en Ciencias Biomédicas, al Centro de Ciencias Genómicas y a DGAPA.

ABSTRACT

Trace concentrations of copper are essential to fulfill the requeriments of aerobic organisms. The ability to switch between two oxidation states Cu^{+} (cuprous) and Cu^{2+} (cupric) make this metal an ideal cofactor for the enzymes that catalyze the transfer of electrons in redox reactions. To maintain the trace concentrations of Cu^{+}/Cu^{2+}, and avoid toxicity organisms have evolved numerous molecular mechanisms involved in uptake, intracelular transit and efflux excess of copper. In Gram negative bacteria the best characterized factors of Cu homeostasis are ATPases, RND-type proteins, two-component regulators and cytoplasmic/ periplasmic chaperones. These proteins participate in the expulsion of Cu^{+}/Cu^{2+} excess. Based on uptake models of other metals and studies in *Mycobacterium tuberculosis*, it is assumed that the entry of Cu can be carried out by outer membrane proteins (porins). These proteins facilitate the translocation of the metal through the membrane in a specific or non-specific way. Bacteria of the order Rhizobiales are a very diverse group of alfa-proteobacteria which groups facultative diazotrophs could form symbiosis with leguminous plants; as well as the intracellular pathogens of plants and animals. It is assumed that there is an excess of copper in their habitat due to the application of fungicides in agricultural crops and to the toxic concentration of copper within the phagosome of animal cells. However, their copper homeostasis mechanisms have been little studied.

In this work we carry out the identification and characterization of the *ropAe* gene that codes for a possible porin involved in the entrance of copper to the bacterium *Rhizobium etli* CFN42.

RESUMEN

El cobre, en concentraciones traza, es un metal esencial para los organismos aeróbicos. La capacidad para cambiar entre sus dos estados de oxidación Cu^{+} (cuproso) y Cu^{2+} (cúprico); hacen de este metal un cofactor ideal para la enzimas que catalizan la transferencia de electrones en reacciones redox. El aumento o disminución en su concentración resulta dañino; por lo que las células poseen mecanismos de homeostasis para mantener un control en la captación, ingreso, tránsito y desintoxicación de cobre y otros metales. Los mecanismos de homeostasis de cobre mejor caracterizados en bacterias Gram negativas son: ATPasas, proteínas de tipo RND, reguladores de dos componentes y chaperonas citoplásmicas/periplásmicas. Estas proteínas participan en la expulsión del exceso de Cu^{+}/Cu^{2+}. Basados en modelos de ingreso de otros metales y estudios en *Mycobacterium tuberculosis* y *Pseudomonas*; se asume que el ingreso de Cu puede llevarse a cabo mediante proteínas de membrana externa (porinas). Estas proteínas facilitan la translocación del metal a través de la membrana de forma específica o inespecífica. Las bacterias del orden *Rhizobiales* son un grupo muy diverso de alfa proteobacterias, el cual agrupa diazótrofos facultativos que forman simbiosis con plantas leguminosas; así como patógenos intracelulares de plantas y animales. Se asume que hay un exceso de cobre en su hábitat debido a la aplicación de fungicidas en cultivos agrícolas y a la concentración toxica del cobre dentro del fagosoma de células animales. Sin embargo, sus mecanismos de homeostasis de cobre no han sido completamente dilucidados.

Para contribuir al avance de este campo del conocimiento, llevamos a cabo la identificación y caracterización del gen *ropAe*; que codifica para una posible porina involucrada en el ingreso de Cu en la bacteria *Rhizobium etli* CFN42.

Índice

INTRODUCCIÓN

Origen y Propiedades químicas del cobre (Cu)

El origen del cobre y otros metales en nuestro planeta, se debe al colapso de estrellas presentes en el universo. Las estrellas están compuestas por diversos elementos, entre ellos metales pesados. Gracias a las altas temperaturas y presiones originadas por la explosión, puede originar la formación de nuevas estrellas e incluso planetas; además, los restos de la explosión son diseminados a lo largo del universo y logran impactar a otros planetas, en forma de meteoritos (Peebles et al 1994, Kirshner R.P. 1994).

El cobre (Cu), es el elemento químico con el número atómico 29. Es el primer elemento del subgrupo IB de la tabla periódica (Fig. 1) y pertenece al grupo de los metales pesados, por tener una densidad de 8.6 g/cm^3. Adicionalmente, se le considera un metal de transición debido a su configuración electrónica; $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^1$. El orbital d está parcialmente lleno de electrones y genera un movimiento de electrones de las capas internas hacia la última capa. A esta propiedad se le denomina transición electrónica. Debido a la baja energía de ionización del electrón que pertenece a la capa $4s^1$, da por resultado una fácil remoción de este mismo y por ende se obtiene el ion cuproso (Cu^+); por otro lado, cuando se remueve un electrón de la capa 3d se obtiene el ion cúprico (Cu^{2+}), siendo este mismo estado de valencia el más común (Matsumoto, P. S. 2005).

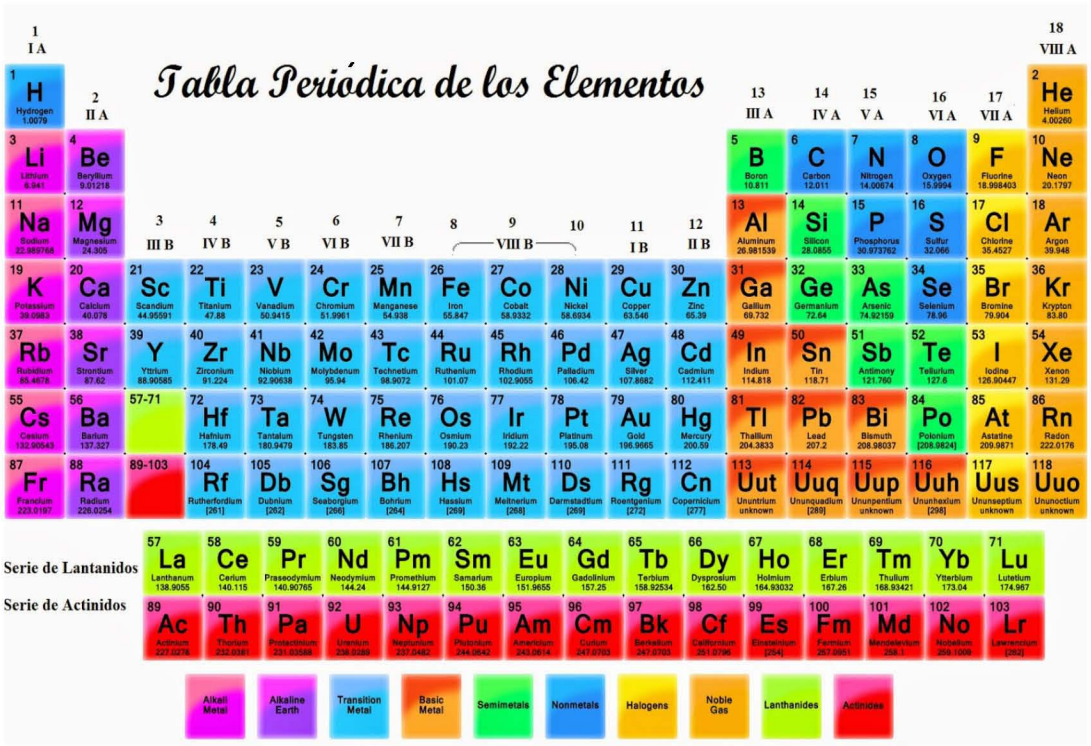

Fig. 1: Tabla periódica de los elementos químicos

Abundancia de cobre en México y el resto del mundo

A lo largo de la historia, el cobre ha sido uno de los metales más importantes en el desarrollo de la civilización humana y actualmente es el tercer metal mayormente utilizado, tan solo detrás del hierro y aluminio. Gracias a sus propiedades físicas y químicas como maleabilidad, resistencia a la corrosión, conductividad térmica y eléctrica, se favorece el uso de cobre en la industria electrónica, arquitectura y química. La industria electrónica representa aproximadamente las tres cuartas partes del uso total de cobre del mundo (Flanagan 2017). Chile es el principal país productor de cobre y el que tiene las reservas más grandes en todo el mundo (Tabla 1).

Tabla 1.- Lista de los 10 principales países productores de cobre en el año 2016 (Flanagan 2017).

País	Producción de Cobre (miles de toneladas)	Reservas (miles de toneladas)
Chile	5500	210000
Perú	2300	81000
China	1740	28000
Estados unidos	1410	35000
Australia	970	89000
Congo	910	20000
Zambia	740	20000
Canadá	720	11000
Rusia	710	30000
México	620	46000

Aplicaciones del cobre en la agricultura

El cobre generalmente se encuentra en la naturaleza asociado con el azufre (calcosina) y el cobre puro se produce a partir de un proceso con múltiples etapas; comenzando con la extracción y concentración de minerales presentes en una mena, el cual contiene minerales de sulfuro de cobre. Después se realiza una fundición y refinación electrolítica para producir cátodos de cobre puro. Una porción creciente de cobre se produce a partir de la lixiviación ácida de minerales oxidados (Flanagan 2017). Por otro lado, están los productos derivados de cobre como el sulfato de cobre ($CuSO_4$), que es el ingrediente activo de diversos productos utilizados ampliamente en la agricultura; por ejemplo:

- Control de enfermedades por hongos
- Corrección de la deficiencia de cobre en suelos y en animales
- Estimulación del crecimiento para cerdos y pollos de engorda
- Control de plagas por babosas y caracoles

El uso de productos derivados de cobre por los humanos, conlleva a la generación de desechos ricos en cobre, siendo elementos de contaminación antropogénica, debido a que el Cu no se degrada y solo se acumula principalmente en el suelo y cuerpos de agua lo que produce efectos negativos en los ecosistemas (Passek T. S. 2017, Solomon F. 2009).

Niveles de cobre en el ambiente

El contenido de Cu en el ambiente es variable, por lo que se ha estudiado en cada ecosistema (aire, agua y suelo) de forma separada encontrando diferentes concentraciones de Cu en cada ecosistema y en diferentes partes del mundo (Tabla 2).

Tabla 2.- Rango de concentraciones de Cu encontradas en los diferentes ecosistemas. Los datos fueron obtenidos la página web de World Health Organization (WHO), estudios publicados en el año 2004. Los datos de las concentraciones de Cu en suelo fueron complementados con el estudio de Kabata-Pendias publicado en el año 2001).

	Contenido de Cu
Aire	0.003 - 7.32 $\mu g/m^3$
Agua	0.0005 – 1 mg/L
Suelo	1 - 140 mg/Kg suelo

Importancia del cobre en los sistemas biológicos

La importancia del cobre como un elemento esencial para la vida, surgió cuando las primeras cianobacterias fotosintéticas antiguas liberaron oxígeno a la atmósfera. Este suceso provocó la precipitación oxidativa de hidróxidos insolubles de hierro (Fe^{3+}), lo cual disminuyó la biodisponibilidad del Fe. Otros metales como el Cu fueron oxidados y las bacterias desarrollaron sistemas para la captación e incorporación de estas moléculas. Un ejemplo es la citocromo oxidasa, enzima clave para el transporte de electrones a través de los citocromos, etapa final en la

respiración aeróbica. La capacidad oxido-reducción (Cu^{1+} - Cu^{2+}) del Cu, favorece a que las enzimas que utilizan moléculas de Cu tengan potenciales redox típicamente entre +0.25 y + 0.75V. Esto permite la eliminación de electrones de diversos sustratos tales como catecoles, superóxido, ascorbato y hierro. En consecuencia, las enzimas dependientes de cobre funcionan en diversos procesos que incluyen fotosíntesis, transporte de electrones, metabolismo de la pared celular, protección al estrés oxidativo etc. (Tabla 3) (Crichton et al 2001, Rubino and Franz 2012, Yruela et al 2009, Ladomersky et al 2015).

Tabla 3: Funciones de algunas proteínas y enzimas que utilizan moléculas de Cu como cofactor encontradas en bacteria, hongos, plantas y animales (Rubino and Franz 2012).

Nombre	Función
Azurina, Citocromo c oxidasa	Transferencia de electrones en la respiración
Galactosa oxidasa	Oxidación de alcoholes primarios a aldehídos en azucares, reducción de O_2 a H_2O_2
Superóxido dismutasa	Reducción de O_2^- a H_2O_2, oxidación de O_2^- a O_2
Tirosinasa, Lacasa	Oxidación de fenoles, reducción de O_2 a H_2O
Plastocianina	Transferencia de electrones en la fotosíntesis
Nitrito reductasa	Reducción de NO_2^- a N_2
Amino oxidasa	Oxidación de aminas primarias, Reducción de O_2 a H_2O_2
Monooxigenasa	Reducción de O_2 a H_2O_2, oxidación de NADH a NAD^+

Toxicidad del cobre

Las actividades de la agricultura y de la industria generan desechos con altos contenidos de metales pesados; estos elementos no se pueden degradar y son persistentes en el ambiente, principalmente en el suelo (Soares et al 2003). Esto provoca que algunas extensiones de suelo se transformen en zonas no cultivables

(Li and Daler 2004, Ochoa-Herrera et al 2011). Entre los metales que pueden provocar estos efectos se encuentra el cobre.

La capacidad del cobre de alternar entre sus estados de oxidación Cu^{+} (cuproso) y Cu^{2+} (cúprico) lo convierte en un cofactor biológico ideal. En algunas condiciones, cuando existen altos niveles de Cu intracelular puede llegar a ser extremadamente tóxico. Se ha propuesto que la toxicidad del Cu bajo condiciones aeróbicas, radica en la formación de radicales hidroxilo libres (OH^{-}) mediante la reacción de Fenton (A). Esto favorece el estrés oxidativo mediante la reacción de Haber-Weis (B); como resultado daña proteínas, lípidos y ADN (Solioz and Stoyanov 2003). Por otro lado, estudios en *E. coli*, se ha descubierto que los iones de Cu^{2+} pueden desplazar las moléculas de Fe en los sitios Fe-S de proteínas y por ende alteran su estructura y/o función. Además, la liberación de átomos de Fe podría acelerar la reacción de Fenton con base en moléculas de Fe (Macomber et al 2009).

$Cu^{+} + H_2O_2 \rightarrow Cu^{2+} + OH^{-} + OH^{+}$ A) Reacción Fenton

$O_2^{-} + H_2O_2 \rightarrow O_2^{+} + OH^{-} + OH^{+}$ B) Reacción Haber-Weiss

Por lo tanto, dada la importancia del manejo intracelular de iones metálicos requeridos biológicamente para la viabilidad celular, existe una maquinaria de regulación dedicada a mantener la homeostasis de los metales en la célula. El concepto de homeostasis se define como el mantenimiento de una concentración óptima, mediada por la captación de metales, el tráfico intracelular y procesos de eflujo y almacenamiento, para satisfacer las necesidades de la célula para ese ion metálico (Ma et al 2010).

Ingreso de Cu en bacterias Gram negativas

Las proteínas de membrana externa (OMP), están involucradas en el ingreso de solutos que no pueden pasar a través de la membrana externa por difusión pasiva. Las diferencias en el tamaño del poro y composición de aminoácidos contribuyen a la especificidad del sustrato (Vollan et al 2016).

Todas las estructuras resueltas de proteínas de membrana han demostrado

pertenecer a dos tipos de plegamientos: el paquete compacto α-helicoidal y el barril β; estos dos motivos estructurales se correlacionan con su ubicación. Por ejemplo las proteínas de membrana que tienen un plegamiento de tipo α-helicoidal se encuentran frecuentemente en receptores y canales de iones en las membranas del retículo endoplásmico. Las proteínas tipo barril β están restringidas en la membrana externa de las bacterias Gram-negativas, mitocondria y cloroplastos (Galdiero et al 2007).

Las proteínas tipo barril β son llamadas así, debido a que estructuralmente están conformadas por cadenas β antiparalelas transmembranales; estas cadenas se encuentran conectadas por giros periplásmicos cortos (T1-T8) y por bucles extracelulares largos (L-L8) (Fig. 2). El modo de transporte de estas proteínas pueden ser dividido en tres clases: porinas (generales y especificas), transportadores de sustrato específico y transportadores activos (Galdiero et al 2007).

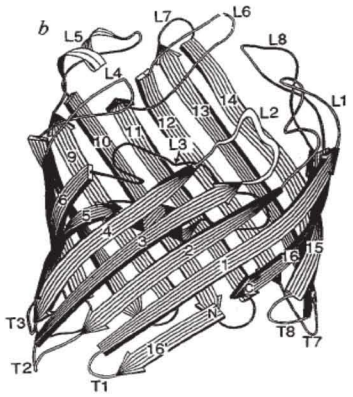

Fig. 2.- Esquema de la proteína β barril OmpF de *E. coli* perteneciente al grupo de las porinas generales. Cadenas β (1-16), giros periplásmicos (T1-T8), bucles extracelulares (L1-L8) (Cowan et al 1992).

Las primeras proteínas de membrana externa caracterizadas fueron OmpC y OmpF de *E. coli* y recibieron el nombre de porinas (Nakae et al 1976). La estructura de las porinas depende de sus propiedades fisicoquímicas y pueden tener entre 8 y 26 cadenas β, las cuales forman un canal estrecho en forma de barril. Las porinas se caracterizan por transportar moléculas menores de 600-700 Daltones y por tener una estructura de tipo homotrimérica o monomérica (Silhavy et al 2010). Una

característica importante de las principales clases de porinas es su alto nivel de síntesis, el cual depende de las condiciones ambientales y puede oscilar entre 10^4-10^6 copias por célula (Schirmer T. 1997, Koebnik et al 2000).

El papel de las proteínas de membrana externa en la adquisición de Cu^+/ Cu^{2+} fue propuesto por primera vez por Lutkenhaus et al 1977. En este estudio se informó sobre el aislamiento de mutantes resistentes a cobre en *E. coli*; el fenotipo de resistencia se asoció a la ausencia de la proteína de membrana externa b (OmpC de acuerdo con la nomenclatura de Lugtenberg and Alphen 1983). Quince años después, la hipótesis de Lutkenhaus fue contrarrestada por estudios con mutantes de OmpC en *E. coli*. Las mutantes *ompC*$^-$ mantuvieron el mismo nivel de resistencia a cobre y a plata, comparadas con la cepa parental (Li et al 1997, Bavoil et al 1977). Sin embargo, la hipótesis de que el ingreso de Cu esta dado por proteínas de membrana tipo porinas se refuerza con estos estudios que a continuación se describen (Tabla 4).

Tabla 4.- Posibles proteínas importadoras de cobre en bacterias.

Nombre del gen	Tipo de proteína	Organismo	Referencias
nosA *oprC*	Proteína de membrana externa	*Pseudomonas*	Lee et al 1989 Yoneyama H. and Nakae T 1996
mspA, mspB, mspC y *mspD*	Proteína de membrana	*Mycobacterium*	Speer et al 2013
hmtA *copA*	ATPasa tipo P	*Pseudomonas* *Enterococcus*	Lewinson et al 2009 Odermatt et al 1993, Solioz and Stoyanov 2003

En la bacteria *Pseudomonas stuzeri* se evidenció por primera vez que el ingreso de Cu pudiera estar dado por una proteína de membrana externa (NosA). Esta proteína es capaz de formar un canal en la membrana externa y permitir el ingreso de cationes entre ellos Cu; la cepa mutante en el gen *nosA* produce enzimas nitrato reductasa carentes de Cu y como resultado la enzima es incapaz de reducir el NO_3^- a NO_2^-. La proteína purificada NosA mostró tener Cu unido a ella. Sin embargo, la

selectividad de la proteína por Cu es baja y parecida a otras porinas no especificas como la proteína F (*P. aeruginosa*) y OmpF (*E. coli*). La presencia del dominio de unión a Cu sugiere que NosA facilita el ingreso de Cu a través de la membrana. Sin embargo, no es una proteína de alta afinidad por Cu (Lee et al 1989). La proteína llamada OprC (65% homologa a NosA) en *P. aeruginosa* también participa en el transporte de Cu a través de la membrana y tiene el mismo papel fisiológico que NosA en *P. stuzeri*. La mutante en *oprC*$^{-}$ no mostró ningún cambio en la tolerancia a Cu. Los autores sugieren que OprC no es esencial para el ingreso de Cu hacia el periplasma y el verdadero papel de la proteína OprC es trasferir el Cu^{2+} eficientemente a la nitrato reductasa (Yoneyama H. and Nakae T 1996).
En la bacteria *Mycobacterium smegmatis* se encontró que las porinas MspA, MspB, MspC y MspD; están involucradas en el ingreso de Cu en la bacteria. Se demostró que bajo condiciones limitantes de Cu en el medio; una cuádruple mutante era incapaz de crecer. Estos resultados demuestran que el ingreso de Cu en *Mycobacterium* está dado por el conjunto de las porinas MspA, MspB, MspC y MspD (Speer et al 2013).

Se ha encontrado que las ATPasas de tipo P, pueden funcionar como importadores de iones de Cu a nivel de membrana externa; como es el caso en *P. aureginosa*, donde un transportador nombrado HmtA puede ingresar iones de Cu y Zn a la célula (Lewinson et al 2009). En *Enteroccoccus hiriae* también se encontró una ATPasa que importa Cu a la célula nombrada CopA (Odermatt et al 1993, Solioz and Stoyanov 2003).
Por otra parte, se desconoce el mecanismo por el cual el Cu puede ser transportado del periplasma hacia el citoplasma.

Mecanismos de expulsión y desintoxicación de Cu en bacterias Gram negativas

En bacterias Gram negativas, los mecanismos de expulsión y desintoxicación están dados principalmente por: ATPasas tipo P1B, multicobre oxidasas y proteínas de la familia RND (sistema CusABCSR) (Fig. 3).

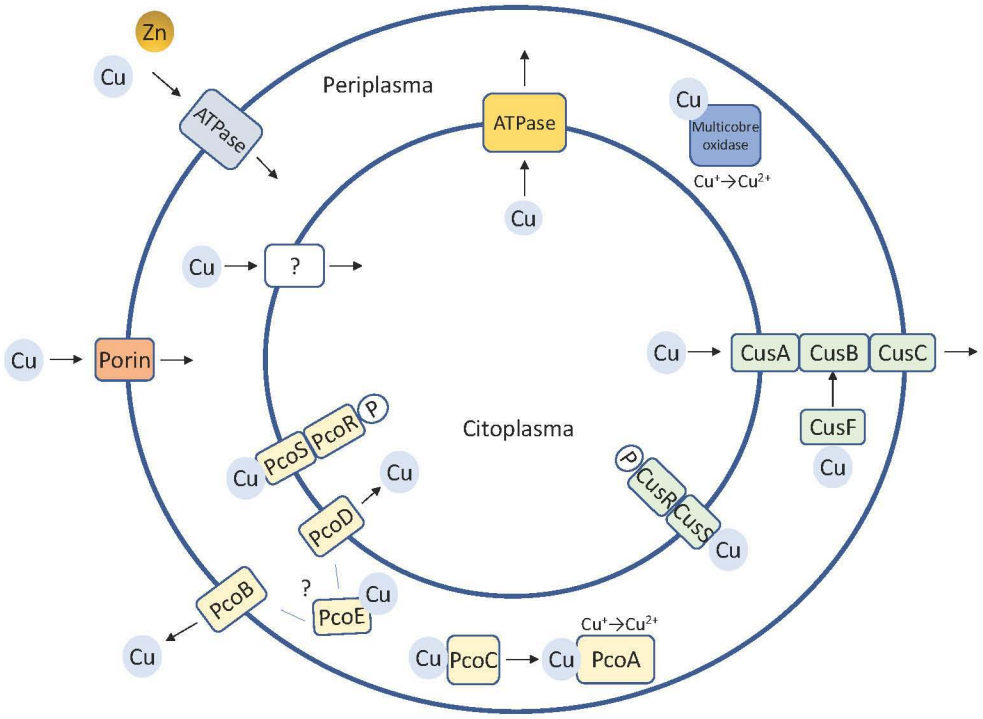

Fig. 3.- Principales mecanismos de homeostasis de Cu en bacterias Gram negativas (Rademacher and Masepohl 2012, modificado)

- **ATPasas de tipo P1B**

Las ATPasas P1B1 son proteínas de membrana interna que utilizan la hidrólisis de ATP para el transporte de cationes metálicos; P1B-1 (Cu^{+}), P1B-2 (Zn^{2+}, Cd^{2+}, Pb^{2+}), P1B-3 (Cu^{2+}), P1B-4 (Co^{2+},Ni^{2+}, Zn^{2+}), P1B-5 (Fe^{2+}) (Sitsel et al 2015). Estas proteínas están presentes en Procariontes, Eucariontes y Archeas. Su función es expulsar metales del citoplasma hacia el periplasma (Argüello et al 2007, Smith et al 2014). La primer evidencia de que una ATPasa P1B1 está involucrada en la expulsión de Cu fue en *E. hiriae* (Odermatt et al 1993). En *E. coli* el gen que codifica para esta proteína es llamado *copA* y su regulación es dependiente de la proteína CueR; un regulador dependiente de cobre, el cual activa la expresión de CopA en presencia de Cu libre en el citoplasma (Outten et al. 2001, Rensing and Grass 2003). Estudios previos, han encontrado los motivos de aminoácidos involucrados en la unión y la translocación del metal, los cuales son: 2 residuos de cisteínas (CXC) ubicado en la hélice transmembranal 6 (TMH6), un residuo de tirosina y asparagina

(YN) (TMH7), residuos de metionina y serina (MXXSS) (TMH8) y el dominio N-terminal de 60-80 aminoácidos. Éste último es rico en residuos con unión a metal (CXXC) y puede estar repetido más de una vez (Fig. 4) (Argüello et al 2003, Smith et al 2014).

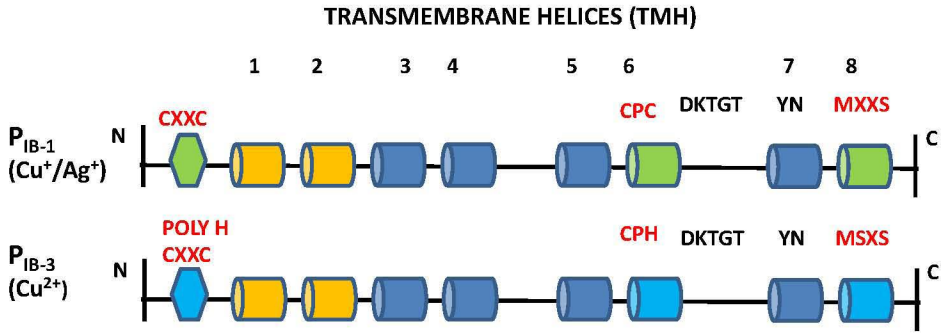

Fig. 4.- Composición de los dominios de las ATPasas pertenecientes a los grupos P1B-1 y P1B-3, propuestos por Argüello et al 2003, 2007.

- **Sistema Cus**

Este sistema pertenece a los transportadores de la superfamilia RND (Resistance-Nodulation-Division). El sistema Cus de *E. coli* está conformado por CusA (proteína de membrana interna), CusB (proteína periplásmica), CusC (proteína de membrana externa) y CusF (proteína periplásmica chaperona). Las proteínas CusABC, interaccionan formando un canal extendido en el periplasma que conecta el citoplasma con la membrana externa y CusF transporta iones de Cu a la proteína CusB. El sistema CusCFBA permite el transporte de cobre y plata mediante la fuerza protón motriz, principalmente en condiciones anaeróbicas (Outten et al. 2001, Grass and Rensing 2001, Rademacher, C and Masepohl, B. 2012). El operón *cus* es regulado por un sistema de dos componentes llamado CusRS, donde CusS es una proteína histidina-quinasa que está unida a la membrana citoplasmática la cual se encarga de detectar el Cu disponible en el citoplasma. Cuando el Cu se une al dominio receptor de CusS, esté fosforila a la proteína CusR y CusR fosforilado activa la transcripción del operón *cusABCF* y *cusRS* (Fig. 3) (Munson et al 2000).

- **Multicobre oxidasa periplásmica**

En *E. coli* el gen *cueO* codifica para una multicobre oxidasa. Esta enzima se encuentra en el periplasma y su función es oxidar el cobre monovalente (Cu^{+}) y transformarlo en cobre divalente (Cu^{2+}). El gen *cueO* también es regulado por CueR de forma positiva; en respuesta al Cu libre del citoplasma (Outten et al. 2001). Las multicobre oxidasas y las ATPasas conforman el mecanismo principal de detoxificación de cobre en condiciones aeróbicas (Fig. 3).

- **Sistema Pco (plasmid-encoded copper resistance)**

El sistema Pco está codificado en el plásmido pRJ1004 en *E. coli,* la cual fue aislada de la flora intestinal de cerdos que eran alimentados con una dieta rica en sulfato de cobre ($CuSO_4$). Esta cepa es portadora de un plásmido que contiene siete genes *pcoABCDRSE* que proporcionan un sistema de resistencia extra a Cu; lo anterior beneficia a la bacteria a contender en altas concentraciones de Cu (Brown et al 1995).

Las proteínas PcoA y PcoC son cruciales para el sistema Pco, debido a que se postula que PcoC se une a Cu^{+} libre en el periplasma y lo transfiere a la multicobre oxidasa PcoA para su oxidación a Cu^{2+} (Lee et al. 2002; Wernimont et al. 2003, Rensing and Grass 2003). PcoB se considera una proteína de membrana externa que participa en la expulsión de Cu fuera de la célula (Rensing and Grass 2003). La función de PcoD no se ha descubierto, pero se cree que participa en el ingreso de Cu hacia el periplasma. PcoE es una proteína periplásmica encargada de secuestrar el Cu libre del periplasma. El grupo de genes *pcoABCDE* es regulado por *pcoRS* el cual codifica para el sistema de dos componentes PcoR-PcoS cuya función es igual al sistema CusRS (Fig. 3) (Lee et al 2002, Bondarczuk and Piotrowska-Seget 2013). Cabe destacar que en *P. syringae* se ha descrito un conjunto de genes (*copABCDRS*) involucrados en el tráfico de Cu y es muy similar al sistema Pco de *E. coli* (Mills et al 1993).

El conjunto de proteínas CopA/CusABCF ó PcoABCD son considerados como mínimos necesarios para la homeostasis de Cu en γ proteobacterias. Sin embargo,

solo el 3% de 268 genomas de γ proteobacterias presentan el conjunto de genes completo (*copA*/*cusABCF* ó *pcoABCD*) (Hernández-Montes et al 2012).

Bacterias pertenecientes al orden *Rhizobiales*

Rhizobios es el nombre común con el que se denomina a un grupo de bacterias Gram negativas, la mayoría de ellas ubicadas taxonómicamente dentro de la división de las alfa proteobacterias. La taxonomía de los rizobios se basa en diferentes enfoques como fisiología, bioquímica, morfología y genética. La agrupación de los rizobios está dada por 17 familias: *Aurantimonadaceae, Bartonellaceae, Beijerinckiaceae, Bradyrhizobiaceae, Brucellaceae, Chelatococcaceae, Cohaesibacteraceae, Hyphomicrobiaceae, Mabikibacteraceae, Methylobacteriaceae, Methylocystaceae, Notoacmeibacteraceae, Phyllobacteriaceae, Rhizobiaceae, Rhodobiaceae, Roseiarcaceae* y *Xanthobacteraceae* (Uniprot 2018). En estas familias se encuentra una gran variedad de bacterias con diferentes características, las más conocidas y estudiadas son miembros de la familia *Rhizobiaceae* y se caracterizan por la capacidad de establecer asociaciones simbióticas con plantas leguminosas y desarrollar el proceso de fijación biológica de nitrógeno. Otras bacterias del genero *Agrobacterium* son patógenos de múltiples plantas en las cuales inducen formación de tumores. Asimismo, existen bacterias que son parásitos intracelulares de animales y parásitos oportunistas de humanos, como los géneros *Brucella* y *Ochrobactrum* (Ormeño-Orrillo et al 2015).

Mecanismos de homeostasis de metales en los rhizobios

El suelo representa el hábitat para un importante número de rizobios diazótrofos que establecen simbiosis con diferentes especies de plantas leguminosas. Experimentos de campo han demostrado que niveles tóxicos de metales disminuyen la población de rizobios del suelo y propician la gradual disminución de la diversidad genética de la población (Laguerre et al, 2006).

El estudio del transporte de metales en miembros del orden *Rhizobiales* es un área poco investigada, encontrandose pocos trabajos en la literatura (Reeve et al 2002, Hohle and O'Brian 2009, 2010, 2016, Li et al 2013, 2014, Sankari and O'Brian 2016). Como resultado existe poca información acerca de los mecanismos moleculares que participan específicamente en la homeostasis de Cu.

En *Rhizobium leguminosarum* se identificó el gen *actP* que codifica para la ATPasa de tipo PIB-1 que transporta el Cu del citoplasma hacia el periplasma y a su posible regulador (HmmR) que está codificado río arriba del gen *actP*. La secuencia de aminoácidos de la proteína ActP comparte un 55% de identidad con 90% de cobertura con respecto a la secuencia de aminoácidos de CopA de *E. coli* (Reeve et al 2002).

Estudios en la cepa *Sinorhizobium. meliloti* CCNWSX0020 nos ha permitido extender el conocimiento de los mecanismos de homeostasis de Cu en los rizobios. Esta cepa fue aislada de suelos pertenecientes a una mina en China. Se han identificado proteínas que están involucradas con la detoxificación de Cu. Por ejemplo: la multicobre oxidasa nombrada McoO que oxida el Cu^{+} a Cu^{2+}, una proteína chaperona de Cu periplásmica nombrada PcoO, cuya función es movilizar el Cu libre del periplasma y llevarla hacia una proteína de membrana externa tipo porina nombrada Omp, cuya función es transportar el Cu hacia el exterior celular (Fig. 5) (Li et al 2013).

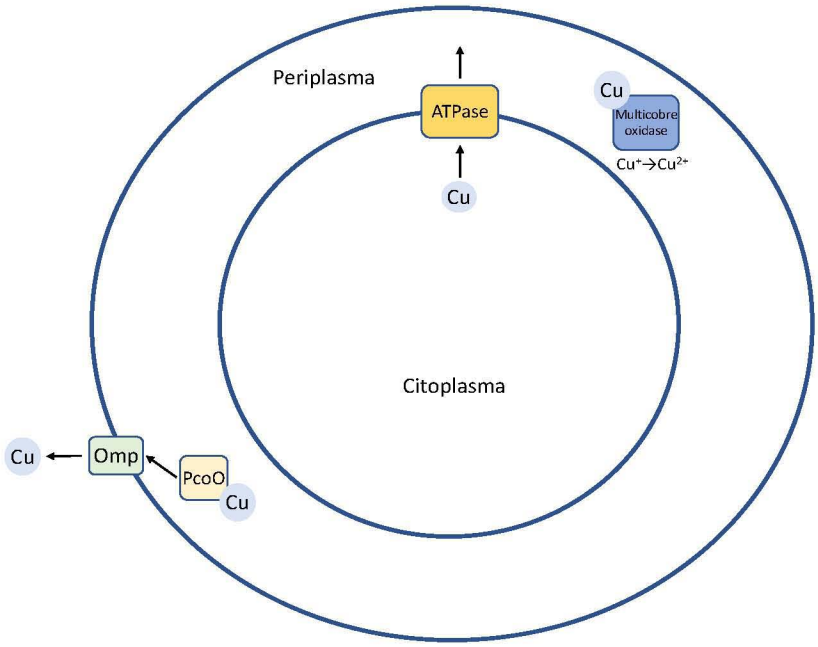

Fig. 5.- Mecanismos de homeostasis de Cu descritos en bacterias del orden *Rhizobiales* (Li et al 2013).

En el trabajo de Cubillas et al 2017, realice una colaboración de investigación de la ocurrencia y diversidad de tres componentes críticos en la maquinaria de expulsión de Cu en bacterias del orden *Rhizobiales*; los cuales fueron la ATPasa CopA P1B-1 y las chaperonas periplásmicas, CopZ/CupA y CusF. Como resultado se obtuvo que el sistema completo CopA/CopZ/CusF solo está presente en 19 de 53 de los genomas de *Rhizobiales* analizados. Estos resultados indican la existencia de mecanismos alternativos y/o proteínas inexploradas involucradas en el tránsito de Cu. Además, en dicho trabajo se analizó la divergencia de las ATPasas P1B. El uso de programas bioinformaticos *ad hoc* para el análisis de secuencias como lo son: perfiles de modelos ocultos de Markov y filogenia de máxima verosimilitud permitieron establecer una nueva subclasificación de las ATPasas P1B (P1B-1, P1B-3).

La ausencia del sistema principal CopZ/CupA y CusF en la expulsión y detoxificación de cobre que presentan otras bacterias, nos sugiere que el sistema de homeostasis de Cu en bacterias presenta interacciones dinámicas de evolución complejas aún inexploradas y resulta difícil proponer un modelo evolutivo y

conservativo de los mecanismos de homeostasis de Cu en bacterias. Estos resultados se pueden observar en el Anexo I.

Identificación del gen *ropAe* que codifica para una posible proteína de membrana externa, involucrada en el tráfico de Cu en *R. etli* CFN42

Rhizobium etli es la especie más abundante en suelos donde se siembra frijol. Una de las cepas de *R. etli* mejor caracterizadas es la CFN42; aislada de raíces de frijol cultivado en Guanajuato, México (Martínez-Romero, 2003). Su genoma, totalmente secuenciado en el Centro de Ciencias Genómicas (CCG) de la UNAM, está constituido por un cromosoma circular de 4.3 Mb y seis plásmidos denominados p42a (194.2 kb), p42b (184.3 kb), p42c (250.9 kb), p42d (371.2 kb) p42e (505.3 kb) y p42f (642.5 kb) (González *et al,* 2006). Con la finalidad de determinar el papel que juegan los plásmidos en la adaptación de la bacteria a diferentes condiciones de estrés ambiental, desde hace varios años se ha llevado a cabo la reducción sistemática del genoma extra cromosomal, seleccionando mutantes derivadas de *R. etli* CFN42 que carecen de cada uno de sus seis replicones (mutantes curadas) o eliminando fragmentos de algunos replicones mediante deleciones sitio-específicas (Landeta et al, 2011). Por medio de la caracterización de estas mutantes nuestro grupo ha demostrado que en los plásmidos están codificadas funciones esenciales para la interacción con plantas, para el crecimiento saprofítico y para la adaptación a diferentes condiciones de estrés ambiental incluyendo la tolerancia a metales (Brom et al, 1992; García-de los Santos and Brom, 1997; García-de los Santos et al, 2008; Landeta et al, 2011; Villaseñor et al, 2011, Cubillas et al 2013, 2014, 2017.)

Una condición de estrés ambiental que enfrentan los rizobios es la presencia de iones metálicos en el ambiente. En el plásmido p42e de *R. etli* CFN42 se han caracterizado dos genes que forman parte de los mecanismos de homeostasis de níquel (RHE_PE00218) y cobre (*actP*), (Cubillas et al, 2013; Landeta et al, 2011). Las mutantes en estos genes generan sensibilidad a níquel y cobre

respectivamente. Cuando analizamos el nivel de tolerancia a cobre de la mutante CFNX185, derivada de *R. etli* CFN42 la cual perdió espontáneamente 200 kb del plásmido p42e, nos sorprendió que su tolerancia a $CuCl_2$ es mayor que la tolerancia de la cepa silvestre. A partir de este resultado se llevó a cabo el análisis genético para contestar la pregunta: ¿Qué genes están involucrados en el incremento de tolerancia a Cu en la mutante CFNX185?

Basados en el fenotipo de aumento a la tolerancia a metales planteamos la siguiente hipótesis.

Hipótesis

La región del plásmido p42e que perdió espontáneamente la cepa CFNX185, contiene genes que codifican proteínas involucradas en el importe de cationes en la bacteria *R. etli* CFN42.

Objetivos

General:

*Identificar los genes que participan en la homeostasis de cobre en *R. etli* CFN42.

Específicos:

1.- Identificar él o los genes involucrados en el fenotipo de resistencia a Cu en la mutante CFNX185 y analizar su función.

2.- Predecir la presencia de otros genes en el resto del genoma, que participen en la homeostasis de cobre en la bacteria *R. etli* CFN42.

Resultados

A continuación se presenta el artículo donde se reporta la identificación y caracterización el gen *ropAe* que codifica para una posible porina involucrada en el transito de cobre en *R. etli* CFN42.

Received: 1 July 2017 | Revised: 19 November 2017 | Accepted: 21 November 2017

DOI: 10.1002/mbo3.573

ORIGINAL RESEARCH

WILEY MicrobiologyOpen Open Access

The *ropAe* gene encodes a porin-like protein involved in copper transit in *Rhizobium etli* CFN42

Antonio González-Sánchez[1] | Ciro A. Cubillas[2] | Fabiola Miranda[3] | Araceli Dávalos[1] | Alejandro García-de los Santos[1]

[1]Programa de Ingeniería Genómica, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

[2]Deparment of Developmental Biology, Washington University School of Medicine, St. Louis, MO, USA

[3]Deparment of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

Correspondence
Alejandro García-de los Santos, Programa de Ingeniería Genómica, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México.
Email: alex@ccg.unam.mx

Funding information
UNAM-DGAPA-PAPIIT, Grant/Award Number: IN209815

Abstract

Copper (Cu) is an essential micronutrient for all aerobic forms of life. Its oxidation states (Cu^{+}/Cu^{2+}) make this metal an important cofactor of enzymes catalyzing redox reactions in essential biological processes. In gram-negative bacteria, Cu uptake is an unexplored component of a finely regulated trafficking network, mediated by protein–protein interactions that deliver Cu to target proteins and efflux surplus metal to avoid toxicity. *Rhizobium etli* CFN42 is a facultative symbiotic diazotroph that must ensure its appropriate Cu supply for living either free in the soil or as an intracellular symbiont of leguminous plants. In crop fields, rhizobia have to contend with copper-based fungicides. A detailed deletion analysis of the pRet42e (505 kb) plasmid from an *R. etli* mutant with enhanced $CuCl_2$ tolerance led us to the identification of the *ropAe* gene, predicted to encode an outer membrane protein (OMP) with a β–barrel channel structure that may be involved in Cu transport. In support of this hypothesis, the functional characterization of *ropAe* revealed that: (I) gene disruption increased copper tolerance of the mutant, and its complementation with the wild-type gene restored its wild-type copper sensitivity; (II) the *ropAe* gene maintains a low basal transcription level in copper overload, but is upregulated when copper is scarce; (III) disruption of *ropAe* in an *actP* (*copA*) mutant background, defective in copper efflux, partially reduced its copper sensitivity phenotype. Finally, BLASTP comparisons and a maximum likelihood phylogenetic analysis highlight the diversification of four RopA paralogs in members of the *Rhizobiaceae* family. Orthologs of RopAe are highly conserved in the *Rhizobiales* order, poorly conserved in other alpha proteobacteria and phylogenetically unrelated to characterized porins involved in Cu or Mn uptake.

KEYWORDS
copper homeostasis, copper uptake, porins, *Rhizobium*, RopA

1 | INTRODUCTION

Copper (Cu) is an essential trace element for aerobic organisms of the three domains of life. Cu can exist as an oxidized cupric ion (Cu^{2+}) or as reduced cuprous ion (Cu^{+}). The capacity of copper to alternate between these two oxidation states, (Cu^{+}/Cu^{2+}), makes this metal the ideal cofactor of key enzymes catalyzing redox reactions in vital biological processes, such as respiration, free radical detoxification, the methane cycle, photosynthesis, the carbon cycle, and the nitrogen cycle (Festa & Thiele, 2011; Rubino & Franz, 2012). Cells have

to maintain intracellular copper in trace concentration; otherwise Cu^{+} can displace iron from iron-sulfur clusters of metalloproteins, resulting in their inappropriate structure and function. In addition, free iron and Cu^{+} may cause hydrogen peroxide to generate an increase at the concentration of hydroxyl radicals by a Fenton type reaction which can damage proteins, lipids, and nucleic acids (Macomber & Imlay, 2009).

As copper is essential for life on earth, it is important to understand how cells fulfill their copper requirements and how trace concentrations are maintained to avoid toxicity.

Although the eukaryotic copper uptake system is well-documented (Boal & Rosenzweig, 2009), little information is available about how prokaryotic cells import this metal.

The role of outer membrane proteins (OMP) in Cu^{+}/Cu^{2+} acquisition was first suggested by Lutkenhaus (1977) who reported the isolation of copper-resistant mutants, in *Escherichia coli* B/r; these appeared with a frequency of 10^{-5} in minimal medium plates with 20 μmol/L $CuSO_4$. This phenotype was associated with the absence of outer membrane protein b (OmpC according to Lugtenberg and Van Alphen's nomenclature published in 1983) (Lugtenberg & Van Alphen, 1983) determined by membrane preparation and gel electrophoresis. Fifteen years later, Lutkenhaus' hypothesis was contradicted by studies with well-characterized *E. coli* B/r isogenic *OmpC* mutants which maintained the same level of copper and silver resistance as the parental strain (Bavoil, Nikaido, & von Meyenburg, 1977; Li, Nikaido, & Williams, 1997).

In *Mycobacterium smegmatis* the uptake of nutrients and beta-lactam antibiotics is mediated by the MspA, MspB, MspC, and MspD porins (Danilchanka, Pavlenok, & Niederweis, 2008; Stephan et al., 2005). Their role in copper transport was experimentally assayed in *Mycobacterium smegmatis* porin-deleted mutants which grew poorly under trace concentrations of $CuSO_4$ and simultaneously increased their copper tolerance when exposed to $CuSO_4$ overload (Speer et al., 2013).

The potential role of multiple porins in the uptake of copper has also been suggested by the transcriptional profiles of copper–adapted cells of *Pseudomonas aeruginosa*. This study revealed that at least eight genes coding for different putative porins were downregulated in copper-adapted cells (Teitzel et al., 2006).

Our research group has been working on the characterization of the transportome of divalent cations in the facultative diazotroph *Rhizobium etli* CFN42 (Cubillas et al., 2013, 2014), an α-proteobacterium belonging to the *Rhizobiales* order that can live as a saprophyte in the rhizosphere of *Phaseolus vulgaris* plants or as a nitrogen fixer in symbiosis with the roots of *P. vulgaris* plants (Segovia, Young, & Martínez-Romero, 1993). Its 6.5 Mb genome is partitioned in one circular chromosome and six plasmids (Gonzalez et al., 2006). The contribution of the *R. etli* pRet42e plasmid to ion balance is due to the presence of NepA, a member of the cation diffusion facilitator family (CDF) required to deal with high nickel concentrations, and the P_{1B}-ATPase, ActP, a copper efflux pump that confers copper resistance (Landeta et al., 2011).

In this study we report that an isogenic mutant of *R. etli* CFN42, named CFNX185, which lacks 200 kb of its 505–kb plasmid pRet42e, is more tolerant to $CuCl_2$ than the wild-type strain. Further analysis of pRet42e led us to the identification and characterization of the *ropAe* gene, coding for a putative outer membrane protein (OMP) whose absence increased copper resistance. The enhanced copper resistance phenotype observed in a double *actP– ropAe–* mutant suggests that disruption of *ropAe* reduces the intracellular copper supply, alleviating the copper toxicity observed in the *actP* single mutant that is defective in its copper efflux. Bioinformatic predictions, genetic experiments and transcriptional analyses allowed us to propose that a basal transcription level of *ropAe* facilitates copper transport across the outer membrane. However, under copper limitation, *ropAe* increased its transcription level, suggesting that it may play an important role in ensuring copper supply when the bacterium faces copper scarcity. A phylogenetic analysis revealed that RopAe belongs to the Porin_2 (PF02530) family and that it is distant from MspA, MspB, and MspC porins in *M. smegmatis*, from OmpC of *E. coli* (Li et al., 1997), from MnoP (*blr0095*) of *Bradyrhizobium japonicum* (Hohle, Franck, Stacey, & O'Brian, 2011), and from the copper-regulated porins of *Pseudomonas aeruginosa* belonging to the OprD family (Teitzel et al., 2006). Three homologs of RopAe were found encoded in the chromosome of *R. etli* CFN42. The four RopA porin-like transporters belong to different monophyletic groups suggesting a functional divergence among them.

2 | EXPERIMENTAL

2.1 | Bacterial strains, media, and growth conditions

The characteristics of the bacterial strains and plasmids used in this study are listed in Table S1. Bacterial growth was started from glycerol stocks (20% and stored at −70°C) propagated in rich PY medium containing 5 g/L peptone, 3 g/L yeast extract, and 15 g/L agar. After sterilization, 10 ml/L 0.7 mol/L $CaCl_2$ was added. Minimal medium (Mm) was prepared from three solutions (A, B, C) and sterilized separately. Solution A contained 1.620 g/L sodium succinate hexahydrate as a carbon source, 0.534 g/L NH_4Cl as a nitrogen source, 0.219 g/L K_2HPO_4, and 0.1 g/L $MgSO_4$. The pH of this solution was adjusted to 6.8 before sterilization in autoclave. Solution B contained filter sterilized 0.025 g/5 ml $FeCl_3 \cdot 6H_2O$ and solution C contained 0.7 mol/L $CaCl_2 \cdot 2H_2O$ (autoclaved). A quantity of 1 ml of B solution and 2 ml of C solution were added to 1 L of A solution.

Antibiotics for *R. etli* were added at the following concentrations (μg/ml): nalidixic acid, 20; streptomycin, 100; gentamicin, 15; and tetracycline, 3. For *E. coli* the antibiotic concentrations were (μg/ml): kanamycin, 30; gentamicin and tetracycline 10.

2.2 | Metal sensitivity assay

Metal sensitivity was determined with a plate assay using square Petri dishes with a grid as follows: 50 mmol/L stock solutions of Cu, Ni, Co, Zn, and Cd, 500 mM Fe, and 45 mM chloride salts (Sigma-Aldrich, St Louis, MO) were prepared in Milli-Q water, filter sterilized and added at increasing concentrations to solid (1.5% wt/vol agar) Mm. The *R. etli* overnight cultures were adjusted to OD_{620} = 0.7, washed twice with

10 mM of $MgSO_4$, serially diluted (10^{-1} – 10^{-6}) and spotted (20 μl) on solid Mm supplemented with or without metal ions. Growth was recorded after 5 days of incubation at 30°C. The total inhibitory concentration of copper for wild-type *R. etli* CFN42 was 20 μmol/L. The Minimal Inhibitory Concentrations (MICs) of metals able to reduce, in at least one log-unit, the growth (CFU) were previously determined (Cubillas et al., 2013) as follows: 100 μmol/L Ni, 100 μmol/L Co, 200 μmol/L Zn, 100 μmol/L Cd, 15 μmol/L Cu, 2.5 mmol/L Fe, and 30 mmol/L Mn.

2.3 | DNA manipulation

Cloning, restriction digest, ligation, transformation, Southern blotting, and hybridization were performed according to standard protocols (Sambrook, Fritsch, & Maniatis, 1989).

2.4 | Generation of site-specific deletions

Deletions of pRet42e plasmid were obtained using the Cre/*loxP* system as previously described (Landeta et al., 2011). Briefly, the region to be eliminated was flanked, in direct orientation, by *loxP* sites present in pVEX1311 plasmid (Ayres, Thomson, Merino, Balderes, & Figurski, 1993) and in pIC20R (Marsh, Erfle, & Wykes, 1984). These suicide vectors, containing a PCR fragment (≥300 bp) of DNA sequences flanking the region to be eliminated were introduced by conjugation into *R. etli* CFN42. Single crossover recombination mediates the plasmids' cointegration at the target sequences. The Cre recombinase, introduced by conjugation, mediated the in vivo recombination of both *loxP* sites and the excision of the target DNA. Plasmid deletions were validated by changes in their electrophoretic mobility using a modification of Eckhardt gel electrophoresis procedure (Hynes & McGregor, 1990) that consists of a gentle lysis, running the SDS solution backwards, followed by the loading of wells with cells and starting the lysis, running the SDS solution forwards. The second validation was done by the absence of amplicons, using the total genome of the deleted strain as PCR template.

2.5 | Site-directed vector integration mutagenesis and mutant complementation

A 400 bp internal fragment of the target genes were amplified by PCR (see Table S2 for primers) and cloned into pK18mob Km^r suicide vector (Schäfer et al., 1994). Plasmids were introduced into *R. etli* CFN42 by conjugation and mutants with vector integration by single-crossover were selected as Km^r clones. Vector integration in the target gene was verified by Southern blot, using the internal 400 bp of the target gene, amplified by PCR as the probe.

To complement the phenotype of the *ropAe* mutant, PCR-amplified fragments containing the complete *ropAe* gene or its *ropAch1*, *ropAch2*, and *ropAch3* paralogs were cloned into the broad–host-range pBBRMCS-5 plasmid (Kovach et al., 1995) or in its derived expression, pSRK vector (Khan, Gaines, Roop, & Farrand, 2008), and introduced into *ropAe* mutant by conjugation (Table S1).

2.6 | Construction of a *ropAe/actP* double mutant

A 1.3 kb PCR-amplified BamH1–XhoI fragment of the *actP* gene was cloned into pBluescript II SK(+) vector. The HindIII–HindIII *ΩSp* interposon (Fellay, Frey, & Krisch, 1987) was inserted into the sole HindIII restriction site of PCR-amplified *actP*, located 900 bp upstream of the TAG stop codon. The BamH1–XhoI fragment containing the *actP*::*ΩSp* was subcloned into pJQ200 SK (Quandt & Hynes, 1993) and then introduced by conjugation into wild type *R. etli* CFN42 using *E. coli* S17-1 (Simon, 1984) as donor. The wild-type *actP* gene was replaced by the *actP*::*ΩSp* by double homologous recombination in most of the transcojugants grown in the presence of sucrose 12.5%. The gene replacement was verified by Southern blot using Redyprime kit for labeling with [α-^{32}P] the DNA fragment used as probe**.** To construct the double mutant *ropAe*::pK18mob/*actP*::*ΩSp;* the wild-type *ropAe* was disrupted by site-directed vector integration of the pK18mob plasmid in the *actP*::*ΩSp* mutant background as described above.

2.7 | Cu-dependent transcriptional response of *ropAe* measured by qRT-PCR

To determine the effect of either copper excess or deficiency on *ropAe* expression, two different experiments were performed. For the copper excess condition, PY cultures (20 ml, $OD_{620\,nm}$ = 0.65–0.75) of wt *R. etli* were exposed for 30 min to 0.5 mmol/L $CuCl_2$. For the copper deprivation treatment, wt *R. etli* was grown overnight in 20 ml of PY with or without 2 mmol/L of the membrane impermeable Cu^+/Cu^{2+} chelator bathocuproine disulfonic acid disodium salt (BCDS, Sigma) (Ding, Xie, & Kang, 2011). These were used to extract the mRNA using the TriPure isolation reagent (Roche). The total RNA (DNA free, 1 μg) was reverse transcribed to cDNA using ReverAid H minus FirstStrand cDNA Synthesis (Fermentas). Quantitative real-time PCR was performed on StepOnePlus (Applied Biosystems) using Maxima Syber Green/ROX qPCR master Mix (Fermentas) and 1 μg of cDNA as template. The *ropAe*, *actP* and *hisCd* genes were amplified by using the primers listed in Table S2. Their expression levels in the presence of $CuCl_2$ or $CuCl_2$ plus the BCDS Cu^+/Cu^{2+} chelator were normalized to the expression level of housekeeping *hisCd* gene (Salazar et al., 2010). The data represent averages of four independent experiments with three technical replicates each. The fold change in gene expression was calculated using the $\Delta\Delta C_T$ method (Schmittgen & Livak, 2008).

2.8 | Bioinformatic prediction of subcellular localization of RopAe

The localization of RopAe was analyzed with two different predictors of subcellular localization with high accuracy for OMP (Bhasin, Garg, & Raghava, 2005). The multimodular PSORT-B (Gardy et al., 2003) examines the query sequence for the presence of 12 different characteristics, such as amino acid composition, similarity to proteins of known localization, signal peptide, alpha helices, motifs, etc. (http://www.psort.org/). The second predictor was CELLO (Yu, Lin, & Hwang, 2004) which uses a single analytical module, a support vector machine based on *n*-peptide composition (http://

cello.life.nctu.edu.tw/). The scores for the different subcellular localization analyses are shown in table S3. Both methods assigned high probability values for the outer membrane localization of RopAe: PSORT-B, probability 9.3 of 10 as maximum score; CELLO, score 4.105 of 5 as maximum score.

2.9 | Searching for RopAe homologs in protein databases and bacterial genomes

Protein Blast program (BlastP) at NCBI with default settings, was used to search databases or bacterial genomes for RopAe homologs. BlastP was also used to align two sequences and estimate identity (%), similarity (%), query cover (%), and E value.

2.10 | Phylogenetic analysis of the Porin_2 family (PF02530)

According to Pfam, the Porin_2 (PF02530) family includes 263 OMP from alpha proteobacteria. To assess the number of rhizobial porins contained in this family, as well as their taxonomic distribution and diversity, the porins of members of the *Rhizobiales* order were downloaded, filtering from other α-proteobacteria with the species distribution sequence search tool included in pfam 30.0. A total of 145 OMP homologs belonging to 35 species of *Rhizobiales*, as well as OmpC (a nonspecific porin) of *E. coli*, MnoP (a manganese transporter) of *B. japonicum* (Hohle et al., 2011), MspA, MspB, MspC, and MspD porins of *M. smegmatis* (Speer et al., 2013) comprised the data set (Table S4) and were aligned against the HMMER profile using hmmalign. The most conserved regions shared among the OMPs were obtained from the alignment of putative RopAe with the HMMER profile. The resultant data set, containing 127 OMPs (Table S4), was used to infer the evolutionary relationships among the rhizobial OMPs belonging to the Porin_2 family (PF02530) relative to the other characterized porins. Due to the low bootstrap values observed in the 127 sequences phylogeny (Figure S2), we decided to select the clades where the RopAe proteins and their closest homologs were located and we run the analysis again including the characterized proteins. The resulting phylogeny (Figure 6) has higher bootstrap values that support the conclusions. The ML-phylogenetic analysis was performed under the LG+G+f model using amino acid alignment. The phylogeny was built using a parallel P threads-based version of RAxML v8.2.4, because this software supports SSE3 vector instructions (Stamatakis, 2014). The ML-search started with 100 random seed trees and the best tree was selected. Rapid bootstrapping was used to assess the branch support (Stamatakis, 2014). The number of necessary replicates was estimated at 100 using the extended majority rule criterion (Pattengale, Aberer, Swenson, Stamatakis, & Moret, 2011).

3 | RESULTS

3.1 | Characterization of an *R. etli* mutant with increased copper tolerance

As a result of previous research conducted to analyze the functional contribution of the six *R. etli* CFN42 plasmids to contend with stress conditions, a collection of derivative mutants was constructed. The derivatives lack each of the following plasmids: pRet42a, pRet42b, pRet42c, pRet42d, and pRet42f, or a 200 Kb fragment of plasmid pRet42e (Brom et al., 1992). We assessed the toxic effect of Cu and other metals on these mutants by determining their growth capacity, as described in Materials and Methods, in minimal medium agar plates supplemented with increased concentrations of $CuCl_2$ (0–25 μmol/L). The assays revealed that 20 μmol/L $CuCl_2$ totally inhibited the growth of the parental strain (CFN42) (Figure 1) as well as that of the other plasmid-cured strains (data not shown). In contrast, the derivative lacking 200 kb of pRet42e (strain CFNX185) showed full growth on 20 μmol/L $CuCl_2$ (Figure 1). To identify the gene(s) responsible for the tolerance phenotype, consecutive small site-directed deletions were constructed using the Cre/*loxP* system (see Materials and Methods) (Landeta et al., 2011). The increased copper tolerance observed in CFNX185 was also observed in strains with deletions peΔ10 (124 kb) and peΔ20 (60 kb); whereas deletion peΔ21 (41 kb) showed a wild-type phenotype (Figure 1). This indicated that the gene(s) involved in copper resistance must be localized on a 19 kb region still present in plasmid peΔ21 (Figure 1, white triangle). The organization of the

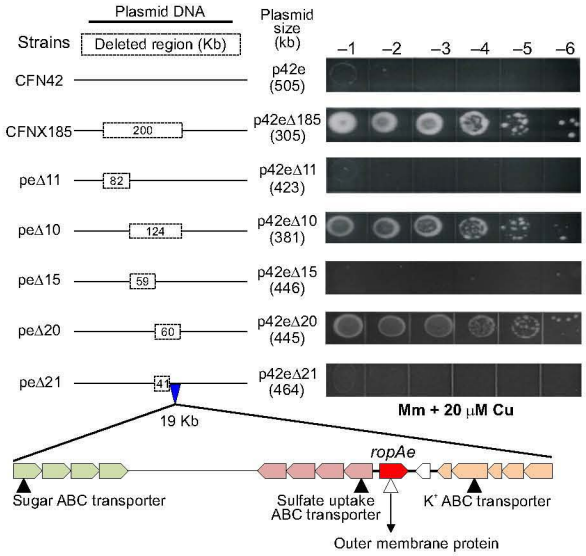

FIGURE 1 Identification of *ropAe* gene through the analysis of the copper-resistance phenotype of wt *R. etli* CFN42 and its isogenic mutants harboring spontaneous (CFNX185) or engineered (Δ11 to Δ21) deletions of plasmid p42e (see Cu sensitivity assay in methods). *R. etli* CFN42 is unable to grow in minimal medium containing 20 μmol/L of $CuCl_2$. Its isogenic mutant, *R. etli* CFNX185, lacking 200-Kb of plasmid p42e, acquired the ability to grow in this medium. The copper-resistance phenotype of mutants harboring sequential deletions of plasmid p42e (peΔ11 to peΔ21), covering the 200-kb region lost in CFNX185, indicates that the absence of a 19-kb fragment increases copper resistance. Four transporter-coding genes contained in this fragment were disrupted by site-directed vector integration (open and filled triangles). Only the mutation of *ropAe* (*RHE_PE00260*) increased Cu resistance (open triangle). Images are representative of five independent experiments

genes localized in this fragment and their predicted functions are listed in Table S5. A schematic representation of some of these genes is shown at the bottom of Figure 1, highlighting the presence of a 4,080 bp four-gene cluster, *RHE_PE00245 to RHE_PE00248*, encoding a putative sugar ABC transport system, a sulphate uptake ABC transport system formed by five proteins *(RHE_PE00259 to RHE_PE00256)*, and a putative OMP (*RHE_PE00260*, 1,017 bp) encoded close to a three-gene cluster, which is part of the KdpA–KdpE (*RHE_PE00266-RHE_PE00262*) potassium transporting system.

3.2 | Disruption of *ropAe* increases copper tolerance in *R. etli* CFN42

The analysis described above indicates that 10 out of 19 proteins encoded in the 19 Kb fragment of pRet42e might be involved in transport of small molecules and cations. To explain the copper resistance phenotype, we hypothesized that these proteins may participate in the uptake of divalent cations, including copper. Under a copper overload the absence of these Cu importers would reduce the transport of metal, resulting in a Cu resistance phenotype. To test this hypothesis, we constructed mutants in genes *RHE_PE00245* and *RHE_PE00259*, coding for solute binding proteins, annotated as components of the sugar and sulfate ABC transport systems, respectively; also in gene *RHE_PE00260*, coding for an uncharacterized OMP, and in *RHE_PE00263*, coding for a putative two-component sensor histidine kinase, belonging to the putative potassium transport system (*kdpA–kdpE*) (Figure 1 and Table S5). The copper resistance phenotype of these mutants was assessed by determining growth of 10-fold serial dilutions on minimal medium plates supplemented with 20 μmol/L $CuCl_2$. The only mutation that produced an increased Cu resistance phenotype was that in the *RHE_PE00260* gene coding for a putative OMP labeled as *ropAe*. The complementation of the *ropAe*$^-$ mutant with the wild-type gene, led to recovery of wild-type levels of Cu sensitivity (Figure 2). These results indicate that the product of *ropAe* gene increases copper sensitivity.

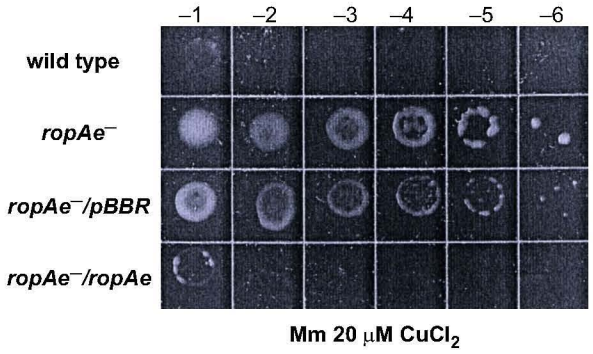

FIGURE 2 The disruption of *ropAe* increases copper resistance. The comparison of the growth presented by wt *R. etli* CFN42, the *ropAe* mutant (*ropAe*−), the mutant complemented with the empty vector (*ropAe*−/pBBR) and with the *ropAe* gene (*ropAe*−/*ropAe*) were assessed with the copper sensitivity plate assay described in the Methods section. Images are representative of five independent experiments

3.3 | Loss of *R. etli ropAe* gene does not enhance Mn^{2+}, Cd^{2+}, Ni^{2+}, Fe^{2+}, Co^2, and Zn^{2+} resistance

According to the Transport Classification Data Base (Saier et al., 2016), most of the 16-strand porins are classified as general or non-specific transporters that allow the diffusion of hydrophilic substrates. To determine if RopAe mediates the sensibility to other divalent metal cations, *R. etli* and its isogenic *ropAe* mutant were grown on plates containing minimal inhibitory concentrations of Mn^{2+}, Cd^{2+}, Ni^{2+}, Fe^{2+}, Co^{2+}, and Zn^{2+} (previously determined for the wild-type strain). We found that the inhibitory effect of all these metals was similar for both strains (Figure S1).

3.4 | The disruption of *ropAe* gene partially reverts the copper sensitivity phenotype of an *R. etli actP* mutant defective in copper detoxification

In the genome sequence of *R. etli* CFN42, ORF *RHE_PE00007* is annotated as an *actP* homolog (*copA* in other bacteria), coding for a $P_{1B\text{-}1}$-copper-transporting ATPase protein, predicted to be located in the inner membrane, whose function is to pump copper out of the cytoplasm (http://genome.annotation.jp/rhizobase/Etli/genes/RHE_PE00007). Landeta et al. (2011) showed that *actP* disruption reduced the copper resistance of this bacterium, suggesting a deficiency in copper efflux. To get insights into the putative porin-like role of *ropAe* in the uptake of copper, we hypothesized that an *actP*::ΩSp mutant, defective in cytoplasmic copper detoxification, should be rescued from a toxic copper overload (7.5 μmol/L $CuCl_2$) if the putative entrance of copper is "closed" by mutation of the *ropAe* gene (*ropAe*::pK18 *mob* Km). As predicted, the *actP*::ΩSp/*ropAe*::pK18mob double mutant increased its copper tolerance in three orders of magnitude, compared to the single *actP* mutant (Figure 3). This result, together with the increased resistance to copper of the *ropAe* mutant, the Cu sensitivity of the complemented *ropAe* mutant, as well as the in silico structural analysis, reinforces the hypothesis that RopAe facilitates copper uptake.

3.5 | The transcription of *ropAe* is upregulated under copper limiting conditions

To determine the effect of copper on the transcription of *ropAe*, its expression was analyzed by qRT-PCR in the exponential growth phase ($OD_{620\,nm}$ = 0.7) of cultures exposed or unexposed for 30 min to a non-inhibitory overload of 0.5 mmol/L $CuCl_2$. These induction conditions were validated by measuring the copper-inducible *actP* gene of *R. etli* CFN42 that encodes a putative $P_{1B\text{-}1}$-Cu^+-ATPase efflux pump (Landeta et al., 2011).

From studies in other bacterial models, such as *Agrobacterium fabrum* C58 (formerly *A. tumefaciens* C58), a close relative of *R. etli*, we assumed that the transcription of *R. etli actP* gene is activated by the

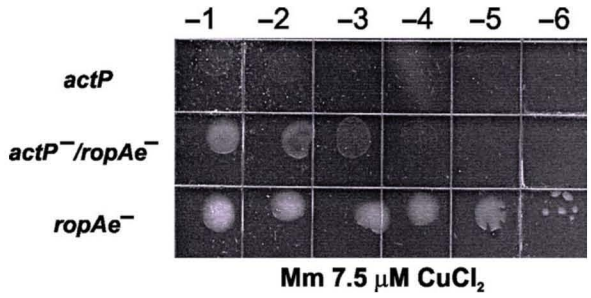

FIGURE 3 The disruption of *ropAe* counteracts the copper sensitivity of an *R. etli actP* mutant. This figure shows the copper resistance assessed as the growth ability of ten-fold serially diluted cultures ($OD_{620\,nm}$ = 0.7) in Mm with 7.5 μmol/L $CuCl_2$.The growth deficiency of an *actP* mutant is assumed to be due to a defective cytoplasmic detoxification of Cu^+ by the inactivation of the P_{1B}-type ATPase encoded in *actP*. The partial growth recovery of the *actP−/ropAe−* double mutant may be due to a reduction in the influx of Cu^+ by disruption of the *ropAe* gene. Images are representative of five independent experiments

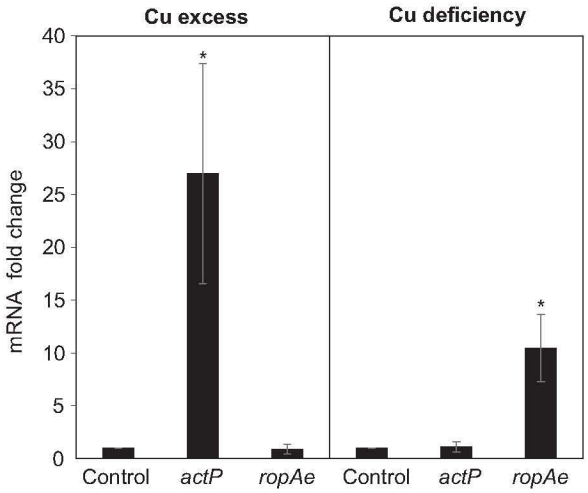

FIGURE 4 The expression of *ropAe* gene is upregulated under copper deficiency. The transcriptional responses of *actP* and *ropAe* to copper excess (PY + 0.5 mM $CuCl_2$) and copper deficiency (PY+ 2mmol/L chelating agent BCDS) were analyzed by qPCR. The copper excess panel shows a 25-fold increase in the mRNA level of *actP* when wt *R. etli* was exposed for 30 min to an overload of 0.5 mmol/L $CuCl_2$. The copper deficiency panel shows a 17-fold increase in the mRNA level of *ropAe* when *R. etli* was culturedovernight in the presence of 2 mmol/L of the Cu^+/Cu^{2+}chelating agent bathocuproine disulfonic acid (BCDS). Both assays were normalized to *hisCd* mRNA level (mean ± *SD.*, *n* = 3) (*t* test, $^{**}p \leq .05$)

monovalent copper-dependent protein CueR that binds to a regulatory region of *actP* and induces its transcription through a conformational DNA change required for transcriptional activation of *actP* (Bittner, Kraus, Schäkermann, & Narberhaus, 2017; Nawapan et al., 2009). As expected, the relative expression of the *actP* gene increased 25-fold after copper treatment (Figure 4). In contrast, no significant difference was found in the expression of *ropAe* in cultures treated similarly (Figure 4).

There are several examples cited in the literature of upregulation of high-affinity porin-encoding genes at lowmetal concentrations (Hohle & O'Brian, 2009; Patzer & Hantke, 1998; Speer et al., 2013). To determine whether *ropAe* is induced under copper limitation, its transcription was assessed in overnight cultures of wt *R. etli* supplemented with 2 mM of the Cu^+/Cu^{2+} chelator agent bathocuproine disulfonic acid (BCDS). BCDS is a charged and membrane impermeable chelator commonly used in studies that require extracellular copper depletion (Asahi et al., 2014; Campos, Guzmán, López-Fernández, & Casado, 2009; Ding et al., 2011; Labbé, Zhu, & Thiele, 1997). As can be seen in Figure 4b, the transcription of *ropAe* increased 10-fold in cultures containing BCDS. The decrease in Cu in overnight cultures was validated by the reduction in *actP* expression in the presence of BCDS (Figure 4b).

3.6 | The *ropAe* gene encodes a putative OMP that shares structural characteristics with porins

According to the annotation of the *R. etli* CFN42 genome, *ropAe* is an ORF of 1,017 nucleotides, located in plasmid p42e, encoding a 338 amino acid protein annotated as an uncharacterized porin predicted to be located in the outer membrane. This predicted location was reinforced using other bioinformatic tools. First, a BlastP search with default settings indicates that the overall sequence of RopAe is a domain shared with members of the Porin_2 family (porins from the α class of Proteobacteria) and that the OM_ channels a superfamily containg 16–18 beta-stranded barrels. To support its outer membrane localization, RopAe was analyzed with two different predictors of subcellular localization (see Methods section for details) with high accuracy for OMPs (Bhasin et al., 2005; Table S3). The multimodular PSORT-B (Gardy et al., 2003) and the CELLO predictor (Yu et al., 2004) are based on *n*-peptide composition (http://cello.life.nctu.edu.tw/). Both methods assigned high probability values to the outer membrane localization of RopAe: the PSORT-B probability was 9.3 of 10 as maximum score; the CELLO probability was 4.105 of 5 as maximum score (Table S3).

The closest homologs are putative OMPs mainly encoded in plasmids of other *R. etli* and *R. leguminosarum* strains, with identities ranging from 96 to 100% with a query coverage ≥95% (Table S6). The identity of RopAe to homologs present in other members of *Rhizobiales* fell in a range from 30% to 50%. To learn more about the structure of RopAe and its closest homologs, their amino acid sequences were analyzed with the transmembrane beta-barrel topology prediction BOCTOPUS2 server (Hayat & Elofsson, 2012) (http://boctopus.bioinfo.se/). All of the analyzed putative OMP porins shared 16 transmembrane β-strands distributed fromamino acids 58 to 338, which form part of the typical β-barrel structure of porins (Table S6). We used RopAe homologs as queries against the CATH/Gene 3D (v4.1) protein domains database to predict the 3D structure adopted by these putative OMPs (Lam et al., 2016; Sillitoe et al., 2015) (http://www.cathdb.info/). This search revealed that all tested RopAe homologs, including the partially characterized OMP from *R. leguminosarum* 248

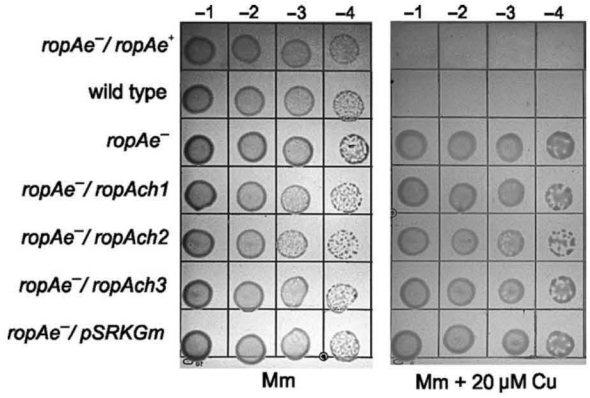

FIGURE 5 The enhanced copper-resistance phenotype of *R. etli* CFN42 *ropAe* mutant (*ropAe*⁻) could not be reduced by genetic complementation with the paralogs genes *ropAch1*, *ropAch2*, and *ropAch3*. The *ropAe*⁻ mutants harboring each one of the paralogs genes were designated *ropAe*⁻/*ropAch1*, *ropAe*⁻/*ropAch2*, and *ropAe*⁻/*ropAch3*. As controls we used the *ropAe*⁻ mutant complemented with the *ropAe* wild type (*ropAe*⁻/ropAe) and *ropAe*⁻ mutant complemented with the empty vector (*ropAe*⁻/pSRKGm). The resistance or sensitivity phenotypes were assessed by growth in the presence or absence of 20 µmol/L Cu with the copper sensitivity plate assay described in the experimental section. Images are representative of three independent experiments

(de Maagd et al., 1992), share a porin domain of 266 amino acids present in the outer membrane proteins grouped in the OMP IIIA family. CATH/Gene 3D (v4.1) predicts "transport" as the associated biological process and "outer membrane" as the cellular component (Table S6). We also identified an additional putative porin present in multiple species of *Mesorhizobium* (NCBI Taxid 3744) that share 31% identity with RopAe. However, CATH/Gene 3D assigned these mesorhizobial porins to the OmpP2 family (Table S6).

3.7 | The plasmid-borne RopAe protein shares structural characteristics with three RopA homologs encoded in the chromosome of *R. etli* CFN42

As the *R. etli* genome is characterized by a high level of gene redundancy (Flores et al., 1987) we searched for RopAe homologs through a BLASTP search using RopAe as query against the proteome of *R. etli* CFN42. This search retrieved three chromosomally encoded RopAe homologs: RHE_CH01349 (RopAch1), RHE_CH02437 (RopAch2) and RHE_CHO3578 (RopAch3) which share 59% and 58% identities respectively with RopAe (Table S7). Although the four putative porins share structural characteristics such as the 16 transmembrane β strands and the predicted 3D model (Table S7), the increase in copper resistance of *ropAe* mutant could not be complemented with the chromosomally encoded *ropA* genes present in *cis*. To get insight into the functional role of RopAch1, RopAch2, and RopAch3 we attempted to obtain mutants by vector integration; however, Southern blot analyses revealed multiple fragments of unexpected sizes suggesting that these mutants contain more than one vector integration (data not shown). This is supported by a pairwise nucleotide sequence alignment that revealed an 80% to 90% identity shared among the three chromosomally encoded *ropA* genes (Table S8).

Alternatively, the functional redundancy of these genes in copper transport was investigated using a genetic complementation approach. The three chromosomally-encoded *ropAch* genes and *ropAe* were cloned into the pSRKGm vector and introduced into the *R. etli* mutant *ropAe*::pK18mob by conjugation. The transconjugants harboring *ropAch1*, *ropAch2* or *ropAch3* maintained the same copper resistance level as the mutant *ropAe*::pK18mob. In contrast, the transconjugant *ropAe* gene was able to reduce the enhanced copper resistance of the *ropAe* mutant. These data support the hypothesis that RopAe and RopAchs play different roles (Figure 5).

Furthermore, the location of the four *ropA* homologs in different gene contex (Table S7) suggests that these genes could be following different evolutionary pathways, as has been observed in genes from other models analyzed with different bioinformatic gene contex tools (Dewey, 2011; Martinez-Guerrero, Ciria, Abreu-Goodger, Moreno-Hagelsieb, & Merino, 2008; Puggioni et al., 2014; Seret & Baret, 2011). In support of the functional divergence hypothesis and outer membrane protein localization, it was recently reported that two *R. etli* RopAe paralogs, RopAch1 and RopAch2, were found to be expressed in the exoproteome of this organism when exposed to the plant flavonoid naringenin (Meneses et al., 2017).

3.8 | RopAe defines a new class of porins distantly related to those involved in Mn and Cu uptake

The BlastP search in the nonredundant database for *R. etli* RopAe homologs also revealed the presence of a conserved domain that spans from amino acid 33 to 322 (E value 1.93^{-72}) and groups these proteins in the Porin_2 (PF02530) Pfam. This family clusters OMPs from α-proteobacteria that share this domain. The compilation of these data revealed that the presence of multiple RopA porins is a widespread characteristic in the genomes of members of the *Rhizobiales*. The number of *ropA* per genome varies from one to nine, with three being the most frequent number of *ropA* per genome (Table S4).

To understand the evolutionary relationships of RopAe and paralogs in relation to other OMPs involved in metal trafficking, we aligned the rhizobial porins grouped in the Porin_2 (PF02530) family, characterized porins involved in Mn (MnoP from *B. japonicum*) and Cu (MspA–D from *M. smegmatis*) uptake and the best-studied porin of γ-proteobacteria, OmpC. The *P. aeruginosa* OprC, OprD, OprE, OprG, and OprN porins, reported to be downregulated in Cu-adapted culture, were excluded from the alignment because they have an e-value higher than the used cut-off ($<1e^{-3}$) (Table S8). The maximum likelihood-based phylogeny inferred from the alignment of 127 amino acid sequences of rhizobial porins was used to define major groups of proteins (Figure S2) including non-RopA protein homologs found in rhizobia. A phylogeny containing the closest RopA homologs as well as multiple out-group characterized proteins (Figure 6) clearly shows that rhizobial RopAe homologs belong to a new class of porins distantly related to those known to be involved in Mn and Cu trafficking.

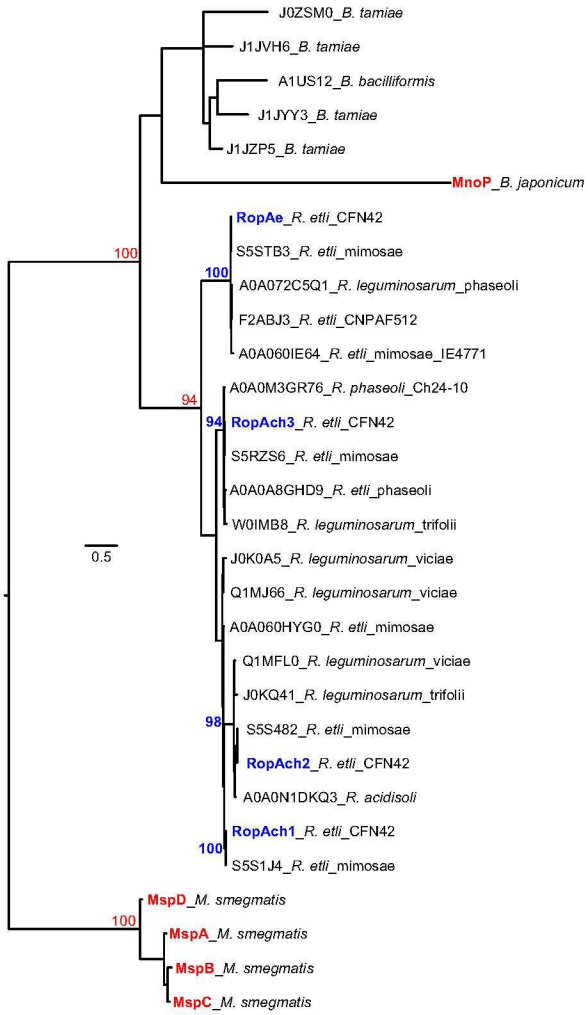

FIGURE 6 The putative RopA OMPs of rhizobia define a new family of porins phylogenetically unrelated to already characterized metal uptake porins. A maximum likelihood phylogenetic tree inferred from a subset of 30 proteins analyzed in a larger phylogeny (Figure S2) showed that RopA proteins clearly define a monophyletic clade of proteins distantly related to characterized metal uptake proteins MnoP (Mn) and MspA-D (Cu), supported by bootstrap values ≥90. The diversification of *R. etli* and *R. leguminosarum* RopAe and its homologs RopAch, is evinced by its separation into two clades. The copper porins of *M. smegmatis*, the general porin OmpC of *E. coli* and the high-affinity Mn^{2+} importer MnoP of *B. japonicum* used as outgroups, are distantly related to *R. etli* porins (Figure S2). The scale bar indicates the expected number of amino acid substitutions per site under the LG + G + f model.

4 | DISCUSSION

As studies on copper uptake are extremely scarce in prokaryotes, the identification and characterization of proteins involved in the acquisition of this metal are a topic of major interest in the study of copper homeostasis in bacteria. The best characterized bacterial proteins involved in copper acquisition are the MspA porin and its paralogs in *M. smegmatis* (Speer et al., 2013). In gram-negative bacteria there is only a study of putative porins which are down-regulated in copper-adapted cultures of *P. aeruginosa* (Teitzel et al., 2006).

As part of a research project focused on understanding copper trafficking in the facultative diazotroph *R. etli* CFN42, we searched for mutants with enhanced copper resistance in a collection of *R. etli* mutants that have lost whole plasmids or fragments of them, both spontaneously or through programmed deletions. Using this approach we identified a plasmid-encoded gene, annotated as *ropAe*, whose disruption enhanced copper resistance, and we further determined that the wt sensitivity could be recovered upon introduction of the wt *ropAe* gene. According to the annotation of the genome at NCBI, the *ropAe* gene (*RHE_PE00260*) is located in plasmid p42e and encodes a protein of 338 amino acids, identified as a porin outer membrane protein. The presence of a putative domain containing a multi-pass transmembrane protein that forms a pore of passive diffusion across the outer membrane is predicted between amino acids 33 and 332. Based on this information we hypothesized that *ropAe* encodes a putative OMP involved in copper uptake. In this study we performed bioinformatics, genetics and gene expression analyses to examine this hypothesis.

The OM localization was supported by an in silico analysis of the amino acid sequence using two different predictors with high accuracy for OMPs: PSORT-B and CELLO. Both predictors assigned high probability values for OM localization. The structural characteristics of RopAe, such as the transmembrane beta-barrel topology and the 3D structure adopted by OMPs were well supported using BOCTOPUS2 server and CATH/gene 3D.

We hypothesized that if *ropAe* encodes a porin-like protein that contributes to the normal supply of copper, we would expect that a *ropAe* mutant will show alterations in copper-dependent functions due to a reduction in copper availability. In support of this proposal we found that an *R. etli* CFN42 *actP* mutant, defective in cytoplasmic detoxification of copper due to a Tn*5* disruption of its P_{1B1}-Cu-transporting ATPase, cannot grow in the presence of 7.5 μmol/L $CuCl_2$. However, the *actP/ropAe* double mutant partially recovered its growth ability in 7.5 μmol/L $CuCl_2$. The results of our experiments suggest that this is due to a deficiency in copper supplementation produced by the disruption of *ropAe*. These data need to be corroborated by determining the transport activity of RopAe.

The inhibitory effect of other divalent metal cations was similar for both strains (Figure S1). This result suggests that RopAe does not facilitate the uptake of metals other than copper, or alternatively, that RopAe might mediate a defective uptake for these metals that does not result in toxic levels in the cytoplasm. Otherwise, there could be a functional redundancy among different porins for these metals, as reported for *M. smegmatis* (Speer et al., 2013).

We also determined that the transcription of *ropAe* did not change when *R. etli* was exposed to a $CuCl_2$ overload. As expected, the expression of *actP*, coding for a copper efflux pump, increased 25-fold after exposure to the same copper overload. In contrast, under copper-limiting conditions, the transcription of *ropAe* is

upregulated and the expression of *actP* is turned off. Low copper conditions were established with the chelator agent BCDS, a charged and membrane impermeable chelator commonly used in studies that require an extracellular copper depletion condition. The BCDS chelates Cu^{+} but is also known to bind Cu^{2+}, forcing a tetrahedral chelator geometry that facilitates the reduction in Cu^{2+} to Cu^{+} (Asahi et al., 2014; Campos et al., 2009; Ding et al., 2011; Labbé et al., 1997). The BLASTP search was restricted to the genome of *R. etli* CFN42 using its own *ropAe* as query; it revealed the presence of three *ropAe* homologs located in distant regions of the *R. etli* CFN42 chromosome. As none of these homologs substitutes for *ropAe*, we proposed a functional divergence between RopAe and its chromosomally encoded homologs RopAch1, RopAch2, and RopAch3, which is also supported by the phylogenetic analysis that separates the plasmid and chromosomally encoded RopA homologs into two different clades.

A complex evolutionary story of porin-encoding genes is assumed by the presence of gene redundancy in both prokaryotic and mitochondrial genomes. The presence of multiple copies may have an important role in the stress response of the cell, maintaining a function in the case of loss of primary genes (Bay, Hafez, Young, & Court, 2012; Pinne et al., 2006; Saccone et al., 2003). For instance, in *M. smegmatis*, the porin-mediated copper uptake requires the cooperation of redundant porin homologs; a phenotype of increased Cu resistance could only be obtained in double and triple porin mutants (Speer et al., 2013).

All the data described above strongly support the hypothesis that *ropAe* encodes a porin-like protein that facilitates copper diffusion by increasing its expression under low copper concentrations. Future experiments will address the transport properties of RopAe and its role, if any, in symbiosis with *Phaseolus vulgaris* plants in the presence of copper.

ACKNOWLEDGMENTS

We gratefully acknowledge the funding received from UNAM-DGAPA-PAPIIT, grant number IN209815. A. González-Sánchez is a doctoral student of the *Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM)* and received a fellowship from CONACYT (No. 290161).

The authors are grateful to Laura Cervantes for her skillful technical assistance, Maria Luisa Tabche (IBT-UNAM) for her valuable help with qRT-PCR experiments and Susana Brom for helpful discussions and for critically reviewing the manuscript.

CONFLICT OF INTEREST

None declared

ORCID

Alejandro García-de los Santos http://orcid.org/0000-0002-3387-9280

REFERENCES

Asahi, H., Tolba, M. E., Tanabe, M., Sugano, S., Abe, K., & Kawamoto, F. (2014). Perturbation of copper homeostasis is instrumental in early developmental arrest of intraerythrocytic *Plasmodium falciparum*. *BMC Microbiology*, *14*(1), 167. Available at: http://bmcmicrobiol.biomedcentral.com/articles/10.1186/1471-2180-14-167 [Accessed June 6, 2017]

Ayres, E. K., Thomson, V. J., Merino, G., Balderes, D., & Figurski, D. H. (1993). Precise deletions in large bacterial genomes by vector-mediated excision (VEX). *Journal of Molecular Biology*, *230*(1), 174–185. https://doi.org/10.1006/jmbi.1993.1134

Bavoil, P., Nikaido, H., & von Meyenburg, K. (1977). Pleiotropic transport mutants of *Escherichia coli* lack porin, a major outer membrane protein. *MGG Molecular & General Genetics*, *158*(1), 23–33. https://doi.org/10.1007/BF00455116

Bay, D.C., Hafez, M., Young, M. J., & Court, D. A. (2012). Phylogenetic and coevolutionary analysis of the β-barrel protein family comprised of mitochondrial porin (VDAC) and Tom40. *Biochimica et Biophysica Acta - Biomembranes*, *1818*, 1502–1519. Available at: https://doi.org/10.1016/j.bbamem.2011.11.027

Bhasin, M., Garg, A., & Raghava, G. P. S. (2005). PSLpred: Prediction of subcellular localization of bacterial proteins. *Bioinformatics*, *21*(10), 2522–2524. https://doi.org/10.1093/bioinformatics/bti309

Bittner, L.-M., Kraus, A., Schäkermann, S., & Narberhaus, F. (2017). The copper efflux regulator *CueR is subject to ATP*-dependent proteolysis in *Escherichia coli*. *Frontiers in Molecular Biosciences*, *4*, 9. Available at: http://journal.frontiersin.org/article/10.3389/fmolb.2017.00009/full [Accessed June 6, 2017].

Boal, A. K., & Rosenzweig, A. C. (2009). Structural biology of copper trafficking. *Chemical Reviews*, *109*(10), 4760–4779. Available at: http://pubs.acs.org/doi/abs/10.1021/cr900104z [Accessed December 20, 2016].

Brom, S., García-de los Santos, A., Stepkowsky, T., Flores, M., Dávila, G., Romero, D., Palacios, R.. (1992). Different plasmids of *Rhizobium leguminosarum* bv. phaseoli are required for optimal symbiotic performance. *Journal of Bacteriology*, *174*(16), 5183–5189.https://doi.org/10.1128/jb.174.16.5183-5189.1992

Campos, C., Guzmán, R., López-Fernández, E., & Casado, Á. (2009). Evaluation of the copper(II) reduction assay using bathocuproinedisulfonic acid disodium salt for the total antioxidant capacity assessment: The CUPRAC–BCS assay. *Analytical Biochemistry*, *392*(1), 37–44. https://doi.org/10.1016/j.ab.2009.05.024

Cubillas, C., Vinuesa, P., Tabche, M. L., García-de los Santos, A. (2013). Phylogenomic analysis of cation diffusion facilitator proteins uncovers Ni2 + /Co2 + transporters. *Metallomics*, *5*(12).

Cubillas, C., Vinuesa, P., Tabche, M. L., Dávalos, A., Vázquez, A., Hernández-Lucas, I., … García-de los Santos, A. (2014). The cation diffusion facilitator protein EmfA of *Rhizobium etli* belongs to a novel subfamily of Mn^{2 +}/Fe^{2 +} transporters conserved in α-proteobacteria. *Metallomics*, *6*(10).

Danilchanka, O., Pavlenok, M., & Niederweis, M. (2008). Role of porins for uptake of antibiotics by *Mycobacterium smegmatis*. *Antimicrobial Agents and Chemotherapy*, *52*(9), 3127–3134. [Accessed June 28, 2017]. https://doi.org/10.1128/AAC.00239-08

de Maagd, R. A., Mulders, I. H., Cremers, H. C., Lugtenberg, B. J. (1992). Cloning, nucleotide sequencing, and expression in *Escherichia coli* of a *Rhizobium leguminosarum* gene encoding a symbiotically repressed outer membrane protein. *Journal of Bacteriology*, *174*(1), 214–221. https://doi.org/10.1128/jb.174.1.214-221.1992

Dewey, C. N. (2011). Positional orthology: Putting genomic evolutionary relationships into context. *Briefings in Bioinformatics*, *12*(5), 401–412. [Accessed June 6, 2017]. https://doi.org/10.1093/bib/bbr040

Ding, X., Xie, H., & Kang, Y. J. (2011). The significance of copper chelators in clinical and experimental application. *The Journal of Nutritional Biochemistry*, *22*(4), 301–310. [Accessed June 6, 2017]. https://doi.org/10.1016/j.jnutbio.2010.06.010

Fellay, R., Frey, J., & Krisch, H. (1987). Interposon mutagenesis of soil and water bacteria: A family of DNA fragments designed for in vitro insertional mutagenesis of gram-negative bacteria. *Gene*, *52*(2–3), 147–154. [Accessed December 21, 2016]. https://doi.org/10.1016/0378-1119(87)90041-2

Festa, R. A., & Thiele, D. J. (2011). Copper: An essential metal in biology. *Current Biology*, *21*(21), R877–R883. Available at: https://doi.org/10.1016/j.cub.2011.09.040

Flores, M., González, V., Brom, S., Martinez, E., Piñero, D., Romero, D., … Palacios, R. (1987). Reiterated DNA sequences in *Rhizobium and Agrobacterium* spp. *Journal of Bacteriology*, *169*(12), 5782–5788. [Accessed December 20, 2016] https://doi.org/10.1128/jb.169.12.5782-5788.1987

Gardy, J. L., et al. (2003). PSORT-B: Improving protein subcellular localization prediction for Gram-negative bacteria. *Nucleic Acids Research*, *31*(13), 3613–3617. https://doi.org/10.1093/nar/gkg602

Gonzalez, V., Santamaría, R. I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., … Dávila, G. (2006). The partitioned *Rhizobium etli* genome: Genetic and metabolic redundancy in seven interacting replicons. *Proceedings of the National Academy of Sciences*, *103*(10), 3834–3839. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0508502103 [Accessed September 10, 2016].

Hayat, S., & Elofsson, A. (2012). BOCTOPUS: Improved topology prediction of transmembrane β barrel proteins. *Bioinformatics*, *28*(4), 516–522. https://doi.org/10.1093/bioinformatics/btr710

Hohle, T. H., Franck, W. L., Stacey, G., & O'Brian, M. R. (2011). Bacterial outer membrane channel for divalent metal ion acquisition. *Proceedings of the National Academy of Sciences*, *108*(37), 15390–15395. https://doi.org/10.1073/pnas.1110137108

Hohle, T. H., & O'Brian, M. R. (2009). The *mntH* gene encodes the major Mn(2 +) transporter in *Bradyrhizobium japonicum* and is regulated by manganese via the Fur protein. *Molecular Microbiology*, *72*(2), 399–409. https://doi.org/10.1111/j.1365-2958.2009.06650.x

Hynes, M. F., & McGregor, N. (1990). Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. *Molecular Microbiology*, *4*, 567–574. https://doi.org/10.1111/j.1365-2958.1990.tb00625.x

Khan, S. R., Gaines, J., Roop, R. M., & Farrand, S. K. (2008). Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Applied and Environmental Microbiology*, *74*(16), 5053–5062. https://doi.org/10.1128/AEM.01098-08

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., & Peterson, K. M. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene*, *166*(1), 175–176. https://doi.org/10.1016/0378-1119(95)00584-1

Labbé, S., Zhu, Z., & Thiele, D. J. (1997). Copper-specific transcriptional repression of yeast genes encoding critical components in the copper transport pathway. *The Journal of Biological Chemistry*, *272*(25), 15951–15958. https://doi.org/10.1074/jbc.272.25.15951

Lam, S. D, Dawson, N. L., Das, S., Sillitoe, I., Ashford, P., Lee, D., … Lees, J. G. (2016). Gene3D: Expanding the utility of domain assignments. *Nucleic Acids Research*, *44*(D1), D404–D409. Available at: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv1231 [Accessed January 13, 2017].

Landeta, C., Dávalos, A., Cevallos, M. Á., Geiger, O., Brom, S., & Romero, D. (2011). Plasmids with a chromosome-like role in rhizobia. *Journal of Bacteriology*, *193*(6), 1317–1326. Available at: http://jb.asm.org/cgi/doi/10.1128/JB.01184-10 [Accessed September 10, 2016].

Li, X. Z., Nikaido, H., & Williams, K. E. (1997). Silver-resistant mutants of *Escherichia coli* display active efflux of Ag+ and are deficient in porins. *Journal of Bacteriology*, *179*(19), 6127–6132. https://doi.org/10.1128/jb.179.19.6127-6132.1997

Lugtenberg, B., & Van Alphen, L. (1983). Molecular architecture and functioning of the outer membrane of *Escherichia coli* and other gram-negative bacteria. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, *737*(1), 51–115. Available at: http://www.sciencedirect.com/science/article/pii/030441578390014X.

Lutkenhaus, J. F. (1977). Role of a major outer membrane protein in Escherichia coli. *Journal of Bacteriology*, *131*(2), 631–637.

Macomber, L., & Imlay, J. A. (2009). The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(20), 8344–8349. Available at: http://www.pnas.org/content/106/20/8344.short.

Marsh, J. L., Erfle, M., & Wykes, E. J. (1984). The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. *Gene*, *32*(3), 481–485. https://doi.org/10.1016/0378-1119(84)90022-2

Martinez-Guerrero, C. E., Ciria, R., Abreu-Goodger, C., Moreno-Hagelsieb, G., & Merino, E. (2008). GeConT 2: Gene context analysis for orthologous proteins, conserved domains and metabolic pathways. *Nucleic Acids Research*, *36*(Web Server issue), W176–W180. https://doi.org/10.1093/nar/gkn330

Meneses, N., Taboada, H., Dunn, M. F., Vargas, M. D. C., Buchs, N., Heller, M., & Encarnación, S. (2017). The naringenin-induced exoproteome of *Rhizobium etli* CE3. *Archives of Microbiology*, *199*(5), 737–755. https://doi.org/10.1007/s00203-017-1351-8

Nawapan, S., Charoenlap, N., Charoenwuttitam, A., Saenkham, P., Mongkolsuk, S., & Vattanaviboon, P. (2009). Functional and expression analyses of the cop operon, required for copper resistance in *Agrobacterium tumefaciens*. *Journal of Bacteriology*, *191*(16), 5159–5168. https://doi.org/10.1128/JB.00384-09

Pattengale, N. D., Aberer, A., Swenson, K., Stamatakis, A., & Moret, B. (2011). Uncovering hidden phylogenetic consensus in large data sets. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *8*(4), 902–911. https://doi.org/10.1109/TCBB.2011.28

Patzer, S. I., & Hantke, K. (1998). The ZnuABC high-affinity zinc uptake system and its regulator Zur in *Escherichia coli*. *Molecular Microbiology*, *28*(6), 1199–1210. https://doi.org/10.1046/j.1365-2958.1998.00883.x

Pinne, M., Denker, K., Nilsson, E., Benz, R., Bergström, S. (2006). The BBA01 protein, a member of paralog family 48 from *Borrelia burgdorferi*, is potentially interchangeable with the channel-forming protein P13. *Journal of Bacteriology*, *188*(12), 4207–4217. https://doi.org/10.1128/JB.00302-06

Puggioni, V., Dondi, A., Folli, C., Shin, I., Rhee, S., & Percudani, R. (2014). Gene context analysis reveals functional divergence between hypothetically equivalent enzymes of the Purine–Ureide Pathway. *Biochemistry*, *53*(4), 735–745. Available at: http://pubs.acs.org/doi/abs/10.1021/bi4010107 [Accessed June 6, 2017].

Quandt, J., & Hynes, M. F. (1993). Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. *Gene*, *127*(1), 15–21. https://doi.org/10.1016/0378-1119(93)90611-6

Rubino, J. T., & Franz, K. J. (2012). Coordination chemistry of copper proteins: How nature handles a toxic cargo for essential function. *Journal of Inorganic Biochemistry*, *107*(1), 129–143. https://doi.org/10.1016/j.jinorgbio.2011.11.024

Saccone, C., Caggese, C., D'Erchia, A. M., Lanave, C., Oliva, M., Pesole, G. (2003). Molecular clock and gene function. *Journal of Molecular Evolution*, *57*(SUPPL. 1), 277–285. https://doi.org/10.1007/s00239-003-0037-9

Saier, M. H., Reddy, V. S., Tsu, B. V., Ahmed, M. S., Li, C., & Moreno-Hagelsieb, G. (2016). The transporter classification database (TCDB): Recent advances. *Nucleic Acids Research*, *44*(D1), D372–D379. https://doi.org/10.1093/nar/gkv1103

Salazar, E., Díaz-Mejía, J. J., Moreno-Hagelsieb, G., Martínez-Batallar, G., Mora, Y., Mora, J., & Encarnación, S. (2010). Characterization of the NifA-RpoN regulon in *Rhizobium etli* in free life and in symbiosis with Phaseolus vulgaris. *Applied and Environmental Microbiology*, *76*(13), 4510–4520. https://doi.org/10.1128/AEM.02007-09

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: A laboratory manual*. NY: Cold spring harbor laboratory press

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., & Pühler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: Selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene*, *145*(1), 69–73. https://doi.org/10.1016/0378-1119(94)90324-7

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols*, *3*(6), 1101–1108. Available at: http://www.nature.com/doifinder/10.1038/nprot.2008.73 [Accessed February 10, 2017].

Segovia, L., Young, J. P., & Martínez-Romero, E. (1993). Reclassification of American *Rhizobium leguminosarum* biovar phaseoli type I strains as *Rhizobium etli* sp. nov. *International Journal of Systematic Bacteriology*, *43*(2), 374–377. https://doi.org/10.1099/00207713-43-2-374

Seret, M.-L., & Baret, P. V. (2011). IONS: Identification of Orthologs by neighborhood and similarity-an automated method to identify orthologs in chromosomal regions of common evolutionary ancestry and its application to hemiascomycetous yeasts. *Evolutionary Bioinformatics Online*, *7*, 123–133.

Sillitoe, I., et al. (2015). CATH: Comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Research*, *43*(D1), D376–D381. https://doi.org/10.1093/nar/gku947

Simon, R. (1984). High frequency mobilization of gram-negative bacterial replicons by the in vitro constructed Tn5-Mob transposon. *Molecular & General Genetics : MGG*, *196*(3), 413–420. https://doi.org/10.1007/BF00436188

Speer, A., et al. (2013). Porins increase copper susceptibility of *Mycobacterium tuberculosis*. *Journal of Bacteriology*, *195*(22), 5133–5140. https://doi.org/10.1128/JB.00763-13

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stephan, J., Bender, J., Wolschendorf, F., Hoffmann, C., Roth, E., Mailänder, C., … Niederweis, M. (2005). The growth rate of *Mycobacterium smegmatis* depends on sufficient porin-mediated influx of nutrients. *Molecular Microbiology*, *58*(3), 714–730. https://doi.org/10.1111/j.1365-2958.2005.04878.x

Teitzel, G. M., Geddie, A., Susan, K., Kirisits, M. J., Whiteley, M., Parsek, M. R. (2006). Survival and growth in the presence of elevated copper: Transcriptional profiling of copper-stressed *Pseudomonas aeruginosa*. *Journal of Bacteriology*, *188*(20), 7242–7256. https://doi.org/10.1128/JB.00837-06

Yu, C.-S., Lin, C.-J., & Hwang, J.-K. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on *n* -peptide compositions. *Protein Science*, *13*(5), 1402–1406. https://doi.org/10.1110/(ISSN)1469-896X

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: González-Sánchez A, Cubillas C, Miranda F, Dávalos A, García-de los Santos A. The *ropAe* gene encodes a porin-like protein involved in copper transit in *Rhizobium etli* CFN42. *MicrobiologyOpen*. 2017;e573. https://doi.org/10.1002/mbo3.573

Material Suplementario

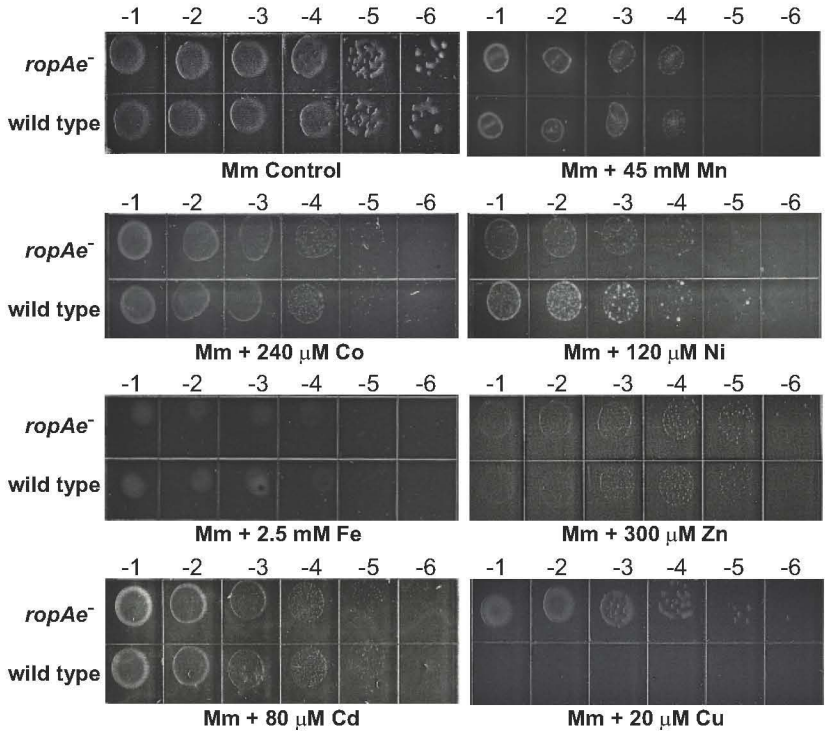

Fig. S1. The absence of RopAe does not increase the resistance of *R. etli* CFN42 to Mn^{2+}, Fe^{2+}, Zn^{2+}, Co^{2+}, Ni^{2+} and Cd^{2+}. The possible role of *ropAe* gene in the transport of Mn^{2+}, Fe^{2+}, Zn^{2+}, Co^{2+}, Ni^{2+} and Cd^{2+} was assessed comparing the minimal inhibitory concentrations (MIC) of these metals that reduce the growth of *R. etli* wt and its isogenic *ropAe* mutant. As observed with copper (bottom of the figure) the higher MIC of *ropAe* mutant is indicative that RopAe facilitate the transport of this metal. The same Mn^{2+}, Fe^{2+}, Zn^{2+}, Co^{2+}, Ni^{2+} and Cd^{2+} MICs for both strains suggest that RopAe does not facilitate the uptake of these metals or there are redundant transporters for these metals.

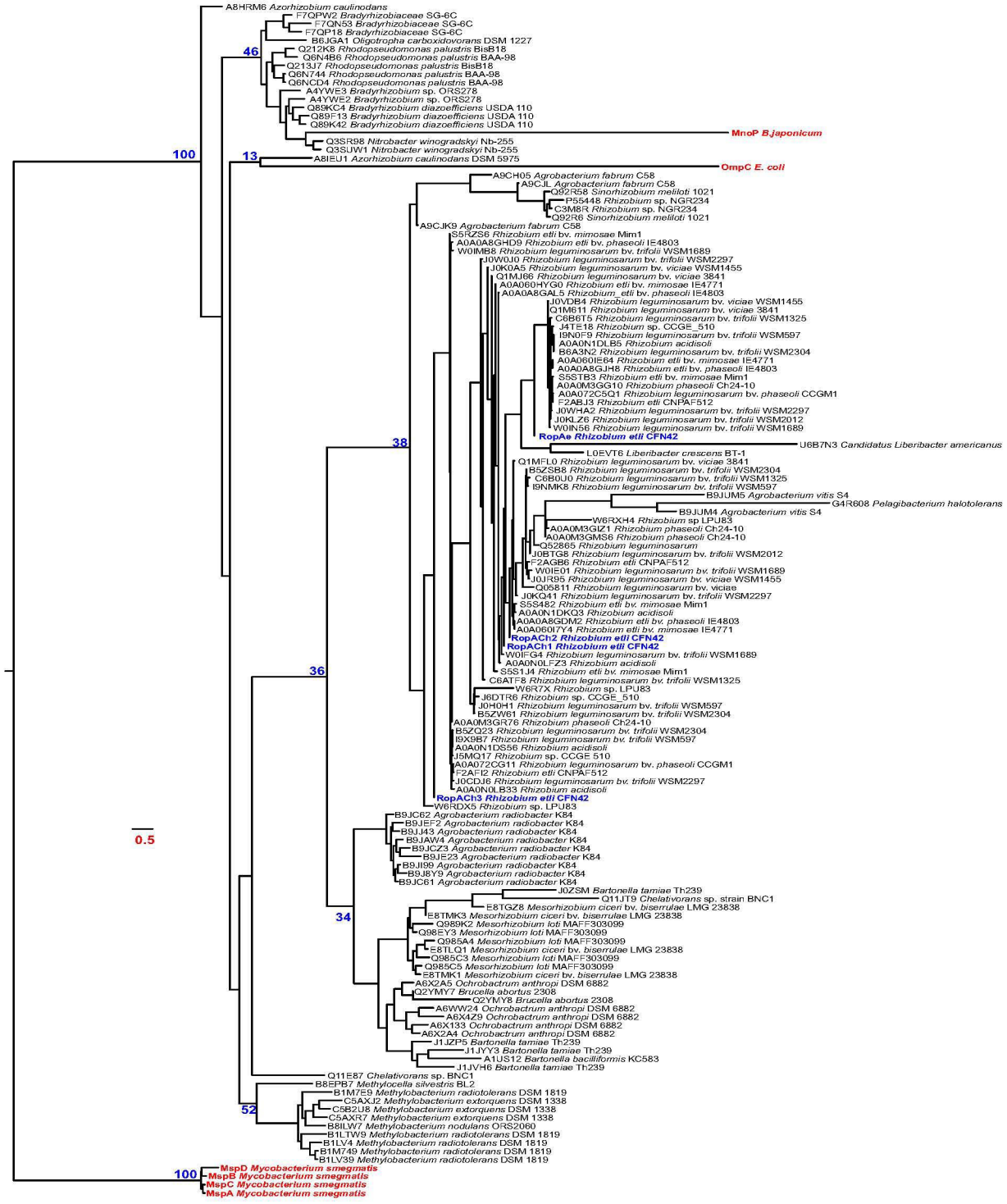

Fig. S2.- Phylogenetical distribution of the Porin_2 family (PF02530). The phylogeny includes 127 aminoacids sequences from aplha proteobacteria belonging to 35 species of Rhizobiales and OmpC sequence of *E. coli*, MnoP sequence of *B. japonicum*, MspA, MspB, MspC, MspD sequences of *M. smegmatis* these included as out group.

Table S1: Bacterial strains and plasmids used in this study.

***Rhizobium etli* strains**	**Relevant Genotype**	**References**
CFN42	Wild type, Nal^r, p42e 505 kb	Segovia et al. 1993
CFNX185	CFN42 p42e lacking 210.097 kb (from 87,541 to 297, 638)	Brom et al. 1992
peΔ11	CFN42 p42eΔ11, lacking 82.028 kb (from 91,132 to 173,160)	Landeta et al. 2011
peΔ10	CFN42 p42eΔ10, lacking 124.160 kb (from 173,900 to 298,060)	Landeta et al. 2011
peΔ15	CFN42 p42eΔ15, lacking 59 kb (from 174,447 to 233,534)	This study
peΔ20	CFN42 p42eΔ20, lacking 60 kb (from 233,531 to 294,302)	This study
peΔ21	CFN42 p42eΔ21, lacking 41.194 kb (from 233,531 to 274,725)	This study
PE00245 Km^r	CFN42 *RHE_PE00245* ::pK18mob Km^r	This study
PE00259 Km^r	CFN42 *RHE_PE00259* ::pK18mob Km^r	This study
CFN42 *ropAe* $^-$	CFN42 *ropAe* ::pK18mob Km^r	This study
CFN42 *kdpD* $^-$	CFN42 *kdpD* ::pK18mob Km^r	This study
CFN42 *actP* $^-$	CFN42 *actP* ::ΩSp^r	This study
CFN42 *ropAe* $^-$ */ropAe*	CFN42 *ropAe* ::pK18mob Km^r complemented with *ropAe* wild type	This study
	cloned into pBBR1MCS5 Gm^r	This study
CFN42 *ropAe* $^-$ *actP* $^-$	CFN42 *ropAe* ::pK18mob Km^r actP::ΩSp^r	This study
CFN42 ropAe-/ropAe	CFN42 *ropAe* ::pK18mob/complemented with *ropAe* into pSRK Gm^r	This study
CFN42 *ropAe -/ropAch1*	CFN42 *ropAe* ::pK18mob/complemented with *ropAch* 1 into pSRK Gm^r	This study
CFN42 *ropAe -/ropAch* 2	CFN42 *ropAe* ::pK18mob/complemented with *ropAch* 2 into pSRK Gm^r	This study
CFN42 *ropAe -/ropAch* 3	CFN42 *ropAe* ::pK18mob/complemented with *ropAch3* into pSRK Gm^r	This study
***E. coli* strains and plasmids**	**Relevant Genotype**	**Reference**
DH5α	host for recombinant plasmids, Nal^r	Stratagene
S17-1	Donor for conjugation	Simon 1984
DH5α/pK18 mob Km^r	pK18 suicide vector, mob, Km^r used for gene disruption	Schäfer et al. 1994
DH5α/pK18mob/*ropAe*	788 bp EcoRI-HindIII fragment of *ropAe* cloned into pK18mob Km^r	This study
DH5α/pK18mob/*PE00245*	463 bp EcoRI-XbaI frag. of *PE00245* gene cloned into pK18mob Km^r	This study
DH5α/pK18mob/*PE00249*	354 bp SmaI-HindIII frag. of *PE00249* gene cloned into pK18 mob Km^r	This study
DH5α/pK18mob/*PE00259*	581 bp EcoRI-HindIII frag. of *PE00259* gene cloned into pK18mob Km^r	This study
DH5α/pK18mob/*kdpD*	547 bp EcoRI-HindIII frag. of *kdpD gene* cloned into pK18mob Km^r	This study
DH5α/pBBR1MCS-5	Broad host range cloning vector, *mob* , Gm^r	Kovach et al. 1995
DH5α/pSRK Gm^r	pBBRMCS-5-derived expression vector lac promoter, *lacI* ^{q}, *lacZ* α^+, Gm^r	Khan et al 2008
DH5α/pSRK Gm^r/*ropAe*	1,211 bp HindIII-KpnI amplicon containing *ropAe* into pSRK Gm^r	This study
DH5α/pSRK Gm^r/*ropAch1*	1,054 bp NdeI-BamHI amplicon containing ropAch1 into pSRK Gm^r	This study
DH5α/pSRK Gm^r/*ropAch2*	1,081 bp BamHI-KpnI amplicon containing *ropAch2* into pSRK Gm^r	This study
DH5α/pSRK Gm^r/*ropAch3*	1,104 bp BamHI-KpnI amplicon containing *ropAch3* into pSRK Gmr	This study
Table S1. Bacterial strains and plasmids used in this study (Continue)		
***E. coli* strains and plasmids**	**Relevant Genotype**	**References**
DH5α/pBBR1MCS-5/*ropAe*	1,249 bp KpnI-XbaI containing *ropAe* gene in pBBR1MCS-5 Gm^r	This study
pVEX1311	VEX derivative with *loxP* , Spr	Ayres et al. 1993
pIC20RDA	pBR322 derivative *loxP* in BglII-ClaI, *oriT* (RK2) in XhoI, Tc^r in HindIII	Marsh et al. 1984
pBBCre	pBBRMCS5 with *cre* gene from pBS7	Ayres et al. 1993
PE00146 /pIC	1,028 bp BamHI-XbaI fragment of *PE00146* cloned into pIC20R	This study
PE00198 /pVEX 1311	884 bp BamHI-XbaI fragment of *PE00198* cloned into pVEX1311	This study
PE00199 /pIC	884 bp BamHI-XbaI fragment of *PE00199* cloned into pIC20R	This study
PE00265 /pVEX1311	1,700 bp BamHI-XbaI fragment of *PE00265* cloned into pVEX1311	This study
PE00199 /pVEX1311	884 bp BamHI-XbaI fragment of *PE00119* cloned into pVEX1311	This study
PE00244 /pIC20R	1,134 bp BamHI-XbaI fragment of *PE00244* cloned into pIC20R	This study

Table S2: Primers used in this study.

Primer name	Sequence 5'-3' (restriction sites used for cloning are underlined)	Restriction sites used for cloning
ropAe-F-mutation	GCGACGGAATTCGGACCGGCTATTTTTA	EcoRI
ropAe-R-mutation	CGATAAGCTTGCCTGTTGTGTAGGCGCT	HindIII
ropAe-F-complementation	ACA GTT GGT ACC CGA AAG CGC CAT GGT G	KpnI
ropAe-R-complementation	GCC GAT CTA GAG GTG TAA CAA AGG CGC GG	XbaI
ropAe-F-qPCR	CGT ACC GAA GTG CGC TTC G	
ropAe-R-qPCR	GGG TGG TCT CGT TGC TGG	
hisCd-F-qPCR	ATC ATC GCA ACG CTA TCT CC	
hisCd-R-qPCR	CGA TGG CGA GAC AGC TAA AT	
actP-F-qPCR	CGA CGG CCA AAC ATT TCT TAA A	
actP-R-qPCR	TCG TCA GCG GGA GGA ATA	
RHE_PE00245-F-mutation	GGC ACG AAT TCC GAC ATC ATC TCG CTG C	EcoRI
RHE_PE00245-R-mutation	TGA AGT TCT AGA TGT CGC CGG CCA TCG T	XbaI
RHE_PE00249-F-mutation	ATT GGC CCG GGT TTA TCA AGT TCG GGG A	SmaI
RHE_PE00249-R-mutation	CGT CGA AGC TTA TCT TGT CGG GAG CCA G	HindIII
RHE_PE00259-F-mutation	TCA TCG GAA TTC CGG TCA TGG CAG GCA G	EcoRI
RHE_PE00259-R-mutation	GGC GCC AAG CTT GAG AAC AGG AAC ATG C	HindIII
RHE_PE00263-F-mutacion	AGG AGA GAA TTC CGC TCG TCG CCG GTT T	EcoRI

RHE_PE00263-R-mutacion	CAG GGG AAG CTT ATG GCA TGA GCC TGC A	HindIII
actP::Ω Sp F BamHI	CTG GAT CCC GCT TCC GTG CCC GTC TAT TTCG	BamHI
actP::Ω Sp R XhoI	CCG CTC GAG GCA AGG GCC GGC GCA TCG T	XhoI
ropAe-F-complemetation-ropAch1	CAT ATG TTG GAG GTC ATT TAT GA	NdeI
ropAe-R-complemetation-ropAch1	GGA TCC AGT CAG ACC AAA TTA	BamHI
ropAe-F-complemetation-ropAch2	GCC CCG GGA TCC AAA GGA ATG GAT TGG T	BamHI
ropAe-R-complemetation-ropAch2	CGA GGA GGT ACC ACA CTC ACC AAG TGG AT	KpnI
ropAe-F-complemetation-ropAch3	CGG TCT GGA TCC TCT CGC ATC TGA ACT G	BamHI
ropAe-R-complemetation-ropAch3	CGG TCT GGT ACC CAG CCT TGA TTA GAA C	KpnI
ropAe-F-complemetation-ropAe+	ACA GTT AAG CTT CGA AAG CGC CAT GGT G	HindIII
ropAe-R-complemetation-ropAe+	GCC GAG GTA CCG GTG TAA CAA AGG CGC GG	KpnI

Table S3: Predictions of subcellular localization of RopA proteins assessed by two different predictors.

	Predictors										
	PSORTb v.3.0 (Probability value: 0 to 10)						**CELLO v2.5 (Probability value: 0 to 5)**				
	Subcellular localization[1]						**Subcellular localization[1]**				
Query	EC	**OM**	P	IM	C		EC	**OM**	P	IM	C
RopAe RHE_PE00260	0.03	**9.3**	0.03	0.01	0		0.561	**4.105**	0.066	0.061	0.206
RopAch1 RHE_CH01349	0	**10**	0	0	0		0.288	**4.52**	0.113	0.025	0.058
RopAch2 RHE_CH0247	0	**10**	0	0	0		0.155	**4.74**	0.049	0.021	0.030
RopAch3 RHE_CH078	0	**10**	0	0	0		0.288	**4.57**	0.071	0.031	0.035

[1]EC, extracellular; OM, outer membrane; P, periplasm; IM inner membrane; C, Cytoplasm.

Table S4: Phylogeny data set (Out group porins are yellow highlighted).

Strains	Porin ID	Total porins
Agrobacterium fabrum C58	A9CH05, A9CJK9, A9CJL0	3
Agrobacterium radiobacter K84	B9JC61, B9JI99, B9JC62, B9J8Y9, B9JAW4, B9JE23, B9JJ43, B9JCZ3, B9JEF2	9
Agrobacterium vitis S4	B9JUM4, B9JUM5	2
Azorhizobium caulinodans DSM 5975	A8IEU1, A8HRM6	2
Bradyrhiobium japonicum JCM 10833	Q89Y60	1
Bartonella bacilliformis KC583	A1US12	1
Bartonella tamiae Th239	J0ZSM0, J1JVH6, J1JZP5, J1JYY3, J0ZSM0	4
Bradyrhizobiaceae SG-6C	F7QPW2, F7QN53, F7QP18	3
Bradyrhizobium diazoefficiens USDA 110	Q89LZ7, Q89KC4, Q89F13, Q89K42	4
Bradyrhizobium sp. ORS278	A4YU63, A4YWE2, A4YWE3	3
Brucella abortus 2308	Q2YMY8, Q2YMY7	2
Candidatus Liberibacter americanus	U6B7N3	1
Chelativorans sp. BNC1	Q11E87, Q11JT9	2
Escherichia coli K12	P06996	1
Liberibacter crescens	L0EVT6	1
Mesorhizobium ciceri biovar biserrulae LMG 23838	E8TMK1, E8TLQ1, E8TMK3, E8TGZ8	4
Methylobacterium extorquens DSM 1338	C5AWY3, C5AXJ2, C5AXR7, C5B2U8	4
Methylobacterium nodulans	B8ILW7	1
Methylobacterium radiotolerans DSM 1819	B1M0H0, B1LV40, B1M7E9, B1LV39, B1M749, B1LTW9	6
Methylocella silvestris BL2	B8EPB7	1
Mycobacterium smegmatis ATCC 700084	A0QPU4, A0QR29, A0R3I3	3
Nitrobacter winogradskyi Nb-255	Q3SUW1, Q3SR98	2
Ochrobactrum anthropi DSM 6882	A6X2A5, A6X133, A6X2A4, A6WW24, A6X4Z9	4
Oligotropha carboxidovorans	B6JGA1	1
Pelagibacterium halotolerans	G4R608	1
Rhizobium etli CFN 42	Q2K0Q0, Q2K7H2, Q2KAI4, Q2K4A3	4
Mesohizobium loti MAFF303099	Q985C5, Q985C3, Q985A4, Q98EY3, Q989K2	5
Rhizobium sp. LPU83	W6RDX5, W6RXH4, W6R7X0	3

Rhizobium sp. NGR234	C3M8R2, C3M8R0, P55448	3
Strains	**Porin ID**	
Rhizobium_acidisoli	A0A0N0LB33, A0A0N0LFZ3, A0A0N1DKQ3, A0A0N1DLB5, A0A0N1DS56	5
Rhizobium_etli CNPAF512	F2ABJ3, F2AFI2, F2AGB6	3
Rhizobium_etli_bv_mimosae IE4771	A0A060HYG0, A0A060I7Y4, A0A060IE64	3
Rhizobium_etli_bv_mimosae Mim1	S5RZS6, S5S1J4, S5S482, S5STB3	4
Rhizobium_etli_bv_phaseoli IE4803	A0A0A8GAL5, A0A0A8GDM2, A0A0A8GHD9, A0A0A8GJH8	4
Rhizobium_leguminosarum	Q52865	1
Rhizobium_leguminosarum_bv_phaseoli CCGM1	A0A072C5Q1, A0A072CG11	2
Rhizobium_leguminosarum_bv_trifolii WSM2304	B5ZQ23, B5ZSB8, B5ZW61, B6A3N2	4
Rhizobium_leguminosarum_bv_trifolii WSM1325	C6ATF8, C6B0U0, C6B6T5	3
Rhizobium_leguminosarum_bv_trifolii WSM597	I9N0F9, I9NMK8, I9X9B7	3
Rhizobium_leguminosarum_bv_trifolii WSM2012	J0BTG8, J0KLZ6	2
Rhizobium_leguminosarum_bv_trifolii WSM597	J0H0H1	1
Rhizobium_leguminosarum_bv_trifolii WSM2297	J0KQ41, J0W0J0, J0WHA2, J0CDJ6	4
Rhizobium_leguminosarum_bv_trifolii WSM1689	W0IE01, W0IFG4, W0IMB8, W0IN56	4
Rhizobium_leguminosarum_bv_viciae WSM1455	J0JR95, J0K0A5, J0VDB4	3
Rhizobium_leguminosarum_bv_viciae	Q05811	1
Rhizobium_leguminosarum_bv_viciae 3841	Q1M611, Q1MFL0, Q1MJ66	3
Rhizobium_phaseoli Ch24-10	A0A0M3GG10, A0A0M3GIZ1, A0A0M3GMS6, A0A0M3GR76	4
Rhizobium_sp CCGE510	J4TE18, J5MQ17, J6DTR6	3
Rhodopseudomonas palustris BAA-98	Q6N744, Q6NCD4, Q6N4B6	3
Rhodopseudomonas palustris BisB18	Q212K8, Q213J7	2
Sinorhizobium meliloti 1021	Q92R60, Q92R58	2
Total porins		145
Total species: 38		

Table S5: Predicted function of 19 genes encoded on plasmid p42eΔ21 and the Cu tolerance phenotype of the mutants.

No.	ID	Gene symbol	Predicted gene product[1]	Nucleotides/Amino acids	Gene disruption[2]	Enhanced Cu resistance[3]	
1	*RHE_PE00245*		sugar ABC transporter, sugar-binding protein	1260/419	YES	NO	
2	*RHE_PE00246*		sugar ABC transporter, permease protein	930/311	NO	nd[4]	
3	*RHE_PE00247*		sugar ABC transporter, permease protein	831/276	NO	nd	
4	*RHE_PE00248*		sugar ABC transporter, ATP-binding protein	1059/352	NO	nd	
5	*RHE_PE00249*		hypothetical protein	900/299	NO	nd	
6	*RHE_PE00250*		hypothetical protein	819/272	NO	nd	
7	*RHE_PE00251*		putative DNA ligase (ATP) protein	1065/354	NO	nd	
8	*RHE_PE00252*	*ligD*	ATP-dependent DNA ligase	2649/882	NO	nd	
9	*RHE_PE00253*		hypothetical protein	315/104	NO	nd	
10	*RHE_PE00254*	*pepT*	peptidase T	1233/410	NO	nd	
11	*RHE_PE00255*		putative cell wall-associated hydrolase protein	819/272	NO	nd	
12	RHE_PE00256		sulfate uptake ABC transporter, ATP-binding protein	1041/346	NO	nd	
13	RHE_PE00257		sulfate uptake ABC transporter, permease protein	873/290	NO	nd	
14	RHE_PE00258		sulfate uptake ABC transporter, permease protein	858/285	NO	nd	
15	*RHE_PE00259*		sulfate uptake ABC transporter, substrate-binding protein	1026/341	YES	NO	
16	*RHE_PE00260*	*ropAe*	porin outer membrane protein	1017/338	YES	YES	
17	*RHE_PE00261*	*ptsNe*	enzyme II of the phosphotransferase system	510/169	NO	nd	
18	*RHE_PE00262*	*kdpE*	two-component response regulator	693/230	NO	nd	
19	*RHE_PE00263*	*kdpD*	two-component sensor histidine kinase	2709/902	YES	NO	

[1]Data taken from RhizoBase (http://genome.microbedb.jp/rhizobase/Etli/genes/RHE_PE00245)

[2]Mutants without increased Cu resistance phenotype are highlighted in green

[3]Mutant with increased Cu resistance is highlighted in yellow

[4]nd: not determined

Table S6: Structural characteristics of *R. etli* RopAe and its closet rhizobial orthologues.

Strain	Size (aa)	Uniprot ID	E value	% I[1]	% S[2]	Genome location	Sub cellular location (Psortb v 3.0.2)	Number of TM[3] β strands (BOCTOPUS 2)	Homology with 3D model CATH gene 3D	Porin Genomic context Potassium/ Sulphate ABC transport system4
R. etli CFN42 RopAe	338	Q2K0Q0	0	100	100	p42e	OM	16	OMP IIIA	YES[4]
R. etli bv. mimosae Mim1 RopA4	338	S5STB3	0	100	100	pRetMIM1d	OM	16	OMP IIIA	YES
R. etli CIAT652 RopAa	338	B3Q1J3	0	99	99	pA	OM	16	OMP IIIA	YES
R. leguminosarum bv viciae 3841 RopA3	338	Q1M611	0	96	98	pRL11	OM	16	OMP IIIA	YES
R. leguminosarum bv viciae WSM1325 Porin	338	C6B6T5	0	96	97	pR132502	OM	16	OMP IIIA	YES
R. phaseoli Ch24-10 RopAa	338	A0A0M3GG10	0	98	99	pRphch24	OM	16	OMP IIIA	YES
R. phaseoli Brasil 5 RopA3	338	A0A1W6GR24	0	98	99	prphabra5c	OM	16	OMP IIIA	YES
Rhizobium tropici CIAT899	366	L0LL14	4E-31	38	54	Chromosome	OM	16	Porin OmpP2b	NO[5]
Sinorhizobium mliloti 1021 RopA1	339	F7X0L1	4E-93	50	67	Chromosome	OM	16	OMP IIIA	NO
Mesorhizobium multiple species	392	NCBI Taxid 3744	1,00E-26	31	43	unavailable	OM	16	OmpP2	NO
[1]Percent identity										
[2]Percent similarity										
[3]TM, transmembrane β strand										
[4]YES, means co-occurrence of RopA porin together with sulphate and potassium ABC transport system.										
[5]NO, means absence of sulphate and potassium ABC transport systems in genome context										

Table S7: Structural characteristics of plasmid-encoded RopAe and its three chromosomally encoded homologues RopAch.

Rhizobium etli CFN42 **RopAe homologues**	**Uniprot ID**	**Length aa**	**% I/S[1]**	**Query Coverage %**	**TM β-strands[2]**	**CATH 3D model[3]**	**Genomic context[4]**
R. etl i CFN42 RopAe	Q2K0Q0	338	100/100	100	16	OMP IIIA	Potassium/Sulphate ABC transport system
R. etli CFN42 RopACh1	Q2KAI4	340	59/72	100	16	OMP IIIA	Cell wall and membrane biogenesis
R. etli CFN42 RopACh2	Q2K7H2	343	59/72	100	16	OMP IIIA	Guanine biosynthesis and lipids metabolism
R. etli CFN42 RopACh3	Q2K4A3	347	58/67	91	16	OMP IIIA	MFS transport system

[1]RopAe was used as query in BlastP 2 sequences searches. I, identity; S, similarity.
[2]Trans membrane β strands were searched in http://boctopus.bioinfo.se/
[3]The most probable 3D structure were searched in http://www.cathdb.info
[4]Genomic context was searched in microbial genomic context viewer at mcgv.embi.ru.nl

Table S8: Comparison of nucleotide sequences of *R. etli* putative porins.

Query sequence	Subject sequence	Identity (%)	Query cover (%)	E value
ropAch1	*ropAch2*	91	93	0.0
ropAch1	*ropAch3*	85	100	0.0
ropAch2	*ropAch3*	81	93	0.0
ropAe	*ropAch1*	68	96	$3e^{-92}$
ropAe	*ropAch2*	66	84	$6e^{-63}$
ropAe	*ropAch*	66	100	$2e^{-82}$

Resultados Adicionales

Para evaluar el efecto de la mutación en el gen *ropAe* sobre la nodulación de raíces de *Phaseolus vulgaris* bajo condiciones de estrés por cobre, se realizaron experimentos de nodulación en presencia y ausencia de cobre. Para estos experimentos, la cepa silvestre y la mutante *ropAe*$^-$ fueron marcadas con proteínas fluorescentes, mediante la inserción de un transposón (Mini Tn7). El transposón se inserta a alta frecuencia en una orientación como una sola copia en un sitio específico llamado att Tn7, en el cual está ubicado en una región justo río abajo de la región de codificación del gen *glmS* (glucosamina sintetasa; presente en muchas bacterias). La inserción ocurre en un sitio cromosómico neutral y no debe afectar la codificación de otros genes del huésped (Lambertsen et al 2004). Este transposón contiene ya sea el gen que codifica para la proteína fluorescente Ds-Red o bien para la Ds-GFP, los cuales fueron insertados mediante cruzas tetraparentales fueron insertados en *R. etli* silvestre (Wt) y *ropAe*$^-$ respectivamente.

Las plantas se crecieron en medio Fahreus semisólido con y sin 20 μM de $CuCl_2$. Después de 21 días post inoculación (dpi), se evaluó de manera cualitativa la formación de nódulos por medio de la fluorescencia emitida por las bacterias (roja–CFN42/MiniT7Ds-Red o verde–*ropAe*$^-$/MiniT7Ds-GFP). Para la visualización y captura de imágenes se utilizó el equipo Typhoon fluorescent imager. Como resultado se observó un número muy reducido de nódulos inducidos por la cepa silvestre (rojos) y abundantes nódulos generados por la cepa *ropAe*$^-$ (verdes) en presencia de 20 μM de $CuCl_2$ (Fig. 6).

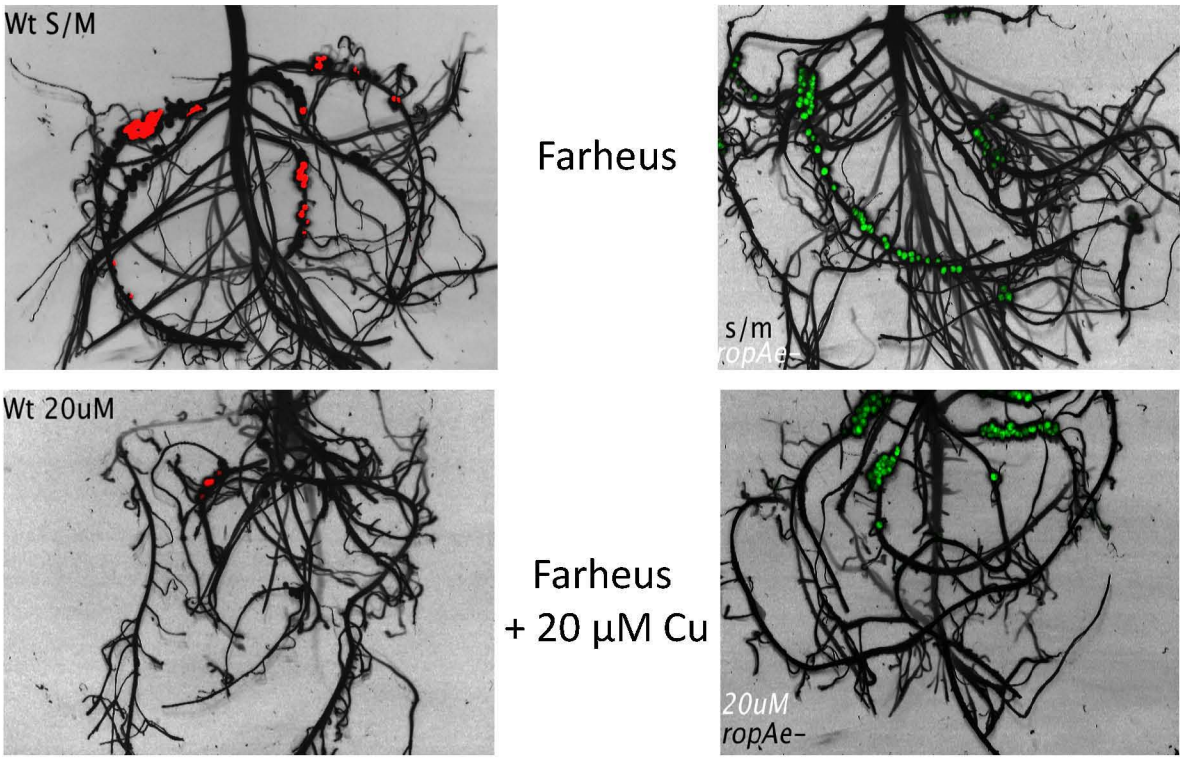

Fig. 6.- Imágenes de raíces de *P. vulgaris* colonizadas por *R. etli* CFN42 (nódulos rojos) y por *R. etli* CFN42 *ropAe*$^-$ (nódulos verdes) 21 dpi, obtenidas con Thypoon fluorescent imager.

Estos resultados nos indican que la mutación en el gen *ropAe* reduce el ingreso de cobre a la célula, lo que favorece la colonización y formación de nódulos en plantas de *P. vulgaris* sometidas a condiciones de estrés por Cu en el medio. Estos resultados son preliminares y requieren de al menos dos repeticiones para validarse.

Búsqueda de genes involucrados en la homeostasis de Cu en *R. etli* CFN42.

Como parte del estudio de los mecanismos de homeostasis de Cu en *R. etli* CFN42, se buscaron en su genoma homólogos de genes para las proteínas involucradas en la homeostasis de Cu descritas en *E. coli y P. aeruginosa.* Para ello, se utilizó la herramienta de búsqueda tipo BlastP (configuración por defecto). Las posibles homólogos en *R. etli*, se enlistan en la Tabla 5.

Tabla 5.- Homólogos de las proteínas que participan en la homeostasis de Cu en *E. coli* y *P. aeruginosa,* encontradas en el genoma de *R. etli* CFN42 utilizando BlastP como herramienta de búsqueda.

	Proteínas	**Posibles homólogos en *R. etli***	**% Cobertura**	**E- value**	**% identidad**
P. aeruginosa	OprC	RHE_CH03076	33%	2.00E-06	25%
	HmtA	RHE_PE00007 (Actp)	89	2E-72	32%
		RHE_CH03719	86	8E-61	32%
		RHE_PF00501 (FIxIF)	88	7E-57	30%
		KdpB RHE_PE00265	63	3E-36	31%
E. coli	CusA	RHE_CH02507	96%	0	55%
	CusB	RHE_CH02506	91%	1E-51	32%
	CusC	RHE_PC00059	96%	1E-41	26%
	CusF	RHE_CH02521	46%	2E-08	33%
	PcoA	RHE_CH02519	39	3E-02	39%
	PcoB	N/A			
	PcoC	N/A			
	PcoD	RHE_PF00512	29	2E-01	33%
	PcoE	N/A			
	PcoF	N/A			

En la Tabla 5 se puede apreciar un bajo porcentaje de identidad de los posibles homólogos de OprC y HmtA de *P. aeruginosa.* ActP y FixIf son ATPasas de tipo P1B conservadas en la mayoría de los seres vivos. Ha sido demostrado que participan en el transporte de Cu a través de la membrana interna. En *R. etli* la mutante *actP*::Ωsp disminuye drásticamente su resistencia a Cu (González-Sánchez et al 2017). Se descarta la existencia de un sistema Pco en *R. etli* CFN42 debido a que, no se encontraron posibles homólogos de PcoB, PcoC, PcoE y PcoF.

Por otro lado, es muy probable que *R. etli* tenga un sistema Cus similar al de *E. coli* debido a que, en el estudio de Li et al 2017 encontraron que el gen *SM0020_RS03640* (cusA-like) está involucrado en la detoxificación de Cu en *S.*

Meliloti. El gen *SM0020_RS03640* comparte un 69% de cobertura y 75% de identidad con el gen *RHE_CH02507* de *R. etli* CFN42. En *E. coli* el sistema Cus está codificado en cromosoma como operon *cusABCFRS*, pero en *R. etli* no se observa dicha conformación. Sin embargo, el contexto genómico de los genes *RHE_CH02505-RHE_CH02507* (Sistema CusABC hipotético) demuestra la existencia de genes que pudieran estar involucrados con el metabolismo de Cu (Fig. 6). Cabe mencionar que, en el trabajo de Li et al 2013 identificaron el gen que codifican para la proteína de membrana externa y el gen que codifica para la multicobre oxidasa. Estas proteínas están involucradas en la detoxificación de Cu en *S. meliloti* y sus posibles homólogos en *R. etli* CFN42 estarían codificados en los genes *RHE_CH02518* (Proteína de membrana externa / eflujo de Cu) y *RHE_CH02519* (McoO) respectivamente (Fig. 7).

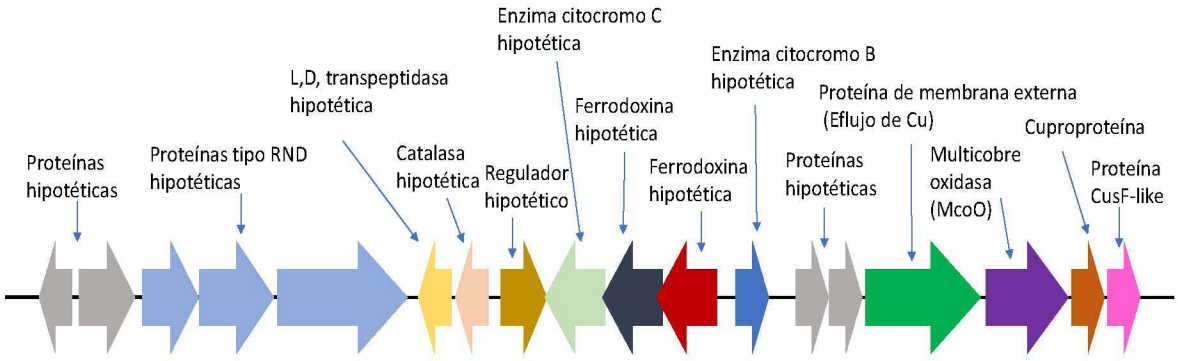

Fig. 7: Región de cromosoma de *R. etli* CFN42 (*RHE_CH02503-RHE_CH02521)*, con posible participación en el metabolismo o homeostasis de Cu.

Estos resultados y el análisis de blastp sugieren la existencia de un conjunto de genes y un sistema Cus-like; localizados en el cromosoma de *R. etli* CFN42, que pudieran codificar para proteínas involucradas en el metabolismo y/o homeostasis de Cu de esta bacteria.

Discusión

El cobre es un micronutriente esencial para todos los organismos aerobios, los

cuales han desarrollado mecanismos moleculares involucrados en la homeostasis de dicho metal, con la finalidad de contender ante condiciones de exceso o limitaciones de cobre. Los mecanismos mejor estudiados son las proteínas que participan en la expulsión del cobre del citoplasma y periplasma (ATPAsa, proteínas RND, multicobre oxidasa); por el contrario los mecanismos de ingreso de cobre están muy poco estudiados. Como parte de un proyecto de investigación enfocado en comprender el tráfico de cobre en *R. etli* CFN42, se buscaron mutantes resistentes a cobre. Para ello se utilizó una colección de mutantes curadas de plásmidos completos y mutantes que presentan deleciones parciales en diferentes regiones de los plásmidos. Usando este enfoque, identificamos un gen codificado en un plásmido anotado como *ropAe*, cuya mutación aumentaba la resistencia a cobre. Además, se demostró que la sensibilidad a Cu podría recuperarse tras la introducción del gen silvestre *ropAe* en el fondo genético de la mutante. De acuerdo con la anotación del genoma en NCBI, el gen *ropAe* (RHE_PE00260) se encuentra en el plásmido p42e y codifica una proteína de 338 aminoácidos, identificada como una proteína de membrana externa tipo porina. Analizando su posible estructura se encontró la presencia de cadenas β, que pueden formar un poro de difusión pasiva (barril β) embebida en la membrana externa. Utilizando esta información se hipotetizó que *ropAe* codifica una supuesta OMP implicada en el transporte de cobre. En este estudio realizamos análisis bioinformáticos y genéticos para examinar esta hipótesis.

La localización en la membrana externa fue respaldada por un análisis in silico de la secuencia de aminoácidos, utilizando programas especializados para predecir con alta precisión proteínas de membrana externa: PSORT-B y CELLO. Ambos programas asignaron valores de alta probabilidad para la localización en la membrana externa. Los resultados para las características estructurales como la topología transmembranal barril-β y la estructura 3D adoptada por las proteínas de membrana externa fueron bien respaldados por los programas especializados BOCTOPUS2 y CATH/gene 3D.

Presumimos que el gen *ropAe* codifica para una proteína de tipo porina y contribuye al suministro normal de cobre en la bacteria, por lo que cabría esperar que una

mutante de *ropAe* muestre alteraciones en las funciones dependientes de cobre, debido a una reducción en la disponibilidad de este metal. En apoyo de esta propuesta, encontramos que la mutante en la ATPasa P1B de cobre de *R. etli* CFN42 (*actP*$^-$), no pudo crecer en presencia de 7.5 µM de $CuCl_2$. Sin embargo, una doble mutante *actP*$^-$/*ropAe*$^-$ recuperó parcialmente su capacidad de crecimiento en 7.5 µM de $CuCl_2$. Estos resultados sugieren que pudiera existir una deficiencia en la suplementación de cobre en la bacteria causada por la interrupción de *ropAe*.

El efecto inhibidor de otros cationes de metales divalentes, fue similar para ambas cepas (Wt/*ropAe*$^-$). Este resultado sugiere que RopAe pudiera no participar en el transporte de otros metales además del cobre. Estos datos deben ser corroborados al determinar la actividad de transporte de iones de Cu por la proteína RopAe.

También determinamos la transcripción del gen *ropAe* en la cepa Wt, la cual no sufre ningún cambio en presencia de $CuCl_2$. Por el contrario, la expresión del gen *actP* que codifica una bomba de flujo de cobre, aumentó 25 veces después de la exposición por $CuCl_2$. Sin emabrgo, bajo condiciones limitantes de cobre la transcripción del gen *ropAe* se ve aumentada hasta 17 veces mientras que la expresión del gen *actP* no sufre ningún cambio significativo. Esto nos indica que la regulación del gen *ropAe* es dependiente de deficiencia de Cu.

La presencia de tres genes homólogos de *ropAe* localizados en regiones distantes del cromosoma de *R. etli* CFN42 (RopAch1, RopAch2, RopAch3) nos hizo pensar que alguno de ellos pudiera suplir la función de la proteína RopAe. Sin embargo, los experimentos de sobre expresión de las RopAch1-2-3 en el fondo genético de la mutante *ropAe*$^-$ no mostraron una complementación funcional y basados en los análisis filogenéticos nosotros proponemos la existencia de divergencia funcional entre RopAe y sus homólogos codificados en el cromosoma.

Conclusiones

Se identificó el gen *ropAe* que codifica para una proteína de membrana externa tipo porina.

La mutación en el gen *ropAe* aumenta el nivel de tolerancia a Cu pero no a otros metales.

La sobreexpresión de los genes homólogos a ropAe; *ropAch1*, *ropAch2* y *ropAch3* en la mutante *ropAe*$^{-}$ no restaura los niveles de tolerancia a Cu.

La expresión del gen *ropAe* aumenta cuando hay deficiencia de Cu en el medio.

La proteína RopAe está altamente conservada en alfa proteobacterias, principalmente en *Rhizobiales*.

El análisis filogenético de las porinas del orden *Rhizobiales*, agrupadas en la familia porin_2 (PF02530) mostró que forman un grupo monofilético distante de las porinas previamente caracterizadas que transportan Mn en *B. japonicum* y Cu en *M. smegmatis*.

La proteína RopAe forma parte de la familia Porin 2 (Pfam02530) y hasta el momento no hay otra proteína caracterizada como un transportador.

Los datos de los experimentos de sobre expresión y la filogenia apoyan la hipótesis de divergencia funcional entre RopAe (localizada en plásmido) y RopAch1, RopAch2 y RopAch3 (cromosomales).

Perspectivas

Los futuros experimentos pretenden purificar la proteína RopAe y caracterizar sus propiedades de transporte de iones metálicos.

Buscar la(s) proteína(s) que participen en la regulación del gen *ropAe*.

Continuar con la investigación del papel que tiene RopAe en simbiosis con plantas de *Phaseolus vulgaris* en presencia y ausencia de cobre en el medio.

Identificar otras proteínas que participen en el ingreso de Cu en la célula, así como en la translocación de moléculas de Cu del periplasma hacia el citoplasma.

Anexo I

Received: 8 August 2016 | Revised: 16 December 2016 | Accepted: 27 December 2016

DOI: 10.1002/mbo3.452

ORIGINAL RESEARCH

WILEY MicrobiologyOpen Open Access

A comprehensive phylogenetic analysis of copper transporting P_{1B} ATPases from bacteria of the *Rhizobiales* order uncovers multiplicity, diversity and novel taxonomic subtypes

Ciro Cubillas | Fabiola Miranda-Sánchez | Antonio González-Sánchez | José Pedro Elizalde | Pablo Vinuesa | Susana Brom | Alejandro García-de los Santos

Programa de Ingeniería Genómica, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

Correspondence
Alejandro García-de los Santos, Programa de Ingeniería Genómica, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México.
Email: alex@ccg.unam.mx

Funding information
UNAM-DGAPA-PAPIIT, Grant/Award Number: IN209815.

Abstract

The ubiquitous cytoplasmic membrane copper transporting P_{1B-1} and P_{1B-3}-type ATPases pump out Cu^+ and Cu^{2+}, respectively, to prevent cytoplasmic accumulation and avoid toxicity. The presence of five copies of Cu-ATPases in the symbiotic nitrogen-fixing bacteria *Sinorhizobium meliloti* is remarkable; it is the largest number of Cu^+-transporters in a bacterial genome reported to date. Since the prevalence of multiple Cu-ATPases in members of the *Rhizobiales* order is unknown, we performed an *in silico* analysis to understand the occurrence, diversity and evolution of Cu^+-ATPases in members of the *Rhizobiales* order. Multiple copies of Cu-ATPase coding genes (2–8) were detected in 45 of the 53 analyzed genomes. The diversity inferred from a maximum-likelihood (ML) phylogenetic analysis classified Cu-ATPases into four monophyletic groups. Each group contained additional subtypes, based on the presence of conserved motifs. This novel phylogeny redefines the current classification, where they are divided into two subtypes (P_{1B-1} and P_{1B-3}). Horizontal gene transfer (HGT) as well as the evolutionary dynamic of plasmid-borne genes may have played an important role in the functional diversification of Cu-ATPases. Homologous cytoplasmic and periplasmic Cu^+-chaperones, CopZ, and CusF, that integrate a CopZ-CopA-CusF tripartite efflux system in gamma-proteobacteria and archeae, were found in 19 of the 53 surveyed genomes of the *Rhizobiales*. This result strongly suggests a high divergence of CopZ and CusF homologs, or the existence of unexplored proteins involved in cellular copper transport.

KEYWORDS
ATPases, CopA, copper, phylogeny, *Rhizobiales*

1 | INTRODUCTION

Copper became bioavailable about one billion years ago when the rise of O_2 caused a loss of soluble iron and an increase in soluble copper (Cu^{2+}) (Rubino & Franz, 2012). Since then, organisms have evolved a diversity of mechanisms to deal with fluctuating copper concentrations in their environments. Copper is a two-edged sword: it is an essential micronutrient for most organisms that synthesize cuproenzymes, but it is highly toxic when intracellular copper levels surpass trace concentrations (Dupont, Grass, & Rensing, 2011). Excessive copper destabilizes

Fe-S clusters, competes with other metals for protein binding sites and may catalyze the formation of reactive oxygen species (Macomber & Imlay, 2009). The cytoplasmic homeostasis of Cu^{+}/Cu^{2+} is maintained by transmembrane proteins belonging to the P_{1B} type ATPase family, widespread in all life domains, which pump Cu^{+}/Cu^{2+} out of the cytoplasm employing active transport (Chan et al., 2010; Palumaa, 2013; Sazinsky et al., 2007; Singleton & Le Brun, 2007). This Cu^{+}-ATPase, known in the vast majority of bacteria as CopA, has also been called CopB, CopF, ActP, or CtpA. In this article, we will use CopA to designate Cu-ATPases, to avoid confusion in the nomenclature. The typical CopA protein contains eight transmembrane helices (TMH), a soluble nucleotide binding domain (N) and a phosphorylation domain (P). The seminal study on the structural and functional diversity of P_{1B} type-ATPases classified CopA transporters into two subtypes P_{1B-1} and P_{1B-3} (Fig. S1) (Argüello, 2003). Subgroup P_{1B-1} contains archaeal, prokaryotic, and eukaryotic Cu^{+}/Ag^{+} transporters. The amino acid motifs involved in metal binding and translocation are two cysteine residues (CXC) in TMH6, a tyrosine and asparagine (YN) residues in TMH7 and methionine and serine residues in TMH8 (MXXSS). P_{1B-1} ATPases also share N-terminal MBDs of 60–80 amino acids, which are usually rich in metal-binding residues. This domain contains a CXXC motif, which can be repeated more than once (Fig.S1) (Argüello, 2003; Argüello, Eren, & Gonzalez-Guerrero, 2007). The P_{1B-3} subtype clusters bacterial and archaeal Cu^{2+} transporters that share the CPH (TMH6) motif, crucial for binding Cu^{2+}, the $GYN(X)_4$ P motif in TMH7 and a MSXST motif in TMH8. They also share an N-MBD of 30–100 amino acids, containing H-rich stretches (Fig. S1). This classification was recently confirmed through similarity-based methods of protein clustering (Smith, Smith, & Rosenzweig, 2014).

Thorough biochemical studies in *Archaeoglobus fulgidus* and *E. coli* have demonstrated that the efflux of cytoplasmic copper to the periplasm requires the interaction of Cu-chaperones-Cu-ATPases, CopZ-CopA, and CopA-CusF, (Gonzalez-Guerrero & Argüello, 2008; Padilla-Benavides, George Thompson, McEvoy, & Argüello, 2014) (Fig. S2). The CopZ copper chaperone is a cytoplasmic protein that binds Cu^{+} and delivers it to CopA. Then CopA delivers copper to the periplasmic CusF chaperone, which finally transfers copper to the CusCFBA efflux system (Franke, Grass, Rensing, & Nies, 2003) that pumps copper out of the cell. The absence of a typical CopZ chaperone in *Streptococcus pneumoniae* led to the identification of an alternative Cu-chaperone, named CupA, capable of transferring Cu^{+} to CopA (Fu et al., 2013).

The *Rhizobiales* order groups a variety of families of alpha-proteobacteria with agronomic and medical relevance. Members of some families are facultative symbiotic diazotrophs that can be found either as free-living organisms in the rhizosphere or as intracellular symbionts in root nodules of leguminous plants, besides being plant or animal pathogens (Carvalho, Souza, Barcellos, Hungria, & Vasconcelos, 2010) (Table 1). Rhizobia, as other soil bacteria, may be exposed to high concentrations of copper ions as a consequence of the Cu-based fungicides and bactericides used in the crop fields, and during the plant or animal-pathogen interaction, as result of the Cu provided by the host in response to infections (Chaturvedi & Henderson, 2014; Cheruiyot, Boyd, & Moar, 2013; Fu, Chang, & Giedroc, 2014).

The multipartite genomes of most species grouped in the *Rhizobiales* order are characterized by the presence of one or two chromosomes, megaplasmids and numerous large plasmids, some of them transmissible by conjugation (Jumas-Bilak, Michaux-Charachon, Bourg, Ramuz, & Allardet-Servent, 1998; MacLean, Finan, & Sadowsky, 2007) (Table 1). This may have played an important role in the spread, acquisition, and evolution of metal resistance genes. Bacterial plasmids have been shown to be critical in bacterial Cu tolerance, and are highly mobile among strains growing in metal polluted soils (Cooksey, 1990; Lakzian, Murphy, & Giller, 2007; Monchy et al., 2007; Tetaz & Luke, 1983). Functional plasmid-encoded CopA transporters have been reported for three rhizobial species (Landeta et al., 2011; Reeve, Tiwari, Kale, Dilworth, & Glenn, 2002). The presence of multiple Cu-ATPase coding genes in *S. meliloti*, highlighted in the pioneer study on functional classification of P-type ATPases, led to the physiological characterization of five plasmid-encoded Cu-ATPases present in this organism (Argüello, 2003; Argüello et al., 2007; Patel, Padilla-Benavides, Collins, & Argüello, 2014; Smith et al., 2014). This study revealed a nonredundant physiological role that is important for saprophytic and symbiotic lifestyles, as well as the presence of a novel CopA regulated by redox stress. A cladogram inferred from a clustalW alignment grouped these five CopA homologs into three different clades but nothing is known about their evolutionary relationships to other rhizobial and nonrhizobial Cu-ATPases (Patel et al., 2014). Also, a genome-wide study on trace element utilization revealed that two *Sinorhizobuim* species contain uncommonly large cuproproteomes, which may be related to the numerous CopA transporters and other trafficking proteins (Zhang & Gladyshev, 2010).

Several questions arise from the studies in *S. meliloti*: Is the model of five plasmid-encoded Cu-ATPases the archetype of Cu-ATPases for the *Rhizobiales* order? How diverse are the multiple Cu-ATPase homologs encoded in the *Rhizobiales* genome? To what extent has horizontal gene transfer (HGT) contributed to Cu-ATPases acquisition? Additionally, we are also interested in analyzing the occurrence of the CopZ/CupA and CusF homologs, cytoplasmic, and periplasmic copper chaperones, respectively, which might be associated with the multiple Cu^{+}-ATPases.

To answer these questions, we surveyed the occurrence of CopA, CopZ/CupA and CusF homologs encoded in members of the *Rhizobiales* order with multi-replicon genomes and analyzed the diversity of CopA homologs through a maximum-likelihood (ML) phylogenetic analysis.

Our study revealed that *copA* multiplicity is a highly conserved characteristic in the genomes of *Rhizobiales*. The diversity of these CopA transporters is higher than previously assessed and is influenced by horizontal gene transfer. We propose a novel phylogeny that distributes bacterial Cu-ATPases in four monophyletic group and six subtypes based in conserved motifs (P_{1B-1}, P_{1B-1a}, P_{1B-1b}, P_{1B-1c}, P_{1B-3}, P_{1B-3a}). The complete CopZ-CopA-CusF set of proteins potentially involved in cytoplasmic copper efflux was only found in 19 of the 53 analyzed genomes of *Rhizobiales*.

Although this bioinformatic analysis was focused on P_{1B} ATPases, our results on multiplicity and diversity of metal transporters have

TABLE 1 Members of the *Rizobiales* order analyzed in this study

Genera	Number of species	Number of strains	Habitat or metabolism	Range of chromosomes number	Range of plasmids number
Agrobacterium	4	4	Plant Pathogen	1–2	2
Aureimonas (formely *Aurantimonas*)	1	1	Potential human pathogen and different environmental sources	1	8
Bartonella	2	2	Mammalian pathogens	1	1–2
Beijerinckia	1	1	Free-living in soil	1	2
Bosea	1	1	Thiosulfate-oxidizing/plant associated	1	1
Bradyrhizobium	1	1	Facultative symbiotic diazotrophs	1	1
Brucella	5	5	Mammalian pathogens	2	0
Chelativorans	2	2	EDTA-degrading	1	3
Ensifer	1	1	Facultative symbiotic diazotrophs	2	2
Hoeflea	1	1	Marine Iron-oxidizing	1	2
Martelella	1	1	Halophyte	1	2
Mesorhizobium	2	2	Facultative symbiotic diazotrophs	1	1–2
Methylobacterium	5	7	Methylotrophic plant-associated	1	1–8
Neorhizobium	1	1	Facultative symbiotic diazotrophs	1	2
Nitrobacter	1	1	Nitrite-oxidizing free-living in soil	1	3
Ochrobactrum	1	1	Opportunistic human pathogen	2	4
Oligotropha	1	2	Wastewater CO-utilizing	1	1–2
Pelagibacterium	1	1	Seawater halotolerant	1	1
Rhizobium	3	7	Facultative symbiotic diazotrophs	1	1–6
Rhodopseudomonas	1	1	Photosynthetic, nitrogen and carbón fixer	1	1
Shinella	1	1	Degradation of toxic aromatic compounds	1	12
Sinorhizobium	3	8	Facultative symbiotic diazotrophs	1	1–4
Xanthobacter	1	1	Chemolitho-autotrophic free-living diazotroph	1	1
23	40	53			

general implications regarding the evolution of resistance, adaptation, and innovation in bacteria.

2 | METHODS

2.1 | Data set, multiple sequence alignments and phylogenetic analysis

As mentioned above, the five Cu-ATPases of *S. meliloti* 2011 are encoded in two extrachromosomal replicons (Patel et al., 2014). Plasmids usually have a higher copy number than chromosomes, which may result in higher expression levels of plasmid-encoded genes, representing an advantage for this bacterium. The increased copy number of plasmids might also accelerate gene diversification (San Millan, Escudero, Gifford, Mazel, & MacLean, 2016; Sano, Maisnier-Patin, Aboubechara, Quiñones-Soto, & Roth, 2014). To widen our knowledge regarding these metal transporters, we surveyed the occurrence of CopA, CopZ/CupA and CusF homologs in 53 fully sequenced multireplicon genomes of rhizobia species with different lifestyle or metabolism (Table 1).

To test if the model of five plasmid-encoded Cu-ATPases reported for *S. meliloti* 2011 is the archetype for the *Rhizobiales* order, we collected 313 P1B-ATPases sequences from the Pfam family PF00122 (Table S1) belonging to 53 fully sequenced multireplicon genomes of species grouped in the *Rhizbiales* order with different lifestyle or metabolism (Table 1). The construction of the data set and the different steps of the phylogenetic analysis are depicted in Fig.S3. The P_{1B}-ATPases and Cu-Chaperones (CopZ/CupA and CusF) homologs were retrieved from their respective Pfam families (CopA, PF00122; CopZ, PF00403; CupA, PF13473; and CusF, PF11604) either filtered by using species–specific tags (e.g., *Rhizobium* copper-translocating/chaperones) or using the profile Hidden Markov Model (HMM) for each family as query in hmmsearch (cut-off value 10^{-3}, under default settings) for those genomes whose proteins have not been assigned yet to specific protein families (i.e., *Rhizobium tropici*). Other sequences were retrieved from PATRIC (Bacterial Bioinformatic Resource Center, Wattam et al., 2014). The P_{1B}-type-ATPases, CopZ/CupA, and CusF homologs found in the 53 genomes belonging to the *Rhizobiales* order are shown in supplementary tables S1–S4. These P_{1B}-type-ATPases may have as substrate one or more of the following metals: Cu^{+}, Ag^{+}, Cu^{2+}, Zn^{2+}, Co^{2+}, Pb^{2+}, Cd^{2+}, and Mn^{2+}.

To infer the phylogeny, the data set of 313 P_{1B} ATPases was complemented with 45 nonrhizobia, characterized P_{1B} ATPases whose

metal(s) specificities have been assessed experimentally (Table S1). Thirty five K^+ transporters grouped in ATPase family P_{1A} were included as out-group. Full-length multiple sequence alignments were done using MUSCLE and hmmalign. Since the profile HMM for the P_{1B} ATPases family available at Pfam lacks most critical motifs required for metal transport, including some involved in metal-ion selectivity (Gonzalez-Guerrero & Argüello, 2008), we used an *ad hoc* profile HMM constructed with 53 characterized P_{1B} ATPases, using hmmbuild under default parameters. This novel alignment and the resultant new profile HMM contained all relevant regions involved in metal transport (Fig. S4).

To infer the highest likelihood phylogenetic tree for P_{1B} ATPases proteins, we started with 100 random seed trees in addition to a BioNJ tree to perform 101 searches. Tree searching under the ML criterion was performed with PhyML v3.026, using the LG+G+f model as the substitution matrix with gamma-correction among-site rate variation. The best tree, shown in Fig. S5, had the highest log-likelihood score from these 101 searches.

2.2 | Detection of phylogenetic incongruence and putative horizontal transfer events

The HGT analysis presented in this study was based on the identification of putative recently transferred genes representing bona fide examples of HGT events. Since these genes still maintainmost of the nucleotide composition of their parental hosts, they were searched for through similarity-based methods, using BLASTP analyses with default parameters.

The 155 Cu-ATPases of *Rhizobiales* defined by the phylogenetic analysis were used as query to find their closest homologs in the nonredundant NCBI database. Thirty rhizobial Cu-ATPases transporters had as closest homologs, Cu-ATPases from species distantly related to members of the *Rhizobiales* order belonging to alpha-, beta- and gamma-proteobacteria and Bacteroidetes (Table S6). Their foreign origin was also assessed by nucleotide compositional methods, searching for deviations between the sequence composition of such genes and their respective genomes. The G+C content was calculated at http://www.genomicsplace.com/gc_calc.html. Genomic G+C content was obtained from: http://www.ncbi.nlm.nih.gov/genome/browse/.The codon adaptation index (CAI) of *copA* and the expected e-CAI were calculated at the E-CAI server (Puigbo, Bravo, & Garcia-Vallve, 2008). E-CAI web server provides a threshold value, named eCAI, to discern if the differences in CAI are statistically significant. The eCAI value is calculated by generating 500 random sequences with the same amino acid composition as the query but with randomly assigned codon usage (Puigbo et al., 2008). The normalized CAI is defined as the quotient between the CAI of a gene and its expected e-CAI value. Values below one indicate a bias in the codon usage preferences of a gene, relative to the codon usage preferences of their respective genomes, used as reference.

The location of *copA* genes in putative genomic islands was predicted by the sequence composition-based method available at IslandView (pathogenomics.sfu.ca/islandviewer/). Their location into a cluster of metal resistance genes was searched for with Microbial Genomic Contex Viewer (Overmars, Kerkhoven, Siezen, & Francke, 2013; http://mgcv.cmbi.ru.nl/).

The hypothesis that these genes arrived through HGT was tested with a phylogenetic approach, comparing the species' phylogeny, inferred from 16S rRNA genes, with the phylogeny inferred from homologous Cu^+-ATPase coding genes (Fig. S6). Aligments for 16S rRNA genes were built using the secondary-structure aware Infernal aligner (Nawrocki, Kolbe, & Eddy, 2009) incorporated in the Ribosomal Database website (https://rdp.cme.msu.edu/). The Cu-ATPases sequences were aligned using an HMMER profile built from 53 characterized Cu-ATPases proteins. These phylogenies were performed with an ML-phylogenetic analysis under GTR+I+G model for 16S rRNA and LG+G+f model for putative CopA proteins, using nucleotides and amino acid alignments, respectively. The phylogenies were inferred using a parallel Pthreads-based version of RAxML v8.2.4, this software supports SSE3 vector instructions (Stamatakis, 2014). The ML-search started with 1000 random seed trees, the best tree was selected (Figure S6). Rapid bootstrapping (Stamatakis, 2014) was uses to assess the branch support in the tree. The number of necessary replicates was estimated using the extended majority rule criterion, 150 and 1,000 replicates for Cu-ATPases and 16S rRNA phylogenies, respectively (Pattengale, Alipour, Bininda-Emonds, Moret, & Stamatakis, 2010).

3 | RESULTS

3.1 | A maximum likelihood phylogeny groups Cu-ATPases into four monophyletic groups

To gain a broader view of the diversity of P_{1B} ATPases present in the 53 multi-replicon members of the *Rhizobiales* order contextualized within an evolutionary framework, we analyzed 313 P_{1B} ATPases encoded in the multireplicon genomes of 53 strains. Fifty three characterized P_{1B} type ATPases known to transport copper or other metals present in eukaryots, archaea, and other bacteria were also included in the phylogenetic analysis (Table S1). Thirty five P1A-type KdpB subunits (potassium-extruding ATPases) and nine putative zinc transport ATPases (ZntA) from different rhizobia were included as out-groups. To improve the accuracy of previous classifications, a profile-HMM was built from the alignment of the previously mentioned characterized P_{1B} ATPases, which contain all the structurally conserved regions (Gourdon et al., 2011) (Table S1, Fig S4). We decided to build this new model instead of using the one available at Pfam because the last one does not include the N-terminal MBD (see Experimental Methods) (Fig S4) and only shows the most conserved region of the P_{1B} TPase family, the actuator and the ATP-binding domains. The alignment of the total data set (410 P_{1B}-ATPases) with our profile-HMM was critical to analyze an increased number of amino acid positions, allowing us to obtain a better description of the P_{1B}-type-ATPase family composition.

In previous studies the P_{1B}-ATPase relationships were inferred from ClustalW2 cladograms (Argüello, 2003; Argüello et al., 2007;

Gonzalez-Guerrero, Raimunda, Cheng, & Argüello, 2010; Patel et al., 2014). The use of *ad hoc* HMM as well as the maximum likelihood method for phylogenetic analysis allowed us to perform a more rigorous and accurate inference of the evolutionary relationships among the 410 P-type-ATPases included in our data set (Fig.S5).

Our ML-phylogenetic analysis classified 181 of the 410 P_{1B} type ATPases as Cu-ATPases and they were assorted in 4 monophyletic groups (VI, VII, VIII and XIV,) with branch bipartitions well supported by Shimodaira-Hasegawa-like ρ value ≥0.9, Each one of the Cu-defined clades contains at least one characterized Cu-ATPase. Bacterial Cu-ATPases not belonging to the *Rhizobiales* order, were assorted in three different clades (I, III and V). Cu-ATPases from eukaryotes were included in clade IV and the Zn/Cd/Mn/Mg/K-ATPases, retrieved during the data mining, were grouped in five well supported (ρ value ≥0.9) monophyletic clades (Fig. S5).

3.2 | The Cu^+-ATPase subtypes differ in signature motifs at the N-terminal MBD and transmembrane domains

Fifty of the 181 putative Cu^+-ATPases were grouped in subtype VIII (Fig. S5, Table 2 and Table S5). These copper transporters maintain the invariant signature motifs for P_{1B-1}-Cu^+-ATPases in the transmembrane helices TM6 (CPC) and TM8 (MXXSS) (Fig. S1) previously described (Argüello, 2003). All these Cu-ATPases conserve two heavy metal associated domains (HMA) and the CASC motif at the cytoplasmic N-MBD, suggesting that they may have functional relevance (Table 2).

The few typical P_{1B-3}-Cu^{2+}-ATPases contained in our data set, identified as Cu^{2+} transporters (Fig. S5), containing an N-terminal rich in His residues and the invariant CPH motif in TM6, were located in group VI (Table 2 and Table S5 Cu-ATPase subtypes). Members of this group could be distributed in three subtypes based in the presence of different motifs and well supported bipartitions (Table S5 Cu-ATPases subtypes). Cluster VIa groups the poorly characterized Cu^{2+}-ATPases, HRA-1 and HRA-2, from *E. coli*, *Staphylococcus aureus* CopB, and *Thermus thermophilus* 1371. The location of *Enterococcus hirae* CopB in this cluster is intriguing because of its ability to transport Cu^+ (Solioz & Odermat, 1995). However, its ion transduction region presents the CPH signature motif invariable in P_{1B-3}-Cu^{2+}-ATPase (Odermatt, Suter, Krapf, & Solioz, 1993). The imidazol from the His residue forms a stronger bond with Cu^{2+} than Cu^+ (Argüello, Gonzalez-Guerrero, & Raymunda, 2011).

Subgroup VIb groups six putative Cu^{2+}-ATPases from the EDTA–degrading *Chelativorans* sp, the osmotolerant *Pelagibacterium halotolerans*, the crude-oil-degrading bacterium *Chelatococcus sp* CO-6 and the seawater halotolerant *Hoeflea sp*. IMCC 20628 together with the characterized archaeal CopB from the hyperthermophilic *Archaeoglobus fulgidus* known to transport Cu^{2+} (Mana-Capelli, Mandal, & Argüello, 2003). Interestingly, their closest homologs, by means of BLAST comparisons, retrieved ATPases from extremophile bacteria with thiosulfate oxidation capabilities as well as methanogenic and osmotolerant *Archaea*. Both clusters contain Cu^+-ATPases with poly His at the N-terminal as well as the invariant CPH at TMH6 (Table 2).

The third subtype, VIc, includes three atypical rhizobial Cu-ATPases from *Xanthobacter autotrophicus*, *Nitrobacter hamburgensis*, and *Bradyrhizobium sp*, which present few histidines at the N-terminal, contain a novel CPD motif instead of the invariant CPH motif in TM6 and are at least 80 kb longer than Cu-ATPases from clusters VIa and VIb. This subtype was named P_{1B}-3a (Table 2 and Table S5 Cu-ATPase subtypes).

Sixty one of the 181 Cu-ATPases were classified in clade VII; interestingly, these transporters share motifs from both P_{1B-3} and P_{1B-1}-Cu^+-ATPases (Table 2 and Table S5 Cu-ATPase subtypes). These hybrid characteristics have not been previously reported. Similar to the characterized Cu-ATPases from *Legionella pneumophila* (Lpg1024) and *Salmonella typhimurium* (SilP), belonging to this subtype, 25 rhizobial homologs contain a His-rich extension at the N-terminal MBD

TABLE 2 Polymorphisms within conserved motifs N-MBD, TM 6, 7, and 8 of rhizobial Cu-ATPases subfamilies

Conserved motifs					
Group (Number of Cu-ATPases)	N-MBD (CXXC)	TM6	TM7	TM8	Subtype
VIII	CASC	CPC	$YN(X)_4P$	MALSS	IB-1[a]
(50)	CASC	CPC	$YN(X)_4P$	MAMSS	
	CASC	CPC	$YN(X)_4P$	MAFSS	
VIa (6)	H-Rich	CPH	$YN(X)_4P$	MSAST	IB-3[a]
VIb (5)	H-Rich	CPH	$YN(X)_4P$	MSAST	IB-3[a]
VIc (3)	Poor H content	CPD	$YN(X)_4P$	MSGSS	IB3-a[b]
VII (25)	H-Rich/ CPIC	CPC	$YN(X)_4P$	MSLSS	IB-1a[b]
VII (14)	CPIC	CPC	$YN(X)_4P$	MALSS	IB-1[a]
VII (15)	H-Rich/CPKC	CPC	$YN(X)_4P$	MSLSS	IB-1b[b]
XIV (56)	CAGC	CPC	$YN(X)_4P$	MSGSS	IB-1c[b]
	CAAC	CPC	$YN(X)_4P$	MSGSS	IB-1c[b]
	CASC	CPC	$YN(X)_4P$	MSGSS	IB-1c[b]
	CAAC	CPC	$YN(X)_4P$	MSLSS	

Parenthesis indicates number of Cu-ATPases in each group.
[a]Established subtype classification (Argüello et al., 2007).
[b]Novel subtypes (This study).

characteristic of subgroup $P_{1B-3.}$ Like the P_{1B-1} subgroup, they present CPIC as the invariant CXXC motif at N-terminal MBD as well as a CPC motif at TM 6, instead of the invariant CPH motif distinctive of P_{1B-3}-Cu^+-ATPases (Table 2 and Table S5 Cu-ATPase subtypes). Following the already established subgroup nomenclature (Argüello, 2003) these transporters were named as P_{1B-1a}-Cu^+-ATPases.

Fifteen of the 61 putative rhizobial Cu-ATPases clustered in group VII have the largest N-MBD found in *Rhizobiales*: 160 amino acids preceding the highly conserved CPKC motif. The first 20 amino acids contain a short H-rich segment of six to 10 residues and two copies of the TRASH/YHS-domain located downstream of this poly-H labeled as Ribo previously reported by Ettema, Huynen, de Vos, and van der Oost (2003), which contains a CXCXC signature that might be involved in copper sensing, trafficking and resistance (Table S5). The Cu-ATPases from *Chelativorans sp* and *M. loti* MAFF303099 lack the CPKC motif; instead, they have a CPIC similar to that from P_{1B-1a}-Cu^+-ATPases. The motifs at TM 6 (CPC), 7 $(YN(X)_4P)$ and 8 (MSLSS) are highly conserved among P_{1B-1}-Cu^+-ATPases of members of the *Rhizobiales* order (Table 2). Since members of this subtype maintain the CPC motif typical of P_{1B-1}-Cu^+-ATPases, these proteins will be named P_{1B-1b}-Cu^+-ATPases (Table 2).

Clade XIV groups FixI proteins (Fig. S1, Table 2 and Table S5 Cu-ATPase subtypes). Analysis of *S. meliloti fixI1* and *fixI2* mutants revealed that these Cu-ATPases do not contribute to the copper tolerance of rhizobia in the free-living state (Patel et al., 2014). Since FixI proteins are encoded close to high oxygen-affinity-cbb3-type cytochrome oxidases expressed under microaerobic conditions during symbiotic nitrogen fixation (Preisig, Zufferey, Thöny-Meyer, Appleby, & Hennecke, 1996), it has been hypothesized that FixI1 plays a key role in bacteroid respiration in the microaerobic environment inside the nodule. On the other hand, FixI2 is a housekeeping Cu^+-ATPase involved in metallation of a constitutive cytochrome oxidase required for respiration in both, symbiosis and in the free-living state (Patel et al., 2014). A cladogram inferred from a ClustalW2 alignment indicated that the FixI transporters belong to different clusters (Patel et al., 2014). In our ML-phylogeny we found small subgroups of FixI, FixI1 and FixI2 spread throughout the clade. Most of the FixI proteins share common N-MB, TM6, TM7 and TM8 motifs. Since these motifs differs from those present in Cu-ATPases of clades VI, VII and VIII these Cu-ATPases subtypes were named $P_{1B-1c.}$ The eukaryotic and prokaryotic ATPases predicted to transport metals other than Cu, were distributed in five different monophyletic clades: IX, (Cd, Co, Zn), X (Ca and Mg), XI (K), XII (Cd and Zn) and XIII (Mn, Zn, and Cd). These transporters will not be described in this study.

3.3 | Occurrence of Cu-ATPase subtypes in the multipartite genomes of rhizobia surveyed

The distribution and abundance of the established and novel CopA subtypes in the 53 genomes analyzed in this study are shown in detail in Table S8 and a summary of the data is presented in Tables 3 and 4. The number of Cu-ATPases per genome is very variable (Table 3), it ranges from zero in two species of the intracellular *Bartonella* pathogens to 8 in the genome of *Rhizobium leguminosarum* bv. viciae 3841. Five Cu-ATPases subtypes, the number of ATPases reported in the genome of *S. meliloti* 2011 and characterized by Patel et al. (2014), was only found in 5 other rhizobia. Forty three percent of Cu-ATPases homologs are encoded in plasmids and fifty seven percent in chromosomes (Table S5).

TABLE 3 Number of Cu-ATPases coded in the 53 multipartite rhizobial genomes surveyed

Number of Cu-ATPases	Occurrence in rhizobia genomes
0	2
1	11
2	8
3	11
4	8
5	5
6	4
8	2

TABLE 4 Occurrence of Cu-ATPases subtypes in the 53 rhizobial genomes surveyed

Cu-ATPases subtypes	Genomes harboring Cu-ATPases subtypes (%)
P1B-1	88
P1B-1a	77
P1B-1b	36
P1B-1c	94
P1B-3	13
P1B-3a	6

We also analyzed the prevalence and multiplicity of the previously established ATPases subtypes P_{1B-1} and P_{1B-3} (Argüello et al., 2007), in the 53 genomes surveyed (Table 4). The P_{1B-1} subtype was the most frequent Cu-ATPase, it was found in 88% of the genomes analyzed. Such genomes contain at least one P_{1B-1}-ATPas. Most of these ATPases (74%) were encoded in the chromosome. In contrast, the P_{1B-3} subtype was only found in 13% of the 53 surveyed genomes. Among the novel subtypes found in this study, the P_{1B-1a} was the most frequent, 77% of the genomes contain at least one copy (Table 4 and Table S8).

3.4 | The diversity of Cu-ATPase subtypes of the *Rhizobiales* order is influenced by horizontal gene transfer events among alpha-, beta-, and gamma-proteobacteria

Attempting to explain the multiplicity and high diversity of Cu-ATPase subfamilies in rhizobia, we examined the role that HGT may have played in the acquisition and divergence of Cu-ATPases in the *Rhizobiales* order. We focused on the identification of recently transferred genes representing bona fide HGT events. Since these genes

still conserve most of the nucleotide composition of their parental hosts, they could be detected through BLASTP-based searches using the 155 Cu-ATPase sequences as query to find their closest homologs in the nonredundant NCBI database (Zhaxybayeva, Gogarten, Charlebois, Doolittle, & Papke, 2006). Genuine events of HGT are difficult to assess when searching for anciently transferred genes, due to amelioration, the process by which the nucleotide composition (G+C, codon bias, nucleotide frequency) of the putative foreign genes change by the mutational pressure of the new host and over time the foreign genes are indistinguishable of the DNA composition of the host genome. Since the signs of amelioration are not easily identified, the search for ancient acquired genes increases the probability of identifying false laterally transferred genes.

30 Cu-ATPases, belonging to different *Agrobacterium*, *Methylobacterium*, *Ochrobactrum*, *Oligotropha*, *Pelagibacterium*, *Chelativorans*, *Aureimonas*, *Bosea*, *Hoeflea*, and *Martelella* species, retrieved as the closest homologs (identity >50%, similarity >70% and coverage >90%) putative Cu-ATPases encoded in the genome of species distantly related to members of the *Rhizobiales* order (Table S6 HGT). These homologs were found in the genomes of gamma-proteobacteria (12), beta-proteobacteria (four), alpha-proteobacteria other than *Rhizobiales* (13) and Bacteroidetes (one). The high similarity values shared between these rhizobial proteins and their counterparts in beta- and gamma-proteobacteria prompted us to further investigate if they concurred with HGT events.

We found that the G+C content of the 30 Cu-ATPase coding genes averaged 4.8% points higher than that of their corresponding genome context (Table S7 G+C, CAI and GI). These differences were supported by a *t*-test (with $p < .00001$; at 99% confidence interval following previously reported methods (Hayek, 2013).

Recent laterally transferred genes are expected to have an atypical codon usage in comparison to the codon usage of their genome context. The CAI is the most prevalent method used for analyzing such differences. We used the E-CAI web server (Puigbo et al., 2008) to quantify the similarity of the codon usage of 30 Cu-ATPase coding genes compared to the codon usage of their respective genomes. Our analysis revealed that twenty two genes had a normalized value of CAI <1, [CAI/eCAI ($p < .01$)] indicating significant differences in codon usage between the Cu-ATPases and their genomes, further supporting that those genes may have arrived through HGT (Table S7 G+C, CAI and GI). Four of the genes were predicted to be part of putative detoxification islands by different web tools (Dhillon, Chiu, Laird, Langille, & Brinkman, 2013). Ten putative foreign Cu-ATPases were localized next to genes coding for diverse metal transporters, transposases and insertion sequences, in clusters whose length ranges from 7 to 75 Kb (Table S7 G+C, CAI and GI).

An ML-phylogenetic analysis was performed to detect incongruity between the species tree, inferred from the alignment of 16S rRNA genes (Fig. S6A), and the CopA tree inferred from the alignment of the Cu-ATPAse homologous sequences from the same bacterial species (Fig. S6B). While the 16S rRNA tree classified alpha-, beta-, and gamma-proteobacteria and bacteroidetes in four clearly defined clusters, in the CopA tree several beta-, gamma-proteobacteria and bacteroidetes were scattered among the *Rhizobiales* (Fig S2). This phylogenetic incongruence strongly points to HGT as a process that has facilitated the exchange of Cu-ATPases among distantly related alpha-, beta-, and gamma- proteobacteria and bacteroidetes.

To get insights into the possible direction of the putative HGT events, the foreing Cu-ATPases found in rhizobia, homologs of gamma-proteobacteria, were used as query in BlastP searches, to compare the occurrence of homologs between orders/families of alpha- and gamma-proteobacteria. Whereas BlastP searches of proteins from alpha-proteobacteria retrieved homologs exclusively from the order *Rhizobiales*, most of them belonging to the *Rhizobiaceae* family, BlastP searches of proteins from gamma- proteobacteria retrieved homologs (60% identity, 80% query cover and e-value = 0) from at least five different orders (*Enterobacterales*, *Alteromonadales*, *Xanthomonadales*, *Vibronales*, *Pseudomonadales*). The broad distribution of homologs in five gamma- proteobacteria orders and the limited occurrence in alpha-proteobacteria (one order) strongly suggests that these genes come from gamma-proteobacteria and have been recently transferred to alpha-proteobacteria.

A similar analysis was carried out with the foreing Cu-ATPases homologs of beta-proteobacteria. However, BlastP searches retrieved homologs from four different beta-proteobacteria orders (*Burkholderiales*, *Nitrosomonadales*, *Methylophilales* and *Rhodocyclales*), BlastP searches into alpha-proteobacteria retrieved homologs exclusively from the *Rhizobiales* order, suggesting that these genes were transferred from beta-proteobacteria to *Rhizobiales*.

A correlation could not be established among the foreing Cu-ATPases of rhizobia and their alpha-proteobacteria homologs other than *Rhizobiales* order. These ATPases were broadly distributed in a variety of genera belonging to *Rhizobiales*, *Rhodobacterales*, *Caulobacterales*, and *Sphingomonadales* orders suggesting exchange among them.

The broad occurrence of the putative Cu-ATPase of *Hoeflea sp* in four of nine orders of alpha-proteobacteria (*Rhodospirilales*, *Rhodobacterales*, *Sphingomonadales*, and *Rhizobiales*) but in two of six orders of *Bacteroidetes* (*Bacteroidetes* order II and *Flavobacterales*) might indicate that this gene has moved from alpha-proteobacteria to *Bacteroidetes*.

3.5 | Low occurrence of cytoplasmic and periplasmic copper chaperones homologous to CopZ/CupA and CusF in genomes of the *Rhizobiales*

Most of the CopZ homologs present in the 53 genomes were downloaded from the HMA (PF00403 family, where the functionally characterized CopZ proteins are located. This family, defined by a profile-HMM, belongs to the Pfam database. The profile HMM and hmmsearch were used to search for CopZ homologs in those genomes whose proteins have not been assigned to a Pfam Database (see Methods for details). These searches are more sensitive than pairwise methods such as BLAST or FASTA, to find distantly related sequences (Eddy, 1998). We found 55 CopZ homologs, 24 coded in plasmids and 31 in chromosomes, distributed in 26 genomes belonging to 10 different genera (Table 5; Table S2 CopZ homologs). The

TABLE 5 Occurrence of putative CopZ,CupA, CopA and CusF proteins in 53 Rhizobial multireplicon genomes[a]

Strains	CopZ	CupA	CusF
1. *Agrobacterium fabrum* str. C58	1C	0	0
2. *Agrobacterium radiobacter* K84	0	0	1C2
3. *Agrobacterium* sp. H13-3	1C	0	0
4. *Agrobacterium vitis* S4	1P	0	0
5. *Aureimonas sp. AU20*	1	0	0
6. *Bartonella grahamii* as4aup	0	0	0
7. *Bartonella tribocorum* CIP 105476	0	0	0
8. *Bosea sp. PAMC 26642*	0	0	0
9. *Beijerinckia indica* subsp. indica ATCC 9039	0	2C	0
10. *Bradyrhizobium* sp. BTAi1	0	2C	0
11. *Brucella abortus* bv. 1 str. 9-941	0	0	0
12. *Brucella canis* ATCC 23365	0	0	0
13. *Brucella melitensis* bv. 1 str. 16M	0	0	0
14. *Brucella ovis* ATCC 25840	0	0	0
15. *Brucella suis* 1330	0	0	0
16. *Chelativorans sp. BNC1*	1C	0	0
17. *Chelatococcus sp. CO-6*	1C	0	1C
18. *Ensifer adhaerens OV14*	1C	1C	1C
19. *Hoeflea sp. IMCC20628*	1C	0	1C
20. *Martelella sp. AD-3*	0	0	0
21. *Mesorhizobium ciceri* biovar biserrulae WSM1271	0	0	0
22. *Mesorhizobium loti* MAFF303099	0	1C	0
23. *Methylobacterium extorquens* AM1	5 C/3P	0	2P
24. *Methylobacterium extorquens* CM4	3C	0	2C
25. *Methylobacterium extorquens* DM4	1C	0	1C
26. *Methylobacterium nodulans* ORS 2060	1C/1P	0	1P
27. *Methylobacterium populi* BJ001	3C	0	2C
28. *Methylobacterium radiotolerans* JCM 2831	3C	0	0
29. *Methylobacterium* sp. 4-46	1C	1C	0
30. *Neorhizobium galegae bv. officinalis bv. officinalis str. HAMBI 1141*	1C	0	C1
31. *Nitrobacter hamburgensis* X14	2P	1C	1C
32. *Ochrobactrum anthropi* DSM 6882	1C1/4C2	1C2	0
33. *Oligotropha carboxidovorans* OM4	1C/1P	0	0
34. *Oligotropha carboxidovorans* OM5	1C/2P	0	1P/3C
35. *Pelagibacterium halotolerans* B2	1C	0	0
36. *Rhizobium etli* bv. mimosae str. Mim1	0	0	0
37. *Rhizobium etli* CFN 42	0	0	1P
38. *Rhizobium etli* CIAT 652	0	0	1P
39. *Rhizobium leguminosarum* bv. trifolii WSM1325	0	0	0
40. *Rhizobium leguminosarum* bv. trifolii WSM2304	0	0	1P
41. *Rhizobium leguminosarum* bv. viciae 3841	0	0	1C
42. *Rhizobium tropici* CIAT 899	0	0	0
43.*Rhodopseudomonas palustris CGA009*	0	0	0
44. *Shinella sp. HZN7*	1	0	0
45. *Sinorhizobium fredii* NGR234	1P	0	0
46. *Sinorhizobium medicae* WSM419	1P	0	1C
47. *Sinorhizobium meliloti* 1021	2P	0	1C
48. *Sinorhizobium meliloti* 2011	1P	0	1C1
49. *Sinorhizobium meliloti* BL225C	2P	0	1C
50. *Sinorhizobium meliloti* GR4	3P	0	1C
51. *Sinorhizobium meliloti* Rm41	2P	0	1C
52. *Sinorhizobium meliloti* SM11	2P	0	1C
53. *Xanthobacter autotrophicus* Py2	2C	0	2P/1C
TOTAL	31C/24P	9C	24C/9P

[a]Replicons: C, Chromosome; C2, Chromosome number two: C3, Chromosome number three; CL, lineal chomosome; P, plasmid. Strains with CopZ, CopA and CusF homologs are yellow highlighted.

largest number of putative CopZ chaperons was found in the genomes of *Methylobacterium extorquens* AM1 and *Ochrobactrum anthropi* DSM6882 with eight and five homologs, respectively. Similar to previously characterized CopZ proteins, the rhizobial homologs are small proteins varying between 67 and 114 amino acids that adopt the ferredoxin-like fold. All of them contain the conserved HMA (PF00403) and the signature MTCXXC motif with the two invariable cysteine residues required for binding and transfer of metal ions (Table S2 CopZ homologs).

Since we could not retrieve CopZ homologs from *Rhizobium*, *Bradyrhizobium*, and *Mesorhizobium* genomes (Table S2 CopZ homologs), we searched for *S. pneumoniae* CupA homologs, which have been reported to functionally substitute CopZ in *Streptococcus* and *Lactobacillus*. CupA is a cell membrane-anchored Cu^{+} chaperone that shares an isostructural cupredoxin-like fold of 100 amino acids with the N-terminus of *Streptococcus pneumoniae* CopA (Fu et al., 2013), which is absent in most CopA homologs. Our hmmsearch of Cupredoxin_1 family, PF13473, revealed that *S. pneumoniae* CupA homologs are mostly absent in the genomes of *Rhizobiales* (Table 5; Table S4 CupA homologs); however, we found eight homologous proteins encoded in the chromosome, five of them in strains without CopZ, including *Bradyrhizobium* and *Mesorhizobium*, and three in strains encoding CopZ. All putative CupA have a similar length (100–150 aa) and carry the cupredoxin_1 domain and the N-terminal single transmembrane helix similar to CupA. The putative rhizobial CupA proteins and their respective CopA partners did not share the four amino acid, Cys....Cys-X-Met-X-Met motif, relevant for Cu(I) coordination and

transfer, as observed in *S. pneumoniae* (Table S4 CupA homologs) (Fu et al., 2013).

The *E. coli* CusF is a small periplasmic copper chaperon protein predominantly found in gamma- and alpha-proteobacteria with copper/silver resistance systems (Kim, Rensing, & McEnvoy, 2010). This protein is able to interact with CusB (part of the CusABC system) (Mealman et al., 2011, 2012) and CopA in *E. coli* (Padilla-Benavides et al., 2014). We explored the presence of CusF in our data set (53 genomes), using the HMM-profile available at the Pfam database for the CusF_Ec (PF11604) family. No CusF homologs were detected in 19 genomes of the *Rhizobiales* order; single copies were found in sixteen genomes, three *Methylobacterium* species contained two CusF copies. Three and four putative CusF were found in *X. autotrophycus* and *O. carboxydovorans*, respectively (Table 5; Table S3 CusF homologs). As members of the Ec_CusF family, most of the rhizobial homologs are small proteins ranging in size between 94 and 112 amino acids, which concurs with the 110 amino acids of Ec_CusF (Table S3 CusF homologs). Although all rhizobial CusF homologs conserve the invariable H36, M47 and M49 residues required for Cu^{+} binding (MBD), they present a significant polymorphism in the five electropositive residues (K23, K30, K31, H35 and R50) suggested as key elements in the interaction of Ec_CusF with Ec_CopA (Padilla-Benavides et al., 2014) and Ec_CusF with CusB (Mealman et al., 2011, 2012) (Table S3 CusF homologs).

4 | DISCUSSION

A previous genome-wide study on the trace element utilization of 540 bacterial genomes (Zhang & Gladyshev, 2010) revealed that two *Sinorhizobium* species, the only two rhizobia included, contained 22 copper-dependent proteins, the largest bacterial cuproproteomes predicted to date. They also contained five Cu-ATPases (CopA) with different functions and regulation (Patel et al., 2014). The high demand for copper suggested by these studies encouraged us to investigate the occurrence and diversity of three critical components of the Cu-trafficking machinery: the P1B-Cu ATPase CopA and the cytoplasmic and periplasmic chaperones, CopZ/CupA, and CusF, respectively, in other members of the *Rhizobiales* order.

4.1 | P_{1B}-type ATPase divergence is higher than previously estimated

The use of profile HMM and maximum likelihood sequence analysis allowed us to refine the diversity of Cu-ATPases beyond the established classification into $P_{1B\text{-}1}$ and $P_{1B\text{-}3}$-ATPases. A similar approach has been proven to be effective for the discrimination of Cu-specific protein subfamilies of sensors (Chang et al., 2014). Our results demonstrated that $P_{1B\text{-}1}$-ATPase proteins have diverged into multiple subtypes. Even though our $P_{1B\text{-}1}$ ATPase dataset (410 sequences) was about half as large as that used in Smith's work employing Transitivity Clustering and Protein Similarity Network (672 sequences), our study displayed twice the number of inferred groups. We think that the higher diversity we revealed is due to the strategy employed: (1) Use of an unbiased selection of P_{1B}-ATPases from *Rhizobiales*, instead of limiting the search to characterized Cu-ATPases; (2) The implementation of *ad hoc* profile HMM that includes the most divergent region of the P_{1B}-type ATPases located at the N-terminal MBD; (3) The use of ML phylogenetic inference for sequence analysis instead of using BLAST- or ClustalW-based methods; and (4) The analysis and allocation of 53 characterized P_{1B}-type ATPases in the phylogenetic tree, which allowed us to predict previously undefined groups of Cu-ATPases with high certainty.

This scheme allowed us to discriminate different Cu-ATPase subtypes and to uncover sequence-specific motif changes, something difficult to achieve with previous data. One example is the subtype-specific polymorphism in the N-terminal MBD of subtypes $P_{1B\text{-}1a}$ and $P_{1B\text{-}1b}$, located in group VII, described in this work. Our results suggest that the N-terminal domains of Cu-ATPases behave as independent evolutionary units that have incorporated a number of new motifs, which potentially could modulate protein function. A similar evolutionary trend has been described in eukaryotic Cu-ATPases (Gupta & Lutsenko, 2012).

4.2 | Contribution of HGT to the diversity of CopA subfamilies found in the *Rhizobiales* order

The presence of *copA* genes on rhizobial plasmids proffers that these genes may be prone to spread through HGT. However, our analysis revealed that 25 of 30 genes predicted to be acquired through HGT were encoded in the chromosome and not in plasmids, as would be expected due to the large amount of extrachromosomal genetic information localized on plasmids (Table S7 G+G, CAI, GI). Interestingly, 12 of 22 chromosomally encoded putative *copA* genes belong to a gene cluster predicted to confer metal resistance; and three of 12 *copA* genes are part of predicted genomic islands. Fifty four percent of these foreign Cu-ATPases belong to family VIII. *Methylobacterium* was the genus showing more foreign gene acquisition (14 putative Cu-ATPases). All these data indicate that horizontal gene transfer has enriched the genomes of members of the *Rhizobiales* order with new Cu-ATPase subtypes. The HGT of Cu-ATPases seems to be independent of the replicon (plasmid or chromosome) where the genes are located. This is in agreement with previous studies that have reported HGT in rhizobia with mono-replicon or multi-replicon genome organization (Finan, 2002).

The HGT of P_{1B}-ATPases has also been detected in metal-rich environments, supporting the contribution of HGT to the evolution of copper resistance (Bouzat & Hoostal, 2013; Martinez et al., 2006; Nongkhlaw, Kumar, Acharya, & Joshi, 2012).

Additionally, we found that 48% of the Cu-ATPases surveyed in the *Rhizobiales* order are encoded in extrachromosomal replicons. This may partially explain the diversification of Cu-ATPases observed in our phylogenetic analysis. According to the model of multipartite genomes, each replicon behaves as an independent evolutionary unit, encoding genes for specialized functions (diCenzo, MacLean, Milunovic, Golding, & Finan, 2014; Galardini, Pini, Bazzicalupo, Biondi, & Mengoni, 2013). Since plasmid copy number is usually higher than that of the chromosome, it may accelerate the evolution of plasmid-borne genes (San

Millan et al., 2016; Sano et al., 2014) In support of this hypothesis, in our ML-phylogeny, we observed that chromosomally-encoded FixI proteins, associated with cytochrome oxidases, form a small cluster in group XIV, whereas plasmid-borne FixI1 and FixI2, with a different physiological role, were grouped in different clusters (Table S5 Cu-ATPase subtypes).

4.3 | What is the function of multiple and diverse Cu-ATPases?

Our results show that the number and type of Cu-ATPases is very variable among different members of the *Rhizobiales* order. Two experimentally testable hypotheses arise from these data. The first one proposes that the multiple Cu-ATPases are functionally redundant and all of them act together to confer resistance when cells face a copper overload. If this hypothesis is true, the strains with high number of Cu-ATPases should be adapted to live in a rich-copper environment. The alternative hypothesis is that the presence of multiple Cu-ATPases is the result of a genetic divergence process and each Cu-ATPase confers copper resistance under different environmental conditions or in the different steps of interactions with their animal or plant host. In this regard, Patel et al. (2014) characterized mutants in each one of the five Cu-ATPases of *S. meliloti* 2011. This phenotypic analysis revealed nonredundant copper resistance, *copA1* encodes the sole Cu-ATPase required to detoxify the excess of copper in the cytoplasm. *Medicago sativa* plants inoculated with this mutant did not show nodule alterations. In contrast, plants inoculated with *S. meliloti* harboring *copA1b, fixI1*, *fixI2*, and *copA3* single mutations showed different abnormalities in nodule development. It is intriguing why other *Sinorhizobium* and *Rhizobium* species perform symbiotic nitrogen fixation with a minor number of Cu-ATPases homologs. Are the differences in nodule organogenesis (determinate or indeterminate nodules) and nitrogen metabolism (the synthesis of amides or ureides) of host plants exerting selective pressure for the different types of ATPases? In the case of human pathogens, they have to deal with intracellular copper synthesized by macrophages to kill engulfed bacteria. The multiplicity of Cu-ATPases may provide a resistance mechanism to escape from the human immune system? (Festa and Thiele, 2011). The answer to these questions will help us to understand the functional diversity of the Cu-ATPases subtypes proposed in this study.

The presence of multiple Cu-ATPases was first highlighted for pathogenic and symbiotic bacteria (Argüello, 2003; Argüello et al., 2011). In our study, we found multiple Cu-ATPases in free-living species of biotechnological interest due to their versatile metabolism as *Methylobacterium extorquens*, *Xanthobacter autotrophicus*, *Chelativorans sp*, and *Oligotropha carboxydovorans*, some of them were grouped together in subtype VIb. The role of multiple and diverse Cu-ATPases in their metabolic properties remain to be elucidated.

4.4 | Restricted occurrence of CopZ/CupA and CusF chaperones in the genomes of *Rhizobiales*

This study revealed the presence of multiple CopZ chaperones distributed in different replicons. In several species, the presence of numerous CopZ homologs correlates with the presence of multiple P1B-Cu ATPases distributed in different replicons; for instance, eight CopZ proteins and five Cu-ATPases are encoded in the genome of *M. extorquens* AM1 (Table S2 CopZ homologs), three *copZ* genes are located close to Cu-ATPase encoding genes, other three are neighbors of mercuric, nucleotide, and metal transporters, and the last two are close to hypothetical proteins. These data suggest that putative CopZ metal binding proteins and CopA metal transporters are functionally linked and some of them may even have co-evolved.

We found CopZ in four rhizobial genomes, previously surveyed in Ridge's work (Ridge, Zhang, & Gladyshev, 2008) but undetected by them; the same study was also unable to detect the presence of the CopZ homolog present in the *A. tumefaciens* plant pathogen which was characterized later on (Nawapan et al., 2009), indicating that Ridge's method, based on BLAST searches, was not sensitive enough to detect distant CopZ homologs.The absence of CopZ homologs in fourteen rhizobial genomes from different species implies that an alternative protein should carry out its function. We found eight putative *S. pneumoniae* CupA homologs, five of them in strains without CopZ. However, putative CupA and their respective CopA partners did not share the Cys….Cys-X-Met-X-Met four amino acid motifs relevant for Cu^{+} coordination and transfer (Fu et al., 2013). These data suggest that different amino acid motifs are involved in coordination and transport, or a novel cytoplasmic chaperone may be present in those rhizobial species lacking CopZ or CupA homologs; both scenarios require to be tested These data are in agreement with a previous genome-wide analysis based on BLAST searches that revealed that CopZ homologs predominate in the Firmicutes phylum, while being very scarce in all other bacterial phyla (Ridge et al., 2008). Ridge et al. (2008) also reported that CusF homologs were absent from the genomes of eighteen surveyed rhizobia. Searching 1,081 bacterial genomes for the presence of HWMM and HMMM metal-binding motifs in a sequence alignment of the best BLASTP hits queried with *E. coli* K12 CusF (Kim et al. 2010), seven CusF-like proteins were found in members of the *Rhizobiales* order. We reported twenty one putative rhizobial CusF proteins encoded in 12 rhizobial genomes, using the HMM-profile of the CusF_Ec (PF11604) family; an unexpected diversity was observed in the protein length, which varies in a range from 94 to 236 amino acids, as well as in the polymorphism of the five electropositive residues (K2, K30, K31, H35 and R50) suggested as being critical for interaction with CopA and CusB. The location in plasmids of 40% of these CusF homologs may also be related to the protein evolution required to interact with the highly diversified group of P1B-type ATPases. The absence of CusF in most rhizobial genomes leads us to hypothezise that alternative unknown chaperones distinct from CusF may be present in those rhizobia lacking CusF homologs.

Although our study is based on P_{1B}-type-ATPases that confer resistance to metals, the analysis of multiplicity and diversity of transporters have general implications for the evolution of resistance, adaptation and innovation in bacteria. For instance, the co-existence of antibiotics and metals in the environment has conducted to the co-evolution of their mechanisms of resistance; the antibiotic resistance induced by exposure to toxic metals is well documented (Berg,

Tom-Petersen, & Nybroe, 2005; Chen et al., 2015; Stepanauskas et al., 2006). Furthermore, metals and antibiotic resistance genes tend to co-occurr (Pal, Bengtsson-Palme, Kristiansson, & Larsson, 2015; Su, Ye, & Zhu, 2012).

5 | CONCLUSION

This study expands the occurrence of multiple and quite diverse Cu-ATPases in the genomes of *Rhizobiales*. We propose a novel classification of Cu-transporting ATPases in seven subfamilies, broadening the previous classifications of Cu^{+}- and Cu^{2+}-ATPases in P_{1B-1} and P_{1B-3} subgroups.

The occurrence of copies of *copA* in plasmids or genomic islands as well as the exchange of Cu-ATPases through lateral gene transfer depicts a genetic divergence process that bacteria have to follow to deal with the selective pressure exerted by copper in the environment.

The complete CopZ-CopA-CusF transport system found only in 37% of the analyzed rhizobial genomes, strongly suggests the existence of an alternative mechanism and/or unexplored proteins involved in copper transit.

This genome-wide study of occurrence, diversity and evolution of Cu^{+}-ATPases in alpha-proteobacteria belonging to the *Rhizobiales* order enriches our current knowledge, which has mainly been based on gamma-proteobacteria analysis.

ACKNOWLEDGMENTS

The research project of A. García-de los Santos was supported by UNAM-DGAPA-PAPIIT Grant number IN209815. F. Miranda-Sánchez, A. González and J.P. Elizalde are doctoral students from the Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and received fellowships from CONACYT (no. 309430, 290161, 354729).

The authors are grateful to Dr. David Romero (Centro de Ciencias Genómicas, UNAM) for helpful discussions.

CONFLICT OF INTEREST

None declared.

REFERENCES

Argüello, J. M. (2003). Identification of ion-selectivity determinants in heavy-metal transport P1B ATPases. *The Journal of Membrane Biology*, *195*, 93–108.

Argüello, J. M., Eren, E., & Gonzalez-Guerrero, M. (2007). The structure and function of heavy metal transport P1B-ATPases. *BioMetals*, *20*, 233–248.

Argüello, J. M., Gonzalez-Guerrero, M., & Raymunda, D. (2011). Bacterial transition metal ATPases Transport mechanism metabolism and roles in virulence. *Biochemistry*, *50*, 9940–9949.

Berg, J., Tom-Petersen, A., & Nybroe, O. (2005). Copper amendment of agricultural soil selects for bacterial antibiotic resistance in the field. *Letters in Applied Microbiology*, *40*, 146–151.

Bouzat, J. L., & Hoostal, M. J. (2013). Evolutionary analysis and lateral gene transfer of two-component regulatory systems associated Argüello, J. M. 2003. Identification of ion-selectivity determinants in heavy-metal transport P_{1B}-type ATPases. *The Journal of Membrane Biology*, *195*, 93–108.

Carvalho, F. M., Souza, R. C., Barcellos, F. G., Hungria, M., & Vasconcelos, A. T. (2010). Genomic and evolutionary comparisons of diazotrophic and pathogenic bacteria of the order *Rhizobiales*. *BMC Microbiology*, *10*, 37.

Chan, H., Babayan, V., Blyumin, E., Gandhi, C., Hak, K., Harake, D., … Liu, H. Y. (2010). The p-type ATPase superfamily. *Journal of Molecular Microbiology and Biotechnology*, *19*, 5–104.

Chang, F. M., Coyne, H. J., Cubillas, C., Vinuesa, P., Fang, X., Ma, Z., … Giedroc, P. D. (2014). Cu(I)-mediated allosteric switching in a copper-sensing operon repressor (CsoR). *Journal of Biological Chemistry*, *289*, 19204–19217.

Chaturvedi, K. S., & Henderson, J. P. (2014). Pathogenic adaptations to host-derived antibacterial copper. *Frontiers in Cellular and Infection Microbiology*, *4*, 3.

Chen, S., Li, X., Sun, G., Zhang, Y., Su, J., & Ye, J. (2015). Heavy Metal Induced Antibiotic Resistance in Bacterium LSJC7. *International Journal of Molecular Sciences*, *16*(10), 23390–23404.

Cheruiyot, D. J., Boyd, R. S., & Moar, W. J. (2013). Exploring lower limits of plant elemental defense by cobalt, copper, nickel, and zinc. *Journal of Chemical Ecology*, *39*, 666–674.

Cooksey, D. A. (1990). Plasmid-determined copper resistance in *Pseudomonas syringae* from impatiens. *Applied and Environmental Microbiology*, *56*, 13–16.

Dhillon, B. K., Chiu, T. A., Laird, M. R., Langille, M. G., & Brinkman, F. S. (2013). IslandViewer update: Improved genomic island discovery and visualization. *Nucleic Acids Research*, *41*, 129–132.

diCenzo, G. C., MacLean, A. M., Milunovic, B., Golding, G. B., & Finan, T. M. (2014). Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLoS Genetics*, *10*, e1004742.

Dupont, C. L., Grass, G., & Rensing, C. (2011). Copper toxicity and the origin of bacterial resistance new insights and applications. *Metallomics: Integrated Biometal Science*, *3*, 1109–1118.

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics*, *14*, 755–763.

Ettema, T. J., Huynen, M. A., de Vos, W. M., & van der Oost, J. (2003). TRASH: A novel metal-binding domain predicted to be involved in heavy-metal sensing, trafficking and resistance. *Trends in Biochemical Sciences*, *28*, 170–173.

Festa, R. A. and Thiele, D. J. 2011. Copper: an essential metal in Biology. *Curr Biol*, *8*, R877–R883. doi:10.1016/j.cub.2011.09.040

Finan, T. M. (2002). Evolving insights: Symbiosis islands and HGT. *Journal of Bacteriology*, *184*, 2855–2856.

Franke, S., Grass, G., Rensing, G., & Nies, D. H. (2003). Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli*. *Journal of Bacteriology*, *185*, 3804–3812.

Fu, Y., Chang, F. M., & Giedroc, D. P. (2014). Copper transport and trafficking at the host bacterial pathogen interface. *Accounts of Chemical Research*, *47*, 3605–3613.

Fu, Y., Tsui, H. C., Bruce, K. E., Sham, L. T., Higgins, K. A., Lisher, J. P., et al. (2013). A new structural paradigm in copper resistance in *Streptococcus pneumoniae*. *Nature Chemical Biology*, *9*, 177–183.

Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G., & Mengoni, A. (2013). Replicon-dependent bacterial genome evolution: The case of *Sinorhizobium meliloti*. *Genome Biology and Evolution*, *5*, 542–558.

Gonzalez-Guerrero, M., & Argüello, J. M. (2008). Mechanism of Cu^{+}-transporting ATPases: Soluble Cu^{+} chaperones directly transfer Cu^{+} to transmembrane transport sites. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 5992–5997.

Gonzalez-Guerrero, M., Raimunda, D., Cheng, X., & Argüello, J. M. (2010). Distinct functional roles of homologous Cu^{+} efflux ATPases in *Pseudomonas aeruginosa*. *Molecular Microbiology*, *78*, 1246–1258.

Gourdon, P., Liu, X. Y., Skjorringe, T., Morth, J. P., Moller, L. B., Pedersen, B. P., & Nissen, P. (2011). Crystal structure of a copper-transporting PIB-type ATPase. *Nature*, *475*, 59–64.

Gupta, A., & Lutsenko, S. (2012). Evolution of copper transporting ATPases in eukaryotic organisms. *Current Genomics*, *13*, 124–133.

Hayek, N. (2013). Lateral transfer and GC content of bacterial resistance genes. *Frontiers in Microbiology*, *4*, 41.

Jumas-Bilak, E., Michaux-Charachon, S., Bourg, G., Ramuz, M., & Allardet-Servent, A. (1998). Unconventional genomic organization in the alpha subgroup of the Proteobacteria. *Journal of Bacteriology*, *180*, 2749–2755.

Kim, E. H., Rensing, C., & McEnvoy, M. M. (2010). Chaperone-mediate copper handling in the periplasm. *Natural Products Reports*, *27*, 711–719.

Lakzian, A., Murphy, P., & Giller, K. E. (2007). Transfer and loss of naturally-occurring plasmids among isolates of *Rhizobium leguminosarum* bv. viciae in heavy metal contaminated soils. *Soil Biology & Biochemistry*, *39*, 1066–1077.

Landeta, C., Davalos, A., Cevallos, M. A., Geiger, O., Brom, S., & Romero, D. (2011). Plasmids with a chromosome-like role in rhizobia. *Journal of Bacteriology*, *193*, 1317–1326. d

MacLean, A. M., Finan, T. M., & Sadowsky, M. J. (2007). Genomes of the symbiotic nitrogen fixing bacteria of legumes. *Plant Physiology*, *144*, 615–622.

Macomber, L., & Imlay, J. A. (2009). The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 8344–8349.

Mana-Capelli, S., Mandal, A. K., & Argüello, J. M. (2003). *Archaeoglobus fulgidus* CopB is a thermophilic Cu-ATPase. *Journal of Biological Chemistry*, *278*, 40534–40541.

Martinez, R. J., Wang, Y., Raimondo, M. A., Coombs, J. M., Barkay, T., & Sobecky, P. A. (2006). Horizontal gene transfer of PIB-type ATPases among bacteria isolated from radionuclide- and metal-contaminated subsurface soils. *Applied and Environment Microbiology*, *72*, 3111–3118.

Mealman, T. D., Bagai, I., Singh, P., Goodlett, D. R., Rensing, C., & Zhou, H. (2011). Interactions between CusF and CusB identified by NMR spectroscopy and chemical cross-linking coupled to mass spectrometry. *Biochemistry*, *50*, 2559–2566.

Mealman, T. D., Zhou, M., Affandi, T., Chacon, K. N., Aranguren, M. E., Blackburn, N. J., … McEvoy, M. M. (2012). N-terminal region of CusB is sufficient for metal binding and metal transfer with the metallochaperone CusF. *Biochemistry*, *51*, 6767–6775.

Monchy, S., Benotmane, M. A., Janssen, P., Vallaeys, T., Taghavi, S., Van der Lelie, D., & Mergeay, M. (2007). Plasmids pMOL28 and pMOL30 of *Cupriavidus metallidurans* are specialized in the maximal viable response to heavy metals. *Journal of Bacteriology*, *189*, 7417–7425.

Nawapan, S., Charoenlap, N., Charoenwuttitam, A., Saenkham, P., Mongkolsuk, S., & Vattanaviboon, P. (2009). Functional and expression analyses of the cop operon, required for copper resistance in *Agrobacterium tumefaciens*. *Journal of Bacteriology*, *191*, 5159–5168.

Nawrocki, E. P., Kolbe, D. L., & Eddy, S. R. (2009). Infernal 1.0: Inference of RNA alignments. *Bioinformatics*, *25*, 1335–1337.

Nongkhlaw, M., Kumar, R., Acharya, C., & Joshi, S. R. (2012). Occurrence of horizontal gene transfer of P(IB)-type ATPase genes among bacteria isolated from the uranium rich deposit of Domiasiat in North East India. *PLoS ONE*, *7*, e48199.

Odermatt, A., Suter, H., Krapf, R., & Solioz, M. (1993). Primary structure of two P-type ATPases involved in copper homeostasis in *Enterococcus hirae*. *Journal of Biological Chemistry*, *268*, 12775–12779.

Overmars, L., Kerkhoven, R., Siezen, R. J., & Francke, C. (2013). MGcV: The microbial genomic context viewer for comparative genome analysis. *BMC Genomics*, *14*(1), 209.

Padilla-Benavides, T., George Thompson, A. M., McEvoy, M. M., & Argüello, J. M. (2014). Mechanism of ATPase-mediated Cu^{+} export and delivery to periplasmic chaperones: The interaction of *Escherichia coli* CopA and CusF. *Journal of Biological Chemistry*, *289*, 20492–20501.

Pal, C., Bengtsson-Palme, J., Kristiansson, E., & Larsson, D. G. J. (2015). Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics*, *16*, 964.

Palumaa, P. (2013). Copper chaperones. The concept of conformational control in the metabolism of copper. *FEBS Letters*, *587*, 1902–1910.

Patel, S. J., Padilla-Benavides, T., Collins, J. M., & Argüello, J. M. (2014). Functional diversity of five homologous Cu+-ATPases present in *Sinorhizobium meliloti*. *Microbiology*, *160*, 1237–1251.

Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R., Moret, B. M., & Stamatakis, A. (2010). How many bootstrap replicates are necessary? *Journal of Computational Biology*, *17*, 337–354.

Preisig, O., Zufferey, R., Thöny-Meyer, L., Appleby, C. A., & Hennecke, H. (1996). A High-affinity cbb3-type cytochrome oxidase terminates the symbiosis-specific respiratory chain of *Bradyrhizobium japonicum*. *Journal of Bacteriology*, *178*, 1532–1538.

Puigbo, P., Bravo, I. G., & Garcia-Vallve, S. (2008). E-CAI: A novel server to estimate an expected value of Codon Adaptation Index (eCAI). *BMC Bioinformatics*, *9*, 65.

Reeve, W. G., Tiwari, R. P., Kale, N. B., Dilworth, M. J., & Glenn, A. R. (2002). ActP controls copper homeostasis in *Rhizobium leguminosarum* bv. viciae and *Sinorhizobium meliloti* preventing low pH-induced copper toxicity. *Molecular Microbiology*, *43*, 981–991.

Ridge, P. G., Zhang, Y., & Gladyshev, V. N. (2008). Comparative genomic analyses of copper transporters and cuproproteomes reveal evolutionary dynamics of copper utilization and its link to oxygen. *PLoS ONE*, *3*(1), e1378.

Rubino, J. T., & Franz, K. J. (2012). Coordination chemistry of copper proteins: How nature handles a toxic cargo for essential function. *Journal of Inorganic Biochemistry*, *107*, 129–143.

San Millan, A., Escudero, J. A., Gifford, D. R., Mazel, D., & MacLean, R. C. (2016). Multicopy plasmids potentiate the evolution of antibiotic resistance in bacteria. *Nature Ecology and Evolution*, *1*, 0010.

Sano, E., Maisnier-Patin, S., Aboubechara, J., Quiñones-Soto, S., & Roth, J. R. (2014). Plasmid copy number underlies adaptive mutability in bacteria. *Genetics*, *198*, 919–933.

Sazinsky, M. H., LeMoine, B., Orofino, M., Davydov, R., Bencze, K. Z., Stemmler, T. L., … Rosenzweig, A. C. (2007). Characterization and structure of a Zn^{2+} and [2Fe-2S]-containing copper chaperone from *Archaeoglobus fulgidus*. *Journal of Biological Chemistry*, *282*, 25950–25959.

Singleton, C., & Le Brun, N. E. (2007). Atx1-like chaperones and their cognate P-type ATPases: Copper-binding and transfer. *BioMetals*, *20*, 275–289.

Smith, A. T., Smith, K. P., & Rosenzweig, A. C. (2014). Diversity of the metal-transporting P1B-type ATPases. *JBIC Journal of Biological Inorganic Chemistry*, *19*, 947–960.

Solioz, M., & Odermat, A. (1995). Copper and silver transport by CopB-ATPase in membrane vesicles of *Enterococcus hirae*. *Journal of Biological Chemistry*, *270*, 9217–9221.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313.

Stepanauskas, R., Glenn, T. C., Jagoe, C. H., Tuckfield, R. C., Lindell, A. H., King, C. J., & McArthur, J. V. (2006). Coselection for microbial resistance to metals and antibiotics in freshwater microcosms. *Environmental Microbiology*, *8*, 1510–1514.

Su, J. Q., Ye, J., & Zhu, Y. G. (2012). Draft genome sequence of a novel bacterial strain, LSJC7, belonging to the family enterobacteriaceae with dual resistance to arsenic and tetracycline. *Journal of Bacteriology*, *194*(24), 7005–7006.

Tetaz, T. J., & Luke, R. K. J. (1983). Plasmid-controlled resistance to copper in *Escherichia coli*. *Journal of Bacteriology*, *154*, 1263–1268.

Wattam, A. R., Abraham, D., Dalay, O., Disz, T. L., Driscoll, T., Gabbard, J. L., Gillespie, J. J., et al. (2014). PATRIC, the bacterial bioinformatics

database and analysis resource. *Nucleic Acids Research*, *42*(D1), D581–D591.

Zhang, Y., & Gladyshev, V. N. (2010). General trends in trace element utilization revealed by comparative genomic analyses of Co, Cu, Mo, Ni, and Se. *Journal of Biological Chemistry*, *285*, 3393–3405.

Zhaxybayeva, O., Gogarten, J. P., Charlebois, R. L., Doolittle, W. F., & Papke, R. T. (2006). Phylogenetic analyses of cyanobacterial genomes: Quantification of horizontal gene transfer events. *Genome Research*, *16*(9), 1099–1108.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Cubillas C, Miranda-Sánchez F, González-Sánchez A, et al. A comprehensive phylogenetic analysis of copper transporting P_{1B} ATPases from bacteria of the *Rhizobiales* order uncovers multiplicity, diversity and novel taxonomic subtypes. *MicrobiologyOpen*. 2017;6:e452. https://doi.org/10.1002/mbo3.452

Material suplementario

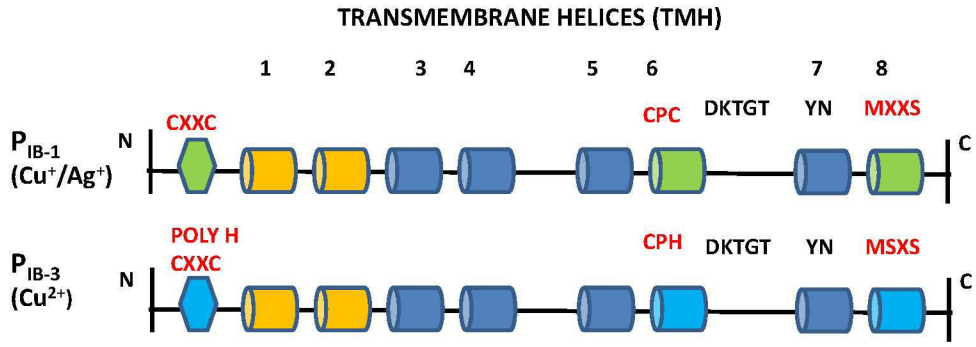

Fig. S1. Domain organization of P_{IB-1} and P_{IB-3} Cu-ATPases subtypes proposed by Argüello (2003, 2007). The N-terminal metal binding domain (MBD) and the metal-binding residues are indicated by a hexagon and amino acids in one letter code. The 8 transmembrane helices (TMH) and their strictly conserved residues are represented by cylinders and amino acids in one letter code. Residues in the TMH6 (CPC/CPH), TMH7 (YN) and TMH8 (MXXS/MSXS) are implicated in Cu^{+} or Cu^{2+} coordination. The aspartic residue (D) in the invariant DKTGT signature is the site of catalytic phosphorylation. Residues in red are conserved signature motifs conserved in most of P_{IB-1} and P_{IB-3} Cu-ATPases.

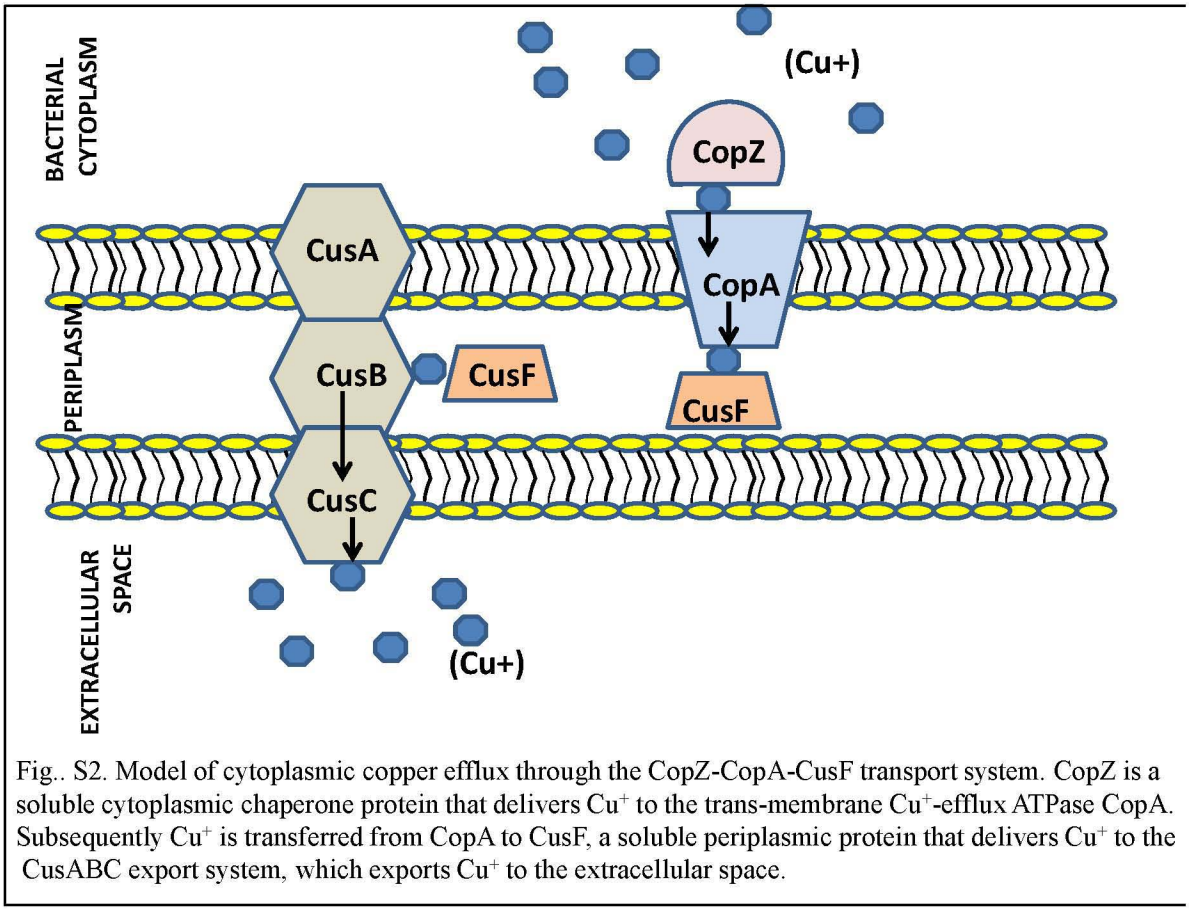

Fig.. S2. Model of cytoplasmic copper efflux through the CopZ-CopA-CusF transport system. CopZ is a soluble cytoplasmic chaperone protein that delivers Cu^{+} to the trans-membrane Cu^{+}-efflux ATPase CopA. Subsequently Cu^{+} is transferred from CopA to CusF, a soluble periplasmic protein that delivers Cu^{+} to the CusABC export system, which exports Cu^{+} to the extracellular space.

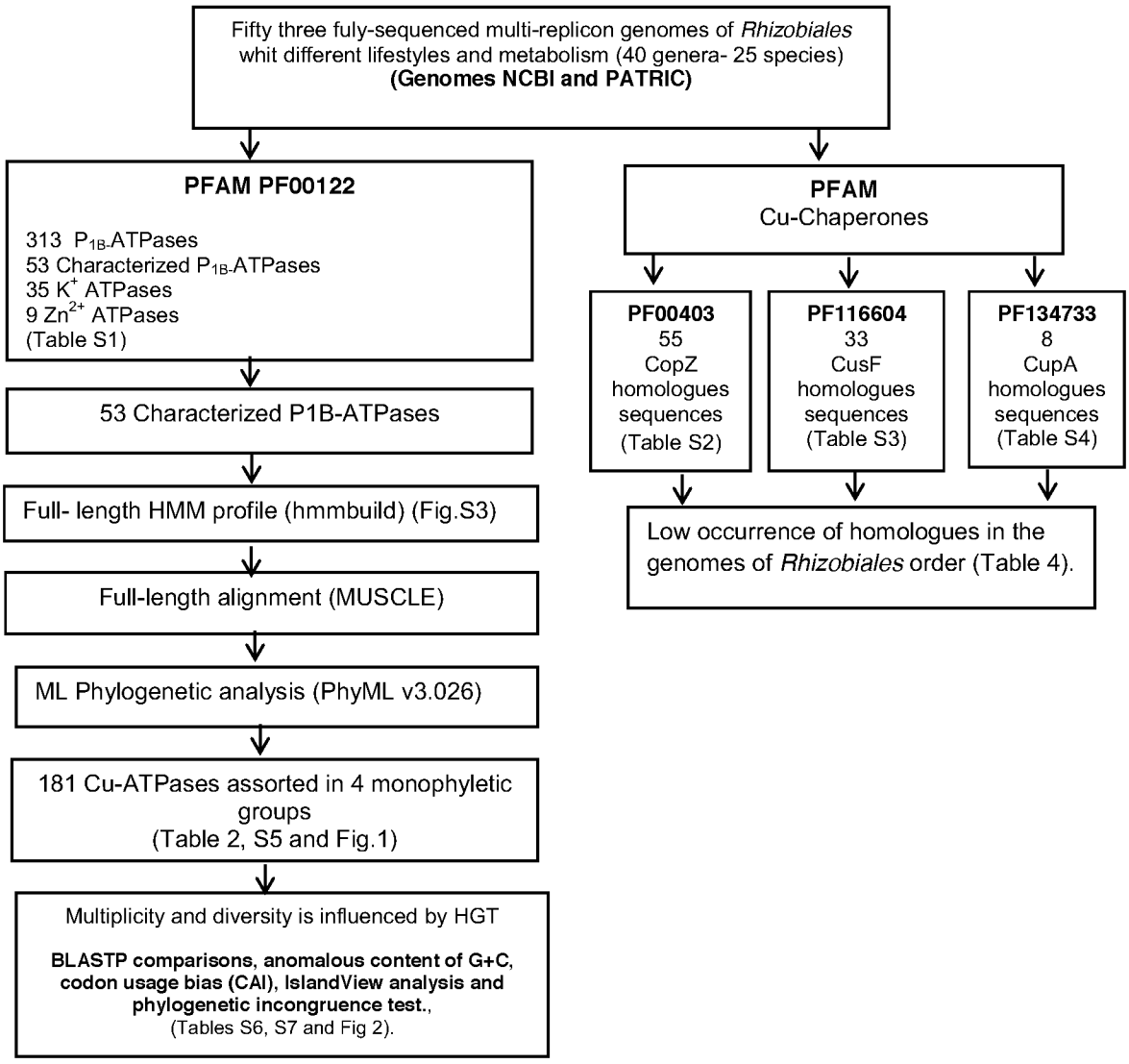

Figure S3. Diagram of data sources (in bold) and phylogenetic analysis.

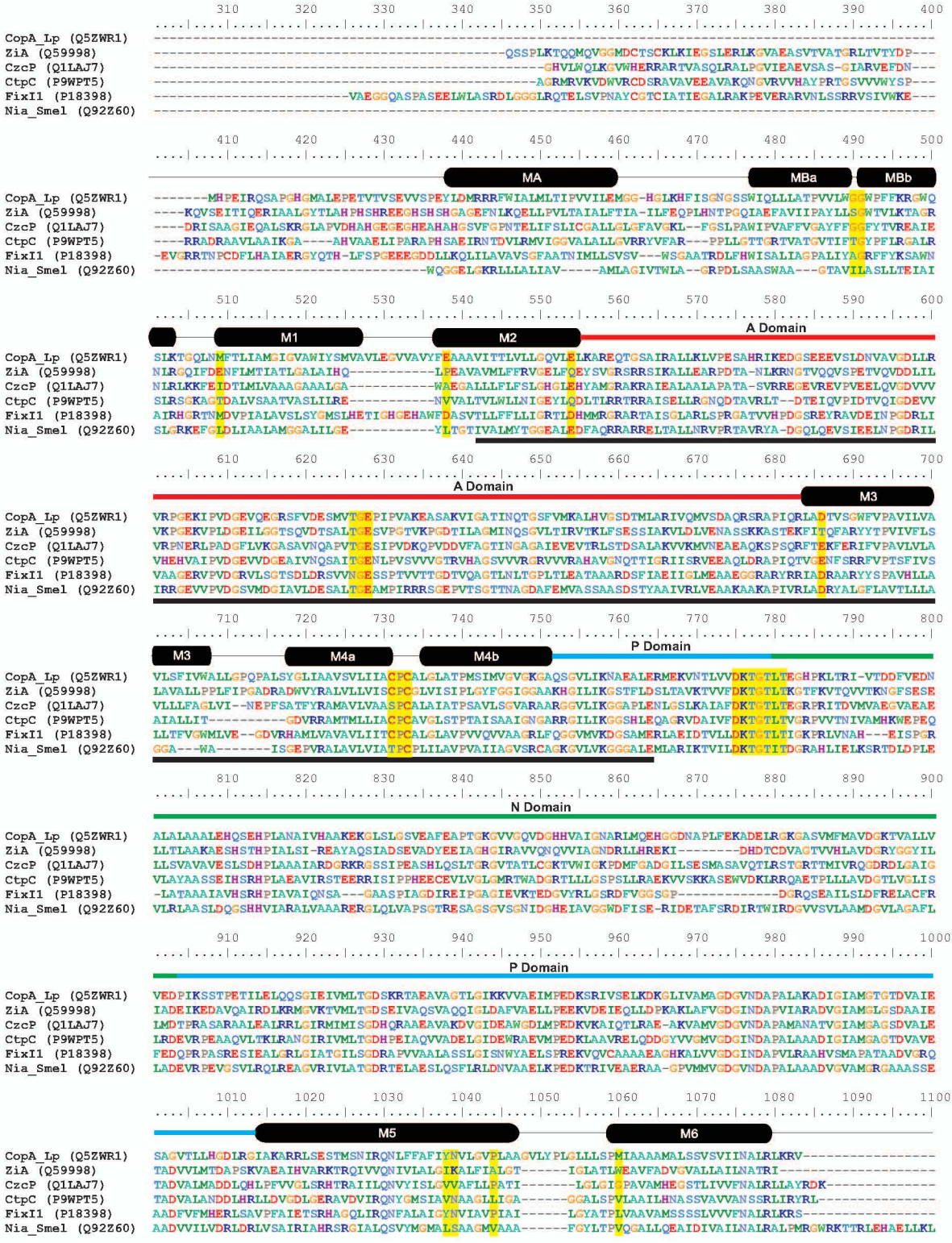

Fig. S4. Alignment of P_{IB}-ATPases used to infer a new HMM profile from 53 characterized P_{IB}-type ATPases. Six of them having Cu (CopA, FixI1), Zn (ZiaA), Co (CzcP), Fe, Ni (Nia) and Mn (CtpC) as substrates are shown.The A- (actuator), P- (phosphorylation), N- (nucleotide binding) and the transmembrane helices (black cylinders) are also indicated based on the crystal CopA structure (Gourdon et al 2011).The black underlined residues represent the profile HMM available at Pfam.

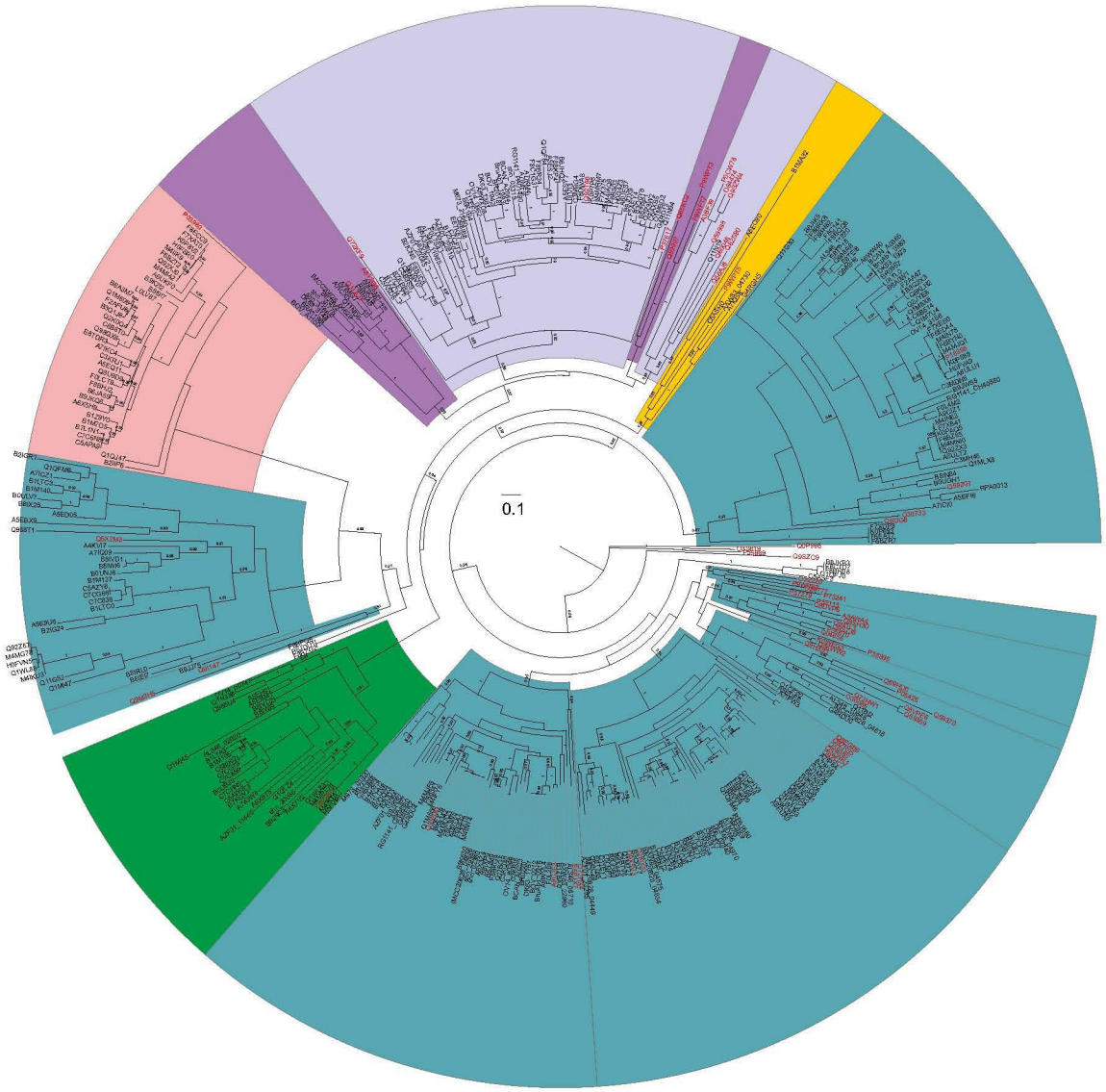

Fig S5. Diversity of P_{IB}-type ATPases revealed by ML-Phylogeny. The ML phylogeny was built by using a 410 P_{IB}-ATPase protein alignment. The fourteen monophyletic groups and their subtypes with a branch support (SH-like) p value ≥ 0.9 are indicated by Roman numbers (I to XIV). Rhizobial Cu-ATPases and their subtypes were distributed in four subfamilies (VI, VII, VIII, XIV). Argüello´s subtypes PIB-1 and PIB-3 are grouped in monophyletic groups VI, VII and VIII. Novel subtypes were distributed in three monophyletic groups (VII, VIII and XIV). The evolutionary relationships of Cu-ATPases with PIB-ATPases for Zn, Cd, Co, Mn and Mg are also shown. The P-ATPases for potassium (K) was used as outgroup.

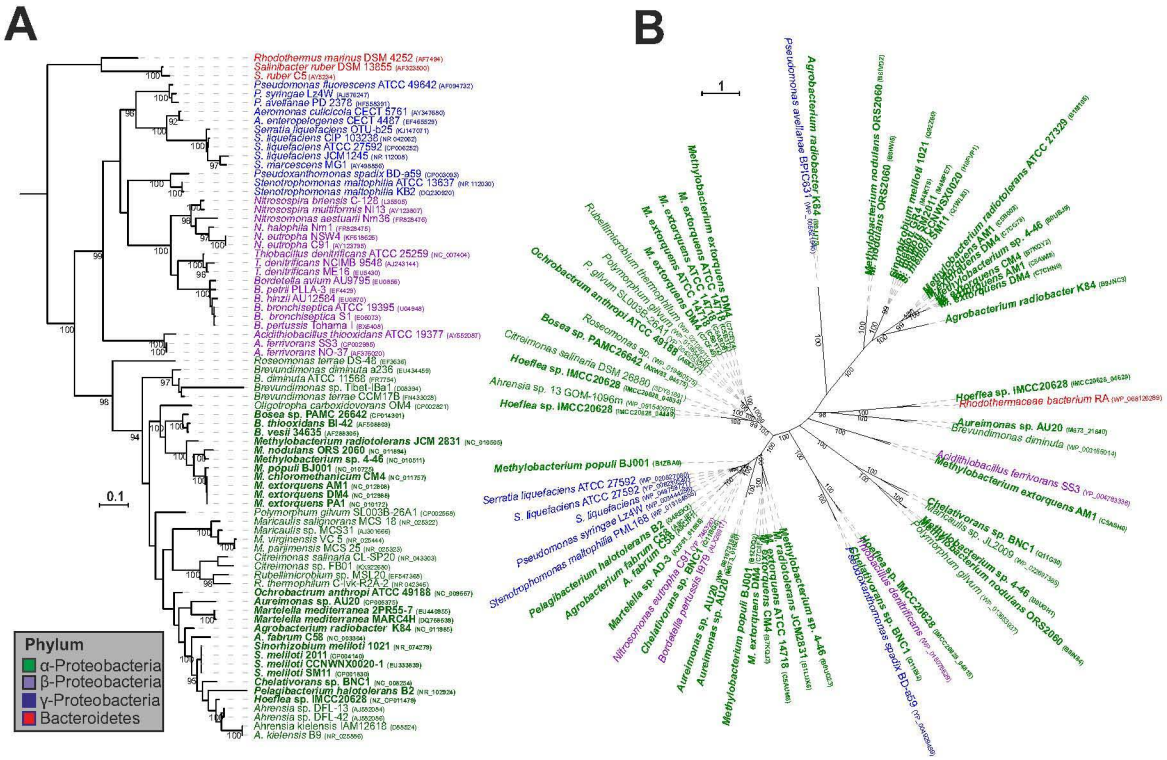

Figure S6. Exchange of Cu-ATPases among several α-, β- and γ- Proteobacteria and Bacteroidetes species. The foreign origin of 30 putative recently transferred genes coding for Cu-ATPases identified in members of *Rhizobiales* order (see text) was tested with a phylogenetic approach comparing the species' phylogeny, inferred from 16S rRNA genes (A) with the phylogeny inferred from homologous Cu^{+}-ATPase coding genes (B). The incongruent phylogenetic relationship amog Cu-ATPases (B) support the hypothesis that those genes have been exchanged among distantly related species. The branch support value (Bootstrap) of each phylogenetic group is indicated at the nodes. The scale bars for *16S rDNA* gene and CopA phylogenies indicate 0.1 and 0.3 expected changes per nucleotide or amino acid under the applied model, respectively.

Table S1: Data set; Occurrence of putative P1B-ATPases in multy-replicon rhizobia genomes (Uncharacterized, FiX homologues and partially Characterized)

Organisms	Uniprot or RefSeq Ids	Gene name	Protein length (aas)
Agrobacterium radiobacter (strain K84)	B9JNC3	actP Arad_7590	644
Agrobacterium radiobacter (strain K84)	B9JJ75	Arad_8781	648
Agrobacterium radiobacter (strain K84)	B9JDX9	Arad_2044	768
Agrobacterium radiobacter (strain K84)	B9JDB0	actP Arad_4551	830
Agrobacterium sp. (strain H13-3)	F0L7S4	AGROH133_04580	906
Agrobacterium sp. (strain H13-3)	F0L206	AGROH133_05421	834
Agrobacterium sp. (strain H13-3)	F0L4M2	fixI AGROH133_06244	763
Agrobacterium tumefaciens (strain C58)	A9CJE3	Atu1195	836
Agrobacterium tumefaciens (strain C58)	A9CJP7	Atu0937	841
Agrobacterium tumefaciens (strain C58)	Q7D0J8	Atu0843	905
Agrobacterium tumefaciens (strain C58)	A9CIZ1	fixI Atu1528	763
Agrobacterium vitis (strain S4)	B9K626	Avi_9859	737
Agrobacterium vitis (strain S4)	B9JWS5	fixI Avi_2393	787
Agrobacterium vitis (strain S4)	B9JWS2	Avi_2388	819
Agrobacterium vitis (strain S4)	B9K1B4	Avi_5561	757
Agrobacterium vitis (strain S4)	B9K630	Avi_9864	826
Aureimonas sp. AU20	M673_01920		825
Aureimonas sp. AU20	M673_21640		734
Beijerinckia indica subsp. indica DSM 1715	B2ICN5	Bind_0269	713
Beijerinckia indica subsp. indica DSM 1715	B2IJD3	Bind_2679	857
Beijerinckia indica subsp. indica DSM 1715	B2IGR1	Bind_2203	849
Beijerinckia indica subsp. indica DSM 1715	B2IDM6	Bind_1836	794
Beijerinckia indica subsp. indica DSM 1715	B2IG24	Bind_2141	839
Bosea sp. PAMC 26642	AXW83_22960		846
Bosea sp. PAMC 26642	AXW83_04575		722
Bosea sp. PAMC 26642	AXW83_04730		681
Bradyrhizobium diazoefficiens (strain USDA 110)	Q59207	fixI blr2769	730
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EH91	BBta_3442	765
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EUA0	actP BBta_p0056	804
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EBM1	BBta_1328	764
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EBX9	mgtA BBta_1451	890
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5ED05	BBta_1846	854
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EFI6	BBta_2797	733
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5E9U6	BBta_0672	890
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5ESM1	BBta_7296	842
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5E8K9	actP BBta_0206	821
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EC90	BBta_1568	703
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EUF6	BBta_p0125	813
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5ESY7	BBta_7421	764

Table S1: Data set, Continue…

Organisms	Uniprot or RefSeq Ids	Gene name	Protein length (aas)
Brucella abortus bv. 1 str. 9-941	BruAb1_1993		804
Brucella abortus bv. 1 str. 9-941	BruAb1_0215		759
Brucella abortus bv. 1 str. 9-941	BruAb1_0383		752
Brucella canis ATCC 23365	BCAN_A2064		814
Brucella canis ATCC 23365	BCAN_A0223		826
Brucella canis ATCC 23365	BCAN_A0365		752
Brucella melitensis bv. 1 str. 16M	DK63_1381		809
Brucella melitensis bv. 1 str. 16M	DK63_1757		826
Brucella melitensis bv. 1 str. 16M	DK63_1923		752
Brucella melitensis bv. 1 str. 16M	DK63_3148		631
Brucella ovis ATCC 25840	BOV_A1100		631
Brucella ovis ATCC 25840	BOV_0212		759
Brucella ovis ATCC 25840	BOV_1942		704
Brucella suis 1330	BR2018		814
Brucella suis 1330	BR0220		826
Brucella suis 1330	BRA1198		567
Brucella suis 1330	BR0357		752
Chelativorans sp. (strain BNC1)	Q11EF0	Meso_2849	734
Chelativorans sp. (strain BNC1)	Q11ND6	Meso_4435	833
Chelativorans sp. (strain BNC1)	Q11BG5	Meso_3893	855
Chelativorans sp. (strain BNC1)	Q11G30	Meso_2255	746
Chelativorans sp. (strain BNC1)	Q11G52	Meso_2232	880
Chelativorans sp. (strain BNC1)	Q11N19	Meso_4212	1022
Chelativorans sp. (strain BNC1)	Q11N94	Meso_4129	955
Chelativorans sp. (strain BNC1)	Q11EY1	Meso_2667	846
Chelativorans sp. (strain BNC1)	Q11G36	Meso_2249	621
Chelativorans sp. (strain BNC1)	Q11BI2	Meso_3876	703
Chelatococcus sp. CO-6	AL346_02020		638
Chelatococcus sp. CO-6	AL346_04970		796
Chelatococcus sp. CO-6	AL346_08710		838
Chelatococcus sp. CO-6	AL346_10545		759
Chelatococcus sp. CO-6	AL346_11220		753
Chelatococcus sp. CO-6	AL346_11965		747
Ensifer adhaerens OV14	OV14_3156		762
Ensifer adhaerens OV14	OV14_a0399		793
Ensifer adhaerens OV14	OV14_b0747		628
Ensifer adhaerens OV14	OV14_0962		770
Hoeflea sp. IMCC20628	IMCC20628_04449		801
Hoeflea sp. IMCC20628	IMCC20628_04629		819
Hoeflea sp. IMCC20628	IMCC20628_04815		808
Hoeflea sp. IMCC20628	IMCC20628_04834		796
Hoeflea sp. IMCC20628	IMCC20628_04836		826
Martelella sp. AD-3	AZF01_01065		825
Martelella sp. AD-3	AZF01_11445		618

Table S1: Data set, Continue…

Organisms	Uniprot or RefSeq Ids	Gene name	Protein length (aas)
Martelella sp. AD-3	AZF01_22810		747
Martelella sp. AD-3	AZF01_23045		710
Martelella sp. AD-3	AZF01_23265		781
Mesorhizobium ciceri bv. biserrulae (strain WSM1271)	E8TP15	Mesci_6043	759
Mesorhizobium ciceri bv. biserrulae (strain WSM1271)	E8T785	Mesci_0186	834
Mesorhizobium ciceri bv. biserrulae (strain WSM1271)	E8TK87	Mesci_2311	737
Mesorhizobium ciceri bv. biserrulae (strain WSM1271)	E8TEI6	Mesci_5504	761
Mesorhizobium loti MAFF303099	Q988T1	mlr6609	896
Mesorhizobium loti MAFF303099	Q98C24	mlr5325	839
Mesorhizobium loti MAFF303099	Q988R8	mll6624	762
Mesorhizobium loti MAFF303099	Q989H6	mlr6417	762
Mesorhizobium loti MAFF303099	Q988U4	mll6590	617
Mesorhizobium loti MAFF303099	Q98IB8	mll2475	749
Methylobacterium extorquens (AM1)	C5AWI8	MexAM1_META1p0915	599
Methylobacterium extorquens (AM1)	C5B002	MexAM1_META1p3832	640
Methylobacterium extorquens (AM1)	C5ASH0	MexAM1_META1p2641 copA	687
Methylobacterium extorquens (AM1)	C5ASF3	MexAM1_META1p2624	820
Methylobacterium extorquens (AM1)	C5AUM6	MexAM1_META1p2883	832
Methylobacterium extorquens (AM1)	C5AZY8	MexAM1_META1p3818	914
Methylobacterium extorquens (AM1)	C5B3T2	MexAM1_META2p0206	806
Methylobacterium extorquens (AM1)	C5B3I7	MexAM1_META2p0089 actP	809
Methylobacterium extorquens (AM1)	C5B1T2	MexAM1_META1p1889 actP	760
Methylobacterium extorquens (AM1)	C5ASD8	MexAM1_META1p2605	806
Methylobacterium extorquens (strain CM4)	B7KPB4	Mchl_1098	806
Methylobacterium extorquens (strain CM4)	B7KQY2	Mchl_1303	646
Methylobacterium extorquens (strain CM4)	B7KPB2	Mchl_1096	731
Methylobacterium extorquens (strain CM4)	B7KNG4	Mchl_1000	809
Methylobacterium extorquens (strain CM4)	B7KQJ2	Mchl_2917	832
Methylobacterium extorquens (strain DM4)	C7CG66	METDI4554	914
Methylobacterium extorquens (strain DM4)	C7CF45	METDI1326	809
Methylobacterium extorquens (strain DM4)	C7C859	METDI0324	641
Methylobacterium extorquens (strain DM4)	C7CG78	METDI4566	640
Methylobacterium extorquens (strain DM4)	C7CHN0	METDI1717	646
Methylobacterium extorquens (strain DM4)	C7C8F4	METDI3449	832
Methylobacterium extorquens (strain DM4)	C7CEY4	actP METDI1264	806
Methylobacterium extorquens (strain DM4)	C7C836	METDI0301	905
Methylobacterium nodulans (ORS2060)	B8IWI5	Mnod_8719	802
Methylobacterium nodulans (ORS2060)	B8IXS2	Mnod_7880	788

Table S1: Data set, Continue…

Organisms	Uniprot or RefSeq Ids	Gene name	Protein length (aas)
Methylobacterium nodulans (ORS2060)	B8IEI7	Mnod_4694	1047
Methylobacterium nodulans (ORS2060)	B8IXP0	Mnod_7843	835
Methylobacterium nodulans (ORS2060)	B8IRL0	Mnod_5730	1032
Methylobacterium nodulans (ORS2060)	B8IVD1	Mnod_6168	945
Methylobacterium nodulans (ORS2060)	B8IVD2	Mnod_6169	776
Methylobacterium nodulans (ORS2060)	B8IX25	Mnod_8078	859
Methylobacterium nodulans (ORS2060)	B8IGD1	Mnod_0801	823
Methylobacterium nodulans (ORS2060)	B8IWI6	Mnod_8720	960
Methylobacterium nodulans (ORS2060)	B8IN84	Mnod_7463	758
Methylobacterium populi (BJ001)	B1ZBA0	Mpop_2539	838
Methylobacterium populi (BJ001)	B1ZCJ6	Mpop_4130	799
Methylobacterium populi (BJ001)	B1ZCE0	Mpop_4071	712
Methylobacterium populi (BJ001)	B1ZBB5	Mpop_2554	786
Methylobacterium populi (BJ001)	B1ZCI4	Mpop_4117	731
Methylobacterium populi (BJ001)	B1ZDQ2	Mpop_2813	830
Methylobacterium radiotolerans (JCM 2831)	B1LTC0	Mrad2831_0411	869
Methylobacterium radiotolerans (JCM 2831)	B1LTC3	Mrad2831_0415	862
Methylobacterium radiotolerans (JCM 2831)	B1MA32	Mrad2831_6437	817
Methylobacterium radiotolerans (JCM 2831)	B1M140	Mrad2831_2604	867
Methylobacterium radiotolerans (JCM 2831)	B1M137	Mrad2831_2601	883
Methylobacterium radiotolerans (JCM 2831)	B1LUX6	Mrad2831_3586	824
Methylobacterium radiotolerans (JCM 2831)	B1M105	Mrad2831_2566	642
Methylobacterium radiotolerans (JCM 2831)	B1LTA3	Mrad2831_0394	641
Methylobacterium sp. (strain 4-46)	B0UNJ6	M446_1604	954
Methylobacterium sp. (strain 4-46)	B0ULV7	M446_6096	874
Methylobacterium sp. (strain 4-46)	B0UGH1	M446_6713	758
Methylobacterium sp. (strain 4-46)	B0UE71	M446_0177	755
Methylobacterium sp. (strain 4-46)	B0UBJ9	M446_3593	637
Methylobacterium sp. (strain 4-46)	B0UQ23	M446_1661	825
Neorhizobium galegae bv. officinalis bv. officinalis str. HAMBI 1141	RG1141_CH40880		763
Neorhizobium galegae bv. officinalis bv. officinalis str. HAMBI 1141	RG1141_PA01490		760
Neorhizobium galegae bv. officinalis str. HAMBI 1141	RG1141_CH08380		808
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QF08	Nham_4630	801
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QFM9	Nham_4377	812
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QFJ0	Nham_4423	817
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QF84	Nham_4535	734
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QFM6	Nham_4380	818
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QLZ4	Nham_1949	618
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QH46	Nham_3727	711
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QHY3	Nham_3434	833
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QFP2	Nham_4364	811
Ochrobactrum anthropi (strain ATCC 49188)	A6WXD8	Oant_0921	844
Ochrobactrum anthropi (strain ATCC 49188)	A6X3W2	Oant_3208	720
Ochrobactrum anthropi (strain ATCC 49188)	A6X3Z4	Oant_3240	852
Ochrobactrum anthropi (strain ATCC 49188)	A6WVL4	Oant_0287	827
Ochrobactrum anthropi (strain ATCC 49188)	A6X873	Oant_4656	616

Table S1: Data set, Continue…

Organisms	Uniprot or RefSeq Ids	Gene name	Protein length (aas)
Ochrobactrum anthropi (strain ATCC 49188)	A6X3Y7	Oant_3233	795
Ochrobactrum anthropi (strain ATCC 49188)	A6X7V2	Oant_4621	599
Ochrobactrum anthropi (strain ATCC 49188)	A6WW40	Oant_0463	752
Oligotropha carboxidovorans (OM5)	F8C1D7	OCA5_pOC1670054 0	817
Oligotropha carboxidovorans (OM5)	F8BZG2	ctpC OCA5_c30670	647
Oligotropha carboxidovorans (OM5)	F8C198	OCA5_pOC1670014 0	813
Oligotropha carboxidovorans (OM5)	B6JDU1	cadA OCAR_4884	654
Oligotropha carboxidovorans (OM5)	B6JJM9	actP	822
Oligotropha carboxidovorans (OM5)	F8BVV1	OCA5_c05050	815
Oligotropha carboxidovorans (OM5)	B6JBV2	cadA4	654
Oligotropha carboxidovorans (OM5)	B6J9V0	cadA3	712
Oligotropha carboxidovorans (OM5)	F8BRK8	fixI OCA5_c09600	753
Oligotropha carboxidovorans (OM5)	B6JKB3	OCAR_7762	823
Oligotropha carboxidovorans (OM5)	F8C141	fixI OCA5_pHCG300770	748
Oligotropha carboxidovorans (OM5)	B6JK76	cadA1	840
Oligotropha carboxidovorans (OM5)	B6JIK5	OCAR_7144	757
Oligotropha carboxidovorans (OM5)	B6JHQ2	cadA2	730
Oligotropha carboxidovorans (strain OM4)	F8BR45	cadA OCA4_pOC167B009 10	840
Oligotropha carboxidovorans (strain OM4)	F8BR08	OCA4_pOC167B005 40	817
Oligotropha carboxidovorans (strain OM4)	F8BJC8	OCA4_c04860	712
Oligotropha carboxidovorans (strain OM4)	F8BN62	fixI OCA4_c09590	753
Oligotropha carboxidovorans (strain OM4)	F8BQQ8	fixI OCA4_pHCG3B0077 0	748
Oligotropha carboxidovorans (strain OM4)	F8BQW9	OCA4_pOC167B001 40	813
Oligotropha carboxidovorans (strain OM4)	F8BJE6	OCA4_c05040	815
Oligotropha carboxidovorans (strain OM4)	F8BPQ6	ctpC OCA4_c30150	647
Oligotropha carboxidovorans (strain OM4)	F8BKG9	actP OCA4_c06180	822
Pelagibacterium halotolerans (strain JCM 15775 / B2)	G4RGH5	KKY_104	647
Pelagibacterium halotolerans (strain JCM 15775 / B2)	G4RDU4	KKY_3874	845
Pelagibacterium halotolerans (strain JCM 15775 / B2)	G4RD13	KKY_2790	626
Pelagibacterium halotolerans (strain JCM 15775 / B2)	G4RDU0	KKY_3870	687
Pelagibacterium halotolerans (strain JCM 15775 / B2)	G4REK2	KKY_813	831
Rhizobium etli (strain CFN 42)	Q2K000	actP RHE_PE00007	840
Rhizobium etli (strain CFN 42)	Q8KLH2	RHE_PD00290	756
Rhizobium etli (strain CFN 42)	Q2K3W2	RHE_CH03719	748
Rhizobium etli (strain CFN 42)	Q2JYE6	fixIf RHE_PF00501	761
Rhizobium etli (strain CIAT 652)	B3Q0Y8	RHECIAT_CH000399 2	747
Rhizobium etli (strain CIAT 652)	B3Q0Z7	actP RHECIAT_PA000000 7	839

Table S1: Data set, Continue…

Organisms	Uniprot or RefSeq Ids	Gene name	Protein length (aas)
Rhizobium etli (strain CIAT 652)	B3Q2X3	fixI RHECIAT_PB000032 5	760
Rhizobium etli CNPAF512	F2AA47	RHECNPAF_290011	760
Rhizobium etli CNPAF512	F2AEY4	RHECNPAF_526007	751
Rhizobium leguminosarum bv. trifolii WSM1325	C6AXB5	Rleg_3796	756
Rhizobium leguminosarum bv. trifolii WSM1325	C6B514	Rleg_4956	761
Rhizobium leguminosarum bv. trifolii WSM1325	C6B7V6	Rleg_7156	841
Rhizobium leguminosarum bv. trifolii WSM1325	B6A445	Rleg2_6008	841
Rhizobium leguminosarum bv. trifolii WSM1325	B5ZRP1	Rleg2_3501	758
Rhizobium leguminosarum bv. trifolii WSM1325	B6A0W1	Rleg2_5022	762
Rhizobium leguminosarum bv. viciae	Q9X5V3	actP	841
Rhizobium leguminosarum bv. viciae 3841	Q1M7U4	fixI1 pRL100210	761
Rhizobium leguminosarum bv. viciae 3841	Q1MI24	RL1892	740
Rhizobium leguminosarum bv. viciae 3841	Q1MI47	RL1869	893
Rhizobium leguminosarum bv. viciae 3841	Q1MI45	RL1871	620
Rhizobium leguminosarum bv. viciae 3841	Q1MBD2	RL4262	756
Rhizobium leguminosarum bv. viciae 3841	Q1M656	pRL110331	824
Rhizobium leguminosarum bv. viciae 3841	Q1M8X6	fixI2 pRL90013	761
Rhizobium leguminosarum bv. viciae 3841	Q1M727	actP pRL110010	841
Rhizobium leguminosarum bv. viciae 3841	Q1MLX8	fixI RL0532	755
Rhizobium sp. (strain NGR234)	C3KQ40	actP NGR_b07400	840
Rhizobium sp. (strain NGR234)	C3KQ24	NGR_b07240	746
Rhizobium sp. (strain NGR234)	C3MH46	fixI2 NGR_c25750	772
Rhizobium sp. (strain NGR234)	C3MBG4	NGR_c34490	747
Rhizobium sp. (strain NGR234)	C3MDN6	fixI1 NGR_c17910	754
Rhizobium sp. (strain NGR234)	C3KRI8	NGR_b12440	830
Rhizobium tropici CIAT 899	L0LNP7	RTCIAT899_CH1757 5	832
Rhizobium tropici CIAT 899	L0LTL0	fixI RTCIAT899_PB0118 0	761
Rhizobium tropici CIAT 899	L0LVI3	RTCIAT899_PC0459 5	783
Rhodopseudomonas palustris CGA009	RPA1660		923
Rhodopseudomonas palustris CGA009	RPA3260		709
Rhodopseudomonas palustris CGA009	RPA0013		732
Shinella sp. HZN7	shn_02340		795
Shinella sp. HZN7	shn_02380		852
Shinella sp. HZN7	shn_03315		749
Shinella sp. HZN7	shn_29395		834
Shinella sp. HZN7	shn_30035		613
Shinella sp. HZN7	shn_30775		659
Shinella sp. HZN7	shn_31665		659
Sinorhizobium medicae (strain WSM419) (Ensifer medicae)	A6UES7	Smed_3336	744
Sinorhizobium medicae (strain WSM419) (Ensifer medicae)	Q9X5X3	actP Smed_4897	827
Sinorhizobium medicae (strain WSM419) (Ensifer medicae)	A6ULT2	Smed_5920	755
Sinorhizobium medicae (strain WSM419) (Ensifer medicae)	A6ULU1	Smed_5929	757
Sinorhizobium meliloti (strain AK83)	AEG57723	Sinme_6240	737

Table S1: Data set, Continue…

Organisms	Uniprot or RefSeq Ids	Gene name	Protein length (aas)
Sinorhizobium meliloti (strain AK83)	AEG56360	Sinme_4694	827
Sinorhizobium meliloti (strain AK83)	AEG55160	Sinme_3455	743
Sinorhizobium meliloti (strain AK83)	AEG57394	Sinme_5876	759
Sinorhizobium meliloti (strain BL225C)	F6BYN0	SinmeB_6163	757
Sinorhizobium meliloti (strain BL225C)	F6BZ85	SinmeB_5734	755
Sinorhizobium meliloti (strain BL225C)	F6C3W5	SinmeB_3653	827
Sinorhizobium meliloti (strain BL225C)	F6BZR7	SinmeB_5077	737
Sinorhizobium meliloti (strain BL225C)	F6BS03	SinmeB_3230	743
Sinorhizobium meliloti (strain SM11)	F7XFT8	atcU2 SM11_pD0582	827
Sinorhizobium meliloti (strain SM11)	F7XB41	fixI2 SM11_pC1367	755
Sinorhizobium meliloti (strain SM11)	Q1WL88		900
Sinorhizobium meliloti (strain SM11)	F7X375	SM11_chr3595	743
Sinorhizobium meliloti (strain SM11)	F7XES0	fixI1 SM11_pC0972	759
Sinorhizobium meliloti (strain SM11)	Q1WLB3		706
Sinorhizobium meliloti (strain SM11)	A4KVI7	orf82	933
Sinorhizobium meliloti (strain SM11)	F7XDP8	actP SM11_pC0600	737
Sinorhizobium meliloti (strain SM11)	Q1WL83		799
Sinorhizobium meliloti (strain Sm2011)	M4MRM7	SM2011_c04128	743
Sinorhizobium meliloti (strain Sm2011)	M4MG78	SM2011_a1155	900
Sinorhizobium meliloti (strain Sm2011)	M4MJQ1	fixI1 SM2011_a1209	757
Sinorhizobium meliloti (strain Sm2011)	M4MNI0	fixI2 SM2011_a0621	755
Sinorhizobium meliloti (strain Sm2011)	M4MJF5	actP SM2011_a1013	826
Sinorhizobium meliloti (strain Sm2011)	M4MG41	SM2011_a1087	733
Sinorhizobium meliloti (strain Sm2011)	M4MUQ3	actP2 SM2011_b21578	827
Sinorhizobium meliloti (strain Sm2011)	M4MPE7	SM2011_a1163	774
Sinorhizobium meliloti 1021) (Ensifer meliloti)	P58342	actP2	827
Sinorhizobium meliloti 1021) (Ensifer meliloti)	Q92Z60	SMa1163	799
Sinorhizobium meliloti 1021) (Ensifer meliloti)	Q92Z67	SMa1155	900
Sinorhizobium meliloti 1021) (Ensifer meliloti)	Q92ZA2	SMa1087 (copA3)	733
Sinorhizobium meliloti 1021) (Ensifer meliloti)	Q92ZX3	fixI2 SMa0621	755
Sinorhizobium meliloti 1021) (Ensifer meliloti)	Q92T56	R00124 SMc04128	743
Sinorhizobium meliloti 1021) (Ensifer meliloti)	P18398	fixI1 RA0659 SMa1209	757
Sinorhizobium meliloti 1021) (Ensifer meliloti)	P58341	actP1	826
Sinorhizobium meliloti CCNWSX0020	H0FVA9	SM0020_05912	757
Sinorhizobium meliloti CCNWSX0020	H0FVP1	SM0020_05862	774
Sinorhizobium meliloti CCNWSX0020	H0FYJ9	SM0020_11415	827
Sinorhizobium meliloti CCNWSX0020	H0FVL4	SM0020_05727	708
Sinorhizobium meliloti CCNWSX0020	H0FVN5	SM0020_05832	883
Sinorhizobium meliloti CCNWSX0020	H0G4Z3	SM0020_22747	743
Sinorhizobium meliloti GR4	M4IN63	C770_GR4pC1192	755
Sinorhizobium meliloti GR4	M4IKT8	C770_GR4pC0848	774
Sinorhizobium meliloti GR4	M4IFY6	C770_GR4pA158	828
Sinorhizobium meliloti GR4	M4IMC4	C770_GR4pC0893	705
Sinorhizobium meliloti GR4	M4IN78	C770_GR4pC0736	757

Table S1: Data set, Continue…

Organisms	Uniprot or RefSeq Ids	Gene name	Protein length (aas)
Sinorhizobium meliloti GR4	M4IAQ6	C770_GR4Chr0131	743
Sinorhizobium meliloti GR4	M4INY4	C770_GR4pD0564	826
Sinorhizobium meliloti GR4	M4IKU3	C770_GR4pC0853	883
Sinorhizobium meliloti GR4	M4IHQ8	C770_GR4pA165	740
Sinorhizobium meliloti Rm41	K0P693	actP BN406_04774	737
Sinorhizobium meliloti Rm41	K0P588	fixI BN406_04379	757
Sinorhizobium meliloti Rm41	K0PTX1	actP2 BN406_05647	827
Sinorhizobium meliloti Rm41	K0PAQ5	BN406_04081	1025
Xanthobacter autotrophicus Py2	A7IGW3	Xaut_2012	629
Xanthobacter autotrophicus Py2	A7IQ58	Xaut_4952	648
Xanthobacter autotrophicus Py2	A7IQ09	Xaut_4898	947
Xanthobacter autotrophicus Py2	A7IP64	Xaut_4589	1020
Xanthobacter autotrophicus Py2	A7IPV0	Xaut_4834	796
Xanthobacter autotrophicus Py2	A7IKM4	Xaut_3338	759
Xanthobacter autotrophicus Py2	A7IJX8	Xaut_3091	868
Xanthobacter autotrophicus Py2	A7IPW3	Xaut_4847	818
Xanthobacter autotrophicus Py2	A7IMG7	Xaut_3987	834
Xanthobacter autotrophicus Py2	A7IQ22	Xaut_4911	793
Xanthobacter autotrophicus Py2	A7ICX0	Xaut_0607	835
Xanthobacter autotrophicus Py2	A7ICI0	Xaut_0465	743
Xanthobacter autotrophicus Py2	A7ICZ1	Xaut_0631	858
Xanthobacter autotrophicus Py2	A7IMF9	Xaut_3978	759

Characterized P1B-ATPases

Organisms	Uniprot or RefSeq IDs	Gene name	Substrate	References
Enterococcus faecium (Streptococcus faecium)	Q8VPE6	tcrB	Cu	Hasman 2005
Enterococcus hirae (strain ATCC 9790 / DSM 20160)	P05425	copB EHR_09090	Cu^{+}	Odermatt et al 2004
Enterococcus hirae (strain ATCC 9790)	P32113	copA EHR_09085	Cu^{+}	Odermatt et al 2004
Escherichia coli	Q59370	HRA-2		Trenor et al 1994
Escherichia coli	Q59369	HRA-1		Trenor et al 1994
Escherichia coli (strain K12)	Q59385	copA (ybaR)	Cu^{2+}/Ag^{+}	Stoyanov et al 2003
Escherichia coli (strain K12)	P37617	zntA (yhhO)	Zn^{2+}/Cd^{2+}	Rensing et al 1997
Escherichia coli (strain K12)	P03960	kdpB	K^{+}	Irzik et al 2011
Helicobacter felis (strain ATCC 49179 / NCTC 12436 / CS1)	O32619	copA	Cu^{2+}	Bayle et al 1998
Helicobacter pylori (strain ATCC 700392 / 26695)	P55989	copA HP_1072	Cu^{2+}	Bayle et al 1998
Homo sapiens (Human)	Q04656	ATP7A MC1 MNK	Cu^{2+}	Bull et al 1993
Homo sapiens (Human)	P35670	ATP7B	Cu^{2+}	Petrukhin et al 1994
Legionella pneumophila (strain Paris)	I7HRD6	ctpA lpp2364	Cu^{+}	Andersson et at 2014
Legionella pneumophila DSM 7513(strain ATCC 33152)	Q5ZWR1	lpg1024	Cu^{+}	Kim et al 2009
Listeria monocytogenes	Q60048	cadA	Cd^{2+}	Lebrun et al 1994

Table S1: Data set, Continue…

Characterized P1B-ATPases

Organisms	Uniprot or RefSeq IDs	Gene name	Substrate	References
Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv)	P9WPT3	ctpD Rv1469 MTV007.16	$Ni^{2+}/Co^{2+}/Zn^{2+}$	Raimunda et al 2012
Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv)	P9WPT7	ctpJ nmtA Rv3743c	$Ni^{2+}/Co^{2+}/Zn^{2+}$	Raimunda et al 2014
Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv)	P9WPT5	ctpC mtaA Rv3270 MTCY71.10	Mn^{2+}	Padilla-Benavideset al 2013
Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv)	P9WPS3	ctpV Rv0969 MTCY10D7.05c	Cu^{+}	Ward et al 2010
Nostoc sp. (strain PCC 7120 / UTEX 2576)	Q8ZS90	alr7622	Cu^{2+}	Verma and Singh 1991
Oryza sativa subsp. japonica (Rice)	A3BF39	OsHMA2 OsJ_22530	Zn^{2+}/Cd^{2+}	Yamaji et al 2013
Oryza sativa subsp. japonica (Rice)	Q655X4	P0473H04.28		
Oryza sativa subsp. japonica (Rice)	A3AWA4	OsHMA5 OsJ_15734	Cu^{+}	Deng et al 2013
Pseudomonas aeruginosa (strain ATCC 15692 / PAO1)	Q9I3G8	PA1549	Zn^{2+}	Teitzel et al
Pseudomonas aeruginosa (strain ATCC 15692 / PAO1)	Q9HX93	PA3920	Cu	Teitzel et al
Pseudomonas aeruginosa (strain ATCC 15692 / PAO1)	Q9I147	PA2435	Cu^{2+}/Zn^{2+}	Lewinson et al 2009
Pseudomonas fluorescens R124	K0WEG3	cueA I1A_000679	Cu	Zhang and Rainy 2007
Pseudomonas putida (Arthrobacter siderocapsulatus)	Q8KWW2	cueA	Cu	Adaikkalam 2002
Ralstonia metallidurans (strain CH34)	Q1LAJ7	czcP Rmet_5970	$Co^{2+}/Zn^{2+}/Cd^{2+}$	Scherer and Nies 2009
Ralstonia metallidurans (strain CH34)	Q58AJ6	pbrA Rmet_5947	Pb^{2+}	Borremans et al 2001
Sinorhizobium medicae WS419	Q9X5X3	ActP	Cu	Reeve 2002
Rhizobium etli	Q2K000	ActP	Cu	Landeta 2011
Rhizobium leguminosarum bv. viciae	Q9X5V3	ActP	Cu	Reeve 2002
Rhodobacter capsulatus (Rhodopseudomonas capsulata)	O30733	ccoI	Cu	Koch et al 2000
Saccharomyces cerevisiae (strain ATCC 204508 / S288c)	P38995	CCC2 YDR270W D9954.6	Cu^{2+}	Fu et al 1995
Salmonella typhimurium	Q9ZHC7	silP	Ag^{+}	Gupta et al 1999
Salmonella typhimurium (strain LT2 /ATCC 700720)	Q8ZR95	copA STM0498	Cu^{+}	Espariz et al 2007
Shewanella oneidensis (strain MR-1)	Q8EGB6	copA SO_1689	Cu^{+}	Toes et al 2008
Staphylococcus aureus	Q69HU0	copB	Cu	Baker et al 2009
Sinorhizobium meliloti	Q92T56	SMc04128	Zn/Cd/Pb	Rossbach 2008
Sinorhizobium meliloti 2011	Q92ZA2	CopA3	Cu^{+}	Patel 2014
Sinorhizobium meliloti 2011	P58341	ActP1	Cu^{+}	Patel 2014
Sinorhizobium meliloti 2011	P58342	ActP2	Cu^{+}	Patel 2014
Sinorhizobium meliloti 2011	Q92ZX3	FixI2	Cu^{+}	Patel 2014
Streptococcus mutans serotype c (ATCC 700610 / UA159)	Q8DVP6	copA SMU_426	Cu^{2+}	Vats and Lee2001
Sulfitobacter sp. NAS-14.1	A3T2G5	NAS141_02821	Co^{2+}	Zielazinski et al 2012
Synechococcus elongatus (PCC 7942) (Anacystis nidulans R2)	P37279	pacS		
Synechocystis sp. (strain PCC 6803 / Kazusa)	P73241	pacS sll1920	Cu	Tottey et al 2001

Table S1: Data set, Continue…

Characterized P1B-ATPases

Organisms	Uniprot or RefSeq IDs	Gene name	Substrate	References
Synechocystis sp. (strain PCC 6803 / Kazusa)	Q59998	ziaA slr0798	Zn^{2+}	Thelwell et al 1998
Synechocystis sp. (strain PCC 6803 / Kazusa)	Q59997	slr0797	Co^{2+}	García-Domínguez et al 2000
Thermus thermophilus (strain HB27)	Q72KF9	TT_C0354	Cu^{+}	Schurig-Briccio et al 2012
Thermus thermophilus (strain HB27)	Q72HW1	TT_C1371 (copB)	Cu^{2+}	Schurig-Briccio et al 2012
Thermus thermophilus (strain HB27)	Q72HX4	pacS TT_C1358 (copA)	Cu^{+}	Schurig-Briccio et al 2012

KdpB sequences used as outgroups for the phylogenetic analysis

Organisms	Uniprot or RefSeq IDs	Gene name
Agrobacterium radiobacter (strain K84)	B9JKQ8	kdpB Arad_9463
Agrobacterium sp. (strain H13-3)	F0LCT9	kdpB
Agrobacterium tumefaciens (strain C58 / ATCC 33970)	Q8U9D9	kdpB
Agrobacterium vitis (strain S4)	B9K2I2	kdpB Avi_6073
Beijerinckia indica subsp. indica DSM 1715	B2IIP6	kdpB Bind_1097
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EQ11	kdpB BBta_6337
Mesorhizobium ciceri bv. biserrulae (strain WSM1271)	E8TDR3	kdpB Mesci_1802
Mesorhizobium loti MAFF303099	Q98GX6	kdpB mll3130
Methylobacterium extorquens (AM1)	C5APA9	kdpB MexAM1_META1p0130
Methylobacterium extorquens (strain CM4)	B7L1N1	kdpB Mchl_0179
Methylobacterium extorquens (strain DM4)	C7C6N8	kdpB METDI0118
Methylobacterium nodulans (ORS2060)	B8I9I7	kdpB Mnod_0195
Methylobacterium radiotolerans (JCM 2831)	B1M7D5	kdpB Mrad2831_0207
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QJ47	kdpB Nham_2993
Ochrobactrum anthropi (strain ATCC 49188)	A6X5H9	kdpB Oant_3777
Oligotropha carboxidovorans (OM5)	B6JA59	kdpB
Oligotropha carboxidovorans (strain OM4)	F8BHJ2	kdpB OCA4_c03370
Rhizobium etli (strain CFN 42)	Q2K0Q4	kdpB RHE_PE00265
Rhizobium etli (strain CIAT 652)	B3Q1J8	kdpB RHECIAT_PA0000210
Rhizobium etli CNPAF512	F2AFU6	kdpB RHECNPAF_590029
Rhizobium leguminosarum bv. trifolii WSM1325	C6B6T0	kdpB Rleg_6749
Rhizobium leguminosarum bv. trifolii WSM1325	B6A3M7	kdpB Rleg2_5837
Rhizobium leguminosarum bv. viciae 3841	Q1M606	kdpB pRL110380
Rhizobium sp. (strain NGR234)	C3KRJ1	kdpB NGR_b12470
Rhizobium tropici CIAT 899	L0LV87	kdpB RTCIAT899_PC04045
Sinorhizobium medicae (strain WSM419) (Ensifer medicae)	A6UKF0	kdpB Smed_5359

Table S1: Data set, continue…

Organisms	Uniprot or RefSeq Ids	Gene name
Sinorhizobium meliloti (strain AK83)	F6ECC9	kdpB Sinme_5541
Sinorhizobium meliloti (strain BL225C)	F6BZT2	kdpB SinmeB_5342
Sinorhizobium meliloti (strain SM11)	F7XAT0	kdpB SM11_pC0041
Sinorhizobium meliloti (strain Sm2011)	M4MI42	kdpB SM2011_a2331
Sinorhizobium meliloti 1021) (Ensifer meliloti)	Q92XJ0	kdpB RA1254 SMa2331
Sinorhizobium meliloti CCNWSX0020	H0FSR0	kdpB SM0020_00865
Sinorhizobium meliloti GR4	M4IIK9	kdpB C770_GR4pC0046
Sinorhizobium meliloti Rm41	K0P8V2	kdpB BN406_03536
Xanthobacter autotrophicus Py2	A7IKC4	kdpB Xaut_3238

ZntA sequences used as outgroup in Phylogenetic analysis

Organisms	Uniprot or RefSeq IDs	Genes name
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5ESJ2	zntA BBta_7259
Methylobacterium extorquens (AM1)	C5ASD5	zntA MexAM1_META1p2601
Methylobacterium extorquens (AM1)	C5B3S9	zntA MexAM1_META2p0203
Methylobacterium extorquens (strain DM4)	C7CFT3	zntA METDI1402
Methylobacterium extorquens (strain DM4)	C7CEY1	zntA METDI1261
Oligotropha carboxidovorans (OM5)	F8C1G3	zntA OCA5_pOC16700800
Oligotropha carboxidovorans (strain OM4)	F8BR34	zntA OCA4_pOC167B00800
Oligotropha carboxidovorans (strain OM4)	F8BKP4	zntA OCA4_c16130
Sinorhizobium meliloti Rm41	K0P7P6	zntA BN406_03258

Table S2: Putative rhizobial CopZ and presence of proteins domains (X)

Strain	Uniprot or RefSeq ID	HMA domain*	CopZ domain	MBS MTC XXC	Predicted Molecular function (Uniprot)	ATPase interaction with STRING score >0.9	Protein lenght (amino acids)	Chromosome or Plasmid location
Agrobacterium sp. H13-3	F0L214_AGRSH	X	X	X	copper chaperone	F0L206_AGRSH	84	C
Agrobacterium tumefaciens C58	A9CJD6_AGRT5	X	X	X	copper chaperone	A9CJP7_AGRFC	84	C
Agrobacterium vitis S4	B9K631_AGRVS	X	X	X	copper binding protein	NO STRING DATA AVAILABLE	67	P
Aureimonas sp. AU20	WP_061934966.1	X	X	X	copper chaperone		67	data not available
Chelativorans sp. BNC1	Q11BG4_MESSB	X	X	X	metal transp/detox protein	Q11BG5_CHESB	73	C
Chelatococcus sp. CO-6	WP_050996748	X	X	X	copper chaperone		67	C
Ensifer adhaerens OV14	WP_025428584	X	X	X	copper chaperone		65	data not available
Hoeflea sp. IMCC20628	WP_047033070.1	X	X	X	copper chaperone		68	P
Methylobacterium extorquens AM1	C5AS47_METEA	X	X	X	copper chaperone	C5B1T2_METEA, C5ASD8_METEA, C5AUM6_METEA	73	C
Methylobacterium extorquens AM1	C5ASF2_METEA	X	X	X	metal transp/detox protein	C5B1T2_METEA, C5ASD8_METEA, C5AUM6_METEA	85	C
Methylobacterium extorquens AM1	C5ASF4_METEA	X	X	X	copper chaperone	NO INTERACTION	114	C
Methylobacterium extorquens AM1	C5ASF6_METEA	X	X	X	copper binding protein	NO INTERACTION	78	C
Methylobacterium extorquens AM1	C5ASG3_METEA	X	X	X	mercury transporter	NO INTERACTION	112	C
Methylobacterium extorquens AM1	C5B3J8_METEA	X	X	X	copper chaperone	NO STRING DATA AVAILABLE	103	P
Methylobacterium extorquens AM1	C5B3U0_METEA	X	X	X	copper binding protein	NO STRING DATA AVAILABLE	70	P
Methylobacterium extorquens AM1	C7C9H7_METED	X	X	X	copper binding protein	NO STRING DATA AVAILABLE	70	C

Table S2: Continue

Strain	Uniprot or RefSeq ID	HMA domain*	CopZ domain	MBS MTC XXC	Predicted Molecular function (Uniprot)	ATPase interaction with STRING score >0.9	Protein lenght (amino acids)	Chromosome or Plasmid location
Methylobacterium extorquens CM4	B7KNH1_METC4	X	X	X	copper chaperone	NO STRING DATA AVAILABLE	86	C
Methylobacterium extorquens CM4	B7KPC2_METC4	X	X	X	copper chaperone	NO STRING DATA AVAILABLE	70	C
Methylobacterium extorquens CM4	B7L0A6_METC4	X	X	X	metal transp/detox protein	NO STRING DATA AVAILABLE	73	C
Methylobacterium extorquens DM4	C7CEZ3_METED	X	X	X	copper binding protein	NO STRING DATA AVAILABLE	70	C
Methylobacterium nodulans ORS2060	B8IF32_METNO	X	X	X	copper chaperone	B8IGD1_METNO	80	C
Methylobacterium nodulans ORS2060	B8IXS0_METNO	X	X	X	metal transp/detox protein	NO STRING DATA AVAILABLE	69	P
Methylobacterium populi BJ001	B1ZBA1_METPB	X	X	X	copper chaperone	NO STRING DATA AVAILABLE	88	C
Methylobacterium populi BJ001	B1ZF58_METPB	X	X	X	metal transp/detox protein	NO STRING DATA AVAILABLE	70	C
Methylobacterium populi BJ001	B1ZJ89_METPB	X	X	X	metal transp/detox protein	NO STRING DATA AVAILABLE	73	C
Methylobacterium radiotolerans JCM 2831	B1M038_METRJ	X	X	X	metal transp/detox protein	B1LUX6_METRJ	70	C
Methylobacterium radiotolerans JCM 2831	B1M184_METRJ	X	X	X	metal transp/detox protein	B1LUX6_METRJ	73	C
Methylobacterium sp. 4-46	B0U7P5_METS4	X	X	X	copper chaperone	B0UE71_METS4, B0UQ23_METS4, B0UGH1_METS4	81	C
Neorhizobium galegae bv. officinalis bv. officinalis str. HAMBI 1141	WP_038541227.1	X	X	X	copper chaperone		68	C1
Nitrobacter hamburgensis X14	Q1QFI9_NITHX	X	X	X	copper chaperone	NO STRING DATA AVAILABLE	71	P
Nitrobacter hamburgensis X14	Q1QFJ6_NITHX	X	X	X	mercury transporter	NO STRING DATA AVAILABLE	94	P
Ochrobactrum anthropi DSM 6882	A6X157_OCHA4	X	X	X	metal transp/detox protein	A6X3Y7_OCHA4, A6X3Z4_OCHA4	65	C
Ochrobactrum anthropi DSM 6882	A6X3R2_OCHA4	X	X	X	mercury transporter	NO INTERACTION	97	C

Table S2: Continue

Strain	Uniprot or RefSeq ID	HMA domain*	CopZ domain	MBS MTC XXC	Predicted Molecular function (Uniprot)	ATPase interaction with STRING score >0.9	Protein lenght (amino acids)	Chromosome or Plasmid location
Ochrobactrum anthropi DSM 6882	A6X3Y2_OCHA4	X	X	X	metal transp/detox protein	NO INTERACTION	71	C
Ochrobactrum anthropi DSM 6882	A6X3Z5_OCHA4	X	X	X	metal transp/detox protein	NO INTERACTION	64	C
Ochrobactrum anthropi DSM 6882	A6X699_OCHA4	X	X	X	copper chaperone	NO STRING DATA AVAILABLE	64	C
Oligotropha carboxidovorans OM4	B6JKB8_OLICO	X	X	–	mercury transporter		94	P
Oligotropha carboxidovorans OM4	B6JKB2_OLICO	X	X	X	copper binding protein		71	P
Oligotropha carboxidovorans OM4	B6JJM8_OLICO	X	X	X	metal transp/detox protein		64	C
Oligotropha carboxidovorans OM4	B6JDU0_OLICO	X	X	–	metal transp/detox protein	F8BPQ7_OLICM	75	C
Oligotropha carboxidovorans OM5	B6JDU0_OLICO	X	X	–	metal transp/detox protein		75	C
Oligotropha carboxidovorans OM5	B6JJM8_OLICO	X	X	X	metal transp/detox protein		64	C
Oligotropha carboxidovorans OM5	B6JKB2_OLICO	X	X	X	metal transp/detox protein		71	P
Oligotropha carboxidovorans OM5	B6JKB8_OLICO	X	X	X	mercury transporter		94	P
Pelagibacterium halotolerans B2	G4RGC3_PELHB	X	X	X	copper binding protein		69	C
Shinella sp. HZN7	A0A021WZD7	X	X	X	copper chaperone		71	data not available
Sinorhizobium fredii NGR234	C3KQE6_RHISN	X	X	X	copper chaperone		73	P
Sinorhizobium medicae WSM419	A6UJH6_SINMW	X	X	X	Heavy Metal Binding Protein		73	P
Sinorhizobium meliloti SM11	F7XH69_SINMM	X	X	X	copper binding protein		73	P
Sinorhizobium meliloti 1021	Q92W08_RHIME	X	X	X	copper chaperone		73	P
Sinorhizobium meliloti 1021	Q92ZE1_RHIME	X	X	X	copper chaperone		73	P
Sinorhizobium meliloti AK83	F6EA90_SINMK	X	X	X	copper binding protein		73	P
Sinorhizobium meliloti BL225C	F6C1B9_SINMB	X	X	X	copper chaperone		73	P

Table S2: Continue

Strain	Uniprot or RefSeq ID	HMA domain*	CopZ domain	MBS MTC XXC	Predicted Molecular function (Uniprot)	ATPase interaction with STRING score >0.9	Protein lenght (amino acids)	Chromosome or Plasmid location
Sinorhizobium meliloti BL225C	F6C3W6_SINMB	X	X	X	copper binding protein		73	P
Sinorhizobium meliloti CCNWSX0020	H0FV63_RHIML	X	X	X	copper chaperone		73	data not available
Sinorhizobium meliloti CCNWSX0020	H0FYK0_RHIML	X	X	–	metal binding protein		91	data not available
Sinorhizobium meliloti SM11	F7XFT9_SINMM	X	X	X	copper binding protein		73	P
Xanthobacter autotrophicus Py2	A7IDF6_XANP2	X	X	X	mercury transporter		98	C
Xanthobacter autotrophicus Py2	A7IJE6_XANP2	X	X	X	mercury transporter		98	C

Table S3: Putative rhizobial CusF and prescence of motifs (X)

Strains	Uniprot or RefSeq ID	Protein lenght (aa)	CusF_ Ec family	MBS H36 M47 M49	Residues for interactions K23 K30 K31 H3 R50	Replicon
Chelatococcus sp. CO-6	WP_019404152	96	X	HMM	K K, H	C
Ensifer adhaerens OV14	WP_025428575.1	90	X	HMM	K, K, K, H	C
Escherichia coli K12	P77214	110	X	HMM	K, K, K, H, R	C
Hoeflea sp. IMCC20628	WP_047031981	95	X	HMM	K, K, K, H	C
Methylobacterium extorquens AM1	C5B3M6	234	X	HMM	K, K, K,	P
Methylobacterium extorquens AM1	C5B3M5	94	X	HMM	K, K, K, H	P
Methylobacterium extorquens CM4	B7KNJ5	234	X	HMM	K, K, K,	C
Methylobacterium extorquens CM4	B7KNJ4	94	X	HMM	K, K, K, H	C
Methylobacterium extorquens DM4	C7CFP8	94	X	HMM	K, K, K, H	C
Methylobacterium nodulans ORS2060	B8IXR4	94	X	HMM	K, K, K, H	P
Methylobacterium populi BJ001	B1ZBD9	236	X	HMM	K, K, K, H, R	C
Methylobacterium populi BJ001	B1ZBE3	94	X	HMM	K, K, K, H	C
Neorhizobium galegae bv. officinalis bv. officinalis str. HAMBI 1141	WP_038543460	94	X	HMM	K, K, K,	C1
Nitrobacter hamburgensis X14	Q1QFL0	132	X	HMM	K, K, K,	P
Oligotropha carboxidovorans OM5	B6JJX4	135	X	HMM	K, K, K, H	C
Oligotropha carboxidovorans OM5	F8C1B8	136	X	HMM	K, K, K, H	P
Sinorhizobium meliloti 1021	Q92S41	94	X	X	K, K, K, R	C
Sinorhizobium meliloti AK83	F6DYU1	94	X	X	K, K, K, R	C1
Sinorhizobium meliloti BL225C	F6BRD9	94	X	X	K, K, K, R	C
Sinorhizobium meliloti GR4	M4I8A9	94	X	X	K, K, K, R	C
Sinorhizobium meliloti Rm41	K0P3C0	94	X	X	K, K, K, R	C
Sinorhizobium meliloti SM11	F7X798	94	X	X	K, K, K, R	C
Sinorhizobium meliloti Sm2011	M4MY42	94	X	X	K, K, K, R	C
Xanthobacter autotrophicus Py2	A7IQ12	104	X	HMM	K, K, K, H	P
Xanthobacter autotrophicus Py2	A7IP79	94	X	HMM	K, K, K, H	C
Xanthobacter autotrophicus Py2	A7IPT9	94	X	HMM	K, K, K, H	P

Table S4: Putative rhizobial CusA homologues and their respective CopA partner. Presence (X) and absence (-) of motifs

Strain	Putative Rhizobial CupA uniprot ID	Cupredoxin_1 domain	Putative Rhizobial CopA	Cupredoxin_1 domain
Streptococcus pneumoniae D39	Q04LG8_STRP2	X	Q04LG7_STRP2	X
			A5MDC7_STREE	X
			M5KA62_STREE	X
Bradyrhizobium sp. BTAi1	A5EG08_BRASB	X	A5E8K9_BRASB	--
Bradyrhizobium sp. BTAi1	A5ET48_BRASB	X	A5EUA0_BRASB	--
			A5EFI6_BRASB	--
Beijerinckia indica DSM 1715	B2IDE2_BEII9	X	B2IDM6_BEII9	--
Beijerinckia indica DSM 1715	B2IE20_BEII9	X	B2IJD3_BEII9	--
Ensifer adhaerens OV14	OV14_a1198	X	OV14_3156	--
			OV14_a0399	--
			OV14_b0747	--
			OV14_0962	--
Mesorhizobium loti MAFF303099	Q983W8_RHILO	X	Q98C24_RHILO	--
Methylobacterium sp. 4-46	B0UGC2_METS4	X	B0UGH1_METS4	--
			B0UQ23_METS4	--
Nitrobacter hamburgensis DSM 10229	Q1QRN2_NITHX	X	Q1QFM6_NITHX	--
			Q1QF08_NITHX	--
			Q1QHY3_NITHX	--
Ochrobactrum anthropi DSM 6882	A6X5Q9_OCHA4	X	A6WW40_OCHA4	--
			A6X3Z4_OCHA4	--
			A6WVL4_OCHA4	--
			A6X3Y7_OCHA4	--

Table S5: Cu-ATPases groups and subtypes predicted from ML phylogeny

Group VIII /P_{1B-1}- Cu+ -ATPases subtype			Prediction of TMH		
STRAINS	Uniprot IDs	LENGHT (aa)	DAS TM filter	TMH MM v2	REPLICON
Acidithiobacillus ferrooxidans (strain ATCC 53993)	B5EJX7	811	nd	8	Ch
Agrobacterium sp. (strain H13-3)	F0L206	834	nd	8	Ch
Agrobacterium tumefaciens (strain C58 / ATCC 33970)	A9CJE3	836	nd	8	Ch
Agrobacterium tumefaciens C58	A9CJP7	841	nd	8	Ch
Agrobacterium vitis strain S4	B9K630	826	nd	8	P
Agrobacterium vitis strain S4	B9JWS2	819	nd	8	Ch
Aureimonas sp. AU20	M673_01920	825	nd	8	Ch
Beijerinckia indica subsp. indica (DSM 1715)	B2IJD3	857	nd	8	Ch
Bosea sp. PAMC 26642	AXW83_22960	846	nd	8	Ch
Brucella abortus bv. 1 str. 9-941	BruAb1_0215	759	nd	8	Ch1
Brucella canis ATCC 23365	BCAN_A0223	826	nd	8	Ch1
Brucella canis ATCC 23365	BCAN_A0365	752	nd	8	Ch1
Brucella melitensis bv. 1 str. 16M	DK63_1757	826	nd	8	Ch1
Brucella ovis ATCC 25840	BOV_0212	759	nd	8	Ch1
Brucella suis 1330	WP_006191273.1	826	nd	8	Ch1
Chelativorans sp. (strain BNC1)	Q11BG5	855	nd	8	Ch
Chelatococcus sp. CO-6	AL346_08710	838	nd	8	Ch
Ensifer adhaerens OV14	OV14_a0399	793	nd	8	Ch2
Hoeflea sp. IMCC20628	IMCC20628_04836	826	nd	8	P
Martelella sp. AD-3	AZF01_01065	825	nd	8	Ch
Methylobacterium extorquens DM4	C7C8F4	832	nd	8	Ch
Methylobacterium extorquens (AM1)	C5AUM6	832	nd	8	Ch
Methylobacterium extorquens CM4	B7KQJ2	832	nd	8	Ch
Methylobacterium nodulans (strain ORS2060)	B8IXP0	835	nd	8	P
Methylobacterium populi (BJ001)	B1ZBA0	838	nd	8	Ch
Methylobacterium populi (strain BJ001)	B1ZDQ2	830	nd	8	Ch
Methylobacterium radiotolerans (JCM 2831)	B1LUX6	824	nd	8	Ch
Methylobacterium sp. (strain 4-46)	B0UQ23	825	nd	8	Ch
Neorhizobium galegae bv. officinalis str. HAMBI 1141	RG1141_CH08380	801	nd	8	Ch1
Ochrobactrum anthropi DSM 6882	A6X3Z4	852	nd	8	Ch
Ochrobactrum anthropi (strain ATCC 49188)	A6WVL4	827	nd	8	Ch2
Oligotropha carboxidovorans (DSM 1227 / OM5)	B6JJM9	822	nd	8	Ch
Oligotropha carboxidovorans (strain OM4)	actP OCA4_c06180	822	nd	8	Ch
Pelagibacterium halotolerans JCM 15775	G4REK2	831	nd	8	Ch
Pelagibacterium halotolerans (strain JCM 15775)	G4RDU4	845	nd	8	Ch
Rhizobium leguminosarum bv. viciae (strain 3841)	Q1M656	824	nd	8	P
Rhizobium sp. (strain NGR234)	C3KQ40	840	nd	8	P

Table S5: Continue

Rhizobium sp. (strain NGR234)	C3KRI8	830	nd	8	P
Shinella sp. HZN7	shn_02380	852	nd	8	Ch1
Shinella sp. HZN7	shn_29395	834	nd	8	P
Sinorhizobium medicae (strain WSM419)	Q9X5X3	827	nd	8	P
Sinorhizobium meliloti 1021 (Ensifer meliloti)	P58342	827	nd	8	P
Sinorhizobium meliloti (strain 1021)	P58341	826	nd	8	P
Sinorhizobium meliloti (strain 2011)(CopA1a)	M4MJF5	826	nd	8	P
Sinorhizobium meliloti (strain BL225C)	SinmeB_3653	827	nd	8	P
Sinorhizobium meliloti (strain SM11)	F7XFT8	827	nd	8	P
Sinorhizobium meliloti (strain 2011)(CopA1b)	M4MUQ3	827	nd	8	P
Sinorhizobium meliloti GR4	C770_GR4pA158 actP2	828	nd	8	P
Sinorhizobium meliloti Rm41	BN406_05647	827	nd	8	P
Thermus thermophilus HB27	Q72HX4	798	nd	8	Ch
TOTAL					P16, Ch34

Group VI / P_{1B-3}-Cu^{+}-ATPase subtype

Via (PIB-3)	**Uniprot IDs**	**LENGHT (aa)**	**DAS TM filter**	**TMH MM v2**	**REPLICON**
Staphylococcus aureus CopB	Q69HU0	672	nd	8	P
Enterococcus hirae (strain ATCC 9790) CopB	P05425	745	nd	8	Ch
Enterococcus faecium (Streptococcus faecium)	Q8VPE6	710	nd	8	P
Escherichia coli HRA-1	Q59370	721	nd	8	Ch
Escherichia coli HRA-2	Q59369	731	nd	8	Ch
Vib (PIB-3)					
Archaeoglobus fulgidus (strain ATCC 49558) CopB	O30085	690	nd	8	Ch
Chelativorans sp. (strain BNC1)	Q11BI2	703	nd	8	Ch
Chelatococcus sp. CO-6	AL346_10545	759	nd	8	Ch
Hoeflea sp. IMCC20628	IMCC20628_04815	808	nd	8	P
Pelagibacterium halotolerans (strain JCM 15775)	G4RDU0	687	nd	8	Ch
Thermus thermophilus HB27	Q72HW1	687	nd	8	Ch
Vic (PIB-3a)					
Xanthobacter autotrophicus (strain Py2)	A7IPW3	818	nd	8	P
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QFP2	811	nd	8	P
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EUF6	813	nd	8	P
TOTAL					P6,Ch8

Group VII / P_{1B-1a}-Cu^{+}-ATPase subtype	**Uniprot IDs**	**LENGHT (aa)**	**DAS TM filter**	**TMH MM v2**	**REPLICON**	
Agrobacterium vitis (strain S4)	B9K626	737	nd	8	P	CPIC
Beijerinckia indica subsp. indica (strain DSM 1715)	B2IDM6	794	nd	8	Ch	H-Rich CPIC
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EUA0	804	nd	8	P	H-Rich CPIC
Bradyrhizobium sp. (strain BTAi1)	A5E8K9	821	nd	8	Ch	CPIC
Bradyrhizobium sp. (strain BTAi1)	RPA1660	821	ND	8	Ch	CPIC
Chelatococcus sp. CO-6	AL346_04970	796	nd	8	Ch	CPIC

Table S5: Continue

Legionella pneumophila (strain Philadelphia 1 DSM 7513)	Q5ZWR1	736	nd	8	Ch	H-Rich CPLC
E coli	Q59385	834	nd	8	Ch	CASC
Methylobacterium extorquens (strain AM1)	C5B3T2	806	nd	8	P	H-Rich CPYC
Methylobacterium extorquens (strain AM1)	C5B3I7	809	nd	8	P	H-Rich CPIC
Methylobacterium extorquens (strain AM1)	C5B1T2	760	nd	8	Ch	CPIC
Methylobacterium extorquens (strain CM4)	B7KPB4	806	nd	8	Ch	H-Rich CPYC
Methylobacterium extorquens (strain CM4)	B7KNG4	809	nd	8	Ch	H-Rich CPIC
Methylobacterium extorquens (strain DM4)	C7CEY4	806	nd	8	Ch	H-Rich CPYC
Methylobacterium extorquens DM4	C7CF45	809	nd	8	Ch	H-Rich CPIC
Methylobacterium nodulans (strain ORS2060)	B8IGD1	823	nd	8	Ch	H-Rich CPIC
Methylobacterium nodulans (strain ORS2060)	B8IXS2	788	nd	8	P	H-Rich CPIC
Methylobacterium populi (strain BJ001)	B1ZCJ6	799	nd	8	Ch	H-Rich CPIC
Methylobacterium populi (strain BJ001)	B1ZBB5	786	nd	8	Ch	H-Rich CPIC
Methylobacterium sp. (strain 4-46)	B0UE71	755	nd	8	Ch	CPIC
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QF08	801	nd	8	P	H-Rich CPIC
Nitrobacter hamburgensis (strain X14 / DSM 10229)	Q1QFM6	818	nd	8	P	H-Rich CPIC
Nitrobacter hamburgensis (strain X14)	Q1QHY3	833	nd	8	Ch	H-Rich CPIC
Ochrobactrum anthropi (strainDSM)	A6X3Y7	795	nd	8	Ch	H-Rich CPIC
Oligotropha carboxidovorans (strain OM4)	OCA4_pOC167B00140	813	nd	8	P	CPIC
Oligotropha carboxidovorans (strain OM4)	OCA4_c05040	815	nd	8	Ch	CPIC
Oligotropha carboxidovorans (strain OM5)	F8C198	813	nd	8	P	CPIC
Oligotropha carboxidovorans (strain OM5)	F8BVV1	815	nd	8	Ch	CPIC
Rhizobium leguminosarum bv. viciae (strain 3841)	Q1MI24	740	nd	8	Ch	H-Rich CPIC
Rhizobium sp. (strain NGR234)	C3KQ24	746	nd	8	P	H-Rich CPIC
Rhodopseudomonas palustris CGA009	RPA1660	923	nd	8	Ch	CPIC
Salmonella typhimurium	Q9ZHC7	824	nd	8	P	H-Rich CPVC
Salmonella typhimurium	Q8ZR95	833	nd	8	Ch	CASC
Shewanella oneidensis	Q8EGB6	753	nd	8	Ch	CAGC
Shinella sp. HZN7	shn_02340	795	nd	8	Ch	H-Rich CPIC
Sinorhizobium meliloti GR4	AGA09593	705	nd	8	p	CPIC
Sinorhizobium meliloti (strain 1021)	Q92ZA2	733	nd	8	P	H-Rich CPIC
Sinorhizobium meliloti (strain SM11)	Q1WLB3	706	nd	8	P	CPIC
Sinorhizobium meliloti 2011)	M4MG41	733	nd	8	P	CPIC
Sinorhizobium meliloti CCNWSX0020	H0FVL4	708	nd	8	Ch	CPIC
Sinorhizobium meliloti GR4	C770_GR4pC0893	705	nd	8	P	CPIC
Sinorhizobium meliloti GR4	C770_GR4pA165	740	nd	8	P	CPIC
Xanthobacter autotrophicus (Py2)	A7IP64	1020	nd	8	Ch	H-Rich CPIC
Xanthobacter autotrophicus (Py2)	A7IJX8	868	nd	8	Ch	H-Rich CPIC
Xanthobacter autotrophicus (strain Py2)	A7IPV0	796	nd	8	P	H-Rich CPIC
Xanthobacter autotrophicus (strain Py2)	A7IQ22	793	nd	8	P	H-Rich CPIC
					P19, Ch27	
Group VII / P_{1B-1b}-Cu^{+}-ATPase subtype						
Rhizobium leguminosarum bv. trifolii WSM2304	B6A445	841	nd	8	P	

Table S5: Continue

Rhizobium etli CFN 42	Q2K000	840	nd	8	P
Agrobacterium radiobacter (strain K84)	B9JDB0	830	nd	8	Ch
Rhizobium etli (strain CIAT 652)	B3Q0Z7	839	nd	8	P
Mesorhizobium loti (strain MAFF303099)	Q98C24	839	nd	8	P
Rhizobium leguminosarum bv. viciae WSM710	Q9X5V3	841	nd	8	P
Mesorhizobium ciceri bv. biserrulae (WSM1271)	E8T785	834	nd	8	Ch
Chelativorans sp. (strain BNC1)	Q11EY1	846	nd	8	Ch
Rhizobium leguminosarum bv. viciae 3841	Q1M727	841	nd	8	P
Rhizobium leguminosarum bv. trifolii (WSM1325)	C6B7V6	841	nd	8	P
Rhizobium tropici CIAT 899	L0LNP7	832	nd	8	Ch
Hoeflea sp. IMCC20628	IMCC20628_04449	801	nd	8	Ch
Hoeflea sp. IMCC20628	IMCC20628_04834	796	nd	8	P
Bosea sp. PAMC 26642	AXW83_04575	722	nd	8	Ch
Martelella sp. AD-3	AZF01_23265	781	nd	8	P
					9P/6Ch
Group XIV FixI/ $P_{1B\text{-}1c}$-Cu^{+}-ATPase subtype					
Agrobacterium vitis	B9JWS5	787	nd	8	Ch1
Bradyrhizobium diazoefficiens (strain JCM 10833 / USDA 110)	Q59207	730	7	7	Ch
Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182) FixI	A5EFI6	733	7	7	Ch
Brucella abortus bv1 16M	DK63_1923	752	ND	8	Ch1
Brucella abortus bv1 9-941	BruAb1_0383	752	nd	8	Ch1
Chelatococcus sp. CO-6	AL346_02020	638	8	8	Ch
Chelatococcus sp. CO-6	AL346_11220	753	nd	7	Ch
Mesorhizobium ciceri	E8TP15	759	nd	7	P
Mesorhizobium ciceri	E8TEI6	761	nd	7	Ch
Mesorhizobium loti MAFF	Q88R8	762	nd	8	Ch
Mesorhizobium loti MAFF	Q989H6	762	nd	8	Ch
Methylobacterium nodulans (strain ORS2060)	B8IN84	758	9	7	Ch
Methylobacterium sp. (strain 4-46)	B0UGH1	758	8	8	Ch
Neorhizobium galegae bv. officinalis bv. str. HAMBI 1141	RG1141_CH40880	763	8	8	Ch1
Ochrobactrum anthropi (strain ATCC 49188)	A6WW40	752	8	8	C1
Ochrobactrum anthropi (strain ATCC 49188)	A6WW40	752	nd	8	Ch1
Oligotropha carboxidovorans OM5	F8BRK8	753	7	8	Ch
Oligotropha carboxidovorans OM5	F8C141	748	8	8	P
Oligotropha carboxidovorans OM5	F8BRK8	753	nd	8	Ch
Oligotropha carboxidovorans OM5	F8C141	748	nd	8	P
Rhizobium etli (strain CFN 42) FixId	Q8KLH2	756	8	8	P
Rhizobium etli (strain CFN 42) FixIf	Q2JYE6	761	8	8	P
Rhizobium etli CFN42	Q8KLH2	756	nd	5	p
Rhizobium etli CIAT 652	B3Q2X3	760	nd	6	P
Rhizobium etli CNPAF 512	F2AA47	760	nd	6	draft
Rhizobium leguminosarum bv. Trifolii SM 2304	B6A0W1	762	nd	6	P

Table S5: Continue

Rhizobium leguminosarum bv. viciae (strain 3841) FixI	Q1MLX8	755	7	8	Ch
Rhizobium leguminosarum bv. viciae 3841	Q1MLX8	755	7	7	Ch
Rhizobium leguminosarum bv. viciae 3841	Q1M8X6	740	8	8	P
Rhizobium leguminosarum bv. viciae 3841	C6B514	740	8	8	Ch
Rhizobium leguminosarum bv. viciae 3841	Q1M7U4	761	8	8	P
Rhizobium tropici CIAT 899	L0LTL0	761	nd	6	P
Rhodopseudomonas palustris CGA009 FixI	RPA0013	732	7	7	Ch
Sinohizobium meliloti (strain 1021) (Ensifer meliloti)	P18398	757	8	8	P
Sinorhizobium fredii (strain NGR234)	C3MH46	772	8	8	Ch
Sinorhizobium fredii NGR234	C3MH46	772	8	8	Ch
Sinorhizobium fredii NGR234	C3MDN6	754	6	6	Ch
Sinorhizobium medicae (strain WSM419)	A6ULT2	755	7	7	P
Sinorhizobium medicae (strain WSM419)	A6ULU1	757	8	8	P
Sinorhizobium meliloti (strain 1021)	Q92ZX3	755	7	6	P
Sinorhizobium meliloti (strain BL225C)	F6BZ85	755	7	6	P
Sinorhizobium meliloti (strain BL225C)	SinmeB_5734	755	7	6	P
Sinorhizobium meliloti (strain BL225C)	F6BYN0	757	8	8	P
Sinorhizobium meliloti (strain SM11)	F7XB41	755	7	6	P
Sinorhizobium meliloti (strain SM11)	F7XES0	759	8	7	P
Sinorhizobium meliloti (strain Sm2011)	M4MNI0	755	7	6	P
Sinorhizobium meliloti (strain Sm2011)	M4MJQ1	757	8	8	P
Sinorhizobium meliloti CCNWSX0020	H0FVA9	757	8	6	Ch
Sinorhizobium meliloti GR4	C770_GR4pC1192	755	7	6	P
Sinorhizobium meliloti GR4	M4IN78	757	8	8	P
Sinorhizobium meliloti Rm41	K0PAQ5	1025	11	11	P
Sinorhizobium meliloti Rm41	K0P588	757	8	8	P
Sinorhizobium meliloti SM11	F7XDP	737	nd	8	P
Xanthobacter autotrophicus (strain ATCC BAA-1158 / Py2)	A7ICI0	743	7	7	Ch
TOTAL					P27, Ch29

DAS TM filter Transmembrane protein prediction. Cserzo et al 2004. Bioinformatics 20:136.

TMHMM v2 Prediction of transmembrane helices in proteins http://www.cbs.dtu.dk/services/TMHMM/

nd Not determined

HMA heavy metal associated domain

Table S6: Rhizobial Cu-ATPases and their closest non-rhizobial

NUMBER	RefSeq or UNIPROT ID	REPLICON	ATPase subtype	RHIZOBIAL STRAIN	BlastP Closest CopA homologs	(% of identity/similarity/coverage/e-value)	LOCUS	PHYLUM/ORDER
1	G4REK2	Ch	P1B-1	*Pelagibacterium halotolerans B2*	*Serratia fonticola*	(64/76/99/0)	WP_021806406	Gammaproteobacteria; Enterobacteriales
2	A9CJP7	Ch	P1B-1	*Agrobacterium fabrum str. C58*	*Serratia fonticola*	(60/73/98/0)	WP_021806406	Gammaproteobacteria; Enterobacteriales
3	B1ZBA0	Ch	P1B-1	*Methylobacterium populi BJ001*	*Halomonas anticariensis*	(65/77/ 98/0)	WP_016415551	Gammaproteobacteria; Oceanospirillales
4	C7C8F4	Ch	P1B-1	*Methylobacterium extorquens DM4*	*Serratia liquefaciens ATCC 27592*	(63/75/97/0)	YP_008230540	Gammaproteobacteria; Enterobacteriales
5	C5AUM6	Ch	P1B-1	*Methylobacterium extorquens AM1*	*Serratia liquefaciens ATCC 27592*	(63/75/97/0)	YP_008230540	Gammaproteobacteria; Enterobacteriales
6	B7KQJ2	Ch	P1B-1	*Methylobacterium extorquens CM4*	*Stenotrophomonas maltophilia*	(65/76/96/0)	WP_019184655	Gammaproteobacteria; Xanthomonadales
7	B1ZDQ2	Ch	P1B-1	*Methylobacterium populi BJ001*	*Serratia liquefaciens*	(62/75/97/0)	YP_008230540	Gammaproteobacteria; Enterobacteriales
8	B1LUX6	Ch	P1B-1	*Methylobacterium radiotolerans JCM 2831*	*Stenotrophomonas maltophilia*	(63/73/ 97/0)	WP_019184655	Gammaproteobacteria; Xanthomonadales
9	B0UQ23	Ch	P1B-1	*Methylobacterium sp. 4-46*	Serratia liquefacien	(62/73/ 99/0)	YP_008230540	Gammaproteobacteria; Enterobacteriales
10	C5B3T2	P	P1B-1a	*Methylobacterium extorquens AM1*	*Polymorphum gilvum SL003B-26A1*	(78/87/93/0)	YP_004305520	Unclassified Alphaproteobacteria
11	C7CEY4	Ch	P1B-1a	*Methylobacterium extorquens DM4*	*Sphingobium*	(74/84/95/0)	YP_004305520	Alphaproteobacteria; Polymorphum
12	C5ASD8	Ch	P1B-1a	*Methylobacterium extorquens AM1*	Rubellimicrobium thermophilum	(78/86/98/0)	WP_021099112	Alphaproteobacteria; Rhodobacterales

Table S6: Continue

NUMBER	RefSeq or UNIPROT ID	REPLICON	ATPase subtype	RHIZOBIAL STRAIN	BlastP Closest CopA homologs	(% of identity/similarity/coverage/e-value)	LOCUS	PHYLUM/ORDER
13	C5B1T2	Ch	P1B-1a	*Methylobacterium extorquens AM1*	*Rubellimicrobium thermophilum*	(79/89/99/0)	WP_021099112	Alphaproteobacteria; Rhodobacterales
14	C7CF45	Ch	P1B-1a	*Methylobacterium extorquens DM4*	*Rubellimicrobium thermophilum*	(74/84/99/0)	WP_021099112	Alphaproteobacteria; Rhodobacterales
15	A6X3Y7	Ch	P1B-1a	*Ochrobactrum anthropi ATCC 49188*	*Rubellimicrobium thermophilum*	(91/92/100/0)	WP_021099112	Alphaproteobacteria; Rhodobacterales
16	Q11BI2	Ch	P1B-3	*Chelativorans sp. BNC1*	*Pseudoxanthomonas spadix BD-a59*	(86/91/95/0)	YP_004929459.1	Gammaproteobacteria; Xanthomonadales
17	B0UGH1	Ch	P1B-1c	*Methylobacterium sp. 4-46*	*Polymorphum gilvum*	(59/72/ 94/0)	YP_004304932	Alphaproteobacteria; Polymorphum
18	B8IN84	Ch	P1B-1c	*Methylobacterium nodulans ORS 2060*	*Polymorphum gilvum SL003B*	(59/72/94/0)	YP_004304932	Alphaproteobacteria; Polymorphum
19	Q11G30	Ch	P1B-1	*Chelativorans sp. BNC1*	*Maricaulis sp. JL2009*	(61/74/98/0)	WP_022697385.1	Alphaproteobacteria; Rhodobacterales
20	F8BKG9	Ch	P1B-1	*Oligotropha carboxidovorans OM4*	*Nitrosomonas eutropha*	(81/89/100/0)	YP_746320	Betaproteobacteria; Nitrosomonadales
21	Q11BG5	Ch	P1B-1	*Chelativorans sp. BNC1*	*Nitrosomonas eutropha C91*	(77/86/ 97/0)	YP_746320	Betaproteobacteria; Nitrosomonadales
22	B9JJ75	Ch	P1B-1c	*Agrobacterium radiobacter K84*	*Pseudomonas avellanae*	(75/85/97/ 0)	WP_005621940	Gammaproteobacteria; Pseudomonadales
23	M673_01920	Ch	P1B-1	*Aureimonas sp. AU20*	*Bordetella bronchiseptica*	61/73/98/0	WP_061973159.1	Betaproteobacteria
24	M673_21640	P		*Aureimonas sp. AU20*	*Brevundimonas diminuta*	72/81/97/0	VP_003165014.1	Alphaproteobacteria; Caulobacterales
25	AXW83_04575	Ch	P1B-1	*Bosea sp. PAMC 26642*	*Roseomonas*	71/84/100/0	WP_019460575.1	*Alphaproteobacteria Rhodospirillales*
26	IMCC20628_04449	CH	P1B-1	*Hoeflea sp. IMCC20628*	*Ahrensia.*	83/91/96/0	WP_051540975	*Alphaproteobacteria Rhodobacterales*
27	IMCC20628_04629	P	P1B-1	*Hoeflea sp. IMCC20628*	*Rhodothermaceae bacterium*	65/76/98/0	WP_068126289	*Bacteroidetes Order II*

Table S6: Continue

NUMBER	RefSeq or UNIPROT ID	REPLICON	ATPase subtype	RHIZOBIAL STRAIN	BlastP Closest CopA homologs	(% of identity/similarity/coverage/e-value)	LOCUS	PHYLUM/ORDER
28	IMCC20628_04815	P	P1B-1	*Hoeflea sp. IMCC20628*	*Thiobacillus denitrificans*	81/90/79/0	WP_018078629	Betaproteobacteria
29	IMCC20628_04834	P	P1B-1	*Hoeflea sp. IMCC20628*	*Citreimonas salinaria*	79/87/99/0	SDY81891	Alphaproteobacteria Rhodobacterales
30	AZF01_01065	Ch	P1B-1	*Martelella sp. AD-3*	*Serratia liquefaciens*	62/76/99/0	WP_048758721	Gammaproteobacteria

Table S7: Characteristics of rhizobial Cu-ATPases encoding genes potentially acquired by horizontal transfer

NUMBER	UNIPROT ID	REPLICON	ATPase subtype	RHIZOBIAL STRAIN	%G+C copA/genomic [1,2,3]	CAI[4]	E-CAI[4] p<0.05	CAI/E-CAI[4] p<0.05	E-CAI[4] p<0.01	CAI/E-CAI[4] p<0.01	Genes located into putatve genomic islands (prediction method)[5]	Genomic context[6]
1	G4REK2	Ch	P1B-1	*Pelagibacterium halotolerans B2*	63.06/61.4	No codon usage available at Codon Usage Database						gene context without functional relation
2	A9CJP7	Ch	P1B-1	*Agrobacterium fabrum str. C58*	63.10/59.1	0.8	0.7	1.1	0.7	1		20 Kb gene cluster encoding type III secretion system and phage packaging proteins
3	B1ZBA0	Ch	P1B-1	*Methylobacterium populi BJ001*	72.86/69.4	0.7	0.762	**0.9**	0.8	**0.9**		gene context without functional relation
4	C7C8F4	Ch	P1B-1	*Methylobacterium extorquens DM4*	73.23/68	0.7	0.7	**0.9**	0.781	**0.9**		gene context without functional relation
5	C5AUM6	Ch	P1B-1	*Methylobacterium extorquens AM1*	73.71/68.5	0.7	0.7	**1**	0.8	**0.9**		gene context without functional relation
6	B7KQJ2	Ch	P1B-1	*Methylobacterium extorquens CM4*	73.07/68.1	0.7	0.7	**1**	0.8	**0.9**		gene context without functional relation
7	B1ZDQ2	Ch	P1B-1	*Methylobacterium populi BJ001*	73.97/69.4	0.7	0.7	**0.9**	0.8	**0.9**		gene context without functional relation
8	B1LUX6	Ch	P1B-1	*Methylobacterium radiotolerans JCM 2831*	77.49/71	0.8	0.831	**1**	0.9	**0.9**		14 kb gene cluster encoding metal transporters
9	B0UQ23	Ch	P1B-1	*Methylobacterium sp. 4-46*	77.16/71.5	0.8	0.8	**1**	0.9	**0.9**		
10	C5B3T2	P	P1B-1a	*Methylobacterium extorquens AM1*	72.45/67.6	0.8	0.8	**0.9**	0.840	**0.9**		75 Kb gene cluster encoding metal efflux proteins
11	C7CEY4	Ch	P1B-1a	*Methylobacterium extorquens DM4*	72.4/68	0.8	0.8	**1**	0.8	**0.9**	IslandPath-DIMOB	20 kb cluster encoding IS,integrase, transposase
12	C5ASD8	Ch	P1B-1a	*Methylobacterium extorquens AM1*	72.41/68.5	0.8	0.8	**0.9**	0.8	**0.9**		50 Kb gene cluster encoding metal efflux proteins
13	C5B1T2	Ch	P1B-1a	*Methylobacterium extorquens AM1*	69.38/68.5	0.7	0.7	**0.9**	0.7	**0.9**		gene context without functional relation
14	C7CF45	Ch	P1B-1a	*Methylobacterium extorquens DM4*	68.93/68	0.6	0.7	**0.9**	0.7	**0.9**		18 Kb cluster encoding four metal transporters and one transposase
15	A6X3Y7	Ch	P1B-1a	*Ochrobactrum anthropi ATCC 49188*	66.12/56.1	0.8	0.8	1	0.8	**1**	SIGI-HMM	45 Kb detoxification island with conugative transfer proteins

Table S7: Continue

NUMBER	UNIPROT ID	REPLICON	ATPase subtype	RHIZOBIAL STRAIN	%G+C copA/genomic [1,2,3]	CAI[4]	E-CAI[4] p<0.05	CAI/E-CAI[4] p<0.05	E-CAI[4] p<0.01	CAI/E-CAI[4] p<0.01	Genes located into putatve genomic islands (prediction method)[5]	Genomic context[6]
16	Q11BI2	Ch	P1B-3	*Chelativorans sp. BNC1*	66.67/61.1	0.8	0.8	1	0.9	**1**		7 kb gene cluster encoding purine and pyrimidine metabolism enzimes
17	B0UGH1	Ch	P1B-1c	*Methylobacterium sp. 4-46*	76.64/71.5	0.8	0.8	**0.9**	0.9	**0.9**	IslandPath-DIMOB	20 kb cluster encoding IS,integrase, transposase
18	B8IN84	Ch	P1B-1c	*Methylobacterium nodulans ORS 2060*	75.58/68.5	0.7	0.8	**0.9**	0.8	**0.9**		gene context without functional relation
19	Q11G30	Ch	P1B-1	*Chelativorans sp. BNC1*	69.57/61.1	0.8	0.8	**1**	0.856	**1**		16 kb cluster encoding copper-dependent proteins
20	F8BKG9	Ch	P1B-1	*Oligotropha carboxidovorans OM4*	69.57/61.1	0.8	0.723	1	0.758	**1**		10 Kb gene cluster encoding metal transporters
21	Q11BG5	Ch	P1B-1	*Chelativorans sp. BNC1*	69.57/61.1	0.9	0.8	1	0.9	**1**		gene context without functional relation
22	B9JJ75	Ch	P1B-1c	*Agrobacterium radiobacter K84*	63.53/59.9	0.7	0.7	1	0.7	0.9		gene context without functional relation

Bibliografía

Argüello, J. M. (2003). Identification of Ion-Selectivity Determinants in Heavy-Metal Transport P1B-type ATPases. *Journal of Membrane Biology*, *195*(2), 93–108. https://doi.org/10.1007/s00232-003-2048-2

Argüello, J. M. (2014). Functional diversity of five homologous Cu+-ATPases present in Sinorhizobium meliloti. *Microbiology (Reading, England)*, *160*, 1237–1251. https://doi.org/10.1099/mic.0.079137-0

Argüello, J. M., Eren, E., & González-Guerrero, M. (2007). The structure and function of heavy metal transport P1B-ATPases. *BioMetals*, *20*(3–4), 233–248. https://doi.org/10.1007/s10534-006-9055-6

Bavoil, P., Nikaido, H., & von Meyenburg, K. (1977). Pleiotropic transport mutants of Escherichia coli lack porin, a major outer membrane protein. *MGG Molecular & General Genetics*, *158*(1), 23–33. https://doi.org/10.1007/BF00455116

Bondarczuk, K., & Piotrowska-Seget, Z. (2013). Molecular basis of active copper resistance mechanisms in Gram-negative bacteria. *Cell Biology and Toxicology*, *29*(6), 397–405. https://doi.org/10.1007/s10565-013-9262-1

Brom, S., De los Santos, A. G., Stepkowsky, T., Flores, M., Davila, G., Romero, D., & Palacios, R. (1992). Different plasmids of *Rhizobium leguminosarum* bv. phaseoli are required for optimal symbiotic performance. *Journal of Bacteriology*, *174*(16), 5183–5189.

Brown, N. L., Barrett, S. R., Camakaris, J., Lee, B. T. O., & Rouch, D. A. (1995). Molecular genetics and transport analysis of the copper-resistance determinant (pco) from Escherichia coli plasmid pRJ1004. *Molecular Microbiology*, *17*(6), 1153–1166. https://doi.org/10.1111/j.1365-2958.1995.mmi_17061153.x

Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., … Rosenbusch, J. P. (1992). Crystal structures explain functional properties of two E. coli porins. *Nature*, *358*(6389), 727–733. https://doi.org/10.1038/358727a0

Crichton, R. R., & Pierre, J. L. (2001). Old iron, young copper: From Mars to Venus. *BioMetals*, *14*(2), 99–112. https://doi.org/10.1023/A:1016710810701

Cubillas, C., Miranda-Sánchez, F., González-Sánchez, A., Elizalde, J. P., Vinuesa, P., Brom, S., & García-de los Santos, A. (2017). A comprehensive phylogenetic analysis of copper transporting P1BATPases from bacteria of the Rhizobiales order uncovers multiplicity, diversity and novel taxonomic subtypes. *MicrobiologyOpen*, *6*(4), 1–13. https://doi.org/10.1002/mbo3.452

Cubillas, C., Vinuesa, P., Luisa Tabche, M., Dávalos, A., Vázquez, A., Hernández-Lucas, I., … Garcia-de los Santos, A. (2014). The cation diffusion facilitator protein EmfA of Rhizobium etli belongs to a novel subfamily of Mn(2+)/Fe(2+) transporters conserved in α-proteobacteria. *Metallomics : Integrated Biometal Science*, *6*(10), 1808–1815. https://doi.org/10.1039/c4mt00135d

Cubillas, C., Vinuesa, P., Tabche, M. L., & García-de los Santos, A. (2013). Phylogenomic analysis of Cation Diffusion Facilitator proteins uncovers Ni2+/Co2+ transporters. *Metallomics*, *5*(12), 1634. https://doi.org/10.1039/c3mt00204g

Galdiero, S., Galdiero, M., & Pedone, C. (2007). beta-Barrel membrane bacterial proteins: structure, function, assembly and interaction with lipids. *Current Protein & Peptide Science*, *8*(1), 63–82. https://doi.org/10.2174/138920307779941541

García-de los Santos, a, & Brom, S. (1997). Characterization of two plasmid-borne lps beta loci of Rhizobium etli required for lipopolysaccharide synthesis and for optimal interaction with plants. *Molecular Plant-Microbe Interactions : MPMI*, *10*(7), 891–902. https://doi.org/10.1094/MPMI.1997.10.7.891

García-De Los Santos, A., López, E., Cubillas, C. A., Noel, K. D., Brom, S., & Romero, D. (2008). Requirement of a plasmid-encoded catalase for survival of Rhizobium etli CFN42 in a polyphenol-rich environment. *Applied and Environmental Microbiology*, *74*(8), 2398–2403. https://doi.org/10.1128/AEM.02457-07

González-Guerrero, M., & Argüello, J. M. (2008). Mechanism of Cu+-transporting ATPases: soluble Cu+ chaperones directly transfer Cu+ to transmembrane transport sites. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(16), 5992–5997. https://doi.org/10.1073/pnas.0711446105

González-Sánchez, A., Cubillas, C. A., Miranda, F., Dávalos, A., & García-de los Santos, A. (2017). The ropAe gene encodes a porin-like protein involved in copper transit in Rhizobium etli CFN42. *MicrobiologyOpen*, (July), 1–11. https://doi.org/10.1002/mbo3.573

Grass, G., & C Rensing. (2001). Genes Involved in Copper Homeostasis in Escherichia coli Genes Involved in Copper Homeostasis in Escherichia coli. *J Bacteriol*, *183*(6), 2145–2147. https://doi.org/10.1128/JB.183.6.2145

Hernández-Montes, G., Argüello, J. M., & Valderrama, B. (2012). Evolution and diversity of periplasmic proteins involved in copper homeostasis in gamma proteobacteria. *BMC Microbiology*, *12*. https://doi.org/10.1186/1471-2180-12-249

Hohle, T. H., & O'Brian, M. R. (2009). The mntH gene encodes the major Mn2+ transporter in Bradyrhizobium japonicum and is regulated by manganese via the fur protein. *Molecular Microbiology*, *72*(2), 399–409. https://doi.org/10.1111/j.1365-2958.2009.06650.x

Hohle, T. H., & O'Brian, M. R. (2016). Metal-specific control of gene expression mediated by Bradyrhizobium japonicum Mur and Escherichia coli Fur is determined by the cellular context. *Molecular Microbiology*, *101*(1), 152–166. https://doi.org/10.1111/mmi.13381

Hohle, T. H., & O'Brian, M. R. (2010). Transcriptional control of the Bradyrhizobium japonicum irr gene requires repression by fur and antirepression by irr. *Journal of Biological Chemistry*, *285*(34), 26074–26080. https://doi.org/10.1074/jbc.M110.145979

Kabata-Pendias, A., & Pendias, H. (2001). *Trace elements in soils and plants. New York* (Vol. 2nd). https://doi.org/10.1201/b10158-25

Kirshner, R. P. (1994). The Earth's elements. *Scientific American*, *271*(4), 58–65. https://doi.org/10.1038/scientificamerican1094-58

Koebnik, R., Locher, K. P., & Van Gelder, P. (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Molecular Microbiology*, *37*(2), 239–253. https://doi.org/10.1046/j.1365-2958.2000.01983.x

Ladomersky, E., Petris, M. J., Physiology, E., Bond, C. S., & Sciences, L. (2016). HHS Public Access, *7*(6), 957–964. https://doi.org/10.1039/c4mt00327f.Copper

Laguerre, G., Courde, L., Nouaïm, R., Lamy, I., Revellin, C., Breuil, M. C., & Chaussod, R. (2006). Response of rhizobial populations to moderate copper stress applied to an agricultural soil. *Microbial Ecology*, *52*(3), 426–435. https://doi.org/10.1007/s00248-006-9081-5

Lambertsen, L., Sternberg, C., & Molin, S. (2004). Mini-Tn7 transposons for site-specific tagging of bacteria with fluorescent proteins. *Environmental Microbiology*, *6*(7), 726–732. https://doi.org/10.1111/j.1462-2920.2004.00605.x

Landeta, C., Dávalos, A., Cevallos, M. Á., Geiger, O., Brom, S., & Romero, D. (2011). Plasmids with a chromosome-like role in rhizobia. *Journal of Bacteriology*, *193*(6), 1317–1326. https://doi.org/10.1128/JB.01184-10

Lee, H. S., Hancock, R. E. W., & Ingraham, J. L. (1989). Properties of a Pseudomonas stutzeri Outer Membrane Channel- Forming Protein (NosA) Required for Production of Copper- Containing N20 Reductase, *171*(4), 2096–2100.

Lee, S. M., Grass, G., Rensing, C., Barrett, S. R., Yates, C. J. D., Stoyanov, J. V., & Brown, N. L. (2002). The Pco proteins are involved in periplasmic copper handling in Escherichia coli. *Biochemical and Biophysical Research Communications*, *295*(3), 616–620. https://doi.org/10.1016/S0006-291X(02)00726-X

Lewinson, O., Lee, A. T., & Rees, D. C. (2009). A P-type ATPase importer that discriminates between essential and toxic transition metals. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(12), 4677–4682. https://doi.org/10.1073/pnas.0900666106

Li, D., & Daler, D. (2004). Ocean pollution from land-based sources: East China Sea, China. *Ambio*, *33*(1–2), 107–113. https://doi.org/10.1579/0044-7447-33.1.107

Li, X. Z., Nikaido, H., & Williams, K. E. (1997). Silver-resistant mutants of *Escherichia coli* display active efflux of Ag+ and are deficient in porins. *Journal of Bacteriology*, *179*(19), 6127–6132. https://doi.org/10.1128/CMR.00043-12

Li, Z., Lu, M., & Wei, G. (2013). An omp gene enhances cell tolerance of Cu(II) in Sinorhizobium meliloti CCNWSX0020. *World Journal of Microbiology and Biotechnology*, *29*(9), 1655–1660. https://doi.org/10.1007/s11274-013-1328-y

Li, Z., Ma, Z., Hao, X., Rensing, C., & Wei, G. (2014). Genes conferring copper resistance in Sinorhizobium meliloti CCNWSX0020 also promote the growth of Medicago lupulina in copper-contaminated soil. *Applied and Environmental Microbiology*, *80*(6), 1961–1971. https://doi.org/10.1128/AEM.03381-13

Lugtenberg, B., & Van Alphen, L. (1983). Molecular architecture and functioning of the outer membrane of Escherichia coli and other gram-negative bacteria. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, *737*(1), 51–115. https://doi.org/10.1016/0304-4157(83)90014-X

Lutkenhaus, J. F. (1977). Role of a Major Outer Membrane Protein in Escherichia coli, *131*(2), 631–637.

Ma, Z., Jacobsen, F. E., & Giedroc, D. P. (2010). Metal Transporters and Metal Sensors: How Coordination Chemistry Controls Bacterial Metal Homeostasis. *October*, *109*(10), 4644–4681. https://doi.org/10.1021/cr900077w.Metal

Macomber, L., & Imlay, J. A. (2009). The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(20), 8344–8349. https://doi.org/10.1073/pnas.0812808106

Martinez-Romero E. 2003. Diversity of Rhizobium-Phaseolus vulgaris symbiosis: overview and perspectives. Plant Soil 252:11–23.

Matsumoto, P. S. (2005). Trends in Ionization Energy of Transition-Metal Elements. *J. Chem. Educ.*, *82*(11), 1660. https://doi.org/10.1021/ed082p1660

Mills, S., Jasalavich, C., & Cooksey, D. (1993). A 2-Component Regulatory System Required for Copper Inducible Expression of the Copper Resistance Operon of Pseudomonas-Syringae. *Journal of Bacteriology*, *175*(6), 1656–1664. Retrieved from http://jb.asm.org/content/175/6/1656.full.pdf

Munson, G. P., Lam, D. L., Outten, F. W., & Halloran, T. V. O. (2000). Identification of a Copper-Responsive Two-Component System on the Chromosome of Escherichia coli K-12 Identification of a Copper-Responsive Two-Component System on the Chromosome of Escherichia coli K-12, *182*(20), 5864–5871. https://doi.org/10.1128/JB.182.20.5864-5871.2000.Updated

Nakae, T., & Ishii, J. (1978). Transmembrane permeability channels in vesicles reconstituted from single species of porins from Salmonella typhimurium. *Journal of Bacteriology*, *133*(3), 1412–1418.

Ochoa-Herrera, V., León, G., Banihani, Q., Field, J. A., & Sierra-Alvarez, R. (2011). Toxicity of copper(II) ions to microorganisms in biological wastewater treatment systems. *Science of the Total Environment*, *412–413*, 380–385. https://doi.org/10.1016/j.scitotenv.2011.09.072

Odermatt, A., Suter, H., Krapf, R., & Solioz, M. (1993). Primary structure of two P-type ATPases involved in copper homeostasis in Enterococcus hirae. *Journal of Biological Chemistry*, *268*(17), 12775–12779.

Ormeño-Orrillo, E., Servín-Garcidueñas, L. E., Rogel, M. A., González, V., Peralta, H., Mora, J., … Martínez-Romero, E. (2015). Taxonomy of rhizobia and agrobacteria from the Rhizobiaceae family in light of genomics. *Systematic and Applied Microbiology*, *38*(4), 287–291. https://doi.org/10.1016/j.syapm.2014.12.002

Outten, F. W., Huffman, D. L., Hale, J. A., & O'Halloran, T. V. (2001). The Independent cue and cus Systems Confer Copper Tolerance during Aerobic and Anaerobic Growth in Escherichia coli. *Journal of Biological Chemistry*, *276*(33), 30670–30677. https://doi.org/10.1074/jbc.M104122200

Padilla-Benavides, T., Thompson, A. M. G., McEvoy, M. M., & Argüello, J. M. (2014). Mechanism of ATPase-mediated Cu+ export and delivery to periplasmic chaperones: The interaction of Escherichia coli CopA and CusF. *Journal of Biological Chemistry*, *289*(30), 20492–20501. https://doi.org/10.1074/jbc.M114.577668

Patel S. J., Padilla-Benavides T., Collins J. M., Argüello, J. M. (2014). Functional diversity of five homologous Cu^{+}-ATPases present in *Sinorhizobium meliloti*. *Microbiology*, (*160*), 1237–1251. https://doi.org/10.1099/mic.0.079137-0

Peebles, P. J., Schramm, D. N., Turner, E. L., & Kron, R. G. (1994). The evolution of the universe. *Scientific American*, *271*(4), 52–57. https://doi.org/10.1038/scientificamerican1094-52

Rademacher, C., & Masepohl, B. (2012). Copper-responsive gene regulation in bacteria. *Microbiology (United Kingdom)*, *158*(10), 2451–2464. https://doi.org/10.1099/mic.0.058487-0

Reeve, W. G., Tiwari, R. P., Kale, N. B., Dilworth, M. J., & Glenn, A. R. (2002). ActP controls copper homeostasis in Rhizobium leguminosarum bv. viciae and Sinorhizobium meliloti preventing low pH-induced copper toxicity. *Molecular Microbiology*, *43*(4), 981–991. https://doi.org/10.1046/j.1365-2958.2002.02791.x

Rensing, C., & Grass, G. (2003). Escherichia coli mechanisms of copper homeostasis in a changing environment. *FEMS Microbiology Reviews*, *27*(2–3), 197–213. https://doi.org/10.1016/S0168-6445(03)00049-4

Rubino, J. T., & Franz, K. J. (2012). Coordination chemistry of copper proteins: How nature handles a toxic cargo for essential function. *Journal of Inorganic Biochemistry*, *107*(1), 129–143. https://doi.org/10.1016/j.jinorgbio.2011.11.024

Sankari, S., & O'Brian, M. R. (2016). The Bradyrhizobium japonicum ferrous iron transporter FeoAB is required for ferric iron utilization in free living aerobic cells and for symbiosis. *Journal of Biological Chemistry*, *291*(30), 15653–15662. https://doi.org/10.1074/jbc.M116.734129

Schirmer, T. (1998). General and specific porins from bacterial outer membranes. *Journal of Structural Biology*, *121*(2), 101–109. https://doi.org/10.1006/jsbi.1997.3946

Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor Perspectives in Biology*. https://doi.org/10.1101/cshperspect.a000414

Sitsel, O., Grønberg, C., Autzen, H. E., Wang, K., Meloni, G., Nissen, P., & Gourdon, P. (2015). Structure and Function of Cu(I)- and Zn(II)-ATPases. *Biochemistry*, *54*(37), 5673–5683. https://doi.org/10.1021/acs.biochem.5b00512

Smith, A. T., Smith, K. P., & Rosenzweig, A. C. (2014). Diversity of the metal-transporting P1B-type ATPases. *Journal of Biological Inorganic Chemistry*, *19*(6), 947–960. https://doi.org/10.1007/s00775-014-1129-2

Soares, F., Moura, N., Sobrinho, A., & Sánchez, S. (2009). Mecanismos de hiperacumulación de metales pesados en plantas.

Solioz, M., & Stoyanov, J. V. (2003). Copper homeostasis in Enterococcus hirae. *FEMS Microbiology Reviews*, *27*(2–3), 183–195. https://doi.org/10.1016/S0168-6445(03)00053-6

Solomon, F. (2009). Impacts of copper on aquatic ecosystems and human health. *Mining.com Magazine*, (January), 25–28. Retrieved from https://yukonwaterboard.ca/registers/quartz/qz08-084/Volumes 9-11/5.0/5.2.1.pdf

Speer, A., Rowland, J. L., Haeili, M., Niederweis, M., & Wolschendorf, F. (2013). Porins increase copper susceptibility of Mycobacterium tuberculosis. *Journal of Bacteriology*, *195*(22), 5133–5140. https://doi.org/10.1128/JB.00763-13

Villaseñor, T., Brom, S., Dvalos, A., Lozano, L., Romero, D., & Los Santos, A. G. De. (2011). Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in Rhizobium etli and Rhizobium leguminosarum. *BMC Microbiology*, *11*. https://doi.org/10.1186/1471-2180-11-66

Vollan, H. S., Tannæs, T., Vriend, G., & Bukholm, G. (2016). In silico structure and sequence analysis of bacterial porins and specific diffusion channels for hydrophilic molecules: Conservation, multimericity and multifunctionality. *International Journal of Molecular Sciences*, *17*(4). https://doi.org/10.3390/ijms17040599

Wernimont, A. K., Huffman, D. L., Finney, L. A., Demeler, B., O'Halloran, T. V., & Rosenzweig, A. C. (2003). Crystal structure and dimerization equilibria of PcoC, a methionine-rich copper resistance protein from Escherichia coli. *Journal of Biological Inorganic Chemistry*, *8*(1–2), 185–194. https://doi.org/10.1007/s00775-002-0404-9

WHO (World Health Organization). (2011). Copper in Drinking-water. *America*, *3*, 23. https://doi.org/10.1016/j.kjms.2011.05.002

Yoneyama, H., & Nakae, T. (1996). Protein C (OprC) of the outer membrane of Pseudornonas aeruginosa is a copper-regulated channel protein. *Microbiology*, *142*(1 996), 2137–2144.

Yruela, I. (2009). Copper in plants: Acquisition, transport and interactions. *Functional Plant Biology*, *36*(5), 409–430. https://doi.org/10.1071/FP08288

Zhang, Y., & Gladyshev, V. N. (2010). General trends in trace element utilization revealed by comparative genomic analyses of Co, Cu, Mo, Ni, and Se. *Journal of Biological Chemistry*, *285*(5), 3393–3405. https://doi.org/10.1074/jbc.M109.071746

Bibliografía electrónica:

Flanagan D. M. 2017. USGS Minerals Information: Copper. (https://minerals.usgs.gov/minerals/pubs/commodity/copper/index.html#mcs) (visto: 31/1/18)

Passek T. S. 2017. Copper Development Association Inc. (https://www.copper.org/resources/properties/compounds/agricultural.html) (visto: 31/1/18)

UniprotKB. 2018. (https://www.uniprot.org/uniprot/?query=rhizobiales&sort=score#orgViewBy) (visto 15/1/18)