



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
CENTRO DE CIENCIAS GENÓMICAS
DOCTORADO EN CIENCIAS BIOMÉDICAS

Estudio del secretoma de *Rhizobium etli* utilizando proteómica

T E S I S

QUE PARA OBTENER EL GRADO DE:
DOCTORA EN CIENCIAS

P R E S E N T A

NIURKA MENESES MORENO

DIRECTOR DE TESIS: DR. SERGIO MANUEL ENCARNACION GUEVARA

CUERNAVACA, MORELOS 2010



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

AGRADECIMIENTOS

Mi primer agradecimiento es para Dios que hizo posible que realizara este proyecto.

Expreso mi más sincero e inmenso agradecimiento al Dr. Sergio M. Encarnación Guevara, por darme la oportunidad de llevar a cabo este proyecto, por su confianza y sobre todo por su inmensa paciencia.

También quiero agradecer al Dr. Guillermo Mendoza Hernández, por confiar en mí y por sus valiosos consejos.

Agradezco también al Dr. Michael Dunn Goelli por sus consejos y su ayuda en la revisión del idioma en el primer artículo que resultó de este proyecto.

De forma muy especial le agradezco a la Dra María de Lourdes Girard por su sincera amistad y sus acertados consejos.

Agradezco también a los doctores Susana Brom, Guillermo Gosset, Humberto Lanz y Mariano Martínez por sus acertadas correcciones para mi tesis y su apoyo para que resultara un mejor manuscrito.

Agradezco al Dr. Pablo Vinuesa por todo su apoyo.

Un agradecimiento especial al Dr. Heller por su apoyo y consejos.

Agradezco a Omar Alejandro por su apoyo en Bioinformática y por su sincera amistad.

Le agradezco a Sandra por todo su apoyo, consejos y su sincera amistad.

Agradezco a Gabriel y a Magda por su apoyo técnico y su ayuda.

Agradezco a Andrés por su apoyo desinteresado.

Agradezco a todos los compañeros de laboratorio, que siempre estuvieron ahí para escucharme cuando lo necesité.

Un agradecimiento sincero para Gladys, Martha y Cinthya por su apoyo desinteresado.

Un agradecimiento a todos los trabajadores del CCG.

Agradezco también a los que no creyeron en mí, porque esto me permitió, buscar la forma de demostrarles que estaban equivocados y me hizo crecer como persona.

Un agradecimiento especial para la UNAM y al CCG por permitirme lograr uno de mis grandes sueños y convertirme en mejor persona.

DEDICATORIAS

A mis queridos y maravillosos David y Leandro que siempre fueron mi inspiración para seguir adelante.

A mis padres y en especial a mi maravillosa madre. Este trabajo es producto de la educación que ellos me dieron.

A mis hermanas Susi e Iris, a mi sobrina Yadira y a toda mi familia.

A todos mis amigos, no escribo sus nombres porque son muchos y no quiero olvidar a ninguno.

A Mani por creer en la magia del amor desde que nos conocimos.

A mi segunda patria, México.

El presente trabajo se desarrollo en el programa de Genómica Funcional de Procariotas del Centro de Ciencias Genómicas, perteneciente a la Universidad nacional autónoma de México, bajo la dirección del Dr. Sergio Manuel Encarnación Guevara.

Para el desarrollo de este trabajo se contó con becas de manutención otorgadas por la Dirección General de Estudios de Postgrado (DGEP).

INDICE

RESUMEN.....	3
ABSTRACT	6
INTRODUCCIÓN.....	8
Proteínas secretadas.....	10
Sistema de secreción de tipo I	11
Sistema de secreción de tipo II.....	12
Sistema de secreción de tipo III.....	12
Sistema de secreción de tipo IV	14
Sistema de secreción de tipo V.....	14
Sistema de secreción de tipo VI	14
Translocación de arginina gemela	15
Liberación de vesículas de la membrana externa	15
Proteómica.....	15
Aplicaciones de la proteómica.....	19
Espectrometría de masas	20
Secuenciación de péptidos por Espectrometría de Masas	22
Secuenciación de una mezcla de péptidos. Espectros MS/MS.....	23
Estrategias para la identificación de las proteínas	24
ANTECEDENTES	30
OBJETIVOS.....	34
MATERIALES Y MÉTODOS.....	35
RESULTADOS	40
RESULTADOS ADICIONALES	41
Proteoma de <i>R. etli</i> en fase de crecimiento exponencial	42
Proteoma de la fracción periplásmica en fase de crecimiento exponencial	42
Proteoma de las OMV en fase de crecimiento exponencial	44
Comparación entre los proteomas de la fracción periplásmica y las OMV en fase exponencial	45
Proteoma de <i>R. etli</i> en fase de crecimiento estacionaria.....	47
Proteoma de la fracción periplásmica en fase estacionaria	47
Proteoma de las OMV en fase estacionaria	48

Comparación entre los proteomas de la fracción periplásmica y las OMV en fase estacionaria	50
Comparación de los proteomas de <i>R. etli</i> las fases de crecimiento exponencial y estacionaria	51
Comparación de los proteomas de las fracciones periplásmica en ambas fases de crecimiento.....	52
Comparación de los proteomas de las OMV en ambas fases de crecimiento	53
Proteoma de <i>R. etli</i> en presencia y ausencia de naringenina	55
Proteoma de la fracción extracelular en presencia de naringenina.....	55
Proteoma de la fracción extracelular en ausencia de naringenina.....	57
Proteoma de las OMV en presencia de naringenina.....	58
Proteoma de las OMV en ausencia de naringenina	59
Comparación de los proteomas extracelulares de <i>R. etli</i> en ausencia y presencia de naringenina.....	60
Comparación de los proteomas OMV de <i>R. etli</i> en ausencia y presencia de naringenina	62
Proteínas identificadas en el proteoma de OMV de <i>R. etli</i> que están presentes en <i>Bradyrhizobium japonicum</i> BTAil.....	63
¿Están las OMV relacionadas con el proceso de simbiosis en <i>R. etli</i> ?.....	71
DISCUSION.....	75
CONCLUSIONES.....	87
REFERENCIAS	89
ANEXOS	97

RESUMEN

En bacterias fijadoras de nitrógeno como *Rhizobium etli* no se han reportado estudios de secretoma, este es un término usado para describir el proteoma de todas las proteínas secretadas, incluyendo las extracelulares, las de membranas y las periplasmáticas. La aplicación de la proteómica al estudio de estas proteínas en cada una de las etapas de crecimiento de la bacteria, podrá elucidar varias incógnitas sobre la función de proteínas extracelulares, así como la vía que utilizan para ser exportadas.

La proteómica proporciona una visión general de los productos de los genes de las bacterias y nos permite realizar un análisis global de diferentes condiciones de crecimiento.

La comunicación entre las bacterias y sus hospederos, generalmente está mediada por las proteínas extracelulares. Un gran número de proteínas secretadas están relacionadas con los procesos de simbiosis que realizan muchas bacterias del suelo como las rhizobiáceas. También algunas de estas proteínas son toxinas que están relacionadas con el proceso de patogenicidad.

En este contexto en el presente trabajo hemos analizado el secretoma de *R.etli* con el objetivo de determinar el papel de las proteínas que lo constituyen en el funcionamiento celular de la bacteria.

Se analizaron las proteínas secretadas en las fases de crecimiento exponencial y estacionario. De cada fase se analizaron las proteínas de la fracción periplásmica, las extracelulares y las que se secretan en las vesículas de membrana externa (OMV). Las proteínas fueron identificadas utilizando cromatografía líquida de doble dimensión y espectrometría de masas, concretamente identificadas utilizando una trampa de iones (LC-MS/MS). Las proteínas identificadas fueron analizadas con ayuda del programa SignalP 3.0 Server, buscando la presencia teórica de un péptido señal que sugiriera un mecanismo de exportación de la proteína.

Adicionalmente se identificaron tanto las proteínas comunes como las específicas en cada una de las condiciones de crecimiento estudiadas. Las proteínas fueron agrupadas por COG (Cluster Orthologous Groups), según NCBI.

Sorprendentemente en la fase de crecimiento exponencial el grupo de proteínas más abundante corresponde a las proteínas ribosomales, a pesar de que estas son las que intervienen en la síntesis de proteínas y se encuentran en citoplasma. Encontramos que en la fase de crecimiento estacionaria el grupo de proteínas más representativo fue el

correspondiente a de función desconocida, esto puede ser porque como las proteínas secretadas han sido las menos estudiadas por lo que aun no se encuentran anotadas en las bases de datos. Estas proteínas, podrían estar involucradas en procesos importantes para la bacteria como: mecanismo de defensa, colonización de nuevos nichos, interacción con otros microorganismos, entre otros. Tomando en cuenta lo anterior la pregunta obligada era mediante que sistema de secreción son exportadas las proteínas ribosomales y demás proteínas componentes del secretoma. Un reporte previo en *E. coli* sugiere que las proteínas ribosomales pudieran ser secretadas por unas estructuras llamadas vesículas de membrana externa (OMV), [Lee et al. 2007] las cuales se sabe son producidas por bacterias Gram negativas.

Observamos que *R. etli* también secreta estas estructuras lo cual sugiere que algún tipo de proteínas incluidas las ribosomales podrían ser secretadas mediante este sistema.

Estudiamos el proteoma de estas estructuras, identificando en ellas, un grupo importante de proteínas entre las cuales se encuentran las proteínas ribosomales.

Para estudiar el posible papel que tienen las OMV de *R. etli*, en el proceso de simbiosis, realizamos experimentos de inoculación de semillas de frijoles con extracto de vesículas. Observamos que las raíces de las plantas inoculadas con vesículas desarrollaron un mayor número de pelos radicales en relación a los controles. Esto sugiere que tanto las OMV como sus productos contenidos, están involucrados en las etapas tempranas del proceso de simbiosis entre la bacteria y la planta.

Para apoyar la idea de que las OMV tienen algún papel en el proceso simbiótico, decidimos estudiar el secretoma en presencia del flavonoide naringenina.

El análisis del secretoma de *R. etli* inducido con naringenina, mostró una cantidad importante de proteínas similares a proteínas involucradas en simbiosis, reportadas previamente en la cepa *Bradyrhizobium* sp. BTAi1. Este resultado sugiere que las OMV en *R. etli*, pueden estar relacionadas con el proceso de simbiosis en etapas tempranas, facilitando el proceso de colonización bacteriana.

Adicionalmente analizamos extractos butanólicos de las OMV purificadas en presencia y ausencia de naringenina, mediante cromatografía líquida de fase reversa y los comparamos con los extractos butanólicos de las OMV de una mutante de *R. etli* que no produce factores de nodulación. Los resultados obtenidos mostraron que en las OMV obtenidas de la bacteria inducida con naringenina encontramos una fracción que no aparecía en los controles, por lo cual analizamos la fracción mediante espectrometría de nodulación. Posteriormente fue realizada una cromatografía de capa fina donde

analizamos los extractos butanólicos obtenidos de las OMV y de la mutante NodA-, en este caso la cepa fue crecida en presencia de acetato de sodio marcado [^{14}C]. Con este experimento demostramos que los factores Nod están siendo exportados en las OMV por *R. etli*, por lo que estas estructuras, están involucradas en las etapas tempranas del proceso de simbiosis entre la bacteria y la planta.

ABSTRACT

For nitrogen-fixing bacteria, such as *Rhizobium etli* (*R.etli*), no reports about the study of their secretome have been published yet. The secretome is a term used to describe the proteome of all secreted proteins, including shed proteins from the extracellular and the periplasmic fraction. The application of proteomics to the study of the secretome in each of the stages of growth of the bacteria may help to clarify many questions about the role of extracellular proteins and the mechanism used to export them. Proteomics can provide an integrated view of expressed gene products from bacteria and allows conducting a comprehensive analysis of different conditions. In this study, we analyzed the secretome *R.etli* in order to determine the possible role of constituent proteins in cell function.

Secreted proteins were analyzed in the exponential and stationary growth phases, and three different fractions were submitted to analysis, namely the periplasmic, the extracellular soluble, and the outer membrane vesicle (OMV) associated proteins, respectively. Proteins were identified using two-dimensional liquid chromatography, SDS-PAGE and mass spectrometry (LC-MS/MS). The identified proteins were analyzed using the SignalP 3.0 Server program to estimate the presence of a signal peptide suggesting a mechanism of protein export. Additionally, the proteins were grouped by COG (Cluster Orthologous Groups), according to NCBI.

Surprisingly, in the exponential growth phase the most abundant group of proteins corresponded to ribosomal proteins, although these are thought to be involved in protein synthesis and located in the cytoplasm. We found that in the stationary growth phase, the most representative group of proteins had no known function, corroborating the fact that secreted proteins have been the least studied. We hypothesized that the proteins secreted by the bacteria could be involved in processes that are essential for the bacteria, such as defense mechanisms, colonization of new niches, and interaction with other organisms, among others. The obvious question was by which secretion system ribosomal proteins and other protein components of the secretome were exported. A previous report suggested that ribosomal proteins were being secreted from *E. coli* in OMV [Lee et al. 2007]. We observed that *R. etli* also secreted OMV containing ribosomal proteins.

To investigate the possible role of OMV from *R. etli* in symbiosis, we carried out inoculation experiments with bean seeds and isolated OMV's. We noted that the roots of plants inoculated with vesicles developed a greater number of root hairs relative to

controls. This suggests that OMV, and its content, are involved in the early stages of the process of symbiosis between bacteria and plant.

Naringenin is released from bean plants to attract *R.etli* and start the symbiotic process.

To find out if the release of OMV were also involved in the symbiosis, we therefore, studied the secretome in the presence of naringenin.

The analysis of the secretome of *R. etli* in the presence of naringenin, showed a significant amount of proteins similar to proteins involved in symbiosis, which were previously reported from the strain *Bradyrhizobium japonicum* sp BTAil. This result suggests that the OMV in *R. etli*, may be related to the process of symbiosis in the early stages, facilitating the process of bacterial colonization.

Additionally, we analyzed butanol extracts of OMV by reverse phase liquid chromatography. OMV's were isolated from wild type and a mutant *R. etli* strain, which does not produce nodulation factor. Furthermore, bacteria were grown in the presence or absence of naringenin. The comparison of the chromatograms revealed a chromatographic peak in the wild type, naringenin-stimulated extract absent in all other conditions. This peak was isolated and analyzed by mass spectrometry, indicating molecular ions similar in mass to those reported for nodulation factors. Subsequently was performed thin layer chromatography which we analyze the n-butanol extracts obtained from OMV, *R. etli* and NodA- mutant, in this case the strain was grown in the presence of labeled sodium acetate [¹⁴C]. This experiment showed that Nod factors are being exported in the OMV, proving that these structures are involved in the early stages of the process of symbiosis between bacteria and plant.

INTRODUCCIÓN

Rhizobium etli es una bacteria Gram negativa presente en la rizósfera que puede formar asociación simbiótica con las raíces de frijol. Durante este proceso la bacteria induce la formación de estructuras en la raíz de la planta llamadas nódulos, en donde se lleva a cabo el proceso de fijación biológica de nitrógeno. Mediante este proceso la planta utiliza el amonio sintetizado por la bacteria, mientras que la bacteria recibe moléculas ricas en carbono, que le proporciona la planta, el cual es utilizado para la generación de energía tan necesaria en un proceso tan demandante como es la fijación de nitrógeno.

Este proceso de fijación biológica de nitrógeno que ocurre en el suelo, hace prescindible el uso de abonos o fertilizantes nitrogenados para las plantas que establecen simbiosis con los rizobios, evitándose así problemas de contaminación ambiental e inversión económica cuantiosa.

En este proyecto analizaremos algunos factores que pueden afectar lo que ocurre en la etapa inicial de la simbiosis, es decir, en la etapa en la cual se desarrolla el diálogo molecular entre la bacteria y la planta (simbiontes). La comunicación entre los simbiontes comienza con la liberación a la rizósfera, por parte de la leguminosa, de metabolitos secundarios contenidos en los exudados de semilla y de raíz, principalmente flavonoides. Estos compuestos son la primera molécula señal emitida dentro del diálogo molecular que se establece entre la planta y la bacteria. Si los flavonoides secretados por la planta son reconocidos por los rizobios, en estos últimos se inicia la transcripción de los genes denominados de nodulación (genes *nod*); a partir de los cuales se producen los factores de nodulación (factores Nod) [Cárdenas et al 1995; Dénarié et al 1996; Soria-Díaz et al 2003; Deakin et al 2009]. Los factores Nod son específicos para cada especie de rizobio, lo cual hace posible, que se pueda o no establecer la simbiosis con una especie determinada de leguminosa. Este hecho nos habla del alto grado de especificidad en esta relación simbiótica. Los factores Nod juegan un papel esencial dentro del establecimiento de la simbiosis y mimetizan muchas de las respuestas que se inducen en los pelos radicales en presencia de los rizobios [Cárdenas et al 1995].

Los factores de nodulación son así una de las señales clave para el reconocimiento simbiótico. Estos factores poseen una estructura genérica constituida por 3-5 residuos de N-acetyl-D-Glucosamina (NAcGlc) unidos por enlaces de β -1,4 en la que el extremo no reductor de la molécula es acilado con varios ácidos grasos. El precursor requerido para la formación de este oligosacárido es UDP-N-acetyl-D-Glucosamina (UDP-NAcGlc), que es también un precursor para la biosíntesis de los lipopolisacáridos y del

peptidogluano. Las mayores diferencias entre los factores Nod de las distintas estirpes de rizobios incluyen:

- Longitud de la estructura básica de NAcGlc, que puede variar de 3 a 5 residuos,
- Naturaleza de cadena lipídica del extremo no reductor, de forma que los ácidos grasos puedan ser saturados, insaturados, α/β insaturados o (ω -1) ácidos grasos hidroxilados, y
- Una variedad de sustituyentes específicos que incluye: grupos acetilos, arabinosilos, carbomoilos, fucosilos, glicerilos, hidroxilos, manosilos y sulfatos, que decoran los extremos reductor y no reductor de la molécula de N-acetil-D-glucosamina [Dénarié *et al.*, 1996].

En la Figura 1 se presenta un esquema general del factor de nodulación, así como la localización de los distintos sustituyentes en los factores sintetizados por los distintos rizobios y los genes responsables de estas decoraciones del esqueleto de NAcGlc.

Algunos experimentos indican que no es tan clara la relación entre factores Nod y la especificidad de nodulación. Así, los factores Nod producidos por *R. etli* y *M. loti* son idénticos, y las dos especies poseen distinto rango hospedador, *Phaseolus* sp. y *Lotus* sp., respectivamente [Cárdenas *et al.* 1995]. Del mismo modo, *R. leguminosarum* bv. *trifolii* produce factores Nod idénticos a los producidos por *R. leguminosarum* bv. *viciae* y estas bacterias poseen distintos rango de hospedador [Spaink *et al.*, 1995]. En contraste, *R. tropici* y *R. etli* nodulan eficientemente *Phaseolus* y producen factores Nod distintos [Poupot *et al.*, 1993 y 1995]. Todos los simbioses capaces de nodular *Glycine max* producen factores Nod fucosilados por *nodZ*, que codifica para una fucosil transferasa, y sin embargo ninguna mutación en *nodZ* altera la nodulación en *Glycine max* u otros hospedadores habituales de *S. fredii* [Stacey *et al.*, 1994; Quesada-Vicens *et al.*, 1997].

Por lo tanto, los factores Nod son esenciales para nodulación y sus modificaciones pueden contribuir a la especificidad de la nodulación pero esta familia de moléculas señal es, probablemente, sólo uno de los varios elementos involucrados en esta especificidad [Broughton y Perret 1999]. Por lo tanto no es posible correlacionar completamente la estructura de los factores Nod con la habilidad para nodular una leguminosa específica.

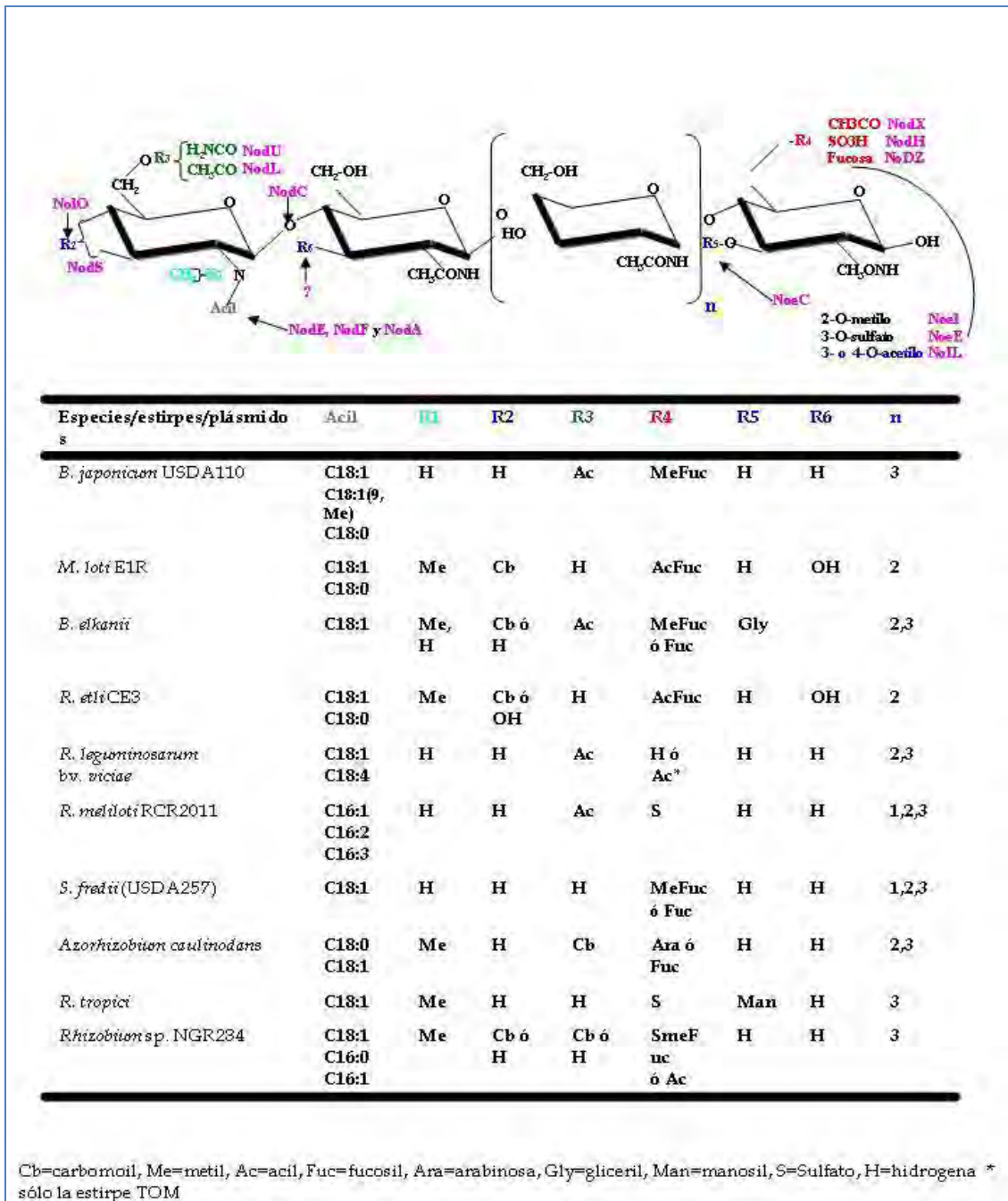


Figura1: Estructura general de los factores de nodulación [Broughton y Perret 1999].

Proteínas secretadas

La secreción proteica no es exclusiva de los organismos eucariotas sino que está presente también en bacterias y Archaea. Las bacterias Gram negativas tienen dos membranas, haciendo así la secreción topológicamente más compleja. Hay al menos seis sistemas de secreción especializados, en bacterias Gram negativas.

A pesar del número, la diversidad y la amplia variedad de funciones que desempeñan las proteínas secretadas (proteólisis, hemólisis, citotoxicidad, reacciones de

fosforilación, etc.), éstas son exportadas utilizando un número limitado de mecanismos. Las vías de secreción en las bacterias Gram-negativas han sido clasificadas en cinco grupos principales: secreción tipo I, II, III, IV y los autotransportadores (figura 2). Además se han reportado otros sistemas de secreción: Dicha clasificación se basa en la naturaleza molecular de las maquinarias de transporte y las reacciones que éstas catalizan [Lee et al 2001]. Sin embargo, estos grupos se pueden aún subdividir en dos grandes clases dependiendo del mecanismo que se utilice para el transporte a través de la membrana plasmática. Las vías **Sec-dependientes**, que utilizan el sistema de secreción denominado Sec, en el que las proteínas a secretarse presentan una secuencia señal o péptido líder en el extremo amino terminal; y las **Sec-independientes** en las que los sustratos se pueden exportar directamente desde el citoplasma hasta el exterior celular sin que exista un intermediario periplásmico, ni una secuencia señal en el amino terminal [Koster et al 2000, Thanassi y Hultgren 2000].

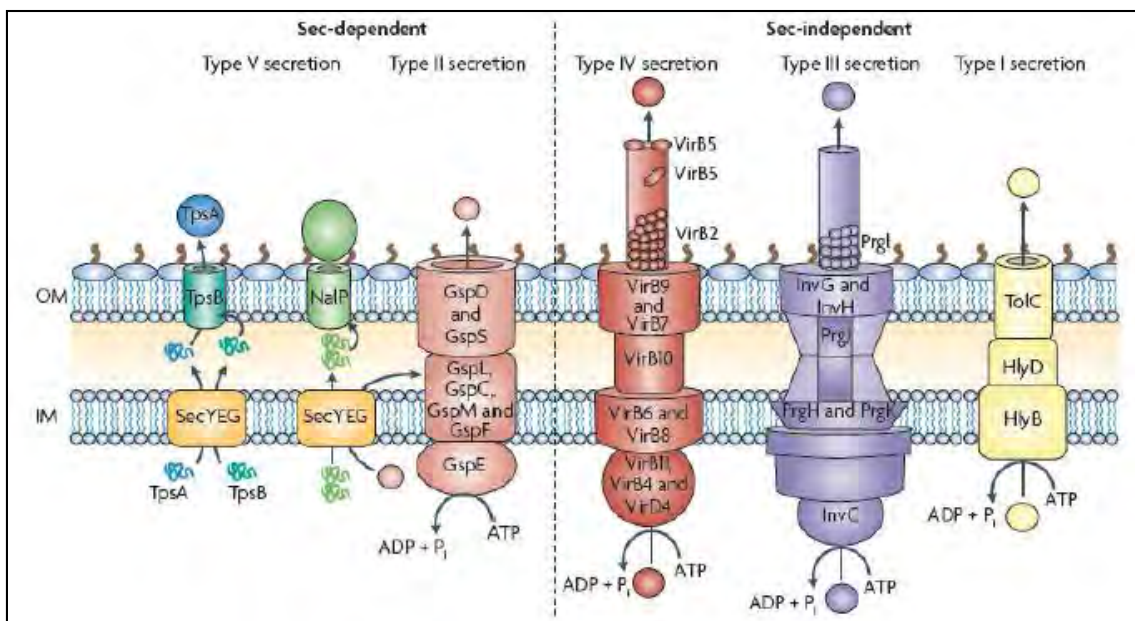


Figura2: Principales sistemas de secreción en bacterias Gram negativas. Nature Reviews Microbiology 7, 703-714, 2009.

Sistema de secreción de tipo I

Este mecanismo es utilizado por una amplia gama de bacterias para la secreción de toxinas, proteasas y lipasas. Es una vía Sec-independiente, por lo que no se requiere del procesamiento de un péptido líder para atravesar la membrana citoplasmática; la secreción proteica se da en un solo paso desde el citoplasma hasta el exterior celular

[Koster et al 2000]. Los sustratos a exportarse por este sistema presentan una señal de secreción en el extremo carboxilo terminal, que sin embargo, no es procesada.

Es el sistema transportador tipo ABC (del inglés ATP binding cassette), pero en este sistema están presentes proteínas adicionales que, junto con las proteínas ABC, forman un canal contiguo que cruza las membranas interiores y exteriores de las bacterias Gram-negativas. Este es un sistema simple, que consiste de sólo tres subunidades de proteínas: las proteínas tipo ABC, proteína de fusión a la membrana y proteína de la membrana externa. Este sistema de secreción transporta diversas moléculas, como iones, antibióticos y proteínas de varios tamaños (de 20 a 100 kDa) [Delepelaire, 2004].

Sistema de secreción de tipo II

El SSTII es responsable de secretar una gran cantidad de enzimas hidrolíticas y toxinas, como la toxina del cólera. Esta vía también se conoce como sistema general de secreción y ocurre en dos etapas. Primero, la maquinaria Sec exporta el sustrato con péptido líder a través de la membrana plasmática, por lo que este sistema es una vía Sec-dependiente. Dicho péptido es generalmente una secuencia corta (de aproximadamente 30 aminoácidos), de los que uno o varios presentan carga positiva, además de una secuencia de 10 a 20 aminoácidos hidrofóbicos.

En una segunda etapa, la proteína pierde el péptido señal y adquiere su conformación nativa en el espacio periplásmico, para posteriormente ser secretada a través de la ME por un complejo sistema multiproteico llamado secretinas [Thanassi y Hultgren 2000, Lee et al 2001]. Además de las secretinas, otras 10-15 proteínas de la membrana interior y exterior forman el aparato de secreción completo, muchas con función aún desconocida. Los pili de las bacterias Gram-negativas usan una versión modificada del sistema tipo II para su biogénesis, y en algunos casos ciertas proteínas se comparten entre un complejo pilus y el sistema tipo II dentro de una misma especie bacteriana. [Sandkvist M, 2001].

Sistema de secreción de tipo III

Este sistema de secreción constituye un área de investigación que ha sido extensamente estudiada en los últimos años. Es una vía Sec-independiente en la que la secreción ocurre en un solo paso desde el citoplasma hasta el exterior celular, que desempeña un papel central en la patogenicidad de muchas bacterias Gram-negativas [Hueck et al 1998].

Este sistema de secreción ha sido identificado en una gran variedad de patógenos de humanos, animales y plantas, incluyendo especies de *Bordetella*, *Chlamydia*, *Erwinia*, *E. coli*, *Pseudomonas*, *Ralstonia*, *Rhizobia*, *Salmonella*, *Shigella*, *Xanthomonas* y *Yersinia*. Mediante este sistema los factores de virulencia se pueden exportar hasta el citoplasma de la célula eucarionte [Galan et al 1999, Cheng et al 2000, Plano et al 2001]. La maquinaria está conservada entre los diferentes patógenos, sin embargo las proteínas secretadas difieren completamente, por lo que el mismo mecanismo de transporte puede generar una amplia gama de enfermedades [Hueck et al 1998]. Además de su papel en la patogénesis, el sistema de secreción tipo III se requiere para la biogénesis flagelar [Macnab et al 1999], y es también esencial para el establecimiento de la relación simbiótica entre *Rhizobium* y plantas leguminosas.

Los sustratos que se exportan por este sistema de secreción carecen de una señal de secreción definida, sin embargo, para algunas proteínas tanto del sistema flagelar, como de virulencia, se ha demostrado que la señal se localiza en el dominio amino terminal [Hueck et al 1998]. A pesar de esto, no ha sido posible identificar secuencias conservadas en esta región y algunos datos sugieren que es más bien una propiedad fisicoquímica, como la anfipaticidad, la que es crítica para la secreción [Lloyd et al 2002]. La hipótesis de la señal de secreción en el amino terminal ha sido cuestionada y se ha propuesto que se localiza en la región 5' codificante del RNA mensajero [Anderson et al 1997]. Sin embargo, no existe evidencia directa de interacción entre el RNA mensajero y los componentes del sistema de secreción.

Por otro lado, en todos los sistemas de secreción tipo III estudiados existen proteínas chaperonas que comparten características comunes como el tamaño pequeño, pI ácido y una alta predicción de estructura α -hélice. La exportación de algunos sustratos es facilitada por dichas chaperonas, las que se propone mantienen a las proteínas en una conformación adecuada para la secreción, previniendo además la agregación prematura en el citoplasma [Bennett et al 2000, Page y Parsot 2002]. Sin embargo, el papel directo de las chaperonas como posible señal de reconocimiento en la secreción, así como su papel en la organización de la jerarquía de la secreción de los efectores, aún no ha sido demostrado.

Este sistema es homólogo al del cuerpo basal flagelar bacteriano. Se parece a una jeringuilla molecular por la cual una bacteria (por ejemplo, ciertos tipos de *Salmonella*, *Shigella*, *Yersinia*) puede inyectar proteínas en células eucarióticas. Una baja

concentración de Ca^{2+} en el citosol abre la puerta que regula el sistema. El sistema Hrp, en patógenos de plantas, inyecta harpinas por mecanismos similares. Este sistema de secreción fue descubierto en la bacteria *Yersinia pestis*, causante de la peste bubónica, y mostró que las toxinas podían ser inyectadas directamente desde el citoplasma bacteriano al citoplasma de las células del anfitrión más que a través del medio extracelular [He et al., 2004].

Sistema de secreción de tipo IV

Es un sistema homólogo a la maquinaria de conjugación de las bacterias (y los flagelos de *Archaea*). Es capaz de transportar tanto ADN como proteínas. Fue descubierto en *Agrobacterium tumefaciens*, que usa este sistema para introducir el plásmido Ti y proteínas en el hospedero, lo que lleva al desarrollo de un tumor. *Helicobacter pylori* usa este sistema de secreción para inyectar CagA en células epiteliales gástricas. *Bordetella pertussis*, bacteria causante de la tosferina, secreta la toxina pertussis en parte por el sistema tipo IV. *Legionella pneumophila*, causante de la legionela, utiliza también este sistema para translocar numerosas proteínas efectoras a su hospedero eucariótico [Christie et al., 2005].

Sistema de secreción de tipo V o autotransportadores

A través de este sistema se exportan proteínas con diferentes funciones incluyendo proteasas, toxinas, adhesinas e invasinas. Los autotransportadores representan una vía Sec dependiente ya que utilizan la maquinaria Sec para atravesar la membrana interna sin embargo, las proteínas no requieren de factores adicionales para transitar del periplasma hacia el exterior celular. Como su nombre lo indica, dirigen su propia exportación. El extremo carboxilo terminal de la proteína dirige la secreción de la región amino terminal a través de la membrana externa. Las proteínas que usan esta ruta tienen la capacidad de formar un barril beta con su terminal C, que se inserta en la membrana externa permitiendo al resto del péptido alcanzar el exterior de la célula. Algunos investigadores sugieren que los remanentes de los autotransportadores dieron lugar a las porinas, que forman estructuras similares de barril beta. [Henderson et al. 2004].

Sistema de secreción de tipo VI

Tipo de la secreción VI es un nuevo mecanismo, descrito para el transporte de proteínas a través de la envoltura celular de las bacterias Gram-negativas. Los componentes, que han sido parcialmente caracterizados incluyen un homólogo de IcmF, de una proteína reguladora del dominio ATPasa de ClpV, y un dominio de la proteína reguladora Fha y las proteínas secretadas VgrG y Hcp. El sistema de secreción Tipo VI es claramente la clave para exportar factores de virulencia para algunas bacterias patógenas importantes y se ha relacionado en la secreción de una proteína efectora potencial en las células eucariotas en *Vibrio cholerae*. La secreción de varias proteínas de este sistema ha sido descrita recientemente en *Vibrio cholerae* y *Pseudomonas aeruginosa*. Estas proteínas carecen de las secuencias señalizadoras como péptido señal [Bingle et al 2008].

Translocación de arginina gemela

Las bacterias, también usan otro sistema de transporte especial llamado desplazamiento de arginina gemela, ruta que, en contraste con la exportación Sec dependiente, transporta proteínas totalmente plegadas a través de la membrana. El nombre del sistema viene del requerimiento de dos argininas consecutivas en la secuencia de señales necesaria para disparar este sistema. [Rose et al. 2002].

Liberación de vesículas de la membrana externa

Además del uso de complejos multiproteínicos indicados anteriormente, las bacterias Gram negativas poseen otro método para la liberación de sustancias: la formación de vesículas en la membrana externa (OMV). Partes de la membrana externa forman estructuras esféricas, compuestas de una bicapa de lípidos, que encierran sustancias periplásmicas. Las vesículas de varias especies bacterianas pueden contener sustancias patógenas, tener efectos inmunomoduladores, y también adherirse directamente e intoxicar las células huésped. Mientras que la liberación de vesículas es una respuesta general a condiciones de estrés, el proceso de cargar proteínas transportadoras parece ser selectivo. [Lauren et al 2005].

Proteómica

La proteómica es la ciencia dedicada al estudio de la expresión de las proteínas y de sus cambios en dependencia del contexto biológico. A diferencia de las técnicas clásicas utilizadas en la bioquímica, la proteómica se basa en la separación y la identificación masiva de proteínas. Muchos de los métodos utilizados en proteómica permiten obtener

un ordenamiento físico de mezclas muy complejas de proteínas, separadas mediante la combinación de dos o más técnicas de separación.

El proteoma es el conjunto de proteínas expresadas por un organismo o parte de él, por ejemplo un tipo de tejido, en un momento dado [Wasinger et al 1995]. La proteómica comprende tanto las técnicas para el estudio en gran escala de las proteínas expresadas como las aplicaciones de estas técnicas al análisis de problemas biológicos. Mientras que el genoma de un organismo es esencialmente constante a lo largo de la vida, el proteoma tiene un carácter dinámico: la expresión de proteínas cambia en diferentes etapas del ciclo celular pero también en respuesta a acciones externas. En particular, las diferencias entre un estado "normal" y uno patológico se traducen también a nivel molecular en cambios en los patrones de expresión de proteínas. Es precisamente esta variabilidad del proteoma lo que lo hace tan atractivo para la investigación en biología.

Por tanto, necesitamos relacionar las características biológicas de un sistema con la expresión de las proteínas y más específicamente, relacionar los cambios en las propiedades biológicas con los cambios en la expresión de proteínas. Para ello es necesario disponer de técnicas que permitan:

- a) separar miles de proteínas;
- b) identificar las proteínas de interés y
- c) cuantificar la magnitud de los cambios en la expresión de proteínas.

La proteómica aborda este complejo problema técnico con variadas herramientas.

Estas técnicas se agrupan en dos grandes campos: las que se basan en la resolución de proteínas intactas y que utilizan principalmente técnicas de electroforesis o de cromatografía de proteínas y las que se basan en la conversión, en una primera etapa, de las mezclas de proteínas en mezclas complejas de sus péptidos de degradación proteolítica.

En el primer grupo están las técnicas que utilizan la electroforesis bidimensional (2DE) atendiendo a una primera separación por carga eléctrica y una segunda separación por peso molecular. También comprende la combinación de dos o tres métodos ortogonales de cromatografía o isoelectroenfoque (IEF) en solución seguido de cromatografía (por ejemplo, separación por peso molecular seguida de separación por hidrofobicidad; separación por carga, por peso molecular y finalmente por hidrofobicidad; separación por IEF de flujo libre seguido de separación por masa o por hidrofobicidad). Las fracciones obtenidas en estas separaciones, tras múltiples etapas, pueden ser resueltas a

continuación por 2-DE o por electroforesis en geles de SDS (SDS-PAGE), para posteriormente, ser digeridas y analizadas mediante espectrometría de masas.

En el segundo grupo (técnicas basadas en un primer paso de conversión de las mezclas de proteínas en mezclas de péptidos obtenidos por proteólisis específica) se sustituye la resolución de proteínas por la resolución cromatográfica de las mezclas de sus péptidos obtenidos por proteólisis específica generalmente con tripsina. Alternativamente, se hace la selección de un conjunto de péptidos representativos de la presencia de las proteínas parentales atendiendo a alguna propiedad.

Si bien la electroforesis en geles de SDS es una técnica de rutina en los laboratorios biológicos, la electroforesis bidimensional de alta resolución requiere de un laboratorio dedicado, tanto por los requisitos de reproducibilidad como por la variedad y complejidad de los pasos del proceso. La electroforesis bidimensional es hasta el presente, la única técnica que permite resolver con alta eficiencia las isoformas que difieren en punto isoelectrónico y por ello, se sigue utilizando ampliamente [Gorg A et al, 2000]. Sin embargo, la electroforesis bidimensional posee algunas limitaciones para analizar proteínas hidrofóbicas ya que por su escasa solubilidad están subrepresentadas en los mapas electroforéticos bidimensionales obtenidos hasta el momento. Por otra parte, aquellas proteínas que poseen puntos isoelectrónicos muy ácidos o básicos son difíciles de focalizar en la mayoría de los geles de isoelectroenfoque disponibles en el mercado. Tampoco nos permite analizar proteínas de muy bajo peso molecular. El amplio rango dinámico es otro problema no resuelto.

Por tales motivos, en los últimos años ha existido la tendencia a trabajar con los péptidos en vez de las proteínas tratando de solucionar las limitaciones señaladas. La idea consiste en marcar diferencialmente con isótopos estables los péptidos generados por la proteólisis de todas las proteínas sintetizadas por una célula o tejido en dos condiciones que se desean estudiar o comparar. Posteriormente, mediante el análisis por cromatografía líquida de fase inversa (o combinada con otro método cromatográfico) acoplada a la espectrometría de masas se puede realizar la identificación de las proteínas en las bases de datos de secuencias. La cuantificación se logra mediante un análisis detallado de las distribuciones isotópicas de los péptidos analizados y así se infiere la expresión diferencial de las proteínas que los contienen.

Aunque la idea parece sencilla, no deja de tener asociada una gran complejidad pues si bien el trabajo con péptidos puede ser más fácil en comparación con las proteínas, no es menos cierto que la mezcla de proteínas que se deriva de una célula o tejido de por sí es

muy compleja por lo que una vez realizada la proteólisis la mezcla de péptidos que se genera debe ser mucho más compleja aún. Por tales motivos se han empleado propiedades ortogonales para lograr la separación de la mezcla de péptidos obtenida. Entre estas propiedades tenemos la talla, carga, hidrofobicidad, e interacción biológica o afinidad. Por lo tanto es de esperar un sinnúmero de posibles combinaciones en la cromatografía multidimensional, pero lo común entre todas radica en que persiguen un mismo objetivo: obtener la mejor resolución posible para posteriormente identificar una mayor cantidad de proteínas en el análisis por espectrometría de masas.

Entre las combinaciones más empleadas se encuentra la cromatografía de intercambio iónico con la cromatografía de fase reversa (RP). Particularmente, en esta combinación, la utilización de columnas de RP en el segundo paso permite que la mezcla a su vez sea desalada y pueda ser analizada mediante espectrometría de masas empleando electronebulización (ESI). También se han reportado otras combinaciones que emplean cromatografías de exclusión por tamaño molecular (SEC) y RP [Opitck et al 1997]. Raida y colaboradores combinaron la cromatografía de intercambio catiónico (SCX) con RP-HPLC acoplada al espectrómetro de masas para medir las masas de alrededor de 3000 péptidos de un hemofiltrado humano [Raida et al 1999]. Aunque los dos pasos cromatográficos fueron desarrollados en experimentos diferentes, este estudio demostró el elevado poder resolutivo de la cromatografía multidimensional y la espectrometría de masas para el fraccionamiento y análisis de un gran número de péptidos.

La inclusión de una cromatografía de afinidad en uno de los pasos de separación multidimensional de péptidos ha sido muy útil, pues permite retener selectivamente péptidos o proteínas basándose en interacciones específicas, lo que simplifica considerablemente la mezcla compleja de péptidos proteolíticos. Así por ejemplo, en la identificación de péptidos fosforilados, que es una de las modificaciones postraduccionales más importantes por estar involucrada prácticamente en todos los mecanismos de transducción de señales, las columnas de IMAC-Fe³⁺ han permitido el aislamiento de dichos fosfopéptidos y permitido el estudio del fosfoproteoma de varios organismos [Anderson et al 1986 y Ficarro et al 2002].

La utilización de columnas con lectinas inmovilizadas ha permitido la captura selectiva de glicopéptidos, que facilita el estudio del glicoma de los organismos [Chakraborty et al 2000]. También se han empleado otras variantes, que incluyen la derivatización química de aminoácidos para que, mediante el empleo de una cromatografía de afinidad,

sea posible el aislamiento selectivo de los péptidos modificados dentro de todos los péptidos proteolíticos generados.

Aplicaciones de la proteómica

Las múltiples aplicaciones de estas técnicas pueden agruparse en tres categorías principales:

1. Caracterización del proteoma de un organismo. El propósito de este estudio es la identificación del mayor número de proteínas expresadas por un organismo en una condición biológica particular. En el caso de organismos unicelulares o de células en cultivo, muchos trabajos han propuesto la separación y la identificación de proteínas en preparaciones de proteínas totales y ya existen mapas en bases de datos para numerosos organismos.

Sin embargo, para lograr un nivel superior de información se prefiere descomponer el estudio del proteoma en el estudio de subproteomas correspondientes a organelos subcelulares.

2. Proteómica Comparativa: consiste en evaluar los cambios en la expresión de proteínas de un organismo sometido a dos o varias condiciones biológicas diferentes, generalmente una de ellas es una condición control que se utiliza como referencia de la expresión de proteínas en condiciones “normales”. Este tipo de trabajo suministra información a nivel molecular de los cambios causados por la acción de un agente externo (por ejemplo un medicamento), por cambios en las condiciones de cultivo, las diferencias entre una línea celular normal y una tumoral, la evaluación de los cambios producidos por un agente infeccioso, el desarrollo de los mecanismos de resistencia a quimioterapéuticos o antimicrobianos. Este tipo de investigación es de alto valor en ciencias biomédicas. Para los trabajos de proteómica comparativa es necesario disponer de un diseño experimental cuidadosamente planificado. Aquí resulta de gran utilidad la existencia de una hipótesis previa que oriente la búsqueda hacia determinado tipo de proteínas que pueden ser seguidas por inmunodetección con anticuerpos específicos o que se suponen localizadas esencialmente en un organelo subcelular. Por ejemplo, algunos proyectos se orientan específicamente hacia la identificación de cambios en los patrones de fosforilación de proteínas. En este caso, la detección del subconjunto de proteínas fosforiladas puede facilitarse mediante el uso de anticuerpos específicos.

3. Interacción de Proteínas. En los últimos años ha quedado evidenciada la importancia de identificar las interacciones entre proteínas y de proteínas con otras moléculas. En

varias publicaciones se han reportado los resultados de aislamiento de complejos de proteínas y la posterior identificación de las proteínas componentes. Estos resultados contribuyen significativamente al establecimiento de mapas de interacciones y al esclarecimiento del mecanismo de determinadas funciones biológicas.

Es posible predecir que las herramientas que se emplean en proteómica se irán perfeccionando e incluso aparecerán nuevos procedimientos y equipos que solucionarán las principales limitaciones actuales. Sin lugar a dudas la proteómica fortalecerá su posición actual en las investigaciones médicas, farmacéuticas, agrícolas y en otros campos y se multiplicarán sus aplicaciones.

Espectrometría de masas

La espectrometría de masas es la herramienta empleada en los estudios de proteómica para la identificación de las proteínas. Su principio se asemeja a lo que ocurre cuando un haz luminoso incide sobre un prisma: la radiación electromagnética es descompuesta o separada de acuerdo a las diferentes longitudes de onda. En un espectrómetro de masas un haz de iones es separado de acuerdo a la relación masa/carga: m/z .

En un espectrómetro de masas existen dos elementos esenciales: *la fuente de ionización*, donde son producidos los iones al suministrarle energía a la muestra que se estudia y el *analizador* (un campo eléctrico, un campo magnético, de tiempo de vuelo o una combinación de ellos), donde los iones son separados de acuerdo a la relación masa/carga (m/z). La energía suministrada puede ser suficiente no solamente para ionizar la molécula y obtener el llamado *ión molecular* (la molécula intacta pero con una carga positiva o negativa) sino también para provocar la fragmentación de la molécula dando lugar a iones fragmentos. Los iones producidos son característicos de un determinado compuesto químico, es decir dependen de la estructura química de este, por tanto es posible su identificación a partir del espectro de masas.

Aunque la espectrometría de masas ya era una técnica establecida para la caracterización de los compuestos orgánicos desde finales de la década de los años 50, los métodos de ionización disponibles hasta 1981 no eran aplicables a las llamadas biomoléculas (péptidos, proteínas, carbohidratos y ácidos nucleicos) ya que era necesaria la evaporación previa de la muestra. Por esta razón, para estudiar los compuestos con elevados puntos de fusión o que se descomponen con la temperatura había que obtener derivados volátiles y en el caso de que esto no fuera posible, simplemente no podían ser estudiados. En 1981, Michael Barber y Richard Bordoli, de

la Universidad de Manchester publican la aplicación de un método de ionización suave, el FAB (*fast atom bombardment*) al estudio de proteínas [Barber et al 1981]. El método, aunque ya no se emplea debido a su baja eficiencia de ionización, fue un resultado de gran importancia pues permitió por primera vez la medición de la masa molecular de los péptidos y la determinación de la secuencia sin necesidad de obtener un derivado volátil.

En 1989, John Fenn publica [Fenn et al 1989] la aplicación de la ionización por electrospray (*electrospray ionization*, ESI) al estudio de péptidos, proteínas y otras biomoléculas. Un año antes Koichi Tanaka había publicado un nuevo método de ionización conocido como MALDI (*Matrix Assisted Laser Desorption Ionization*) [Tanaka et al 1988]. Poco después Michael Karas y Franz Hillenkamp reportan varias aplicaciones del MALDI en el campo de las proteínas [Karas et al 1988]. Por las aportaciones mencionadas, Fenn y Tanaka recibieron el premio Nobel de Química en el 2002.

Estos dos métodos revolucionaron completamente la aplicación de la espectrometría de masas a las biomoléculas y en apenas diez años la espectrometría de masas se convirtió en el método más poderoso para el análisis y la caracterización de péptidos y proteínas.

El MALDI emplea un láser ultravioleta para suministrar la energía capaz de ionizar el compuesto. Previamente la muestra es mezclada íntimamente con la matriz. La matriz es un elemento muy importante y está constituida por compuestos insaturados (ácido sinapínico, ácido 2,5 dihidroxibenzoico (DHB), ácido 4-hidroxi- α -cianocinámico, etc.) capaces de absorber energía a la longitud de onda del láser. De esta forma la matriz absorbe la mayor parte de la energía suministrada que de otra forma destruiría la muestra totalmente y facilita la ionización del compuesto que se desea analizar.

El ESI se había desarrollado como un procedimiento para la interfase entre la cromatografía líquida y la espectrometría de masas [Whitehouse et al 1985]. Así la muestra es transportada en solución (puede ser desde un cromatógrafo líquido) a través de un capilar en cuyo extremo se encuentra un alto voltaje (entre 3-4 kV). De esta forma se produce una nebulización de pequeñas gotas formadas por el solvente y la muestra y que portan un número considerable de cargas eléctricas. El solvente es eliminado con la ayuda de una corriente de gas (usualmente nitrógeno) y finalmente queda la muestra ionizada con varias cargas.

Tanto el MALDI como el ESI son métodos de ionización *suave*, esto es, la energía suministrada a la muestra es relativamente baja, y los compuestos son ionizados pero no

ocurre la fragmentación de las moléculas como en el caso de la ionización por impacto electrónico, el método de ionización más usado para el estudio de moléculas orgánicas. En ambos métodos de ionización, MALDI y ESI, un péptido adquiere una carga positiva o negativa por adición o eliminación de protones respectivamente, dando lugar a los iones pseudo-moleculares que corresponden a la masa molecular del péptido más (o menos en el caso de iones negativos) la masa de uno o varios protones. Si se adiciona un protón la carga del péptido será unitaria, y si se adicionan varios protones se obtendrán iones multicargados. En el MALDI los iones son fundamentalmente monocargados mientras que en el ESI los iones son usualmente multicargados. En principio es posible estudiar los iones positivos o negativos, aunque es mucho más frecuente la espectrometría de masas de iones positivos.

Al introducir una mezcla de péptidos en el espectrómetro de masas usando MALDI o ESI como método de ionización, se obtiene un grupo de señales que corresponden a los iones moleculares de cada uno de los péptidos, por cuanto no se producen fragmentaciones de las moléculas.

Secuenciación de péptidos por Espectrometría de Masas

Al no producirse fragmentaciones durante la ionización por MALDI o ESI se obtiene solamente información sobre la masa molecular de un péptido, pero no es posible obtener información de su secuencia. Para lograr este propósito se emplea un proceso llamado Disociación Inducida por Colisiones (DIC) o Disociación Activada por Colisiones (DAC). Este consiste en introducir el haz de iones de la muestra en una región que contiene un gas químicamente inerte (usualmente argón), para que por las colisiones generadas entre ambos, una parte de la energía cinética que poseen las moléculas se transforme en energía vibracional de sus enlaces y ocasione la disociación en dependencia de su fortaleza [Hayes 1990]. Aunque la eficiencia de la fragmentación de un péptido está muy influenciada por las características de su secuencia, se ha observado que también depende del gas de colisión que se emplee [Biemann et al 1988]. La nomenclatura aceptada por los especialistas en este campo para clasificar los iones fragmentos de un péptido es la que propusieron Roepstorff y Fohlmann en 1984. Se plantea que la ruptura del esqueleto carbonado de un péptido puede originar seis tipos diferentes de fragmentaciones (a_n , b_n , c_n , x_n , y_n , z_n) que se clasifican en series del extremo *N*-terminal (a_n , b_n , c_n) o series del extremo *C*-terminal (x_n , y_n , z_n) en dependencia de cuál de los dos extremos del péptido original se conservan en sus

estructuras. El subíndice que acompaña cada tipo de fragmentación corresponde con la cantidad de residuos aminoacídicos que posee el ión fragmento.

Aunque las fragmentaciones de las cadenas laterales son de gran importancia para la diferenciación de aminoácidos que poseen igual masa (aminoácidos isobáricos) y suministran más información estructural, es necesario destacar que en proteómica estas fragmentaciones no resultan de utilidad para la identificación de las proteínas pues la mayoría de los espectrómetros que se emplean actualmente realizan la disociación inducida por colisiones de los péptidos en un régimen de baja energía y las fragmentaciones de las cadenas laterales están ausentes del espectro de masas. Sin embargo, el desarrollo reciente de los espectrómetros de masas MALDI-TOF/TOF [Medzihradzky et al 2000 y Yergey et al 2002] pudiera vislumbrar una aplicación en el campo de la proteómica al observarse fragmentaciones de las cadenas laterales que suministran mayor cantidad de información estructural. Además se obtienen espectros de masas con muy buena calidad en el rango de los femtomoles. La abundancia relativa de estas series *N* y *C*-terminales en el espectro está determinada generalmente por la posición del residuo básico en la secuencia. Así por ejemplo, cuando residuos de Arginina están ubicados en el extremo *N* o en el extremo *C* del péptido, abundan en el espectro las series *N* o *C*-terminal respectivamente. Este fenómeno es conocido como “*charge remote fragmentation*” [Adams et al 1986] y plantea que la fragmentación ocurre a lo largo de la cadena peptídica inducida por la carga a pesar de que ésta se mantiene en un lugar fijo (alejado en muchos casos), presumiblemente en la posición en que está ubicado el residuo básico. Por este motivo es usual observar que los espectros de masas de los péptidos trípticos son abundantes en iones *C*-terminales (series *yⁿ*).

Para secuenciar un péptido, cada uno de los iones fragmentos observados en el espectro se debe clasificar de acuerdo a la nomenclatura antes mencionada, pues de esta manera las diferencias en masas de series consecutivas de un mismo tipo (*b*₅, *b*₄, *b*₃, *b*₂...) pueden ser asignadas a uno de los veinte aminoácidos más comunes y así obtener la secuencia peptídica. Por otra parte, un péptido se puede secuenciar simultáneamente a partir de sus dos extremos con un solo espectro pues la información suministrada por las series *N* y *C*-terminales son complementarias. Se obtienen secuencias parciales desde ambos extremos y al superponerlas se determina una secuencia altamente confiable.

Secuenciación de una mezcla de péptidos. Espectros MS/MS

Lo referido en el párrafo anterior implica que debemos tener un péptido puro, pues si tuviéramos una mezcla de péptidos se producirían fragmentos de todos ellos y sería imposible la interpretación de los espectros. Por esta razón se requiere aislar o seleccionar un péptido determinado para proceder a su secuenciación. Esto es posible mediante la espectrometría de masas en tándem o MS/MS que consiste en dos analizadores separados por la cámara de colisiones donde el gas neutro provoca las fragmentaciones. Cuando se introduce una mezcla de péptidos en la fuente de ionización, se produce la ionización de todos y al penetrar en el primer analizador es posible seleccionar solo uno de ellos, los restantes simplemente chocan con las paredes del analizador. El péptido seleccionado pasa a la cámara de colisiones donde se producen los fragmentos que son posteriormente analizados en el segundo analizador, con lo que se obtiene la secuencia del péptido.

Una vez concluida esta operación para un péptido, se selecciona otro péptido y se procede de forma similar. Este procedimiento es rápido y automático, por lo que se pueden medir los espectros MS/MS de varios péptidos presentes en una mezcla en un tiempo relativamente breve y obtener sus secuencias. Esta posibilidad técnica también tiene otra repercusión importante en proteómica y es que las proteínas que se van a identificar no tienen que estar totalmente puras. Es posible identificar varias proteínas en una mezcla como ocurre en ocasiones en una banda de electroforesis donde puede existir más de una proteína.

Estrategias para la identificación de las proteínas

Como ya fue mencionado anteriormente, las proteínas separadas por electroforesis bidimensional son digeridas con una enzima, usualmente tripsina, con lo que se obtiene una mezcla de péptidos. Los péptidos son extraídos del gel e introducidos directamente en el espectrómetro de masas para su identificación (Fig 3). En la actualidad existen tres estrategias fundamentales para la identificación de proteínas basadas en la información primaria suministrada por la espectrometría de masas:

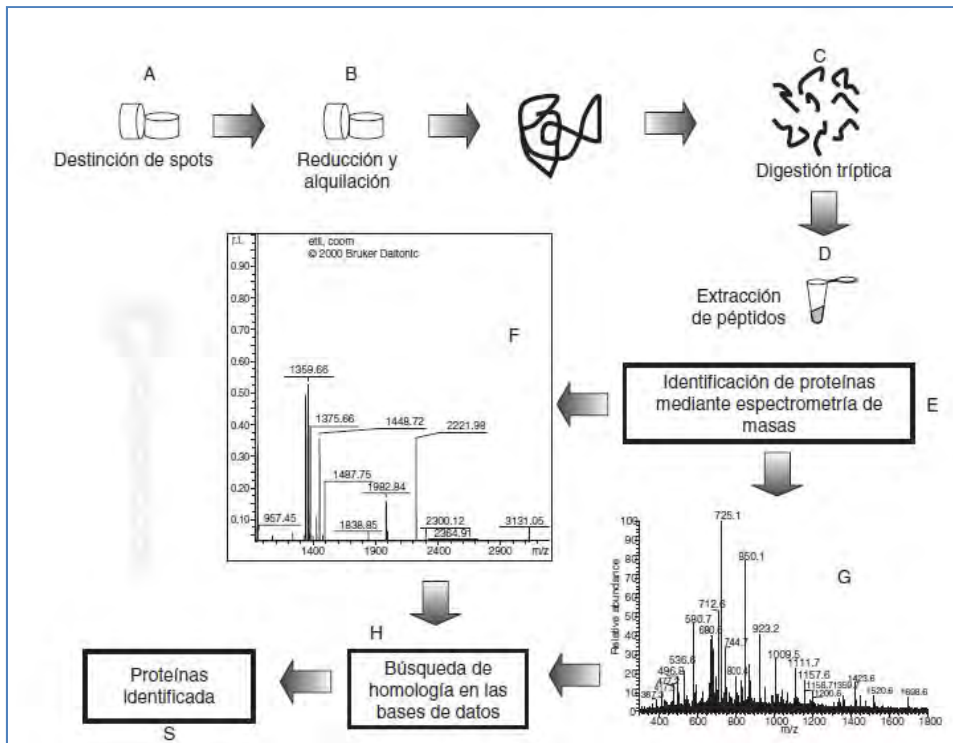


Figura3: Diagrama esquemático de la identificación de proteínas mediante espectrometría de masas. Después de la separación bidimensional, las proteínas son recortadas y sujetas a distinción (A), reducción y alquilación (B), previo a la digestión con tripsina (C). El volumen fue ajustado a $\sim 20 \mu\text{l}$ (D), en este punto la muestra está lista para su identificación mediante espectrometría de masas (E). Una alícuota del sobrenadante conteniendo los péptidos tripticos es analizada mediante MALDI-TOF-MS, del cual resulta en mapa peptídico de la proteína; el panel muestra los picos del espectro de masas (F). El remanente del sobrenadante es analizado mediante ESI-MS. La figura muestra la secuencia derivada de la fragmentación del mismo péptido (G) [Encarnación et al 2006].

La primera de las estrategias para la identificación de proteínas en las bases de datos se basa en que la digestión proteolítica de una proteína con una proteasa altamente específica, origina de manera reproducible un conjunto de péptidos que es característico para cada una de las proteínas.

Por esta razón, es posible predecir mediante una digestión *in silico*, el conjunto de péptidos que se debe generar mediante el tratamiento proteolítico específico de cada una de las proteínas almacenadas en las bases de datos de secuencias y por lo tanto se puede calcular teóricamente las masas moleculares de cada uno de los péptidos esperados.

La masa molecular de cada uno de los péptidos proteolíticos se puede determinar experimentalmente mediante ESI-MS o MALDI-MS. Esta información se conoce como huella de masas de los péptidos de una proteína o *peptide mass fingerprint* (PMF). Por tanto, la proteína analizada puede ser identificada exitosamente cuando se produce una

correspondencia entre los valores experimentales de las masas moleculares de cada uno de sus péptidos con los valores de masas teóricas esperadas para una de las proteínas existentes en las bases de datos.

Con la instrumentación disponible en la actualidad la medición de las masas moleculares puede realizarse de manera rutinaria con un error inferior a los 50 ppm por lo que las identificaciones por esta metodología son muy confiables y es previsible que con la introducción masiva de espectrómetros más acuciosos la calidad de las identificaciones sea aún superior pues está directamente relacionada con la exactitud de la medición de las masas moleculares de los péptidos proteolíticos [Clauser et al 1999; Wool y Smilansky 2002].

Hoy en día existen varios programas accesibles vía Internet que brindan el servicio de identificación de proteínas en línea utilizando la estrategia del PMF y entre los más empleados se encuentra el MASCOT, (Matrix Science, London, UK) y el Phenyx. De todas las estrategias para la identificación de proteínas en las bases de datos de secuencia, ésta fue la que inicialmente ganó más popularidad, debido a su gran sencillez por ello, numerosos grupos de trabajo la implementaron en sus laboratorios y desarrollaron los programas de cómputo apropiados para este propósito [Mann et al 1993 y Gevaert et al 1996].

Sin embargo, cuando se digieren proteínas poco abundantes en la banda de gel, incluso empleando los protocolos optimizados, con frecuencia se recuperan bajas cantidades de los péptidos proteolíticos y es frecuente encontrar que solo aparezcan las señales de masas correspondientes a uno, dos o tres péptidos proteolíticos. En estos casos la estrategia de PMF puede brindar resultados ambiguos o identificaciones erróneas. En tales situaciones es necesario obtener información de la secuencia de los péptidos analizados.

Esta limitante fue resuelta con el surgimiento de otra estrategia conocida como etiqueta de secuencia (sequence tag) para la identificación de proteínas. Ella plantea que para lograr la identificación solo es necesario extraer una pequeña secuencia parcial (3 a 4 aminoácidos) del espectro ESI-MS/MS de un péptido [Mann et al 1994]. Este método revolucionó la identificación de las proteínas a partir de la secuencia deducida de los espectros de masas, por cuanto se requería una interpretación exhaustiva del espectro MS/MS para obtener la secuencia aminoacídica lo más completa posible de algunos de los péptidos proteolíticos, para poder identificar la proteína correspondiente. El procedimiento del sequence tag es particularmente útil en los espectros ESI-MS/MS de

los péptidos trípticos pues esa pequeña secuencia parcial es muy fácil de extraer pues la región de mayor masa es muy limpia en cuanto a la relación señal/ruido y es rica en los fragmentos y^+n (el aminoácido básico se conserva en el C-terminal).

Aunque el método es muy confiable y hay programas que extraen de manera automática la secuencia parcial a partir de los espectros ESI-MS/MS, en algunos casos se requiere de la intervención del especialista, pues el éxito de la identificación se basa en una correcta interpretación de espectros para extraer la secuencia parcial.

Otra estrategia conocida como huella de masas del espectro MS/MS, se basa en que cada péptido que se fragmenta mediante un proceso de disociación inducida por colisiones produce un espectro de iones fragmentos (MS/MS), que es muy característico de su secuencia y que lo diferencia del resto de los péptidos almacenados en las bases de datos.

El espectro MS/MS de un péptido está mayoritariamente representado por fragmentos originados por la ruptura del enlace peptídico (y^+n o b_n). Si la secuencia de la proteína que contiene al péptido se conoce, es posible predecir las masas de los fragmentos iónicos que se originarán. Por ello, el programa basa su identificación en buscar primeramente el conjunto de todos los péptidos (dentro de las proteínas en las bases de datos) que, originados por un corte específico de la proteasa empleada, coincidan con la masa molecular del péptido que se analiza. Posteriormente, a cada uno de los péptidos que posean esa característica se le calcula las masas moleculares de los iones fragmentos posibles y se comparan con las masas moleculares de los fragmentos obtenidos experimentalmente en el espectro MS/MS. Se selecciona el péptido con mejor coincidencia y así queda identificada la proteína.

Resulta evidente que para todos estos programas mencionados es indispensable disponer de la secuencia de las proteínas en las bases de datos para una correcta identificación.

Particularmente, para la identificación de proteínas que pertenezcan a un organismo de genoma desconocido, solo se podrá obtener un resultado exitoso mediante la estrategia de PMF, si la proteína de interés comparte una identidad de secuencia elevada con respecto a alguna de las proteínas reportadas en las bases de datos. Mientras menor sea la identidad de secuencia, menos confiable será la identificación realizada.

Si la similitud de secuencia es muy baja o si la secuencia del gen no es conocida, entonces los péptidos trípticos tienen que ser secuenciados completamente a partir de la interpretación manual o automática de los espectros MS/MS de cada péptido. La identificación se realiza mediante el alineamiento de las secuencias determinadas

experimentalmente contra bases de datos de secuencias conocidas, para encontrar secuencias similares. No obstante, si la secuencia de la proteína en cuestión no es muy similar a las reportadas en las bases de datos esta estrategia puede fallar. Esto es debido a que es poco probable que la secuencia de genes de una especie sea idéntica a la otra y las sustituciones de nucleótidos pueden provocar sustituciones de aminoácidos que implican cambios en las masas de los péptidos analizados y por tanto, repercute negativamente en la identificación.

Los programas de alineamiento de secuencias de proteínas y genes más empleados son el BLAST www.ncbi.nlm.nih.gov/blast y el FASTA www.ebi.ac.uk/searches/fasta.html. La identificación a través de estos programas resulta difícil, ya que ambos algoritmos se han optimizado para comparar secuencias de proteínas y nucleótidos de mayor longitud que los péptidos secuenciados por espectrometría de masas. Por otro lado, la espectrometría de masas no permite diferenciar aminoácidos isobáricos ya que presentan masas muy similares (Lys y Gln) o idénticas (Leu e Ile) y la calidad del alineamiento puede estar determinada por cuál aminoácido de los antes mencionados se introduzca en la secuencia a alinear ya que estos programas contienen una matriz diferente para cada uno.

El auge alcanzado en la aplicación de la espectrometría de masas a la secuenciación de proteínas, ha motivado que los programas de alineamientos hayan sido recientemente modificados para adaptarlos al alineamiento óptimo de secuencias cortas de péptidos. Particularmente, el FASTA modificado (FASTS) [Mackey et al 2002] posee aún una limitante pues las búsquedas son lentas y la puntuación final de los envíos depende no sólo del número de péptidos sino que disminuye con el número de secuencias de péptidos en cada solicitud enviada. Shevchenko *et al.*, [Shevchenko et al 2001 y Liska et al 2003] propusieron el programa BLAST optimizado para análisis de péptidos secuenciados por espectrometría de masas (MS BLAST). Éste se encuentra disponible mediante un servicio web <http://dove.emblheidelberg>. A diferencia del BLAST y el FASTA, el MS BLAST permite introducir algunas modificaciones a los péptidos secuenciados que aumentan la confiabilidad de la identificación, como por ejemplo las puntuaciones individuales para los aminoácidos isobáricos (Leu/Ile y Gln/Lys) son sustituidas por sus valores promedios y así la puntuación del péptido T-S-L-V-K-M es igual a la del péptido T-S-I-V-G-M).

En marzo del 2003 surgió una nueva alternativa para la identificación eficiente de péptidos provenientes de organismos de genoma desconocido. Esta herramienta es la

secuenciación mediante MultiTag [Sunyaev et al 2003] que constituye una versión modificada de la estrategia de identificación mediante etiqueta de secuencia (*sequence tag*). Ésta permite identificar proteínas de muy baja abundancia y de las que sólo se pueden secuenciar de tres a cuatro aminoácidos, cuya identificación no sería confiable por MS BLAST.

ANTECEDENTES

La proteómica puede proporcionar una visión integrada de los productos de los genes de las bacterias y nos permite realizar un análisis global de diferentes condiciones. Muchas bacterias patógenas secretan factores de virulencia, por lo que el análisis del proteoma extracelular de estas bacterias puede facilitar los estudios de los procesos de patogénesis.

Las bacterias secretan proteínas que cumplen una variedad de funciones para asegurar la supervivencia de éstas en sus respectivos ambientes. Proteínas secretadas están involucrados en la absorción de nutrientes, comunicación de célula a célula, la desintoxicación y la eliminación de microorganismos que compiten en sus respectivos nichos. Las bacterias patógenas tienen la necesidad de secretar, además, un arsenal de diferentes factores necesarios para la adhesión y la invasión, así como la proliferación y la supervivencia en el huésped. El análisis de estas proteínas secretadas permite la identificación de nuevos factores de virulencia, que son cruciales para la supervivencia del patógeno durante la infección del huésped.

La comunicación entre las bacterias y sus hospederos, generalmente está mediada por las proteínas extracelulares. Un gran número de proteínas secretadas son toxinas que están relacionadas con el proceso de patogenicidad y con los procesos de simbiosis que realizan muchas bacterias del suelo como las rhizobiáceas. En consecuencia, diferentes grupos de investigación se han dedicado a analizar el sobrenadante de cultivo de estas bacterias.

En *E. coli* enteropatógena se ha reportado que varias proteínas en el proteoma extracelular pueden ser utilizadas como biomarcadores de la patogenicidad, entre estas encontramos: EspA, EspB, EspD, y Tir que son comunes para las cepas EHEC EDL933 y EPEC E2348/69, mientras que EspP y TagA son específicas para las cepas EHEC EDL933 y EspC es específica para EPEC E2348/69. También reportaron otras proteínas en el proteoma extracelular como chaperonas, proteínas de membrana externa, proteínas periplásmicas y proteínas citoplasmáticas [Li et al. 2004].

En *Herbaspirillum seropedicae* que es una bacteria endofítica se observó que el proteoma extracelular está constituido por un 45.4% de las proteínas involucradas en la motilidad de la célula, 36% involucradas en el metabolismo de la bacteria, de estas el 18,8% corresponden a proteínas de transportes, el 14.1% correspondió a las proteínas hipotéticas y las proteínas menos abundantes fueron las relacionadas con el procesamiento y almacenamiento de la información [Seixas et al. 2009].

En bacterias que realizan simbiosis con plantas, como los rhizobios, se ha sugerido que después que los factores de nodulación mantienen una actividad continúa al menos en las primeras 24 horas, se estimula la expresión de otras proteínas que también intervienen en el proceso de simbiosis [Deakin et al. 2009]. En correlación con lo anterior existen reportes de proteínas moduladoras del proceso de nodulación, clasificadas como polisacáridos de superficie (polisacáridos extracelulares, polisacáridos capsulares y lipolisacáridos) [Deakin et al. 2009].

En el proteoma extracelular de *Xanthomonas campestris* pv. *campestris*, fueron identificadas proteínas metabólicas, proteínas que se unen a ADN, proteínas de plegamiento, y proteínas citosólicas, se desconoce el papel que estas proteínas tienen en el secretoma de esta bacteria. También fueron identificadas peptidasas, enzimas para la degradación de la pared celular, estas son muy importantes porque ayudan a la bacteria a realizar su proceso de infección. [Watt et al. 2005].

En estudios realizados en *Mycobacterium tuberculosis*, se identificaron proteínas extracelulares que pueden ser útiles como posibles antígenos para el diagnóstico de la tuberculosis [Mattow et al. 2003].

Varias proteínas han sido identificadas en el proteoma extracelular de diferentes especies de *Listeria* las cuales han sido reportadas como proteínas que participan en la patogenicidad de *Listeria*, como IPEA (Lmo1847) que se ha sugerido está involucrada en el proceso de invasión, Oppa (Lmo2196), la cual es parte de una proteína de unión a oligopéptido que participa en la supervivencia intracelular y Auto (Lmo1076) que presenta actividad autolítica siendo necesaria para lograr una infección eficaz [Trost et al. 2005].

Erwinia chrysanthemi segrega varias proteínas en el medio de cultivo. Aproximadamente 100 proteínas fueron visualizados en geles 2-D del sobrenadante de *E. chrysanthemi* en condiciones de cultivo diferentes. Algunas proteínas secretadas; celulasa Cel5, las proteasas PrtA, PrtB, y proteínas AvrL, no parecen requerir componentes específicos de la planta como inductores de su síntesis. Por otro lado, en presencia de extracto de galacturonato y extractos de plantas se sabe que se induce la producción de pectinasas, de las cuales en este trabajo se reportan catorce proteínas inducibles. Todas correspondían a enzimas previamente descritas por su acción sobre polímeros pépticos. También fue demostrado en este trabajo que otros compuestos

presentes en paredes de células vegetales, tales como ramnosa y ácido ferúlico, inducen la producción de proteínas extracelulares [Kazemi-Pour et al. 2004].

Los primeros reportes relacionados con OMV fueron publicados en la década de los 80s, en estos se mencionaba que estas estructuras en bacterias eran las responsables de la transferencia lateral que ocurría entre las bacterias [Dorward et al. 1989, Lorentz et al. 1994, Yaron et al. 2000, Renelli et al. 2004]. Por otro lado Jan Wensink y Bernard Witholt en 1981, reportaron el posible mecanismo de formación para las vesículas de membrana externa en bacterias Gram negativas. Su hipótesis consistía en que las OMV se forman cuando la membrana externa se expande más rápido que la capa de peptidoglicano subyacente, en cuyo caso la membrana externa tenderá a separarse localmente desde la capa de peptidoglicano. Si este exceso de péptidoglicano en la membrana externa permanece durante algún tiempo, el resultado será la expansión de la membrana externa hasta que se repliega sobre sí misma y se libera en el medio [Wensink y Withold 1981]. Más tarde en 1998, Zhou et al., reportaron una hipótesis similar para la formación de las vesículas con otros resultados experimentales, (figura 4), demostrando adicionalmente que en el interior de las vesículas se secretaba ADN y proteínas.

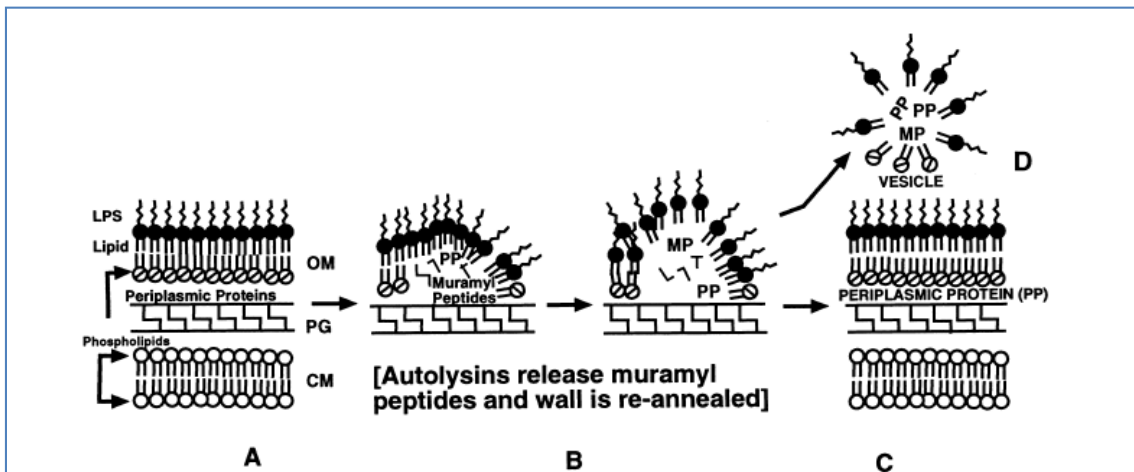


Figura4: Origen de las vesículas de membrana extracelular de las bacterias Gram-negativas. Durante el crecimiento, la pared celular se suprime y se libera el peptidoglicano. Los péptidos muramílicos liberados crean una presión de turgencia en la membrana externa. Si la presión no es aliviada por la absorción de los péptidos muramílicos, la membrana seguirá creciendo y finalmente se liberará al medio de cultivo. A: sección transversal de bacterias Gram-negativa que muestra la membrana citoplasmática, espacios periplásmico, peptidoglicano y la membrana externa. B: Turnover de peptidoglicano y el crecimiento inicial de la membrana externa. C: Liberación de vesículas que contienen los componentes de la membrana externa, el espacio periplásmico y componentes de peptidoglicano. CM, la membrana

citoplasmática, PG, peptidoglicano; PP, proteínas periplásmicas; OM, membrana externa; LPS, lipopolisacárido [Zhou et al. 1998].

Posteriores estudios demostraron que en las OMV las bacterias secretan ADN, ARN, proteínas y otras moléculas, que ayudan a la bacteria al proceso de colonización o infección de las células hospederas. Entre las proteínas identificadas en estas estructuras encontramos enzimas como hidrolasas [Li et al. 1998] y toxinas [Kuehn y Kesty 2005]. En consecuencia se han asignado varias funciones a las OMV como vehículos de transporte intercelular, ayudando en la digestión de nutrientes y la eliminación de organismos competidores así como en la respuesta a stress [McBroon y Kuehn 2007]. También se han reportado los proteomas de OMV para algunas bacterias Gram negativas, donde se han identificado proteínas citoplasmáticas, ribosomales entre ellas, proteínas de membrana externa, proteínas periplasmáticas así como proteínas de función desconocida [Vaughan et al. 2006, Lee et al. 2007, Nally et al. 2005, Ferrari G et al. 2006]

Hasta la fecha no se han reportado estudios relacionados con el proteoma de las proteínas secretadas en *R. etli* ni estudios relacionados con la identificación del proteoma de OMV en rizobiaceas. Por estos motivos consideramos que este estudio es muy importante porque sentará las pautas para nuevos trabajos de investigación y podremos conocer de forma global que está sucediendo en el medio extracelular de la bacteria.

OBJETIVOS

General

- Definir el secretoma de *Rhizobium etli* y contribuir a determinar el papel de las proteínas que lo constituyen en el funcionamiento celular de la bacteria.

Específicos

- Estudiar mediante proteómica el secretoma de *R. etli*, en vida libre (en condiciones de laboratorio) durante la fase logarítmica y estacionaria de crecimiento.
- Estudiar mediante proteómica el secretoma de *R. etli*, en etapas tempranas de la fijación de nitrógeno específicamente ante la presencia de inductores producidos por *Phaseolus vulgaris* como el flavonoide naringenina, esto con la finalidad de conocer los posibles productos proteicos sintetizados por la bacteria los cuales podrían tener un papel en el reconocimiento inicial.
- Comparar el secretoma de *R. etli*, en las diferentes etapas propuestas, tratando de sugerir la participación de los componentes proteicos en cada una de ellas.

MATERIALES Y MÉTODOS

En esta sección solo explicaremos los materiales y métodos que no están expuestos en el artículo.

Obtención de proteínas de la fracción periplasmática

Una vez centrifugado el cultivo, lavamos la pastilla dos veces con 1M de NaCl frío, se resuspendió la pastilla en una solución de Sacarosa fría (Sacarosa 20%, EDTA 1mM, inhibidores de proteasa 100µl), se incubó 5 minutos a temperatura ambiente, se centrifugó a 6000 xg por 10 minutos y resuspendimos la pastilla en agua bidestilada fría, incubamos 5 minutos a temperatura ambiente, centrifugamos como en el caso anterior pero durante 30 minutos, tomamos el sobrenadante, posteriormente precipitamos con 2.5 volúmenes de acetona al 100%, incubamos a -20°C durante toda la noche, centrifugamos a 8000 x g por 45 minutos, lavamos el pellet con acetona al 80 % y resuspendimos en buffer de solubilización (7 M urea; 2% CHAPS; 1 mM DTT). Se almacenó a -20°C hasta su uso.

Determinación de la concentración de proteínas

Se utilizó el método de Bradford para la determinación de la concentración de las proteínas [Bradford 1876].

Obtención de las vesículas de membrana externa

Después de crecido el cultivo hasta la fase exponencial o estacionaria, se centrifugó a 6000 xg por 10min. El sobrenadante fue filtrado con una membrana millipore de 0.45µm. Se liofilizó en una liofilizadora FreeZone 4.5, Labconco por 24h, para concentrar la muestra. Se ultracentrifugó en tubos transparentes (14x89mm; Beckman) a 100000xg por 2h a 4°C, se utilizó el rotor SW41. El botón obtenido se resuspendió en 500µl de Tris-HCl 1mM pH 8. La muestra se almacenó hasta uso a -20°C.

Microscopia electrónica de transmisión

Se analizaron muestras del cultivo de células y de las vesículas de membranas. Se colocaron 10µl del cultivo de células sobre una rejilla de cobre (EMS, Matfield, PA, USA) durante 5 min. , se agregó 5µl de ácido fosfotunstico al 2% por 1 min. Para el análisis de las vesículas de membrana se realizó el mismo procedimiento. La

microscopía electrónica fue realizada en un microscopio JEM1011 (JEOL, Japón) a 100 kilovoltios de voltaje de aceleración.

Inducción con naringenina

R. etli CE3, se cultivó en 2L de medio mínimo (0,2mM sulfato de magnesio, 0,1mM de hidrogeno fosfato de potasio, 0,2 mM de ácido succínico y 0,1 mM de cloruro de amonio) con agitación a 30°C. Cuando el cultivo alcanzó una densidad óptica de 0.5-0.6 se añadió naringenina a una concentración final de 1.5µM y se dejó creciendo por 12 horas más. Luego las proteínas fueron extraídas tanto del sobrenadante como de las OMV.

SDS-PAGE y digestión de las proteínas

La electroforesis fue llevada a cabo en geles de acrilamida SDS-PAGE 12,5%. Los geles fueron teñidos con Azul de Coomassie y fueron cortadas 10 bandas para cada carril del gel. Las bandas fueron transferidas a tubos eppendorf y cubiertas con 20 µl de Etanol al 20%. Las bandas fueron lavadas con una solución de Tris/HCl pH8 (100mM) y acetonitrilo 50/50, reducidas con 50mM DTT en 50 mM Tris/HCl pH 8 durante 30 min at 37 °C, alkyladas con 50mM iodoacetamida en 50mM Tris/HCl pH8 durante 30 min at 37°C en la oscuridad, e incubadas 5h a 37°C con 10ng/µl de tripsina (sequencing grade, Promega, Switzerland) en 20 mM de Tris/HCl pH8. Los fragmentos trópticos fueron extraídos con 20 µl de ácido fórmico al 20% (v/v) y analizados directamente por nano-LC-MS/MS.

Espectrometría de masas

La secuencia de los péptidos fue realizada en un espectrómetro de masas LTQ XL-Orbitrap (Thermo Fisher Scientific, Bremen; Germany) equipado con un sistema de nanoflujo (Flux Instruments, Reinach; Switzerland) y una fuente de ionización de electro spray operado a un voltaje de 1.6kV. La separación de los péptidos fue realizada en una nano columna magic C18 (5µm, 100Å, 0.075 x 70 mm) usando una velocidad de flujo de ~400nl/min y un gradiente lineal de 60min de 5 a 40% de acetonitrilo en H₂O y 0.1% ácido fórmico. La adquisición de datos fue realizada en el modo data dependant. Los espectros MS fueron adquiridos en el rango de masas de 300 a 1800m/z, con una resolución R=60000 y la fragmentación fue llevada a cabo por disociación inducida por colisión con gas helio en LTQ.

Identificación de proteínas

Fueron generados archivos mgf (Mascot generic files) por medio de pearl usando el software Hardklor, v1.25 (Copyright 2007-2009, Universidad de Washington).

Los datos de MS/MS (archivos mgf) fueron analizados usando el algoritmo PHENYX (Geneva Bioinformatics) mediante dos corridas contra la base de datos de proteínas de Uniprot_KB *Rhizobium etli* (cepas CFN42 y CIAT), SwissProt entries, se analizó también la secuencia inversa con el objetivo de eliminar la mayor cantidad posible de falsos negativos. Los parámetros para la primera corrida fueron: tolerancia de error 20ppm, modo de corte normal con un sitio perdido, fueron consideradas las siguientes modificaciones en amino ácidos (fijas en Cys_CAM, variable Oxidation_M). Los parámetros de la segunda corrida fueron: tolerancia de error 20ppm, modo de corte medio con 4 sitios perdidos, fueron consideradas las siguientes modificaciones en los amino ácidos (variable Cys_CAM, variable Deamid, variable phos, variable Oxidation_M). Las proteínas identificadas solo fueron aceptadas con un score mínimo de 10, lo que significa que al menos la secuencia de dos péptidos diferentes está presente en la proteína identificada. Este algoritmo nos permite realizar análisis semicuantitativos, ya que en los resultados que se exportan del PHENYX nos muestran la expresión relativa de cada proteína (PMSS, sumatoria de la calificación de los péptidos identificados) [Heller et al. 2007].

Las proteínas identificadas fueron agrupadas por COG (Cluster Orthologous Groups), según NCBI, cuya nomenclatura mostramos a continuación:

C	Producción y conversión de energía
D	Control del ciclo celular, división celular, división cromosomal
E	Metabolismo y transporte de aminoácidos
F	Metabolismo y transporte de nucleótidos
G	Metabolismo y transporte de carbohidratos
H	Metabolismo y transporte de Coenzimas
I	Transporte y metabolismo de Lípidos
J	Traducción, estructura ribosomal y biogénesis
K	Transcripción
L	Replicación, recombinación y reparación
M	Pared celular/membrana/biogénesis
N	Motilidad de la célula
O	Modificación postransduccional, plegamiento de proteína , chaperonas
P	Transporte y metabolismo de iones inorgánicos

Q	Transporte y catabolismo de metabolitos secundarios
R	Predicho solo como función general
S	Función desconocida
T	Mecanismo de transducción de señales
U	Transporte intracelular, secreción, transporte vesicular
V	Mecanismo de defensa
SC	Sin COG

Para el análisis de proteínas mediante geles de doble dimensión son necesarias concentraciones proteicas de 500 µg, por gel, esto hace que por la naturaleza de la muestra (productos de secreción celular colectados en los sobrenadantes), sean necesarios una gran cantidad de litros de cultivo celular lo cual hace poco operante el uso de esta metodología sobre todo en las etapas tempranas del crecimiento bacteriano o en los estudios de inducción mediante inductores como naringenina adicionalmente las limitaciones metodológicas como la imposibilidad de resolver proteínas de peso molecular bajo, así como las de pI extremo (por debajo de pH 3 ó por encima de pH 10). Por esta razón y después de realizar unos experimentos exploratorios decidimos utilizar una estrategia alterna utilizamos cromatografía líquida acoplada a espectrometría de masas, ya que mediante esta estrategia las cantidades de proteínas requeridas son mucho menores (nanogramos) en comparación con la electroforesis bidimensional. En todos los casos fueron analizadas dos replicas experimentales.

Análisis de la extracción butanólica de las OMV

La cepa fresca fue crecida durante toda la noche a 30 °C con agitación, en medio rico PY suplementado con ácido nalidíxico. Utilizando este cultivo se inocularon 2L de medio mínimo a una densidad óptica de 0.5 para purificar las OMV tanto de la cepa silvestre de *R. etli* como de la mutante en Nod A (CE3 NodA⁻) luego fueron extraídas con butanol. El extracto de butanol fue secado al vacío y resuspendido en agua destilada estéril y analizado mediante HPLC usando una columna de fase reversa C-18 (Nova-Pack C18 3.9x150mm Waters), usando un detector (Waters 996 a 215 nm) y una corrida isocrática con CH₃CN al 20%. La fracción de interés fue colectada y liofilizada. Posteriormente fue disuelta en agua destilada estéril y se preparó una mezcla de la muestra con 50% de metanol en la proporción 1:1. Se tomaron 5µl de esta nueva mezcla y se inyectó al espectrómetro de masas. La corrida fue realizada en modo positivo mediante DIC usando un LTQ XL-Orbitrap (Thermo Fisher Scientific, Bremen; Germany). Las masas obtenidas fueron analizadas con la ayuda de la base de

datos Glycomod, mediante este software podemos obtener las posibles estructuras a partir de las masas experimentales obtenidas por MS/MS. (http://www.expasy.org/tools/glycomod/glycomod_masses.html).

Ensayo de nodulación de las plantas

Los ensayos fueron realizados con semillas de *Phaseolus vulgaris* Negro Xamapa, las cuales fueron tratados y germinados como describió previamente Martínez et al 1985. Las semillas germinadas fueron sembradas en frascos con medio mínimo sólido e inoculadas con 50 µl de OMV purificadas, para el control positivo las plántulas de frijol fueron inoculadas con células de *R. etli*, cultivadas durante 12 hrs., en medio rico PY (10^5 bacteria por planta aproximadamente).

Cromatografía de capa fina TLC

Cultivos de *R. etli* fueron crecidos en presencia de [^{14}C] acetato de sodio (1µCi/ml) en medio mínimo. Cuando el cultivo alcanzó la fase exponencial se le adicionó Naringenina (1,2µM). El sobrenadante fue liofilizado y fueron purificadas las OMV. Posteriormente se extrajo el sobrenadante y las OMV con butanol. La TLC fue realizada como describe Roche et al en 1991.

RESULTADOS

En este capítulo mostraremos los resultados que corresponden al análisis de los proteomas de *R. etli* en las fases de crecimiento exponencial y estacionario, para las fracciones del sobrenadante. Estos resultados fueron publicados en la revista *Proteome Science* 2010; 8:51. Ver Anexos 1.

RESULTADOS ADICIONALES

Con la intención de obtener más información acerca de las proteínas secretadas por *R. etli* y además para tratar de explicar por qué en fase de crecimiento exponencial encontramos muchas proteínas ribosomales, decidimos realizar otros estudios. En reportes previos encontramos que las bacterias gram negativas secretan al medio vesículas de membrana externa (OMV del inglés outer membrane vesicles). Sin embargo, hasta el momento no existen reportes sobre la presencia de OMV en Rhizobiaceas, por esta razón primero decidimos investigar si *R. etli* produce y secreta estas estructuras. Para la búsqueda y posterior purificación de las vesículas de membrana utilizamos el protocolo descrito previamente [Amanda et al 2007] con las siguientes modificaciones. Después de la separación mediante centrifugación del sobrenadante del paquete celular, liofilizamos el sobrenadante, posteriormente lo filtramos con un filtro de 0.22 μ m con la finalidad de separar las posibles células que permanecieran en él. Centrifugamos a 25 000xg durante 2 hrs., dicho sobrenadante, con la intención de precipitar las posibles vesículas. La pastilla fue resuspendida en buffer fosfato, tomando 5 μ l para realizar la microscopía electrónica de transmisión, fig 5. Observando en el sobrenadante de *R. etli* la presencia de estas estructuras. Estas vesículas tienen un tamaño aproximado entre 40 y 100 nm de diámetro, aunque está reportado que su tamaño puede oscilar entre 30 y 200nm aproximadamente [Lee et al. 2007]. Después de comprobar que *R. etli* secreta OMV, el siguiente objetivo a cumplir fue el análisis e identificación de las proteínas presentes en dichas vesículas.

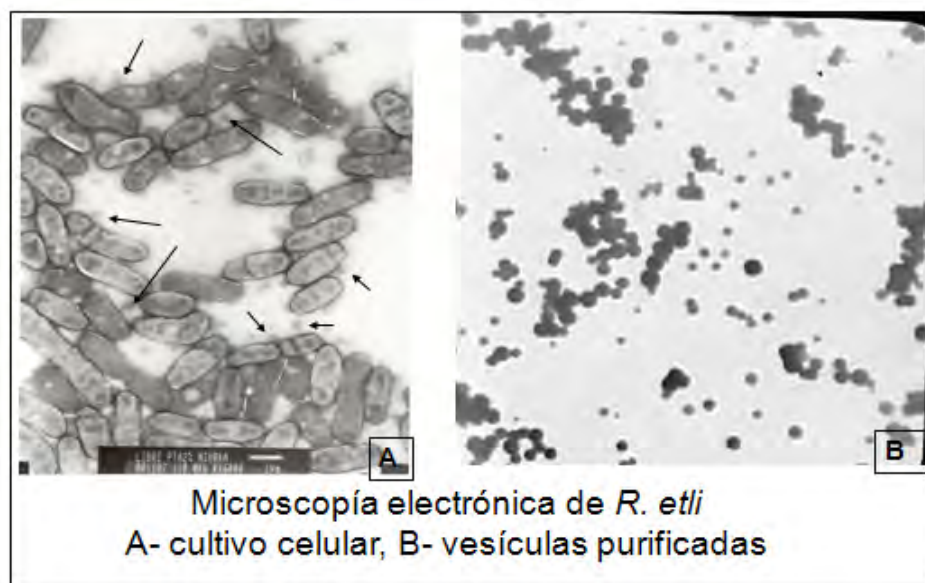


Figura 4: Microscopía electrónica de *R. etli*.

Para realizar la identificación de las proteínas presentes en la OMV, colectamos 4 litros del sobrenadante de medio mínimo por cada etapa de crecimiento a analizar, fase de crecimiento exponencial (6H) y estacionaria (24H). Así mismo, analizamos las fracciones periplasmáticas correspondientes a estas fases de crecimiento. De cada una de las muestras obtenidas se procedió extraer el complejo proteico para la posterior determinación de su concentración por el método de Bradford. Una vez determinada la concentración proteica se procedió a la separación del complejo proteico mediante electroforesis de una sola dimensión (SDS-PAGE), para lo cual se cargaron 20 microgramos por carril de cada muestra proteica proveniente tanto de las OMV, como de las fracciones periplásmicas, la electroforesis se realizó como se reporta en la sección de Material y Métodos. Después de realizada la electroforesis los geles, fueron teñidos con azul de Coomassie. Los geles fueron cortados en zonas de 5 milímetros de grosor. Las bandas fueron digeridas con la enzima tripsina a una concentración final de 10 ng/μl. Los péptidos producto de la digestión fueron extraídos con la ayuda de 20% de ácido fórmico.

Cada una de la mezcla de péptidos obtenidos de cada porción o zona del gel fueron secuenciadas mediante espectrometría de masas (LC-MS/MS). Para este propósito se utilizó un espectrómetro de masas ESI-LTQ-Orbitrap (CID_LTQ_scan_LTQ), los parámetros usados se muestran en Materiales y Métodos. Los resultados obtenidos se muestran en las tablas 1 y 2 (ver además anexo 2). En todos los casos se realizaron dos replicas experimentales, obteniéndose una reproducibilidad entre las replicas entre 80 y 85%.

Proteoma de *R. etli* en fase de crecimiento exponencial

Proteoma de la fracción periplásmica en células colectadas en fase de crecimiento exponencial

Las proteínas de la fracción periplásmica fueron extraídas según el protocolo reportado por Watt et al 2005. La separación proteica mediante electroforesis, la tinción e identificación de las proteínas se realizó como se menciona en el párrafo anterior. Mediante este procedimiento se identificaron un total de 701 proteínas, las cuales fueron agrupadas por COG según NCBI.

Las proteínas fueron agrupadas en subgrupos o categorías funcionales (tabla 1, anexo 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 94 proteínas, de las cuales 69 pertenece al grupo funcional “traducción, estructura ribosomal y biogénesis”, 18 al grupo funcional “transcripción” y 7 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas 109 proteínas. De estas 20 proteínas pertenecen al grupo funcional “mecanismo de traducción de señales”, 24 proteínas pertenecen al grupo funcional “pared celular/membrana/biogénesis”, 43 al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas”, 3 al grupo funcional “control del ciclo celular, división celular, división cromosomal”, 13 al grupo funcional “motilidad de la célula” y 6 al grupo funcional “transporte intracelular, secreción, transporte vesicular”, no se identificó ninguna proteína perteneciente al grupo funcional “mecanismo de defensa”.
- Metabolismo: en este grupo fueron identificadas 344 proteínas. De estas 43 pertenecen al grupo funcional “producción y conversión de energía”, 102 al grupo funcional “metabolismo y transporte de aminoácidos”, 29 al grupo funcional “metabolismo y transporte de nucleótidos”, 89 al grupo funcional “metabolismo y transporte de carbohidratos”, 22 al grupo funcional “metabolismo y transporte de coenzimas”, 21 al grupo funcional “transporte y metabolismo de lípidos”, 26 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 12 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.
- Pobremente caracterizadas: en este grupo fueron identificadas 84 proteínas correspondientes a las proteínas con función desconocidas y a las proteínas que fueron pobremente caracterizadas.
- Sin COG: en este grupo fueron identificadas 70 proteínas.

Es interesante resaltar que las proteínas identificadas en esta condición correspondientes al grupo “proteínas ribosomal y biogénesis” son muy abundantes, representando 9,84% del total de proteínas identificadas en esta fracción. El gran número de proteínas ribosomales identificadas como las 30S (16) , 50S (19), factores de elongación (Efp, FusA, Tsf, Tuf1) y otras proteínas pertenecientes a este grupo funcional coincide con

los resultados obtenidos para el proteoma del sobrenadante de *R. etli* en fase exponencial, aunque no sabemos por qué y para qué estas proteínas son secretadas.

Para la categoría funcional “procesos celulares y señalización” fueron identificadas un gran número de proteínas que pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas”, éste fue el grupo funcional más representado en esta categoría funcional. También identificamos un gran número de transportadores. Resultados esperados ya que en el espacio periplásmico ocurre el plegamiento de muchas proteínas, el transporte de elementos esenciales para la bacteria así como se encuentran un gran número de enzimas necesarias para la defensa de la bacteria.

Proteoma de las OMV en fase de crecimiento exponencial

En el extracto proteico obtenido a partir de las vesículas de membrana externa identificamos 454 proteínas. Estas proteínas fueron agrupadas primero por COG y luego agrupadas por categorías funcionales como se muestra a continuación (tabla 1, anexo 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 79 proteínas. 59 proteínas pertenecen al grupo funcional “estructura ribosomal y biogénesis”, 12 pertenecen al grupo funcional “transcripción” y 8 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 92 proteínas. 2 proteínas pertenecen al grupo funcional “control del ciclo celular, división celular, división cromosomal”, 5 pertenecen al grupo funcional “mecanismo de transducción de señales”, 25 al grupo funcional “pared celular/membrana/biogénesis”, 20 pertenecen al grupo funcional “motilidad de la célula”, 31 pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” y 9 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. No se identificó ninguna proteína perteneciente al grupo funcional “mecanismo de defensa”.
- Metabolismo: en este grupo fueron identificadas un total de 198 proteínas, de ellas 43 pertenecen al grupo funcional “producción y conversión de energía”, 60 al grupo funcional “metabolismo y transporte de aminoácidos”, 14 al grupo funcional “metabolismo y transporte de nucleótidos”, 39 al grupo funcional “metabolismo y transporte de carbohidratos”, 7 al grupo funcional “metabolismo

y transporte de coenzimas”, 15 al grupo funcional “transporte y metabolismo de lípidos”, 14 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 6 al grupo funcional “ transporte y catabolismo de metabolitos secundarios”.

- Pobremente caracterizadas: en este grupo fueron identificadas 36 proteínas pertenecientes a los grupos funcionales “función desconocida” y “pobremente caracterizadas.
- Sin COG: en este grupo fueron identificadas 49 proteínas.

Debemos destacar que en esta fracción las proteínas correspondientes al grupo funcional “traducción, estructura ribosomal y biogénesis” representan el 12,9% del total de proteínas identificadas para esta fracción y el 74,6% para la categoría funcional “almacenamiento y procesamiento de la información”. Estos resultados son similares a los obtenidos en la fracción periplásmica para la fase exponencial, lo que sugiere que las OMV se forman a partir del periplasma.

Comparación entre los proteomas de la fracción periplásmica y las OMV en fase exponencial

Cuando comparamos los proteomas de la fracción periplásmica y el de las vesículas de membrana externa, encontramos 304 proteínas en común. Estas están distribuidas de la siguiente manera según las categorías funcionales (tabla 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 56 proteínas comunes a las dos fracciones. 40 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 11 pertenecen al grupo funcional “transcripción” y 5 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 57 proteínas comunes a las dos fracciones. 3 pertenecen al grupo funcional “mecanismo de transducción de señales”, 16 al grupo funcional “pared celular/membrana/biogénesis”, 11 pertenecen al grupo funcional “motilidad de la célula”, 23 pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” y 4 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. No se encontraron proteínas comunes para el grupo funcional “control del ciclo celular, división celular, división cromosomal”.

- **Metabolismo:** en este grupo fueron identificadas un total de 143 proteínas comunes a las dos fracciones, de ellas 26 pertenecen al grupo funcional “producción y conversión de energía”, 48 al grupo funcional “metabolismo y transporte de aminoácidos”, 7 al grupo funcional “metabolismo y transporte de nucleótidos”, 32 al grupo funcional “metabolismo y transporte de carbohidratos”, 4 al grupo funcional “metabolismo y transporte de coenzimas”, 11 al grupo funcional “transporte y metabolismo de lípidos”, 12 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 3 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.
- **Pobremente caracterizadas:** en este grupo fueron identificadas 23 proteínas comunes a las dos fracciones, pertenecientes a los grupos funcionales “función desconocida” y “pobremente caracterizadas”.
- **Sin COG:** en este grupo fueron identificadas 25 proteínas comunes para las dos fracciones analizadas.

De las proteínas comunes en ambas condiciones de estudio las más abundantes fueron las identificadas para la categoría funcional “metabolismo”, representando el 47% del total de proteínas comunes para los proteomas de ambas condiciones. El grupo funcional “traducción, estructura ribosomal y biogénesis” representa el 13,2 % del total de las proteínas identificadas que son comunes a ambas fracciones y el 71,4 % de la categoría funcional “almacenamiento y procesamiento de la información”.

De acuerdo al análisis anterior podemos decir que para la fase de crecimiento exponencial se identificaron 397 proteínas que son específicas para fracción periplásmica en esta fase de crecimiento lo que significa que las proteínas que son secretadas en las OMV utilizan un mecanismo específico a través del periplasma, el cual, aun no ha sido descrito. Las proteínas de la categoría funcional “almacenamiento y procesamiento de la información” representan el 9,82 % del total de proteínas específicas para la fracción periplásmica. Las proteínas de la categoría funcional “procesos celulares y señalización” representan el 12,8% del total de proteínas específicas para la fracción periplásmica, la categoría funcional “metabolismo” está representada por el 50,6% y la categoría funcional “pobremente caracterizada” representa el 15,4% del total de proteínas específicas identificadas para esta fracción. Las proteínas sin COG en este caso representan el 11,3% del total de proteínas específicas identificadas para la fracción periplásmica (tabla 2).

Para la fracción de OMV se identificaron 150 proteínas específicas para esta fracción. El 15,3% de estas proteínas pertenecen a la categoría funcional “almacenamiento y procesamiento de la información”, en esta categoría la mayor cantidad de proteínas identificadas pertenece al grupo funcional “traducción, estructura ribosomal y biogénesis” representando un 82,6%. El 23,3 % del total de proteínas específicas para esta fracción pertenece a la categoría funcional “procesos celulares y señalización”. El 36.6% pertenece a la categoría funcional “metabolismo”, el 8,6% pertenecen a la categoría funcional “pobremente caracterizadas” y sin COG encontramos el 16% del total de proteínas específicas de las vesículas de membrana externa en la fase de crecimiento exponencial (tabla 2). En las dos fracciones comparadas aquí podemos apreciar que presentan una distribución similar en cuanto a la composición de los proteomas. Adicionalmente estos datos sugieren que las proteínas presentes en las OMV son secretadas a través del periplasma.

Proteoma de *R. etli* en fase de crecimiento estacionaria

Proteoma de la fracción periplásmica en fase estacionaria

En la fracción periplásmica obtenida de células cosechadas en la fase de crecimiento estacionaria fueron identificadas 730 proteínas. Las proteínas fueron agrupadas por COG y en categorías funcionales (tabla 1, anexo 2); cómo se explica a continuación:

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 78 proteínas. 55 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 18 pertenecen al grupo funcional “transcripción” y 5 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 142 proteínas. 6 pertenecen al grupo funcional “control del ciclo celular, división celular, división cromosomal”, 35 pertenecen al grupo funcional “mecanismo de transducción de señales”, 33 al grupo funcional “pared celular/membrana/biogénesis”, 14 pertenecen al grupo funcional “motilidad de la célula”, 51 pertenecen al grupo funcional “modificación postranducciónal, plegamiento de proteínas, chaperonas” y 2 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. En el grupo funcional “mecanismo de defensa” solo fue identificada una proteína.

- **Metabolismo:** en este grupo fueron identificadas un total de 326 proteínas, de ellas 33 pertenecen al grupo funcional “producción y conversión de energía”, 97 al grupo funcional “metabolismo y transporte de aminoácidos”, 30 al grupo funcional “metabolismo y transporte de nucleótidos”, 74 al grupo funcional “metabolismo y transporte de carbohidratos”, 27 al grupo funcional “metabolismo y transporte de coenzimas”, 26 al grupo funcional “transporte y metabolismo de lípidos”, 23 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 16 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.
- **Pobrementemente caracterizadas:** en este grupo fueron identificadas 95 proteínas, pertenecientes a los grupos funcionales “función desconocida” 41 y al grupo de proteínas “pobrementemente caracterizadas” 54 proteínas.
- **Sin COG:** en este grupo fueron identificadas 89 proteínas.

Para la fracción periplásmica en la fase de crecimiento estacionaria el número de proteínas identificadas correspondientes a la característica funcional “pobrementemente caracterizadas” sumado a las proteínas sin COG representan el 25,2% del total de proteínas identificadas. Mientras que las características funcionales “almacenamiento y procesamiento de la información” representan el 10,7%, “procesos celulares y señalización” representa el 19,5 %, en esta categoría funcional el grupo más representado son las chaperonas, resultado esperado ya que en el periplasma ocurre el plegamiento de un gran número de proteínas. La categoría funcional “metabolismo” representa el 44,7% del total de proteínas identificadas para la fracción periplásmica en esta fase de crecimiento, este resultado es muy lógico si tenemos en cuenta que en el periplasma y anclados en membrana, encontramos muchos transportadores responsables del suministro de elementos esenciales para la bacteria como es el caso de azúcares, iones metálicos, aminoácidos, entre otros (tabla 2, anexo 2). También fueron identificadas una gran cantidad de enzimas clasificadas como peptidasas, hidrolasas, fosfatasas, fosfodiesterasas, así como proteínas de membrana externa.

Proteoma de las OMV en fase estacionaria

En el proteoma de las vesículas de membrana externa para la fase de crecimiento estacionaria fueron identificadas un total de 400 proteínas. Al realizar la distribución por grupos y características funcionales, estas proteínas quedaron distribuidas de la siguiente manera (tabla 1, anexo 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 38 proteínas. 30 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 5 pertenecen al grupo funcional “transcripción” y 3 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 78 proteínas. 9 pertenecen al grupo funcional “mecanismo de transducción de señales”, 20 al grupo funcional “pared celular/membrana/biogénesis”, 18 pertenecen al grupo funcional “motilidad de la célula”, 28 pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” y 2 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. En el grupo funcional “mecanismo de defensa” solo fue identificada una proteína, no encontramos proteínas para el grupo funcional “control del ciclo celular, división celular, división cromosomal”.
- Metabolismo: en este grupo fueron identificadas un total de 177 proteínas, de ellas 19 pertenecen al grupo funcional “producción y conversión de energía”, 58 al grupo funcional “metabolismo y transporte de aminoácidos”, 8 al grupo funcional “metabolismo y transporte de nucleótidos”, 51 al grupo funcional “metabolismo y transporte de carbohidratos”, 7 al grupo funcional “metabolismo y transporte de coenzimas”, 4 al grupo funcional “transporte y metabolismo de lípidos”, 22 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 8 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.
- Pobremente caracterizadas: en este grupo fueron identificadas 35 proteínas, pertenecientes a los grupos funcionales “función desconocida” 20 y “pobremente caracterizadas” 17.
- Sin COG: en este grupo fueron identificadas 70 proteínas.

Para esta fracción estudiada, podemos apreciar que la característica funcional mejor representada corresponde a “metabolismo” ya que representa el 44,3 % del total de proteínas identificadas, correspondientes a un gran número de transportadores de aminoácidos, transportadores de carbohidratos y transportadores de iones inorgánicos. Seguidas por las “pobremente caracterizadas” y sin COG que en su totalidad representan el 26,8% del total de proteínas identificadas para la fracción de OMV en fase de crecimiento estacionario.

Comparación entre los proteomas de la fracción periplásmica y las OMV en fase estacionaria

Cuando comparamos los proteomas de la fracción periplásmica y el de las vesículas de membrana externa, para la fase estacionaria de crecimiento encontramos 274 proteínas en común. Estas están distribuidas de la siguiente manera según las categorías funcionales (tabla 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 28 proteínas comunes a las dos fracciones. 22 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 3 pertenecen al grupo funcional “transcripción” y 3 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 60 proteínas comunes a las dos fracciones. 7 pertenecen al grupo funcional “mecanismo de transducción de señales”, 13 al grupo funcional “pared celular/membrana/biogénesis”, 10 pertenecen al grupo funcional “motilidad de la célula”, 27 pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” y 2 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. Para el grupo funcional “mecanismo de defensa solo se encontró una proteína en común en los proteomas de las dos fracciones estudiadas”.
- Metabolismo: en este grupo fueron identificadas un total de 127 proteínas comunes a las dos fracciones, de ellas 12 pertenecen al grupo funcional “producción y conversión de energía”, 48 al grupo funcional “metabolismo y transporte de aminoácidos”, 6 al grupo funcional “metabolismo y transporte de nucleótidos”, 34 al grupo funcional “metabolismo y transporte de carbohidratos”, 3 al grupo funcional “metabolismo y transporte de coenzimas”, 3 al grupo funcional “transporte y metabolismo de lípidos”, 16 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 5 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.
- Pobremente caracterizadas: en este grupo fueron identificadas 28 proteínas comunes a las dos fracciones, pertenecientes a los grupos funcionales “función desconocida” y “pobremente caracterizadas”.

- Sin COG: en este grupo fueron identificadas 31 proteínas comunes para las dos fracciones analizadas.

Las proteínas más abundantes que son comunes a las fracciones periplásmicas y OMV en fase de crecimiento estacionario corresponde a la categoría funcional “metabolismo” y estas representan el 46,4% del total de proteínas identificadas que son compartidas por ambas fracciones en fase de crecimiento estacionario.

Cuando realizamos esta comparación entre las proteínas identificadas para las fracciones periplásmicas y OMV en fase de crecimiento estacionario, encontramos que para la fracción periplásmica fueron identificadas un total de 451 proteínas específicas, para la fracción OMV fueron identificadas 129 proteínas específicas, que solo están presentes en esta fracción. Si analizamos las categorías funcionales observamos que para la fracción periplásmica el 11,1% del total de proteínas encontradas específicas para esta fracción pertenece a la categoría “almacenamiento y procesamiento de la información”, el 17,1% pertenece a la categoría “procesos celulares y señalización”, en esta categoría no se encontraron proteínas pertenecientes a los grupos funcionales “mecanismo de defensa” y “transporte intracelular, secreción, transporte vesicular”; la categoría funcional “metabolismo” está representada en un 44,1% con respecto al total de proteínas específicas para esta fracción y las categorías “pobremente caracterizadas” y sin COG en conjunto representan el 27,7% del total de proteínas específicas identificadas para la fracción periplásmica en fase de crecimiento estacionario. En el caso de las proteínas específicas para la fracción OMV las categorías funcionales están distribuidas de la siguiente manera: 7,8% de estas proteínas pertenecen a la categoría funcional “almacenamiento y procesamiento de la información”, no estando representado el grupo funcional “replicación, recombinación y reparación”; el 14% corresponde a la categoría funcional “procesos celulares y señalización”, no encontrándose proteínas para los grupos funcionales “control del ciclo celular, división celular, división cromosomal”, “mecanismo de defensa” y “transporte intracelular, secreción, transporte vesicular”; la categoría funcional metabolismo se encuentra representada en un 38,8% y el 39,6 % corresponde a las categorías pobremente caracterizadas y sin COG.

Comparación de los proteomas de *R. etli* en las fases de crecimiento exponencial y estacionaria

Comparación de los proteomas de las fracciones periplásmica en ambas fases de crecimiento.

Al comparar los proteomas de las fracciones periplásmicas en ambas fases de crecimiento, analizamos tanto las proteínas que son comunes para ambas fases así como las proteínas que solo fueron identificadas en la fase exponencial y la estacionaria. Para la fracción periplásmica encontramos 474 proteínas comunes entre los dos proteomas para las dos fases de crecimiento; encontramos 227 proteínas específicas para el periplasma en fase exponencial y 256 proteínas específicas para el periplasma en fase estacionaria. A continuación se muestra un análisis detallado de estos resultados.

Las 474 proteínas que son comunes a las dos fases de crecimiento para la fracción periplásmica fueron agrupadas por COG y por categorías funcionales (tabla1, anexo 2), como sigue:

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 55 proteínas comunes. 43 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 8 pertenecen al grupo funcional “transcripción” y 4 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 85 proteínas comunes a las dos fases. 7 pertenecen al grupo funcional “control del ciclo celular, división celular, división cromosomal”, 17 pertenecen al grupo funcional “mecanismo de transducción de señales”, 18 al grupo funcional “pared celular/membrana/biogénesis”, 10 pertenecen al grupo funcional “motilidad de la célula”, 35 pertenecen al grupo funcional “modificación postranducción, plegamiento de proteínas, chaperonas” y 2 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. Para el grupo funcional “mecanismo de defensa no se encontraron proteínas comunes a los proteomas de la fracción periplásmica para ambas fases de crecimiento”.
- Metabolismo: en este grupo fueron identificadas un total de 229 proteínas comunes a las dos fases de crecimiento, de ellas 24 pertenecen al grupo funcional “producción y conversión de energía”, 70 al grupo funcional “metabolismo y transporte de aminoácidos”, 20 al grupo funcional “metabolismo y transporte de nucleótidos”, 59 al grupo funcional “metabolismo y transporte de carbohidratos”, 10 al grupo funcional “metabolismo y transporte de coenzimas”, 17 al grupo funcional “transporte y metabolismo de lípidos”, 18

al grupo funcional “transporte y metabolismo de iones inorgánicos” y 11 al grupo funcional “ transporte y catabolismo de metabolitos secundarios”.

- Pobremente caracterizadas: en este grupo fueron identificadas 58 proteínas comunes a las dos fracciones, pertenecientes a los grupos funcionales “función desconocida” y “pobremente caracterizadas”.
- Sin COG: en este grupo fueron identificadas 47 proteínas comunes para las dos fases de crecimiento analizadas.

En el estudio de las proteínas específicas para periplasma en fase de crecimiento exponencial fueron identificadas 40 proteínas para la categoría funcional “almacenamiento y procesamiento de la información”, lo cual representa el 17,6 % del total de proteínas específicas para el periplasma en esta fase de crecimiento. Con relación al periplasma en fase de crecimiento estacionario fueron encontradas 23 proteínas, que representan el 8,9% del total de proteínas específicas identificadas para periplasma en fase de crecimiento estacionario. Correspondientes a la categoría funcional “procesos celulares y señalización”, fueron identificadas 23 y 57 proteínas específicas para el periplasma en fase exponencial y estacionaria respectivamente, lo que representa el 10,1% y el 22,3% del total de proteínas identificadas específicas para el periplasma en cada fase de crecimiento respectivamente. Cuando analizamos la categoría funcional “metabolismo” encontramos para la fase exponencial 115 proteínas identificadas lo que representa el 50,7 % de las proteínas identificadas específicas para la fase de crecimiento exponencial y encontramos 97 proteínas para la fase estacionaria lo que representa el 37,9% del total de proteínas específicas identificadas para la fase estacionaria. Para las categorías “pobremente caracterizada” y sin COG, fueron identificadas 49 proteínas para la fase exponencial y 79 proteínas para la fase estacionaria; lo que representa el 21,6% y el 30,9% de las proteínas específicas identificadas para la fracción periplásmica en las fases de crecimiento exponencial y estacionario respectivamente.

Comparación de los proteomas de las OMV en ambas fases de crecimiento

Cuando comparamos los proteomas de las OMV para las fases de crecimiento exponencial y estacionario, encontramos un total de 219 proteínas comunes para los proteomas en ambas fases de crecimiento. Las proteínas fueron agrupadas por COG y por categorías funcionales, las proteínas quedaron distribuidas como sigue (tabla 1, anexo 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 27 proteínas comunes. 21 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 3 pertenecen al grupo funcional “transcripción” y 3 al grupo funcional “replicación, recombinación y reparación”. Es interesante destacar que algunas de las proteínas presentes en esta categoría también se encuentran presentes en las fracciones periplásmicas y la del sobrenadante. Pero lo más interesante es que para la fase exponencial las proteínas ribosomales se encuentran en mayor concentración.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 52 proteínas comunes a las dos fases. 3 pertenecen al grupo funcional “mecanismo de transducción de señales”, 15 al grupo funcional “pared celular/membrana/biogénesis”, 13 pertenecen al grupo funcional “motilidad de la célula”, 19 pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” y 2 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. Para los grupos funcionales “mecanismo de defensa” y “control del ciclo celular, división celular, división cromosomal” no se encontraron proteínas comunes en los proteomas de las vesículas de membrana externa para ambas fases de crecimiento.
- Metabolismo: en este grupo fueron identificadas un total de 96 proteínas comunes a las dos fases de crecimiento, de ellas 13 pertenecen al grupo funcional “producción y conversión de energía”, 36 al grupo funcional “metabolismo y transporte de aminoácidos”, 3 al grupo funcional “metabolismo y transporte de nucleótidos”, 24 al grupo funcional “metabolismo y transporte de carbohidratos”, 2 al grupo funcional “metabolismo y transporte de coenzimas”, 3 al grupo funcional “transporte y metabolismo de lípidos”, 12 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 3 al grupo funcional “transporte y catabolismo de metabolitos secundarios”. La gran cantidad de proteínas metabólicas identificadas en vesículas sugiere que estas estructuras no se secreta por la bacteria por azar, pero desconocemos la función de estas proteínas en estas estructuras.
- Pobremente caracterizadas: en este grupo fueron identificadas 16 proteínas comunes a las dos fracciones, pertenecientes a los grupos funcionales “función desconocida” y “pobremente caracterizadas”.

- Sin COG: en este grupo fueron identificadas 28 proteínas comunes para las dos fases de crecimiento analizadas.

Cuando comparamos los proteomas de las fracciones OMV para las fases de crecimiento exponencial y estacionario, identificamos como proteínas específicas para las OMV producidas por células en fase de crecimiento exponencial un total de 150 proteínas, mientras que para la fase estacionaria fueron identificadas como proteínas específicas un total de 129 proteínas. El 22,1% del total de proteínas específicas identificadas para las OMV en fase exponencial corresponde a la categoría funcional “almacenamiento y procesamiento de la información”, donde la mayor cantidad de proteínas fue aportada por el grupo funcional “traducción, estructura ribosomal y biogénesis, con 38 proteínas. Mientras que para la fase estacionaria el 6,1% pertenecen a la categoría funcional “almacenamiento y procesamiento de la información, en esta categoría no se encontraron proteínas pertenecientes al grupo funcional “replicación, recombinación y reparación”. En la categoría funcional procesos celulares y señalización fue clasificado el 17% de proteínas identificadas en las OMV de células cosechadas en la fase exponencial mientras que el 14,4% fue identificado para la fase estacionaria. El 43,4% del total de proteínas específicas para las OMV en fase exponencial correspondió a la categoría funcional “metabolismo”, para la fase estacionaria fueron identificadas para la misma categoría funcional el 44,8%. En la categoría funcional pobremente caracterizadas y sin COG fueron identificadas para la fase exponencial el 17,4% del total de las proteínas específicas para la fase exponencial mientras que para la fase estacionaria las proteínas identificadas representan el 34,8% del total de proteínas específicas identificadas para las OMV en fase estacionaria de crecimiento.

Proteoma de *R. etli* en presencia y ausencia de naringenina

Con la intención de cumplir con el objetivo específico “Estudiar mediante proteómica el secretoma de *R. etli*, en etapas tempranas de la fijación de nitrógeno específicamente ante la presencia de inductores producidos por *Phaseolus vulgaris* como el flavonoide naringenina, se analizaron los secreteomas de *R. etli* en presencia y ausencia de naringenina en el sobrenadante y OMV.

Proteoma de la fracción extracelular en presencia de naringenina

Para el proteoma extracelular de *R. etli* en presencia de naringenina fueron identificadas un total de 628 proteínas; estas fueron clasificadas o agrupadas por COG y por categorías funcionales, quedando distribuidas de la siguiente manera (tabla 3, anexo 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 86 proteínas. 63 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 12 pertenecen al grupo funcional “transcripción” y 11 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 124 proteínas. 3 proteínas pertenecen al grupo funcional “control del ciclo celular, división celular, división cromosomal”, 9 pertenecen al grupo funcional “mecanismo de transducción de señales”, 35 al grupo funcional “pared celular/membrana/biogénesis”, 22 pertenecen al grupo funcional “motilidad de la célula”, 44 pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” y 10 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. En el grupo funcional “mecanismo de defensa” solo fue identificada una proteína.
- Metabolismo: en este grupo fueron identificadas un total de 266 proteínas, de ellas, 51 pertenecen al grupo funcional “producción y conversión de energía”, 78 al grupo funcional “metabolismo y transporte de aminoácidos”, 15 al grupo funcional “metabolismo y transporte de nucleótidos”, 58 al grupo funcional “metabolismo y transporte de carbohidratos”, 12 al grupo funcional “metabolismo y transporte de coenzimas”, 16 al grupo funcional “transporte y metabolismo de lípidos”, 25 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 11 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.
- Pobremente caracterizadas: en este grupo fueron identificadas 57 proteínas, pertenecientes a los grupos funcionales “función desconocida” 25 y “pobremente caracterizadas” 32.
- Sin COG: en este grupo fueron identificadas 95 proteínas.

Para el proteoma extracelular de *R. etli* en presencia de naringenina la categoría funcional más representada fue “metabolismo” que corresponde a un 42,4% del total de proteínas identificadas para esta condición. Las proteínas correspondientes a las

categorías funcionales pobremente caracterizadas y sin COG representan el 24,2% del total de proteínas identificadas en esta condición.

Proteoma de la fracción extracelular en ausencia de naringenina

Cuando analizamos el proteoma extracelular de *R. etli* en ausencia de naringenina fueron identificadas un total de 625 proteínas; estas fueron clasificadas por COG y categorías funcionales, quedando clasificado de la siguiente manera (tabla 3, anexo 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 77 proteínas. 62 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 10 pertenecen al grupo funcional “transcripción” y 5 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 119 proteínas. 2 proteínas pertenecen al grupo funcional “control del ciclo celular, división celular, división cromosomal”, 7 pertenecen al grupo funcional “mecanismo de transducción de señales”, 30 al grupo funcional “pared celular/membrana/biogénesis”, 23 pertenecen al grupo funcional “motilidad de la célula”, 44 pertenecen al grupo funcional “modificación postranducción, plegamiento de proteínas, chaperonas” y 11 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. En el grupo funcional “mecanismo de defensa” solo fueron identificadas 2 proteínas.
- Metabolismo: en este grupo fueron identificadas un total de 270 proteínas, de ellas 47 pertenecen al grupo funcional “producción y conversión de energía”, 74 al grupo funcional “metabolismo y transporte de aminoácidos”, 16 al grupo funcional “metabolismo y transporte de nucleótidos”, 67 al grupo funcional “metabolismo y transporte de carbohidratos”, 9 al grupo funcional “metabolismo y transporte de coenzimas”, 18 al grupo funcional “transporte y metabolismo de lípidos”, 27 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 12 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.
- Pobremente caracterizadas: en este grupo fueron identificadas 64 proteínas, pertenecientes a los grupos funcionales “función desconocida” 29 y “pobremente caracterizadas” 35.
- Sin COG: en este grupo fueron identificadas 95 proteínas.

En ausencia de naringenina la categoría funcional más representada fue “metabolismo” correspondiendo a un 43,2 % del total de proteínas identificadas para esta condición. Las proteínas correspondientes a las categorías pobremente caracterizadas y sin COG corresponden a un 25,4% del total de proteínas identificadas para esta condición.

Proteoma de las OMV en presencia de naringenina

Fueron identificadas un total de 462 proteínas para el proteoma de OMV en presencia de naringenina; estas fueron agrupadas por COG y categorías funcionales, quedando distribuidas como sigue (tabla 3, anexo 2).

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 48 proteínas. 37 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 7 pertenecen al grupo funcional “transcripción” y 4 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 98 proteínas. 2 proteínas pertenecen al grupo funcional “control del ciclo celular, división celular, división cromosomal”, 7 pertenecen al grupo funcional “mecanismo de transducción de señales”, 27 al grupo funcional “pared celular/membrana/biogénesis”, 20 pertenecen al grupo funcional “motilidad de la célula”, 36 pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” y 5 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. En el grupo funcional “mecanismo de defensa” solo fue identificada una proteína.
- Metabolismo: en este grupo fueron identificadas un total de 187 proteínas, de ellas 21 pertenecen al grupo funcional “producción y conversión de energía”, 55 al grupo funcional “metabolismo y transporte de aminoácidos”, 10 al grupo funcional “metabolismo y transporte de nucleótidos”, 54 al grupo funcional “metabolismo y transporte de carbohidratos”, 7 al grupo funcional “metabolismo y transporte de coenzimas”, 12 al grupo funcional “transporte y metabolismo de lípidos”, 20 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 8 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.
- Pobremente caracterizadas: en este grupo fueron identificadas 45 proteínas, pertenecientes a los grupos funcionales “función desconocida” 23 y “pobremente caracterizadas” 22.

- Sin COG: en este grupo fueron identificadas 84 proteínas.

Las categorías funcionales más representadas para el proteoma de las OMV en presencia de naringenina fueron “metabolismo” y “pobrementemente caracterizadas” y sin COG, representando el 40,5% y el 27,9% respectivamente. Encontramos 16 proteínas similares a las reportadas en *Bradyrhizobium japonicum* BTAil, (esta es una cepa que fija nitrógeno en ausencia de factores de nodulación) que intervienen en la fijación de nitrógeno estas fueron: una posible oxidoreductasa, proteína de polimerización de polisacáridos, una posible transportador tipo ABC, aldosa-1-epimerasa, un transportador de ribosa tipo ABC, una proteína exportadora de exopolisacárido, la proteasa Clp subunidad proteolítica, sistema transportador de hemina y dos proteínas hipotéticas (tabla 4).

Proteoma de las OMV en ausencia de naringenina

Para el proteoma de OMV en ausencia de naringenina fueron identificadas un total de 391 proteínas; las cuales fueron agrupadas por COG y categorías funcionales, quedando distribuidas de la siguiente forma (tabla 3, anexo 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 52 proteínas. 39 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 7 pertenecen al grupo funcional “transcripción” y 6 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 82 proteínas. Solo una proteína pertenece al grupo funcional “control del ciclo celular, división celular, división cromosomal”, 5 pertenecen al grupo funcional “mecanismo de transducción de señales”, 28 al grupo funcional “pared celular/membrana/biogénesis”, 14 pertenecen al grupo funcional “motilidad de la célula”, 27 pertenecen al grupo funcional “modificación postranducción, plegamiento de proteínas, chaperonas” y 6 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. En el grupo funcional “mecanismo de defensa” solo fue identificada una proteína.
- Metabolismo: en este grupo fueron identificadas un total de 159 proteínas, de ellas 26 pertenecen al grupo funcional “producción y conversión de energía”, 50 al grupo funcional “metabolismo y transporte de aminoácidos”, 10 al grupo

funcional “metabolismo y transporte de nucleótidos”, 39 al grupo funcional “metabolismo y transporte de carbohidratos”, 2 al grupo funcional “metabolismo y transporte de coenzimas”, 9 al grupo funcional “transporte y metabolismo de lípidos”, 18 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 5 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.

- Pobremente caracterizadas: en este grupo fueron identificadas 35 proteínas, pertenecientes a los grupos funcionales “función desconocida” 17 y “pobremente caracterizadas” 18.
- Sin COG: en este grupo fueron identificadas 63 proteínas.

Para esta condición estudiada la categoría funcional mejor representada fue “metabolismo” que corresponde a un 40,7 % del total de proteínas identificadas para esta condición. Seguidas en abundancia por las proteínas pertenecientes a las categorías funcionales pobremente caracterizadas y sin COG que representan el 25,1% del total de proteínas identificadas para esta condición. Sorprendentemente encontramos en esta condición 18 proteínas que son similares a las reportadas en una cepa de *Bradyrhizobium japonicum* que fija nitrógeno en ausencia de factores de nodulación, solo encontramos dos diferentes a las encontradas en las OMV en presencia de naringenina estas fueron: un posible transportador tipo ABC y la manosa-1-fosfato guanidiltransferasa (tabla 4).

Comparación de los proteomas extracelulares de R. etli en ausencia y presencia de naringenina

Cuando comparamos los proteomas extracelulares de *R. etli* en ausencia y presencia de naringenina obtuvimos un total de 516 proteínas comunes para ambos proteomas. La clasificación de estas proteínas por COG y grupos funcionales quedó de la siguiente manera (tabla 3, anexo 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 70 proteínas comunes. 58 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 7 pertenecen al grupo funcional “transcripción” y 5 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 103 proteínas comunes a las dos fases. Solo 2 proteínas fueron identificadas en el grupo funcional “control del ciclo celular, división celular, división

cromosomal”, 6 pertenecen al grupo funcional “mecanismo de transducción de señales”, 26 al grupo funcional “pared celular/membrana/biogénesis”, 21 pertenecen al grupo funcional “motilidad de la célula”, 37 pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” y 10 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. Para el grupo funcional “mecanismo de defensa” solo se encontró una proteína común a los proteomas de las fracciones extracelular y OMV en presencia y ausencia de naringenina.

- **Metabolismo:** en este grupo fueron identificadas un total de 215 proteínas comunes a las dos fases de crecimiento, de ellas 41 pertenecen al grupo funcional “producción y conversión de energía”, 60 al grupo funcional “metabolismo y transporte de aminoácidos”, 11 al grupo funcional “metabolismo y transporte de nucleótidos”, 49 al grupo funcional “metabolismo y transporte de carbohidratos”, 8 al grupo funcional “metabolismo y transporte de coenzimas”, 14 al grupo funcional “transporte y metabolismo de lípidos”, 23 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 9 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.
- **Pobremente caracterizadas:** en este grupo fueron identificadas 48 proteínas comunes a las dos fracciones, pertenecientes a los grupos funcionales “función desconocida” y “pobremente caracterizadas”.
- **Sin COG:** en este grupo fueron identificadas 80 proteínas comunes para los dos proteomas analizados, en presencia y ausencia de naringenina.

Al comparar estos proteomas también pudimos identificar las proteínas específicas a cada condición. Para el caso del proteoma extracelular en presencia de naringenina fueron identificadas un total de 112 proteínas, de estas el 14,3 % corresponde a la categoría funcional “almacenamiento y procesamiento de la información”, el 18,8% corresponde a la categoría funcional “procesos celulares y señalización”, en esta categoría no fueron identificadas proteínas en los grupos funcionales “mecanismo de defensa” y “transporte intracelular, secreción, transporte vesicular”; el 44,6% corresponde a la categoría funcional “metabolismo” y el 22,3% corresponde a las categorías funcionales “pobremente caracterizadas” y sin COG. Es interesante señalar que las proteínas ribosomales están ligeramente más abundantes o su expresión relativa es mayor para el proteoma en presencia de naringenina.

Para el proteoma extracelular en ausencia de naringenina fueron identificadas un total de 109 proteínas, de estas el 6,4% corresponde a la categoría funcional “almacenamiento y procesamiento de la información”, el 14,7% corresponde a la categoría funcional “procesos celulares y señalización”, en esta categoría no encontramos proteínas identificadas para el grupo funcional “control del ciclo celular, división celular, división cromosomal”; el 50,5% corresponde a la categoría funcional “metabolismo” y el 28,4 % corresponde a las categorías funcionales “pobremente caracterizadas” y sin COG.

Comparación de los proteomas de OMV de R. etli en ausencia y presencia de naringenina

Al comparar los proteomas de las OMV en ausencia y presencia de naringenina encontramos un total de 328 proteínas comunes en ambos proteomas, estas fueron clasificadas por COG y categorías funcionales, quedando distribuidas de la siguiente manera (tabla 3, anexo 2):

- Almacenamiento y procesamiento de la información: en este grupo fueron identificadas 39 proteínas comunes. 30 proteínas pertenecen al grupo funcional “traducción, estructura ribosomal y biogénesis”, 6 pertenecen al grupo funcional “transcripción” y 3 al grupo funcional “replicación, recombinación y reparación”.
- Procesos celulares y señalización: en este grupo fueron identificadas un total de 73 proteínas comunes a las dos fases. Solo una proteína fue identificada en el grupo funcional “control del ciclo celular, división celular, división cromosomal”, 5 pertenecen al grupo funcional “mecanismo de transducción de señales”, 24 al grupo funcional “pared celular/membrana/biogénesis”, 14 pertenecen al grupo funcional “motilidad de la célula”, 24 pertenecen al grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” y 4 al grupo funcional “transporte intracelular, secreción, transporte vesicular”. Para el grupo funcional “mecanismo de defensa” solo se encontró una proteína común a los proteomas de las fracciones extracelular y OMV en presencia y ausencia de naringenina.
- Metabolismo: en este grupo fueron identificadas un total de 128 proteínas comunes a las dos fases de crecimiento, de ellas 17 pertenecen al grupo funcional “producción y conversión de energía”, 41 al grupo funcional

“metabolismo y transporte de aminoácidos”, 7 al grupo funcional “metabolismo y transporte de nucleótidos”, 37 al grupo funcional “metabolismo y transporte de carbohidratos”, solo una proteína fue identificada en el grupo funcional “metabolismo y transporte de coenzimas”, 7 al grupo funcional “transporte y metabolismo de lípidos”, 14 al grupo funcional “transporte y metabolismo de iones inorgánicos” y 4 al grupo funcional “transporte y catabolismo de metabolitos secundarios”.

- Pobremente caracterizadas: en este grupo fueron identificadas 31 proteínas comunes a las dos fracciones, pertenecientes a los grupos funcionales “función desconocida” y “pobremente caracterizadas”.
- Sin COG: en este grupo fueron identificadas 57 proteínas comunes para los dos proteomas analizados, en presencia y ausencia de naringenina.

En la comparación entre los proteomas de las OMV en ausencia y presencia de naringenina pudimos identificar las proteínas específicas para cada condición. Para el proteoma de OMV en presencia de naringenina fueron identificadas un total de 134, de esta el 44,0% corresponde a las proteínas pertenecientes a la categoría funcional “metabolismo” y el 29,5 % a las categorías funcionales “pobremente caracterizadas” y sin COG.

Para el proteoma de OMV en ausencia de naringenina fueron identificadas un total de 63 proteínas específicas para esta condición, de estas el 49,2 % corresponden a la categoría funcional “metabolismo”. En la categoría funcional “procesos celulares y señalización” no encontramos proteínas en los grupos funcionales “control del ciclo celular, división celular, división cromosomal”, “mecanismo de transducción de señales”, “mecanismo de defensa” y “motilidad de la célula”.

Proteínas identificadas en el proteoma de OMV de *R. etli* que están presentes en *Bradyrhizobium japonicum* BTail

Cuando realizamos la comparación entre las proteínas de los proteomas de las OMV en presencia y ausencia de naringenina encontramos un total de 18 proteínas comunes con las reportadas para *Bradyrhizobium japonicum*, [Giraud et al en el 2007] que se propone están involucradas en el proceso de simbiosis en esta bacteria, lo que representa un 41,9% del total de proteínas reportadas para *Bradyrhizobium japonicum*. Estas fueron agrupadas en los grupos funcionales “producción y conversión de energía”, “control del ciclo celular, división celular, división cromosomal”, “metabolismo y

transporte de aminoácidos”, “metabolismo y transporte de carbohidratos”, “pared celular/membrana/biogénesis”, “modificación postranduccional, plegamiento de proteínas, chaperonas”, “transporte y metabolismo de iones inorgánicos”, “transporte y catabolismo de metabolitos secundarios”, “pobremente caracterizadas” y “sin COG” (tabla 4).

Tabla 1: Proteínas identificadas de *R. etli* en fase exponencial (6H) y en fase estacionaria (24H). Fracciones periplasmáticas (PER) y Vesículas de membrana externa (OMV).

COG	Proteínas específicas 6H OMV	Proteínas específicas 24H OMV	Proteínas Comunes 6 y 24H OMV	Proteínas totales 6H OMV	Proteínas totales 24H OMV	Proteínas específicas 6H PER	Proteínas específicas 24H PER	Proteínas comunes 6 y 24H PER	Proteínas totales 6H PER	Proteínas totales 24H PER
Almacenamiento y procesamiento de la información										
J	38	9	21	59	30	26	12	43	69	55
K	9	2	3	12	5	10	10	8	18	18
L	5	0	3	8	3	4	1	4	7	5
sub totales	52	11	27	79	38	40	23	55	94	78
Procesos celulares y señalización										
D	2	0	0	2	0	0	3	3	3	6
T	2	6	3	5	9	3	18	17	20	35
V	0	1	0	0	1	0	1	0	0	1
M	10	5	15	25	20	5	15	18	24	33
N	7	5	13	20	18	3	4	10	13	14
O	12	9	19	31	28	8	16	35	43	51
U	7	0	2	9	2	4	0	2	6	2
sub totales	40	26	52	92	78	23	57	85	109	142
Metabolismo										
C	30	6	13	43	19	19	9	24	43	33
E	24	22	36	60	58	32	27	70	102	97
F	11	5	3	14	8	9	10	20	29	30
G	15	27	24	39	51	30	15	59	89	74
H	5	5	2	7	7	12	17	10	22	27
I	12	1	3	15	4	4	9	17	21	26
P	2	10	12	14	22	8	5	18	26	23
Q	3	5	3	6	8	1	5	11	12	16

sub totales	102	81	96	198	177	115	97	229	344	326
Pobrememente caracterizada										
R	12	9	8	20	17	20	24	30	50	54
S	8	12	8	16	20	6	13	28	34	41
sub totales	20	21	16	36	37	26	37	58	84	95
Sin COG	21	42	28	49	70	23	42	47	70	89
Totales	235	181	219	454	400	227	256	474	701	730

Tabla 2: Comparación de los proteínas identificadas de *R. etli* en fase exponencial (6H) y en fase estacionaria (24H). Fracciones periplasmáticas (PER) y Vesículas de membrana externa (OMV).

COG	Proteínas totales 6H PER	Proteínas totales 6H OMV	Proteínas Comunes 6H PER Y 6H OMV	Proteínas específicas 6H PER	Proteínas específicas 6H OMV	Proteínas totales 24H PER	Proteínas totales 24H OMV	Proteínas comunes 24H PER y 24H OMV	Proteínas específicas 24H PER	Proteínas específicas 24H OMV
Almacenamiento y procesamiento de la información										
J	69	59	40	29	19	55	30	22	33	8
K	18	12	11	7	1	18	5	3	15	2
L	7	8	5	3	3	5	3	3	2	0
sub totales	94	79	56	39	23	78	38	28	50	10
Procesos celulares y señalización										
D	3	2	0	3	2	6	0	0	6	0
T	20	5	3	17	2	35	9	7	28	2
V	0	0	0	0	0	1	1	1	0	0
M	24	25	16	7	9	33	20	13	20	7
N	13	20	11	2	9	14	18	10	4	8
O	43	31	23	20	8	51	28	27	19	1

U	6	9	4	2	5	2	2	2	0	0
sub totales	109	92	57	51	35	142	78	60	77	18
Metabolismo										
C	43	43	26	17	17	33	19	12	21	7
E	102	60	48	54	12	97	58	48	49	10
F	29	14	7	22	7	30	8	6	24	2
G	89	39	32	57	7	74	51	34	40	17
H	22	7	4	18	3	27	7	3	24	4
I	21	15	11	10	4	26	4	3	23	1
P	26	14	12	14	2	23	22	16	7	6
Q	12	6	3	9	3	16	8	5	11	3
sub totales	344	198	143	201	55	326	177	127	199	50
Pobremente caracterizadas										
R	50	20	14	36	6	54	17	11	43	6
S	34	16	9	25	7	41	20	17	24	6
sub totales	84	36	23	61	13	95	37	28	67	12
SC	70	49	25	45	24	89	70	31	58	39
Totales	701	454	304	397	150	730	400	274	451	129

Tabla 3: Proteínas identificadas de *R. etli*, en presencia y ausencia de naringenina. Proteínas de la fracción extracelular y OMV.

COG	Proteínas totales CE3 CON NAR	Proteínas totales CE3 SIN NAR	Proteínas comunes CE3 CON y SIN NAR	Proteínas especificas CE3 CON NAR	Proteínas especificas CE3 SIN NAR	Proteínas totales CE3 CON NAR OMV	Proteínas totales CE3 SIN NAR OMV	Proteínas comunes OMV (CON y SIN) NAR	Proteínas especificas CE3 CON NAR OMV	Proteínas especificas CE3 SIN NAR OMV
Almacenamiento y procesamiento de la información										
J	63	62	58	5	4	37	39	30	7	9
K	12	10	7	5	3	7	7	6	1	1
L	11	5	5	6	0	4	6	3	1	3

sub totales	86	77	70	16	7	48	52	39	9	13
Procesos celulares y señalización										
D	3	2	2	1	0	2	1	1	1	0
T	9	7	6	3	1	7	5	5	2	0
V	1	2	1	0	1	1	1	1	0	0
M	35	30	26	9	4	27	28	24	3	4
N	22	23	21	1	2	20	14	14	6	0
O	44	44	37	7	7	36	27	24	12	3
U	10	11	10	0	1	5	6	4	1	2
sub totales	124	119	103	21	16	98	82	73	25	9
Metabolismo										
C	51	47	41	9	6	21	26	17	4	9
E	78	74	60	18	14	55	50	41	14	9
F	15	16	11	4	5	10	10	7	3	3
G	58	67	49	9	18	54	39	37	17	2
H	12	9	8	4	1	7	2	1	6	1
I	16	18	14	2	4	12	9	7	5	2
P	25	27	23	2	4	20	18	14	6	4
Q	11	12	9	2	3	8	5	4	4	1
sub totales	266	270	215	50	55	187	159	128	59	31
Pobremente caracterizadas										
R	32	35	29	3	6	22	18	15	7	3
S	25	29	19	7	10	23	17	16	7	1
sub totales	57	64	48	10	16	45	35	31	14	4
sin COG	95	95	80	15	15	84	63	57	27	6
Totales	628	625	516	112	109	462	391	328	134	63

Tabla 4: Proteínas identificadas en *R. etli* en OMV en presencia y ausencia de naringenina presentes en *Bradyrhizobium japonicum* BTail. [Giraud et al en el 2007]

AC	condición	nombre	COG	PMSS (expresión relativa)	MW	PI	descripción	BTail
B3PUT8	CE3 Snar VES	RHECIAT_CH0001509	C	39,391763	38491,619	6,2	Putative oxidoreductase protein	BBta4127- 5628
Q2K589	CE3 Snar VES	pssP	D	96,00918	84127,828	5,373	Exopolysaccharide polymerization protein	BBta1034
Q2K589	CE3 Cnar VES	pssP	D	167,85235	84127,828	5,373	Exopolysaccharide polymerization protein	BBta1034
Q2JZA9	CE3 Snar VES	RHE_PF00186	E	52,317848	29689,204	5,856	Probable amino acid ABC transporter, substrate-binding protein	BBta 5670
Q2JZA9	CE3 Cnar VES	RHE_PF00186	E	131,451152	29689,204	5,856	Probable amino acid ABC transporter, substrate-binding protein	BBta 5670
Q2K1G2	CE3 Snar VES	RHE_PC00118	E	136,253285	29585,074	5,625	Probable amino acid ABC transporter, substrate-binding protein	BBta 5670
Q2K1G2	CE3 Cnar VES	RHE_PC00118	E	589,448277	29585,074	5,625	Probable amino acid ABC transporter, substrate-binding protein	BBta 5670
Q2K5U2	CE3 Snar VES	RHE_CH03027	E	21,538606	26434,771	7,625	Putative amino acid ABC transporter, substrate-binding protein	BBta 5670
Q2K5U2	CE3 Cnar VES	RHE_CH03027	E	85,65954	26434,771	7,625	Putative amino acid ABC transporter, substrate-binding protein	BBta 5670
Q2KA68	CE3 Snar VES	RHE_CH01465	E	29,929382	28713,754	6,197	Probable amino acid ABC transporter, substrate-binding protein	BBta 5670
Q2JZ78	CE3 Snar VES	RHE_PF00217	G	24,148296	33832,81	5,48	Probable D-ribose ABC transporter, substrate-binding protein	BBta 5634
Q2JZ78	CE3 Cnar VES	RHE_PF00217	G	52,024248	33832,81	5,48	Probable D-ribose ABC transporter, substrate-binding protein	BBta 5634
Q2K308	CE3 Snar VES	galM	G	13,982723	37651,149	5,607	Aldose-1-epimerase protein	BBta 5621
Q2K308	CE3 Cnar VES	galM	G	35,511917	37651,149	5,607	Aldose-1-epimerase protein	BBta 5621
Q2KCK8	CE3 Snar VES	rbsBch1	G	148,424239	34847,378	8,623	Ribose ABC transporter, substrate-	BBta 5634

							binding protein	
Q2KCK8	CE3 Cnar VES	rbsBch1	G	206,545211	34847,378	8,623	Ribose ABC transporter, substrate-binding protein	BBta 5634
Q2K578	CE3 Snar VES	noeJ	M	41,773701	53084,711	6,44	Mannose-1-phosphate guanylyltransferase (GDP) protein	BBta 5620
Q2K587	CE3 Snar VES	pssN	M	461,111651	43028,577	5,807	Exopolysaccharide export protein	BBta1029-28
Q2K587	CE3 Cnar VES	pssN	M	370,529355	43028,577	5,807	Exopolysaccharide export protein	BBta1029-28
Q2KC60	CE3 Snar VES	noeL1	M	29,099606	41056,422	5,678	GDP-mannose 4,6-dehydratase protein	BBta 1071
Q2KC60	CE3 Cnar VES	noeL1	M	32,190812	41056,422	5,678	GDP-mannose 4,6-dehydratase protein	BBta 1071
B3PVY4	CE3 Cnar VES	clpP2	O	45,3146	23345,845	6,084	ATP-dependent Clp protease proteolytic subunit	BBta 5674
Q2K553	CE3 Snar VES	hmuT	P	122,764106	32580,644	5,704	Hemin transport system, hemin-binding protein	BBta1073
Q2K553	CE3 Cnar VES	hmuT	P	167,187399	32580,644	5,704	Hemin transport system, hemin-binding protein	BBta1073
B3PQY3	CE3 Cnar VES	RHECIAT_CH0004208	Q	15,427811	31320,608	5,469	Probable oxidoreductase protein	BBta4127
Q2K695	CE3 Snar VES	RHE_CH02873	R	55,509453	60971,187	5,458	Probable ABC transporter, ATP-binding protein	BBta 5666
Q2K695	CE3 Cnar VES	RHE_CH02873	R	49,673878	60971,187	5,458	Probable ABC transporter, ATP-binding protein	BBta 5666
Q2K1J3	CE3 Cnar VES	RHE_PC00149	SC	13,067282	38040,241	5,73	Putative sugar ABC transporter, substrate-binding protein	BBta 5634
Q2KBS8	CE3 Snar VES	RHE_CH00897	SC	100,387288	120610,665	4,652	Hypothetical conserved protein	BBta 4171
Q2KBS8	CE3 Cnar VES	RHE_CH00897	SC	124,527841	120610,665	4,652	Hypothetical conserved protein	BBta 4171
Q8KL06	CE3 Snar VES	RHE_PD00077	SC	115,348534	16073,312	9,579	Putative uncharacterized protein	BBta 5647
Q8KL06	CE3 Cnar VES	RHE_PD00077	SC	70,863618	16073,312	9,579	Putative uncharacterized protein	BBta 5647

PMSS (expresión relativa): sumatoria de la calificación de los péptidos identificados

¿Están las OMV relacionadas con el proceso de simbiosis en R. etli?

Con la finalidad de conocer el posible papel de las OMV en las primeras etapas del proceso simbiótico realizamos varios experimentos incubando plántulas de frijol en presencia de OMV sin inducir, como se describe en Materiales y Métodos.

Los resultados obtenidos muestran de manera sorprendente un incremento considerable en la cantidad de pelos radiculares en las plántulas de frijol cultivadas en presencia de las OMV purificadas, en comparación con los controles positivos (plantas inoculadas con *R. etli*) y negativos (plantas sin inocular) figura 6.

Por lo anterior decidimos buscar la presencia de los factores de nodulación dentro de las vesículas de membrana externa de *R. etli* o alguna molécula que favoreciera esta respuesta.

Por esta razón realizamos inducciones con naringenina. El procedimiento experimental a grandes rasgos es el siguiente; se cultivo a la bacteria en medio mínimo y se le añadió naringenina a una concentración final 1,2 μ M a las 6h de cultivo, posteriormente a las 12 horas de agregar el inductor se cosecharon los sobrenadantes mediante centrifugación, en este momento fueron analizadas las proteínas que habían sido exportadas al medio por la bacteria. El estudio del secretoma del sobrenadante de *R. etli* en presencia y ausencia de naringenina lo comparamos con el proteoma de las OMV en presencia y ausencia del inductor, los resultados son presentados en la tabla 4.

Previamente en *Bradyrhizobium japonicum* BTAil la cual es una cepa bacteriana que fotosintetiza y es capaz de realizar simbiosis con la leguminosa acuática *Aeschynomene*, en ausencia de factores nod [Giraud et al. 2007] fueron reportadas un grupo de proteínas implicadas en el proceso de simbiosis, por lo que decidimos comparar estos resultados con los obtenidos en el proteoma de las OMV inducidas con naringenina, encontrando varias de estas proteínas presentes en las OMV de *R. etli*, tabla 4.

Gracias a estos resultados ya contábamos con una posible explicación, que involucra las proteínas identificadas, para el comportamiento de las plantas de frijol *Phaseolus vulgaris* Negro Xamapa cuando son inoculadas con OMV. Sin embargo decidimos analizar si existía otro tipo de moléculas en dichas OMV inducidas con naringenina Por lo cual realizamos extracciones con 1-butanol de las vesículas purificadas en presencia y en ausencia de naringenina [Soria-Díaz et al. 2003]. Posteriormente analizamos el extracto butanólico mediante cromatografía líquida de fase reversa como se describe en Materiales y métodos.

Los espectros obtenidos de la cromatografía realizada correspondiente al extracto proveniente de las OMV en presencia de naringenina, presentan un “pico” que no se observa en los controles utilizados fig 7. Esta fracción fue colectada y liofilizada, para su análisis mediante espectrometría de masas tratando de identificar su composición. El procedimiento es explicado en Materiales y métodos; los resultados obtenidos se muestran en la tabla 5. En esta tabla se muestran las masas obtenidas que pudieran corresponder a los factores Nod, estas fueron 1111.45, 898.27, 760.23 y 571.26.



Figura 6: Pelos radiculares en las raíces de las plantas de frijol 4 días después de la infección. A-sin inocular. B-inoculados con *R. etli*. C- inoculados con OMV purificadas.

Con la ayuda del software glycomod analizamos los resultados obtenidos y definiendo las posibles estructuras para las masas experimentales obtenidas de la fragmentación en el espectrómetro de masas [Cooper et al. 2001], tabla 5, figuras en anexos. Cuando analizamos estos resultados pudimos encontrar varias estructuras que presentan N-acetylglucosamina, que es una parte integrante de la estructura de los factores Nod lo que sugiere que posiblemente estos factores de nodulación están siendo secretados en OMV por *R. etli*.

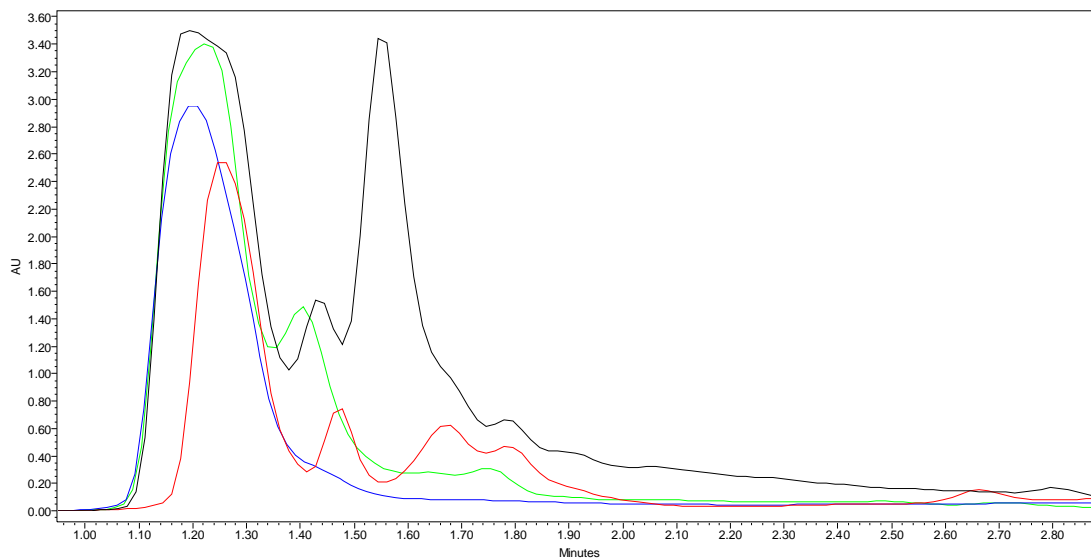


Figura 7: Espectro de HPLC donde se muestra el pico obtenido en la extracción realizada de las OMV con 1-butanol. Línea negra: *R. etli* CE3 con naringenina, línea verde CE3 sin naringenina, línea azul mutante de *R. etli* CE3 NodA- con naringenina, línea roja mutante de *R. etli* CE3 NodA- control negativo.

Tabla 5: Probable estructuras obtenidas por el Glycomod a partir de las masas experimentales obtenidas por espectrometría de masas.

Masas experimentales	Masas glycoform	Δ mass (Dalton)	estructura
1111.45	1092.361	0.071	(HexNAc) ₁ (NeuAc) ₂ (NeuGc) ₁
	1092.361	0.071	(HexNAc) ₃ (NeuGc) ₁ (HexA) ₁
898.27	879.249	0.003	(Hex) ₂ (HexNAc) ₁ (HexA) ₂
	879.249	0.003	(NeuGc) ₁ (Pent) ₃ (HexA) ₁
	879.286	-0.033	(Hex) ₁ (Deoxyhexose) ₁ (NeuGc) ₁ (Pent) ₂
	879.286	-0.033	(Hex) ₂ (NeuAc) ₁ (Pent) ₂
760.23	741.281	-0.068	(HexNAc) ₃ (Pent) ₁
571.26	552.217	0.025	(HexNAc) ₂ (Deoxyhexose) ₁

Hexose (Hex) (galactosa, manosa, glucosa)

HexNAc (N-acetylglucosamina)
NeuAc (N-acetyl ácido neuraminico (ácido siálico))
NeuGc (N-glycolyl ácido neuraminico)
Aducto ($[M+H]^+$): 1.00727
Masas derivadas (extreme reductor libre): 18.0105546

Finalmente para demostrar el posible papel de las OMV en el proceso de simbiosis, realizamos una cromatografía de capa fina (TLC), fig 8; podemos observar como en los carriles 2 (extracto butanólico del sobrenadante de *R. etli* en presencia de naringenina) y 4 (extracto butanólico de las OMV en presencia de naringenina) aparecen bandas que no aparecen en los demás carriles (señaladas con flechas). Por lo que podemos apreciar que algunos factores de nodulación están presentes en las OMV cuando realizamos la inducción con naringenina. Este resultado demuestra el papel que tienen las OMV en el proceso de simbiosis entre la bacteria y la planta.

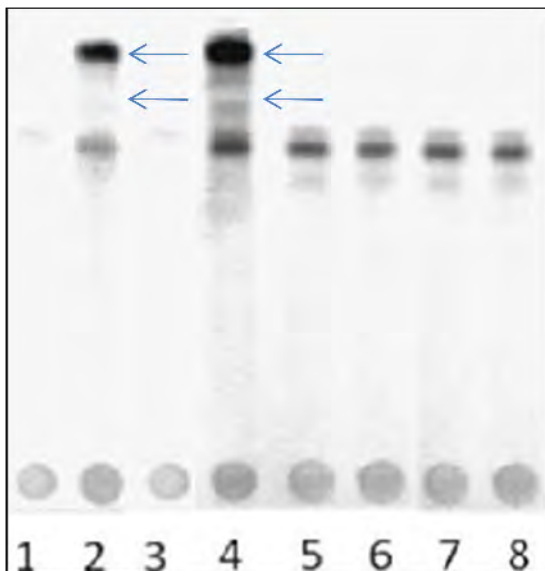


Figura 8. TLC, que demuestra la presencia de los factores Nod en las OMV. Líneas 1 y 2 SN de CE3, Línea 3 y 4 OMV de CE3, Líneas 5 y 6 SN de CE3 NodA-, Líneas 7 y 8 OMV de CE3 Nod A-. Líneas 1, 3, 5 y 7 en ausencia de naringenina. Líneas 2, 4, 6 y 8 en presencia de naringenina. SN-sobrenadante. Las flechas indican la presencia de los factores Nod, que no están presentes en los demás carriles.

DISCUSION

En este trabajo se ha realizado un análisis global de las proteínas que secreta *R. etli* en las fases de crecimiento exponencial y estacionario, incluyendo los proteomas de las fracciones periplásmicas, extracelular y OMV. También se analizó el conjunto de proteínas que se secretan en presencia y ausencia de naringenina, ya que este es el flavonoide que se encuentra en el exudado de frijol y es reconocido por la bacteria para iniciar el proceso de simbiosis.

El objetivo central del proyecto fue estudiar las proteínas secretadas por *R. etli*, en fase de crecimiento exponencial y estacionario. Estos resultados ya fueron publicados y son presentados en Anexos 1. Para ello utilizamos la proteómica como herramienta, pero a medida que comenzamos a obtener resultados fueron surgiendo nuevas preguntas y por esta razón analizamos también las proteínas secretadas por la bacteria en las OMV. En todos los casos realizamos el análisis de la expresión relativa de las proteínas identificadas para cada proteoma analizado.

Del análisis de los resultados obtenidos para el proteoma de la fracción periplásmica en fase de crecimiento exponencial consideramos necesario distinguir algunas proteínas que dentro de cada categoría funcional son expresadas en mayor cantidad o son más abundantes con respecto a las demás proteínas. En este caso tenemos que para la categoría funcional “almacenamiento y procesamiento de la información” encontramos que las proteínas más expresadas corresponden a los factores de elongación G (FusA) con una expresión relativa de 2394.34 y Tu (Tuf1) con una expresión relativa de 894.27, para la categoría funcional “metabolismo” encontramos que las Succinyl-CoA ligasa [ADP-forming], subunidad beta (SucC) con una expresión relativa de 682.41 y la Superóxido dismutasa (SodB) con una expresión relativa de 2217.23; fueron las proteínas más expresadas, para la categoría funcional “procesos celulares y señalización” fueron identificadas chaperonas como GroEL, DnaK, transportadores y flagelinas como las proteínas que presentaron mayor expresión relativa con valores por encima de 1000 para esta categoría. Un gran número de proteínas pertenecientes a los grupos funcionales “función desconocida” y “pobremente caracterizadas”, las cuales han sido reportadas previamente en otras bacterias como en el caso de *Synechocystis* [Kurian et al. 2006], en donde un grupo importante de proteínas en el secretoma de esta bacteria, estudiado mediante geles de doble dimensión y MALDI-TOF, fueron, posibles transportadores y proteínas de función desconocida. También han sido reportadas proteínas como la SodB en periplasma de *E. coli* en condiciones de pH ácido

[Stancik et al 2002], esta proteína juega un papel importante en el mecanismo de defensa ya que protege a la célula de la acción dañina del superóxido. Cabe destacar el gran número de proteínas ribosomales identificadas en la fracción periplásmica en fase de crecimiento exponencial, lo cual podría justificar el gran número de estas proteínas presentes en el proteoma extracelular de *R. etli* en fase exponencial [Meneses et al 2010].

En el proteoma de las OMV en fase exponencial fueron identificadas un gran número de proteínas que presentan una alta concentración; para la categoría funcional “almacenamiento y procesamiento de la información”, entre ellas encontramos el factor de elongación Tu (Tuf1) presenta una expresión relativa de 2517.83, resultado que coincide con el obtenido para el periplasma en esta fase de crecimiento, también encontramos un gran número de proteínas ribosomales, que presentan una expresión relativa elevada, estos resultados nos sugieren que encontramos estas proteínas en el secretoma de *R. etli* porque son exportadas desde el periplasma a las OMV, siendo este el mecanismo mediante el cual son exportadas al exterior de la célula, aunque no sabemos qué función podrían tener estas proteínas en el secretoma de *R. etli*. Vaughan et al 2006 reportaron un estudio en donde analizaban la composición de OMV de *Neisseria meningitidis* y *Neisseria lactamica*, usadas como vacunas para prevenir la enfermedad meningococcica invasiva, ellos encontraron gran número de proteínas ribosomales, proteínas de membrana externa, transportadores, entre otros, resultados muy parecidos a los obtenidos para las OMV de *R. etli*.

Entre las proteínas abundantes en las OMV en fase exponencial, encontramos proteínas pertenecientes a la categoría funcional “procesos celulares y señalización”, flagelinas como la FlaCch1 y FlaCch2, que presentan una expresión relativa de 5741.28 y 10055.88; proteínas de membrana externa como la RopB1, RopB1, con una expresión relativa de 9666,37; chaperonas como la GroEL1, que presentan una expresión relativa elevada por encima de 1000; estos resultados coinciden con los reportados [Lee et al 2007 y Scorza et al 2008], estas proteínas podrían estar relacionadas con el mecanismo de defensa de la bacteria. Además presentan expresión relativa elevada las proteínas pertenecientes a la categoría funcional “metabolismo” en las que se incluyen algunos transportadores (BraC2, DppAch1, AapJ, ChvE, pstS), se desconoce la función que desempeñan estas proteínas al ser exportadas al exterior mediante las OMV, estas proteínas han sido reportadas en otros trabajos y no le han asignado una posible función. Adicionalmente, se identificaron en las OMV enzimas como FbaB, la cual es una enzima glicolítica que participa en la condensación de fructosa1,5

bifosfato en dihidroxyacetona fosfato y gliceraldehido 3 fosfato [Perham 1990], Gap, es una enzima glicolítica [Kim et al. 1995], desconocemos la función de estas enzimas en las OMV. Fue identificada en OMV la toxina RzcA, está relacionada con el proceso de patogénesis, es considerada un importante factor de virulencia en bacterias Gram-negativas [Czuprynski y Welch 1995], por lo que en OMV pudiera estar participando en el mecanismo de defensa de la bacteria. Fueron identificadas además un gran número de proteínas que pertenecen a las categorías “pobremente caracterizadas” y “sin COG” (ver tabla en Anexos). Todos estos resultados son comunes con los obtenidos en la fracción periplásmica lo que sugiere que llegan a las OMV a través del periplasma lo que concuerda con la hipótesis de la formación de las OMV, que plantea que las OMV se forman cuando la membrana externa se expande más rápido que la capa de peptidoglucano subyacente, en cuyo caso la membrana externa tenderá a separarse localmente desde la capa de peptidoglucano. Si este exceso de péptidoglucano en la membrana externa permanece durante algún tiempo, el resultado será la expansión de la membrana externa hasta que se repliega sobre sí misma y se libera en el medio [Zhou et al 1998]. Cuando comparamos los proteomas de las OMV y el periplasma encontramos proteínas que son específicas para cada uno de los proteomas, es decir, no se comparten, nos surgieron las siguientes preguntas ¿cómo llegaron esas proteínas a las OMV?, ¿tendrá *R. etli* un mecanismo diferente para producir las OMV? Al analizar un reporte realizado en *Legionella pneumophila* sobre los proteomas de esta bacteria patógena para el sobrenadante y las OMV ellos muestran como las OMV comienzan a secretarse por la bacteria desde la fase de crecimiento exponencial, lo que podría sugerir un mecanismo específico de producción que aun no es bien conocido, este resultado concuerda con nuestros resultados ya que también en *R. etli* se producen las OMV desde la fase de crecimiento exponencial. La presencia de proteínas específicas en el periplasma indica que no todas las proteínas son exportadas a través de las OMV, estas proteínas que no se exportan a través de las OMV utilizan algunos de los ya conocidos sistemas de secreción.

Cuando comparamos los proteomas de las fracciones periplásmicas y OMV en fase de crecimiento exponencial nos dimos cuenta que hay muchas proteínas que son comunes para ambos proteomas, pero algo interesante es que la expresión de muchas de estas proteínas es diferente para cada proteoma. Si analizamos la categoría funcional “almacenamiento y procesamiento de la información” podemos apreciar que las proteínas ribosomales como las 30S, 50S y el factor de elongación Tu, se encuentran más abundantemente en OMV (Anexos

2), lo que nos indica que estas proteínas podrían tener algún papel importante en las OMV, ya que son expresadas con mayor abundancia en estas estructuras. Lee et al en el 2007 reportaron un gran número de proteínas citoplásmicas (chaperonas, proteínas ribosomales y factores de elongación), en el proteoma de OMV para *E. coli*, estos investigadores sugieren que si en las OMV también encontramos ADN y ARN, es hasta cierto punto coherente encontrar también estas proteínas y proponen que deben ser objeto de futuros estudios, también existen reportes en donde se les ha asignado una segunda función a los factores de elongación, ya que estas proteínas se asocian a proteínas de membrana externa y ayudan a mantener la forma de las bacterias como en *Bacillus subtilis* [Soufoa et al 2010]. Las proteínas ProS, FusA, ArgS, Tsf, LeuS se encuentran en mayor concentración en el periplasma. No hemos encontrado reportes científicos relacionados con estos resultados. También identificamos proteínas como las flagelinas y las chaperonas que son más abundantes en las OMV comparado con el periplasma en fase exponencial (ver tabla en Anexos).

Para el proteoma de la fracción periplásmica de *R. etli* en fase de crecimiento estacionario encontramos algunas proteínas que son muy abundantes en la categoría funcional “almacenamiento y procesamiento de la información”, como son el factor de elongación Tu (Tuf1), el factor de elongación G (FusA), factor de reciclaje del ribosoma (Frr), resultados como estos coinciden con los reportados por otros investigadores en *Yersinia pestis* [Pieper et al 2008]. Para la categoría funcional “procesos celulares y señalización” encontramos que las chaperonas presentan una expresión relativa muy elevada, ya que es posible que estén participando en el plegamiento de algunas proteínas importantes en esta fase de crecimiento como son las proteínas de membrana externa. En el caso de las proteínas metabólicas, presentes en periplasma, como algunos transportadores tipo ABC, y algunas enzimas como la SodB, las cuales son muy abundantes en el periplasma para esta fase de crecimiento; sugerimos que los transportadores son necesarios para transportar elementos importantes para la bacteria como aminoácidos, iones metálicos, azúcares entre otros y la proteína SodB debe tener un importante papel para contrarrestar el estrés producido por sustancias tóxicas en el exterior de la célula (ver tabla en anexos 2) [Pieper et al. 2008].

Al analizar el proteoma de las OMV en la fase estacionaria encontramos que para la categoría funcional “almacenamiento y procesamiento de la información” solo se encuentra en mayor abundancia el factor de elongación Tu, y las proteínas ribosomales que se identificaron en esta condición están poco abundantes, lo que concuerda con los resultados obtenidos para el

secretoma de *R. etli* para la fracción del sobrenadante, ya que encontramos en la fase estacionaria un pequeño número de proteínas ribosomales presentes en esta fracción. En el caso de la categoría funcional “procesos celulares y señalización” encontramos varias proteínas que presentan una expresión relativa muy alta con respecto a las demás; como es el caso de la proteína de membrana externa (RopB1), flagelinas (FlaCe, FlaCch1, FlaCch2, FlaCch3) y chaperonas. Estudios llevados a cabo en bacterias fitopatógenas con patrones moleculares asociados a microbios (MAMP del inglés *microbe associated molecular patterns*) y su percepción por las plantas [Aslam et al. 2009], revelaron que las plantas tienen sistemas desarrollados para percibir diferentes proteínas bacterianas como flagelinas, lipopolisacáridos (LPS), factor de elongación Tu (EF-Tu), proteína de estrés por frío (CSP) o peptidoglicano (PGN) [Gómez-Gómez et al. 2000, Silipo et al. 2005, Zipfel et al. 2006]

Curiosamente, muchas de estas respuestas también se han detectado en las células epidérmicas de las raíces de las leguminosas, poco después de la aplicación de los factores Nod [Soto et al. 2009]. Todos estos resultados indican que en *R. etli* estas proteínas podrían estar involucradas en las etapas iniciales de la simbiosis.

Otras proteínas muy abundantes para esta fracción son algunas de las que pertenecen a la categoría funcional “metabolismo” entre ellas encontramos transportadores como PstS, SfuA, AgIE, ChvE, FbaB, AapJ; así mismo, encontramos algunas proteínas de función desconocida como proteínas muy abundantes (ver tabla en anexos 2). Estos resultados coinciden con los reportados para el proteoma de las OMV de *E. coli*, *Xantomonas campestris* y *Legionella pneumophila* [Lee et al. 2007, Sidhu et al. 2008, Galka et al. 2008], en estos trabajos no se ha reportado una explicación para la presencia de estas proteínas en las OMV, pero si todos coinciden de que las OMV intervienen en los procesos de defensa de las bacterias, de adaptación a diferentes nichos y son responsables de la transferencia lateral de material biológico entre las bacterias.

En los resultados obtenidos del proteoma del sobrenadante de *R. etli* encontramos en fase estacionaria la proteína TolB [Meneses et al. 2010], esta proteína está relacionada en *E. coli* con el proceso de producción de las OMV, según McBroom 2006. Estos investigadores reportaron que para *E. coli* las proteínas de membrana externa y lipoproteínas de membrana externa también están involucradas en el proceso de producción de las OMV, en el sobrenadante de *R. etli* para ambas fases de crecimiento se identificaron estas proteínas,

[Meneses et al. 2010] por lo que en *R. etli* podrían estar involucradas también en el proceso de producción de las OMV.

Al comparar los proteomas de las fracciones periplásmicas y OMV en la fase de crecimiento estacionario encontramos 274 proteínas comunes, observamos que lo que cambia es la expresión relativa de las proteínas en algunos casos para las OMV y en otras para el periplasma. Para la categoría funcional “almacenamiento y procesamiento de la información” encontramos más abundantes para el periplasma las siguientes proteínas RplS, RHE_CH02879 y NusB; mientras que para la fracción de OMV son más abundantes las proteínas PrfA, RpsC, RpsA y Hrm; las demás proteínas identificadas comunes para ambos proteomas presentan aproximadamente la misma abundancia. Para la categoría funcional procesos celulares y señalización, para el periplasma las proteínas más abundantes son CheY3, ChvI, SecB, RkpK, PpiB; mientras que para el proteoma de las OMV las proteínas más abundantes fueron TolB, RHE_CH02869, RHE_CH02417, RopB1, FlaCch1, FlaCch2, FlaCe, RHE_CH02084, FlaCch3, RHE_CH00187, GroELc, PpiA, GroEL1. Para la categoría funcional metabolismo, las proteínas más abundantes corresponden casi en su totalidad al proteoma de las OMV encontrándose solo algunas abundantes para el periplasma como es el caso de SodB. No encontramos reportes científicos que comparen los proteomas de las OMV y el periplasma. Estos resultados sugieren que las OMV podrían estar involucradas en el proceso de resistencia a estrés de la bacteria o en el proceso de adaptación a condiciones adversas por parte de la bacteria.

En *Legionella pneumophila* y *Helicobacter pylori*, se ha reportado la presencia de chaperonas tipo Hsp60 y Hsp70, en el proteoma extracelular, sugiriendo que están involucradas en el proceso de colonización, invasión y adhesión de estas bacterias [Bumann et al. 2002]. Nosotros observamos en el secretoma de *R. etli* estas proteínas chaperonas, sugiriendo un involucramiento en el proceso de colonización e invasión de la bacteria.

Al analizar los proteomas en las fases de crecimiento exponencial y estacionario para las fracciones periplásmicas encontramos que las proteínas que presentan una mayor expresión relativa en la fase de crecimiento exponencial corresponden a algunos transportadores, las proteínas ribosomales también son más abundantes para esta fase de crecimiento, excepto el factor de elongación Tu que es más abundante en la fase estacionaria, este resultado es muy lógico si tenemos en cuenta que ha sido reportado que el factor de elongación Tu y las flagelinas tienen una participación activa en los procesos de señalización y defensa, aunque

no se conoce como ocurre exactamente este proceso, se han encontrado en plantas receptores para estas proteínas, facilitando el proceso de infección por parte de las bacterias [Zipfel et al. 2006].

Así mismo, el grupo funcional “modificación postranduccional, plegamiento de proteínas, chaperonas” presentan una mayor expresión relativa para la fase estacionaria excepto algunas proteínas como Tig, DnaK, TypA y RopB1 que están más expresadas en la fase de crecimiento exponencial. En el caso de las proteínas de función desconocida las proteínas más abundantes fueron identificadas para la fase de crecimiento exponencial. En *Yersinia pestis* un análisis del proteoma del periplasma para las fases de crecimiento exponencial y estacionario [Pieper et al 2008], mostró diferencias significativas entre la abundancia de las proteínas con respecto a ambas fases de crecimiento, estos autores reportaron que la mayoría de las proteínas transportadoras relacionadas con la función de proteínas de unión a solutos estaban fuertemente reguladas en la fase de crecimiento estacionario, ya que en esta fase las células incrementan su capacidad de importar nutrientes como aminoácidos, azúcares e iones metálicos. También fueron encontradas enzimas como glicosilasas, fosfodiesterasas y peptidasas que incrementaron su actividad en la fase de crecimiento estacionario [Pieper et al 2008]. En nuestro trabajo identificamos una gran cantidad de transportadores que al igual que en otras bacterias también son necesarios en la fase estacionaria para importar nutrientes, como iones metálicos y azúcares que tienen como finalidad contender con diferentes tipos de estreses. (Ver tabla en anexos 2). Las proteínas mencionadas anteriormente fueron identificadas también en el sobrenadante, lo que sugiere que pueden estar siendo desprendidas de la membrana durante el proceso de purificación o que están siendo secretadas en las OMV, resultados que analizaremos más adelante.

Al comparar los proteomas de la OMV en fase de crecimiento exponencial y estacionario encontramos que un total de 219 proteínas son compartidas por ambos proteomas. Un gran número de proteínas fueron más abundantes para la fase de crecimiento estacionario entre ellas tenemos: transportadores de aminoácidos, de azúcares, PotF, ArgD1, AglE, DppAch1, enzimas como las, CysK, PhbAch, proteínas de membrana externa como RopB1, transportadores de iones metálicos como SfuA y algunas proteínas de función desconocida. Sin embargo para la fase exponencial encontramos que en las OMV se identificaron proteínas enzimáticas como Mdh y PckA, también tienen una mayor abundancia un gran número de las proteínas ribosomales, lo que explica por qué fueron identificadas en el proteoma del

sobrenadante para la fase exponencial [Meneses et al. 2010]; aun se desconoce qué papel podrían tener estas proteínas cuando son secretadas en las OMV, similares resultados fueron reportados en *E. coli* en el proteoma de OMV encontrando un gran número de proteínas ribosomales [Lee et al. 2007]. Adicionalmente, presentan una expresión relativa muy alta también para la fase exponencial en OMV, las flagelinas, FlaCch1, FlaCch2, FlaCch3, FlaCch4, FlaCch6 y FlgKch con valores por encima de las 1000 unidades y en algunos casos como la FlaCch2 alcanzan niveles de 10000. La proteína Tig fue identificada con un valor de expresión relativa de 728.16 y toxinas como RzcA y RHE_CH01892 (una posible toxina RTX) su expresión relativa se mantuvo elevada con un valor de 1250.93 y 765.78 respectivamente, en ambas condiciones (ver tabla en anexos 2).

Al analizar el incremento de la formación de pelos radiculares, producido por las OMV en las raíces de las plantas de frijol, nos dimos a la tarea de buscar que ocasiona este comportamiento, por lo que tratamos de definir un posible papel de las OMV en *R. etli* en las etapas tempranas de la simbiosis decidimos estudiar los proteomas de la bacteria en presencia y ausencia de naringenina. Cuando analizamos los proteomas del sobrenadante identificamos en presencia de naringenina 628 proteínas mientras que en ausencia de naringenina fueron identificadas 625 (ver anexo 2); de estas 516 fueron comunes a las dos condiciones de las cuales muy pocas tenían una expresión incrementada para uno de los dos proteomas por ejemplo PdhB se expresa alrededor de 2 veces más para el proteoma en presencia de naringenina, esta enzima en *Pseudomonas aeruginosa* participa en la activación de factores de virulencia que se secretan por el sistema de secreción tipo III, [Dacheux et al. 2002], posiblemente esta enzima en *R. etli* esté relacionada con el mecanismo de defensa de la bacteria. También tiene una mayor expresión la enzima Adk en presencia de naringenina, esta enzima es esencial para el mantenimiento y el crecimiento de la célula [Goldberg et al. 2004], no sabemos qué implicación pueda tener en presencia de naringenina. La proteína GlnB (<http://www.uniprot.org/uniprot/Q2K8E9>) también se encuentra más abundante en el proteoma en presencia de naringenina, esta es una proteína reguladora de la utilización de nitrógeno. Adicionalmente encontramos que en el caso de las proteínas ribosomales están siendo ligeramente más abundantes y en algunos casos hasta dos veces más en presencia de naringenina, a lo cual no tenemos una explicación posible y solo sugerimos un posible papel en el exterior de la célula. Para el caso de las proteínas en ausencia de naringenina encontramos que algunos transportadores se encuentran en mayor abundancia o están más

expresados que en presencia de naringenina (ver anexo 2). En ambos proteomas encontramos posibles transportadores de nitratos, sulfonatos y bicarbonatos, con una abundancia semejante para ambas condiciones, lo que nos dice que no son dependientes de la inducción con el flavonoide.

Después de analizar las proteínas que son comunes para ambos proteomas buscamos las que son específicas para cada condición, y encontramos en presencia de naringenina a la proteína NifUch1, esta proteína está relacionada con el proceso de la fijación de nitrógeno, ya que participa en el ensamblaje de la nitrogenasa [Hu et al. 2008], así mismo en una de las replicas experimentales a las proteínas RosR (Q2KAT2) que es un regulador transcripcional del proceso de nodulación (<http://www.uniprot.org/uniprot/Q2KAT2>), la proteína GlnK (Q2K367) (<http://www.uniprot.org/uniprot/Q2K367>) proteína reguladora de la utilización de nitrógeno y una posible nitroreductasa (<http://www.uniprot.org/uniprot/Q2K861>). Es conveniente mencionar que estas proteínas no fueron encontradas para las dos replicas realizadas esto probablemente se debe a que se encuentran en baja concentración y fueron identificadas solo con 2 péptidos. Para el caso de las proteínas específicas en ausencia de naringenina encontramos que las proteínas más abundantes pertenecen al grupo funcional “transporte y metabolismo de carbohidratos”, si consideramos que el metabolismo de carbohidratos es necesario para la producción de energía en la mayoría de las reacciones metabólicas, pudiéramos inferir que en las OMV en estas condiciones se están llevando a cabo procesos que no conocemos pero que necesitan energía para su realización.

Continuamos con nuestro análisis en presencia y ausencia de naringenina pero ahora analizando las OMV, con la intención de definir si estas estructuras estaban involucradas en las etapas tempranas de la simbiosis; para el caso de las proteínas de OMV en presencia de naringenina encontramos la proteína NifUch1. Ya habíamos mencionado en anteriores párrafos que esta proteína relacionada con la fijación de nitrógeno, la habíamos identificado en el sobrenadante del cultivo en presencia de naringenina. Este resultado nos indica que esta proteína está siendo exportada en las OMV y que estas podrían tener algún papel importante durante el proceso de simbiosis de *R. etli*. Al analizar las replicas en una de ellas encontramos la proteína NtrX una proteína reguladora de la respuesta del sistema de dos componentes NtrY/NtrX el cual se sugiere está relacionado en la regulación de los genes de fijación de nitrógeno [Gregor et al. 2007]. Además encontramos las proteínas RosR, GlnK y una posible

nitroreductasa al igual que en el proteoma del sobrenadante en presencia de naringenina, un resultado que sugiere que estas proteínas son secretadas a través de las OMV.

Para *Mesorhizobium loti* fue reportado un análisis de las interacciones proteína-proteína para la fijación de nitrógeno en [Shimoda et al. 2008]; utilizando el sistema de dos híbridos (YTH), en este trabajo realizaron una búsqueda entre los genes cuya expresión es regulada en bacteroides o en microarobiosis y los genes localizados en la isla simbiótica, también estudiaron los genes que tienen homólogos conservados en otros rizobios. Encontraron interacciones con chaperonas y con la endopeptidasa Clp, así como proteínas relacionadas con el stress térmico y muchas proteínas de función desconocida. Además reportaron ferredoxinas, similar a O-linked GlcNAc transferasa, transportadores ABC, proteína de unión a ATP, transketolasa, transportadores de aminoácidos, proteína que polimeriza exopolisacáridos, entre otras muchas proteínas, lo interesante para nosotros es que estas que hemos citado están presentes en el proteoma de las OMV en presencia de naringenina, lo cual soporta que las OMV podrían estar involucradas en el proceso de simbiosis (ver anexo 2).

Adicionalmente, recientemente fue publicado un trabajo por Koch et al. 2010, en el cual se estudiaron los posibles genes involucrados en el proceso de simbiosis en *Bradyrhizobium japonicum*, realizando estudios de transcriptoma y proteómica. En este trabajo reportaron que dependiendo de la leguminosa con la cual se asocia la bacteria podrían estar involucrados en la simbiosis las siguientes proteínas: 4-Hydroxyphenylpyruvato dioxygenasa, una posible monooxygenasa, transportador ABC de unión a sustrato, Acyl-CoA dehydrogenasa, transportador ABC de unión a sulfonato, Enoyl-CoA hidratasa, posible subunidad α de la transketolasa, 2-Nitropropano dioxygenasa, además de diversas proteínas hipotéticas. En las OMV nosotros identificamos, los transportadores, la transketolasa, la Enoyl-CoA las que pueden estar relacionadas con el proceso de simbiosis en *R. etli*.

En la tabla 5 se muestran las posibles estructuras del análisis realizado al extracto butanólico de las OMV, que sugieren que en las OMV en presencia de naringenina se encuentran presentes los factores de nodulación, ya que en estas posibles estructuras está presente la N-acetylglucosamina (HexNAc), se han reportado las estructuras para muchos factores de nodulación de diferentes cepas de rizobios, en común en su estructura presentan entre 2 y 6 β - (1-4) residuos de HexNAc unidos a una amida unida a un residuo acyl graso (saturado o insaturado) en el terminal no reducido de un residuo de glucosamina (HexN) [Soria-Díaz et al 2003]. Esta estructura básica presenta variaciones que dependen de la cepa o de la especie de

Rhizobio. Nosotros analizamos un extracto butanólico de las OMV en presencia de naringenina, como no encontramos reportes de la existencia de factores Nod secretados en las OMV, realizamos el análisis por MS utilizando el extracto completo sin realizar ninguna modificación para buscar masas que nos indicaran posibles estructuras correspondientes a los factores Nod.

Así mismo, analizamos los extractos de n-butanol, mediante TLC, de las OMV y el sobrenadante para la cepa silvestre de *R. etli* CE3 y para la mutante en el gen *nodA* (CE3 NodA-), la cual no produce factores de nodulación. Este resultado nos confirma que las OMV intervienen en el proceso de simbiosis entre la bacteria y la planta, al participar en el transporte del factor de nodulación. De esta manera podemos sugerir que *R. etli* utiliza las OMV en el inicio del proceso de comunicación con la planta.

Se sabe que la presencia de los factores Nod en concentraciones del orden 10^{-12} M son suficientes para que en la planta se produzca la deformación de los pelos radiculares [Lerouge et al., 1990] pero se necesitan niveles mayores, del orden de 10^{-7} a 10^{-9} M para provocar la formación de los pre-canales de infección, la división de las células corticales y la inducción de genes implicados en las fases previas a la nodulación, las nodulinas tempranas [Truchet et al. 1991]. Esta elevada sensibilidad a los factores Nod hace suponer que debe existir un mecanismo mediado por receptores aunque aún no se han podido ni determinar el número ni identificarlos.

Nosotros proponemos que con la ayuda de las OMV los factores Nod son excretados, facilitando el proceso de infección de la bacteria. Cuando la planta secreta en sus exudados el flavonoide naringenina, la bacteria inicia la síntesis de factores Nod [Cárdenas et al. 1995; Dénarié et al. 1996; Soria-Díaz et al. 2003; Deakin et al. 2009], siendo las OMV el mecanismo de excreción de dichos factores; no hemos encontrado reportes que mencionan el sistema de secreción mediante el cual son secretados los factores Nod. Por lo que proponemos que *R. etli* utiliza al menos la vía de OMV para secretar los factores Nod y lograr el dialogo molecular inicial con la planta. Sin duda es necesario realizar experimentos que nos confirmen que las OMV es el sistema de excreción de dichos factores.

Adicionalmente, como se mencionamos anteriormente, Giraud et al en el 2007 publicaron un artículo donde demostraban que la cepa fotosintética de *Bradyrhizobium japonicum* BTAil realiza simbiosis en ausencia de factores de nodulación, ellos reportados las posible proteínas implicadas en este proceso, que corresponden a genes que codifican para proteínas

relacionadas con la síntesis o degradación de fitohormonas, modificación del antígeno O, ó lipopolisacáridos y en la biosíntesis de exopolisacáridos; pero ellos no pudieron explicar cuál es el mecanismo mediante el cual se produce la simbiosis entre la bacteria y la planta utilizando estas proteínas. Cuando realizamos una comparación de estas proteínas con las presentes en el proteoma de las OMV encontramos que en ellos están presentes el 41, 9% de estas proteínas tanto en las OMV extraídas de *R. etli* en ausencia o en presencia de naringenina. Este resultado adicionado al anterior, confirma nuestra propuesta de que las OMV deben ser consideradas como un importante enlace entre la bacteria y la planta, participando activamente en el proceso inicial de la simbiosis facilitando luego el proceso de colonización de la bacteria. En las OMV nosotros encontramos proteínas que polimerizan exopolisacáridos, proteínas exportadoras de polisacáridos y un gran número de transportadores tipo ABC, que son necesarios para que la bacteria pueda adherirse y colonizar.

CONCLUSIONES

1. El uso de la proteómica y la espectrometría de masas nos permitió realizar un análisis global de las proteínas secretadas por *R. etli* en fase de crecimiento exponencial y estacionaria.
2. En la fase crecimiento exponencial se identificaron entre otras proteínas, un gran número de proteínas ribosomales, en el secretoma, en la fracción periplásmica y en las OMV. No conocemos el papel de estas proteínas en el secretoma de *R. etli*. Aunque ya ha sido reportado para bacterias patógenas que el factor de elongación Tu interviene en el proceso de infección y que este se asocia a proteínas de membrana externa.
3. En la fracción periplásmica, en la fase de crecimiento exponencial de *R. etli*, fueron identificadas un gran número de transportadores y chaperonas. Resultado esperado ya que en el periplasma ocurre el plegamiento de muchas proteínas y además en el periplasma se realiza el transporte de elementos esenciales para la bacteria a través de la membrana externa.
4. Cuando comparamos los proteomas de las fracciones periplásmica y OMV podemos apreciar que presentan una composición similar. Estos datos sugieren que las proteínas presentes en las OMV son secretadas a través del periplasma, aunque se desconoce cómo está regulado el proceso de formación de las OMV.
5. En el secretoma de la fase de crecimiento estacionario identificamos como el grupo de proteínas más abundante, el de proteínas de función desconocida. Algunas de estas proteínas presentan dominios de enzimas que intervienen en el proceso de defensa de la bacteria como toxinas. Este grupo de proteínas nos sugiere que muchas de estas proteínas pueden tener papeles específicos en el secretoma bacteriano y al ser este de manera general poco estudiado, podría ser la razón por la cual aun no se les ha asignado una función. Por lo anterior estas proteínas podrían ser una fuente de nuevos estudios, para esclarecer su papel como proteínas exportadas.
6. En la fracción periplásmica en la fase de crecimiento estacionario fueron identificados una gran cantidad de transportadores tipo ABC, que han sido reportados como factores importantes en el mantenimiento y defensa de la célula. Adicionalmente fueron encontradas un gran número de enzimas como peptidasas, hidrolasas, fosfatasas, fosfodiesterasas, así como proteínas de membrana externa.

7. En la fracción de las OMV para la fase de crecimiento estacionario de *R. etli* las proteínas más abundantes fueron transportadores de aminoácidos, carbohidratos e iones inorgánicos. Se desconoce la función de estas proteínas en estas estructuras.
8. En las OMV en presencia y ausencia de naringenina identificamos 18 proteínas que están presentes en la cepa de *Bradyrhizobium japonicum* BTAil la cual fija nitrógeno en ausencia de factores de nodulación, estas proteínas fueron reportadas como responsables del proceso de fijación de nitrógeno. Lo que indica que las OMV están involucradas en el proceso de fijación de nitrógeno.
9. Las OMV son capaces de incrementar la formación de pelos radiculares en las raíces de las plantas de frijol.
10. En el extracto butanólico de las OMV en presencia de naringenina, identificamos una fracción que no está presente en el extracto butanólico de las OMV en ausencia de naringenina. Cuando analizamos la fracción encontrada en las OMV mediante espectrometría masas en presencia de naringenina, encontramos varias estructuras que presentan N-acetylglucosamina, que es una parte integrante de la estructura de los factores Nod lo que sugiere que posiblemente estos factores de nodulación están siendo secretados en OMV por *R. etli*.
11. Las OMV tienen un papel importante en el proceso de simbiosis ya que proteínas responsables del establecimiento de la simbiosis en *B. japonicum* BTAil y los factores Nod son secretados a través de las OMV. Quedando aun por identificar que otros metabolitos son secretados mediante estas estructuras y que puedan jugar un papel en el establecimiento del proceso simbiótico.
- 12.-En el presente proyecto se identificaron en las diferentes condiciones estudiadas un total de 1659 proteínas lo cual constituye el análisis más completo del secretoma de una bacteria fijadora de nitrógeno realizado hasta el momento.

REFERENCIAS

- Adams J, Gross M L. Energy requirement for remote charge site ion decompositions and structural information from collisional activation of alkali metal cationized fatty alcohols. *J. Am. Chem. Soc.* 1986, 108(22), 6915.
- Amanda J. McBroom and Meta J. Kuehn. Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response. *Molecular Microbiology* 2007, 63(2), 545.
- Anderson DM, y Schneewind O. A mRNA signal for the type III secretion of Yop proteins by *Yersinia enterocolitica*. *Science* 1997, 278:1140.
- Anderson, L, and Porath J. Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography. *Anal. Biochem.* 1986, 154(1), 250.
- Aslam SN, Erbs G, Morrissey KL, Newman M-A, Chinchilla D, Boller T, Molinaro A, Jackson W, Cooper RM. Microbe-associated molecular pattern (MAMP) signatures, synergy, size and charge: influences on perceptionor mobility and host defence responses. *Molecular Plant Pathology* 2009,10:375.
- Barber M, Bordoli R S, Sedgwick R D, Tyler A N. Fast Atom Bombardment: A New Ion Source for Mass Spectrometry. *J. Chem. Soc., Chem. Commun.* 1981, 325.
- Bennett JCQ, y Hughes C. From flagellum assembly to virulence: the extended family of type III export chaperones. *Trends in Microbiol* 2000, 8:202.
- Biemann, K. Contributions of mass spectrometry to peptide and protein structure, Biomed Environm. *Mass Spectrom.* 1988, 16, 99.
- Bingle EH L, Bailey CM and Pallen MJ. Type VI secretion: a beginner's guide. *Current Opinion in Microbiology* 2008, 11:3.
- Bradford M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-Dye binding. *Analytical Biochemistry* 1976,72:248.
- Broughton W, Perret X. Genealogy of legume-Rhizobium symbiosis. *Current Opinion in Plant Biology* 1999, 2: 305.
- Bumann D, Aksu S, Wendland M, Janek K, Zimny-Arndt U, Sabarth N, Meyer TF, Jungblut PR: Proteome analysis of secreted proteins of the gastric pathogen *Helicobacter pylori*. *Infection and Immunity* 2002, 70:3396.
- Cárdenas, L. *et al.*, Isolation, chemical structures and biological activity of the lipo-chitin oligosaccharide nodulation signals from *Rhizobium etli*. *Plant Mol. Biol.* 1995, 29:453.
- Chakraborty JA, Geng M, Zhang X, Amini A, Bina M and Regnier F. Strategy for qualitative and quantitative analysis in proteomics based on signature peptides. *J. Chromatogr.* 2000, 745(1), 197.
- Cheng LW, y Schneewind O. Type III machines of Gram-negative bacteria: delivering the goods. *Trends Microbiol* 2000, 8:214.
- Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S & Cascales E. Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu Rev Microbiol* 2005, 59: 451.

- Clauser KR, Baker P, Burlingame AL. Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem.* 1999, 71(14):2871.
- Cooper C A, Harrison M J, Wilkings M R, Packer N H. GlycoSuiteDB: a new curated relational database of glycoprotein glycan structures and their biological sources. *Nucleic Acids Res.* 2001 January 1; 29(1): 332.
- Czuprynski C J, Welch R A. Biological effects of RTX toxins: the possible role of lipopolysaccharide. *Trends Microbiol* 1995, 3:480.
- Dacheux D, Epaulard O, de Groot A, Guery B, Leberre R, Attree I, Polack B, and Toussaint B. Activation of the *Pseudomonas aeruginosa* Type III Secretion System Requires an Intact Pyruvate Dehydrogenase *aceAB* Operon. *Infection and Immunity*, 2002, 70: 3973.
- Seixas Chaves D F, de Souza E M, Monteiro R A, de Oliveira Pedrosa F. A two-dimensional electrophoretic profile of the proteins secreted by *Herbaspirillum seropedicae* strain Z78. *Journal of Proteomics* 2009, 73: 50.
- Deakin W J and Broughton W J. Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nature Reviews Microbiology* 2009, 7: 312.
- Delepelaire P: Type I secretion in gram-negative bacteria. *Biochim Biophys Acta* 2004, 1694(1-3):149.
- Dénarié J, Debelle F, Promé J-C. *Rhizobium* lipo-chitooligo-saccharide nodulation factors: Signalling molecules mediating recognition and morphogenesis. *Annu. Rev. Biochem.* 1996, 65: 503.
- Dorward D W, Garon C F, Judd R C. Export and Intercellular Transfer of DNA via Membrane Bleds of *Neisseria gonorrhoeae*. *Journal of Bacteriology* 1989, 171: 2499.
- Encarnación S, Salazar E, Martínez-Batallar A G, M Hernández, Reyes-Pérez A, Contreras S, Vargas M C, Domínguez-Vidaña R, Gonzaga-Jáuregui C, Mora Y, Rivero M R, Mora J. Genómica funcional de Rizobiaceas. *Rev Latinoam Microbiol* 2006; 48 (2): 131.
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM, Electrospray ionization for mass spectrometry of large biomoléculas. *Science* 1989, 246(4926), 64.
- Ferrari G, Garaguso I, Adu-Bobie J, Doro F, Taddei A R, Biolchi A, Brunelli B, Giuliani M M, Pizza M, Norais N, Grandi G. Outer membrane vesicles from group B *Neisseria meningitidis* *Δgna33* mutant: Proteomic and immunological comparison with detergent-derived outer membrane vesicles . *Proteomics* 2006, 6, 1856.
- Ficarro SB, McClelland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 2002, 20(3), 301.
- Galán JE, y Collmer A. Type III secretion machines: bacterial devices for protein delivery into host cells. *Science* 1999, 284:1322.
- Galka F, Wai S N, Kusch H, Engelmann S, Hecker M, Schmeck B, Hippenstiel S, Uhlin B E, and Steinert M. Proteomic Characterization of the Whole Secretome of *Legionella pneumophila* and Functional Analysis of Outer Membrane Vesicles. *Infection and Immunity*, 2008, 76: 1825.

- Gevaert K, Verschelde J, Puype M, Van Damme J, Goethals M, De Boeck S, Vandekerhove J. Structural analysis and identification of gel-purified proteins, available in the femtomole range, using a novel computer program for peptide sequence assignment, by matrix-assisted laser desorption ionization-reflectron time-of-flight-mass spectrometry. *Electrophoresis* 1996, 17(5), 918.
- Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E, Avarre J-C, Jaubert M, Simon D, Cartieaux F, Prin Y, Bena G, Hannibal L, Fardoux J, Kojadinovic M, Vuillet L, Lajus A, Cruveiller S, Rouy Z, Mangenot S, Segurens B, Dossat C, Franck W L, Chang W-S, Saunders E, Bruce D, Richardson P, Normand P, Dreyfus B, Pignol D, Stacey G, Emerich D, Verméglio A, Médigue C, Sadowsky M. Legumes Symbioses: Absence of Nod Genes in Photosynthetic Bradyrhizobia. *Science* 2007, 316: 1307.
- Goldberg RN, Tewari YB, Bhat TN. Thermodynamics of enzyme-catalyzed reactions--a database for quantitative biochemistry. *Bioinformatics*, 2004; 20(16): 2874.
- Gómez-Gómez L, Boller T. FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Mol Cell* 2000, 5:1003.
- Gorg A, Obermaier GB, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 2000, 21: 1037.
- Gregor J, Zeller T, Balzer A, Habertzettl K, Klug G. Bacterial regulatory networks include direct contact of response regulator proteins: interaction of RegA and NtrX in *Rhodobacter capsulatus*. *J Mol Microbiol Biotechnol*. 2007; 13(1-3):126.
- Hayes, R. N.; Gross, M. L. Collision-induced dissociation. *Methods Enzymol*. 1990,193: 237.
- He SY, Nomura K & Whittam TS. Type III protein secretion mechanism in mammalian and plant pathogens. *Biochim Biophys Acta* 2004, 1694: 181.
- Heller M, Schlappritzi E, Stalder D, Nuoffer J-M, Haerberli A. Compositional Protein Analysis of High Density Lipoproteins in Hypercholesterolemia by Shotgun LC-MS/MS and Probabilistic Peptide Scoring. *Molecular & Cellular Proteomics* 2007, 6:1059.
- Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D: Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* 2004,68(4):692.
- Hu Y, Fay A W, Lee C C, Yoshizawa J, and Ribbe M W. Assembly of Nitrogenase MoFe Protein. *Biochemistry* 2008, 47, 3973.
- Hueck CJ. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* 1998, 62:379.
- Jan Wensink y Bernard Witholt. Outer-Membrane Vesicles Released by Normally Growing *Escherichia coli* Contain Very Little Lipoprotein. *Eur. J. Biochem*. 1981,116: 331.
- Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem*. 1988, 60(20):2299.
- Kazemi-Pour N, Condemine G, Hugouvieux-Cotte-Pattat N. The secretome of the plant pathogenic bacterium *Erwinia chrysanthemi*. *Proteomics* 2004, 4, 3177.
- Kettman JR, Coleclough C, Frey JR, Lefkovits I. Clonal proteomics: One gene-family of proteins. *Proteomics* 2002, 2, 624.

- Kim H, Feil IK, Verlinde CL, Petra PH, Hol WG. Crystal structure of glycosomal glyceraldehyde-3-phosphate dehydrogenase from *Leishmania mexicana*: implications for structure-based drug design and a new position for the inorganic phosphate binding site. *Biochemistry* 1995; 34 :14975.
- Koch M, Delmotte N, Rehrauer R, Vorholt J A, Pessi G, and Hennecke H. Rhizobial Adaptation to Hosts, a New Facet in the Legume Root-Nodule Symbiosis. *MPMI* 2010, 23, 6: 784.
- Koster M, Bitter W, y Tommassen J . Protein secretion mechanisms in Gram-negative bacteria. *Int J Med Microbiol* 2000, 290:325.
- Kuehn M J Kesty NC. Bacterial outer membrane vesicles and the host-pathogen interaction. *Genes Dev* 2005, 19: 2645.
- Kurian D, Phadwal K and Mäenpää P. Proteomic characterization of acid stress response in *Synechocystis* sp. PCC 6803. *Proteomics* 2006, 6: 3614.
- Lauren M. Mashburn & Marvin Whiteley. Membrane vesicles traffic signals and facilitate group activities in a prokaryote. *Nature* 2005, 437:422.
- Lee E-Y, Bang J Y, Park G W, Choi D-S, Kang J S, Kim H-J, Park K-S, Lee J-O, Kim Y-K, Kwon K-H, Kim K-P, Gho Y G. Global proteomic profiling of native outer membrane vesicles derived from *Escherichia coli*. *Proteomics* 2007, 7, 3143.
- Lee VT, y Schneewind O. Protein secretion and the pathogenesis of bacterial infections. *Genes and Develop* 2001, 15:1725.
- Lerouge P, Roche P, Faucher C, Maillet F, Truchet G, Promé JC, Dénarié J. Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* 1990, 19; 344:781.
- Li M, Rosenshine I, Tung S L, Wang X H, Friedberg D, Hew C L and Leung K Y. Comparative Proteomic Analysis of Extracellular Proteins of Enterohemorrhagic and Enteropathogenic *Escherichia coli* Strains and Their *ihf* and *ler* Mutants. *Applied and Environmental Microbiology*, 2004, 70, 9: 5274.
- Li Z, Clarke A J, Beveridge T J. Gram-Negative Bacteria Produce Membrane Vesicles Which Are Capable of Killing Other Bacteria. *Journal of Bacteriology* 1998, 180: 5478.
- Liska A J, Schevchenko A. Expanding the organismal scope of proteomics: cross-species protein identification by mass spectrometry and its implications. *Proteomics* 2003, 3(1), 19.
- Lloyd SA, Sjöström M, Andersson S, y Wolf-Watz H. Molecular characterization of the type III secretion signals via analysis of synthetic N-terminal amino acid sequences. *Mol Microbiol* 2002, 43:51.
- Lorenz M G y Wackernagel W. Bacterial Gene Transfer by Natural genetic Transformation in the Environment. *Microbiological Reviews* 1994, 58: 563.
- Mackey A J, Haystead TA, Pearson WR. Getting more from less: algorithms for rapid protein identification with multiple short peptide sequences. *Mol Cell Proteomics* 2002, 1(2):139.
- Macnab R M. The bacterial flagellum: reversible rotary propellor and type III export apparatus. *J Bacteriol* 1999, 181:7149.

- Mann M and Wilm M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem.* 1994, 66(24): 4390.
- Mann M, Hojrup P, Roepstorff P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol. Mass Spectrom.* 1993, 22(6): 338.
- Martínez-Romero, E., Pardo, M.A., Cevallos, M.A. and Palacios, R. "Reiteration of Nitrogen Fixation Gene Sequences and Specificity of *Rhizobium* in Nodulation and Nitrogen Fixation in *Phaseolus vulgaris*". *J. Gen. Microbiol.* 1985, 131:1779.
- Mattow J, Schaible U E, Schmidt F, Hagens K, Siejak F, Brestrich G, Haeselbarth G, Müller E-C, Jungblut P R, Kaufmann S H E. Comparative proteome analysis of culture supernatant proteins from virulent *Mycobacterium tuberculosis* H37Rv and attenuated *M. bovis* BCG Copenhagen. *Electrophoresis* 2003, 24, 3405.
- McBroom A J, Johnson A P, Vemulapalli S, and Kuehn M J. Outer Membrane Vesicle Production by *Escherichia coli* Is Independent of Membrane Instability. *Journal of Bacteriology* 2006, 188: 5385.
- McBroom A J, Kuehn M J. Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response. *Molecular Microbiology* 2007, 63(2): 545.
- Medzihradzky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, Vestal ML, Burlingame AL. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal Chem.*, 2000, 72(3): 552.
- Meneses N, Mendoza-Hernandez G, Encarnación S. The Extracellular Proteome of *Rhizobium etli* CE3 in Exponential and Stationary Growth Phase. *Proteome Science* 2010, 8:51.
- Nally J E, Whitelegge J P, Aguilera R, Pereira M M, Blanco D R, Lovett M A. Purification and proteomic analysis of outer membrane vesicles from a clinical isolate of *Leptospira interrogans* serovar Copenhageni. *Proteomics* 2005, 5: 144.
- Opitck G J and Jorgenson JW. Two-dimensional SEC/RPLC coupled to mass spectrometry for the analysis of peptides. *Anal.Chem.*, 1997,69(13): 2283.
- Page A-L, y Parsot C. Chaperones of the type III secretion pathway: jacks of all trades. *Mol Microbiol* 2002, 46:1.
- Perham RN. The fructose-1,6-bisphosphate aldolases: same reaction, different enzymes. *Biochem Soc Trans.* 1990 ;18(2):185.
- Pieper R, Huang S-T, Clark D J, Jeffrey M. Robinson J M, Parmar P P, Alami H, Bunai C L, Perry R D, Fleischmann R D and Peterson S N. Characterizing the dynamic nature of the *Yersinia pestis* periplasmic proteome in response to nutrient exhaustion and temperature change. *Proteomics* 2008, 8, 1442.
- Plano GV, Day JB, y Ferracci F. Type III export: new uses for an old pathway. *Mol Microbiol* 2001, 40: 284.
- Poupot R, Martínez-Romero E, Gautier N, Promé J-C. Wild-type *Rhizobium etli*, a bean symbiont, produces acetyl-fucosylated, N-methylated, and carbamoylated nodulation factors. *J. Biol. Chem.* 1995, 270: 6050.
- Poupot R, Martínez-Romero E, Promé JC. Nodulation factors from *Rhizobium tropici* are sulfated or nonsulfated chitopentasaccharides containing an N-terminal – N-acylglucosaminyl terminus. *Biochemistry* 1993, 32: 10430.

- Quesada-Vincens D, Fellay R, Nasim T, Viprey V, Burger U, Promé J-C, Broughton W J and Jabbouri S. *Rhizobium* sp. strain NGR234 NodZ protein is a fucosyltransferase. *J. Bacteriol.* 1997, 179: 5087.
- Raida M, Schulz-Knappe P, Heine G, Forssmann WG. Liquid chromatography and electrospray mass spectrometric mapping of peptides from human plasma filtrate. *J Am SocMass Spectrom.* 1999, 10(1): 45.
- Renelli M, Matias V, Lo R Y, Beveridge T J. DNA-containing membrane vesicles of *Pseudomonas aeruginosa* PAO1 and their genetic transformation potential. *Microbiology*, 2004, 150: 2161.
- Roche P, Debelle Â F, Maillet F, Lerouge P, Faucher C, Truchet G, et al. Molecular basis of symbiotic host specificity in *Rhizobium meliloti* : nodH and nodPQ genes encode the sulfation of lipo-oligosaccharide signals. *Cell* 1991, 67: 1131.
- Roepstorff, P.; Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* 1984,11(11): 601.
- Rose RW, Bruser T, Kissinger JC, Pohlschroder M: Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Mol Microbiol* 2002, 45(4):943.
- Sandkvist M: Biology of Type II secretion. *Mol Microbiol* 2001, 40(2):271.
- Schevchenko A, Sunyaev S, Loboda A, Bork P, Ens W, Standing K G. Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. *Anal. Chem.* 2001, 73(9): 11917.
- Scorza F B, Doro F, Rodri'guez-Ortega M J, Stella M, Liberatori S, Taddei A R, Serino L, Gomes Moriel D, Nesta B, Fontana M R, Spagnuolo A, Pizza M, Norais N, and Grandi G. Proteomics Characterization of Outer Membrane Vesicles from the Extraintestinal Pathogenic *Escherichia coli* *_tolR* IHE3034 Mutant. *Molecular & Cellular Proteomics* 2008, 7:473.
- Shimoda Y, Shinpo S, Kohara M, Nakamura Y, Tabata S, and Sato S. A Large Scale Analysis of Protein-Protein Interactions in the Nitrogen-fixing Bacterium *Mesorhizobium loti*. *DNA Research* 2008, 15: 13.
- Sidhu V K, Vorhölter F-J, Niehaus K and Watt S A. Analysis of outer membrane vesicle associated proteins isolated from the plant pathogenic bacterium *Xanthomonas campestris* pv. *Campestris*. *BMC Microbiology* 2008, 8:87.
- Silipo A, Molinaro A, Sturiale L, Dow JM, Erbs G, Lanzetta R, Newman MA, Parrilli M. The elicitation of plant innate immunity by lipooligosaccharide of *Xanthomonas campestris*. *J Biol Chem* 2005, 280(33):660.
- Soria-Díaz M E, Tejero-Mateo P, Espartero J L, Rodríguez-Carvajal M A, Morón B, Sousa C, Megías M, Amarger N, Thomas-Oates J, Gil-Serrano A M. Structural determination of the lipo-chitin oligosaccharide nodulation signals produced by *Rhizobium giardinii* bv. *Giardinii* H152. *Carbohydrate Research* 2003,338: 237.
- Soto MJ, Domínguez-Ferreras A, Pérez-Mendoza D, Sanjuán J, Olivares J. Mutualism versus pathogenesis: the give-and-take in plant-bacteria interactions. *Cellular Microbiology* 2009, 11:381.

- Soufoa H J D, Reimolda C, Linneb U, Knusta T, Geschera J, and Graumann P L. Bacterial translation elongation factor EF-Tu interacts and colocalizes with actin-like MreB protein. *PNAS* 2010, 107 (7): 3163.
- Spaink H P, Wijfjes A H M, Lugtenberg B J J. *Rhizobium* NodI and NodJ proteins play a role in the efficiency of secretion of lipochitin oligo-saccharides. *J. Bacteriol.* 1995, 177: 6279.
- Stacey G, Luka S, Sanjuán J, Banfalvi Z, Nieuwkoop A J, Chun J Y, Forsberg L S, Carlson R. *nodZ*, a unique host-specific nodulation gene, is involved in the fucosylation of the lipooligosaccharide nodulation signal of *Bradyrhizobium japonicum*. *J. Bacteriol.* 1994, 176: 620.
- Stancik LM, Stancik D M, Schmidt B, Barnhart D M, Yoncheva Y N, and Slonczewski J L. pH-Dependent Expression of Periplasmic Proteins and Amino Acid Catabolism in *Escherichia coli*. *Journal of Bacteriology* 2002, 184: 15, 4246.
- Sunyaev S, Liska A J, Golod A, Schevchenko A. MultiTag: multiple error-tolerant sequence tag search for the sequence-similarity identification of proteins by mass spectrometry. *Anal. Chem.* 2003, 75(6): 1307.
- Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y and Yoshida T. Protein and polymer analysis of up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.*, 1988, 2: 151.
- Thanassi DG, y Hultgren SJ . Multiple pathways allow protein secretion across the bacterial outer membrane. *Curr Op Cell Biol* 2000, 12:420.
- Trost M, Wehmhöner D, Kärst U, Dieterich G, Wehland J, Jänsch L. Comparative proteome analysis of secretory proteins from pathogenic and nonpathogenic *Listeria* species. *Proteomics* 2005, 5: 1544.
- Truchet G, Roch P, Lerouge P, Vasse J, Camut S, de Billy J, Prome J C & Denarie J. *Nature* 1991 (London) 351: 670.
- Vaughan T E, Skipp P J, O'Connor C D, Hudson M J, Vipond R, Elmore M J, Gorringe A R. Proteomic analysis of *Neisseria lactamica* and *Neisseria meningitidis* outer membrane vesicle vaccine antigens . *Vaccine* 2006, 24: 5277.
- Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphery-Smith I. Progress with gene-product mapping of *Mycoplasma genitalium*. *Electrophoresis* 1995, 16:1090.
- Watt S A, Wilke A, Patschkowski T and Niehaus K. Comprehensive analysis of the extracellular proteins from *Xanthomonas campestris* pv. *campestris* B100. *Proteomics* 2005, 5:153.
- Wensink J y Witholt B. Outer-Membrane Vesicles Released by Normally Growing *Escherichia coli* Contain Very Little Lipoprotein. *European Journal of Biochemistry* 1981, 116 (2):331.
- Whitehouse CM, Dreyer RN, Yamashita M, Fenn JB., Electrospray interface for liquid chromatographs and mass spectrometers. *Anal Chem.* 1985, 57(3): 675.
- Wool A, Smilansky Z. Precalibration of matrix-assisted laser desorption/ionization-time offlight spectra for peptide mass fingerprinting. *Proteomics* 2002, 2(10):1365.

Yaron S, Kollins G L, Simon L, Matthews K R. Vesicle-Mediated Transfer of Virulence Genes from *Escherichia coli* O157:H7 to Other Enteric Bacteria. *Applied and Environmental Microbiology*, 2000, 66: 4414.

Yergey AL, Coorsen JR, Backlund PS Jr, Blank PS, Humphrey GA, Zimmerberg J, Campbell JM, Vestal ML. De novo sequencing of peptides using MALDI/TOF-TOF. *J Am Soc Mass Spectrom*. 2002, 13(7):784.

Zhou L, Srisatjaluk R, Justus D E, Doyle R J. On the origin of membrane vesicles in Gram-negative bacteria. *FEMS Microbiology Letters* 1998, 163: 223.

Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones J D G, Boller T, and Felix G. Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts *Agrobacterium*-Mediated Transformation. *Cell* 2006, 125: 749.

ANEXOS

Anexo 1

Meneses N, Mendoza-Hernández G and Encarnación S. The Extracellular Proteome of *Rhizobium etli* CE3 in Exponential and Stationary Growth Phase. *Proteome Science* 2010, 8:51.

Anexo 2

Proteínas identificadas para *R. etli* en fase de crecimiento exponencial y estacionario.

RESEARCH

Open Access

The extracellular proteome of *Rhizobium etli* CE3 in exponential and stationary growth phase

Niurka Meneses¹, Guillermo Mendoza-Hernández², Sergio Encarnación^{1*}

Abstract

Background: The extracellular proteome or secretome of symbiotic bacteria like *Rhizobium etli* is presumed to be a key element of their infection strategy and survival. Rhizobia infect the roots of leguminous plants and establish a mutually beneficial symbiosis. To find out the possible role of secreted proteins we analyzed the extracellular proteome of *R. etli* CE3 in the exponential and stationary growth phases in minimal medium, supplemented with succinate-ammonium.

Results: The extracellular proteins were obtained by phenol extraction and identified by LC-ESI MS/MS. We identified 192 and 191 proteins for the exponential and stationary phases respectively. Using the software Signal P, we predicted signal peptides for 12.95% and 35.60% of the proteins identified in the exponential and stationary phases, respectively, which could therefore be secreted by the Sec pathway. For the exponential growth phase, we found in abundance proteins like the ribosomal proteins, toxins and proteins belonging to the group "defence mechanisms". For the stationary growth phase, we found that the most abundant proteins were those with unknown function, and in many of these we identified characteristic domains of proteases and peptidases.

Conclusions: Our study provided the first dataset of the secretome of *R. etli* and its modifications, which may lead to novel insights into the adaptive response of different stages of growth. In addition, we found a high number of proteins with unknown function; these proteins could be analyzed in future research to elucidate their role in the extracellular proteome of *R. etli*.

Background

Nitrogen-fixing symbioses between plants belonging to the family *Leguminosae* and soil bacteria collectively called rhizobia, contribute substantially to plant productivity. *Rhizobium etli* is a Gram-negative soil bacterium that carries out symbiosis with bean plants, specifically with *Phaseolus vulgaris*. The plant offers the bacteria a source of carbon (product of photosynthesis) in exchange for fixed nitrogen. This association allows legumes, like bean, to grow in nitrogen poor soil. To establish symbiosis the two partners exchange a series of molecular signals. In general, symbiotic or pathogenic processes are mediated by elicitors that are perceived by receptors on the cell membrane of the plant. These elicitors may be directly synthesized by bacteria and

included in cell wall components, like lipopolysaccharides, proteins or peptides, or they may be low molecular weight compounds released from the plant cell wall by bacterial enzymes during the infection process [1]. These effectors are commonly referred to as microbe-associated molecular patterns (MAMPs). MAMPs are mainly surface-associated molecules that mostly cannot be readily altered and are typical of classes of microorganisms [2]. They are usually related with virulence factors and another group of secreted proteins. These are usually degrading enzymes that are secreted actively and support the early and late stages of infection [1].

The secreted proteins are very important in the adaptation of the bacteria to the environment, as well as in the interaction with other organisms. These proteins have different functional roles such as degradation of substrates as well as cell wall turnover and sensing [3]. However, the role, of many of the proteins secreted by *R. etli* is unknown.

* Correspondence: encarnac@ccg.unam.mx

¹Programa de Genómica Funcional de Procaríotes, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Apdo. Postal 565-A, Cuernavaca, Morelos, CP 62210, México

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

In this study, we used a proteomic approach in order to display the extracellular proteome (secretome) of *R. etli* CE3 cultivated in minimal medium.

The extracellular proteins were extracted from the culture supernatant of *R. etli* CE3 in exponential and stationary growth phases. These extracts were analyzed by LC-ESI MS/MS, which allowed us to classify the proteins identified in functional groups. We also carried out, a bioinformatics analysis to identify proteins having signal peptides. The collected data from the extracellular proteome of *R. etli* CE3 cultivated in minimal medium, should give an insight into which proteins are secreted during two of the most important phases of growth (exponential and stationary phases), and how these proteins could participate in the survival of the bacteria in its environment and during the interaction between *R. etli* and its host plant (*P. vulgaris*).

Materials and methods

Cell culture and preparation of culture supernatant

Rhizobium etli strain CE3 was grown as previously described [4], at 30°C in 2 L of minimal medium (0.2 mM magnesium sulphate; 0.1 mM potassium monoacid phosphate; 0.2 mM succinic acid, 0.1 mM ammonium chloride, 1 mM calcium chloride) for 6 h and 24 h for exponential and stationary phases, respectively. Cells were pelleted by centrifugation (7500 × g at 4°C, 30 min). The liquid phase was centrifuged for 45 min at 6500 × g at 4°C. The cell-free supernatant was frozen at 80°C and lyophilized to complete dryness.

Extraction of extracellular proteins

The lyophilized supernatant was rehydrated in 5 ml extraction buffer (0.7 M sucrose; 0.5 M tris-base; 0.1 M potassium chloride; 50 mM EDTA; 2% β-mercaptoethanol; 12 mg/ml Polyvinylpyrrolidone PVPP; 30 mM hydrochloric acid), 5 ml water-saturated phenol was added, mixed thoroughly and incubated at RT for 45 min on a tilting table. The mixture was centrifuged at 3000 × g at 4°C for 30 min and the phenol phase was transferred to a fresh tube, supplemented with 200 μl of 1 M 1,4-dithiothreitol (DTT) and 300 μl of 8 M-ammonium acetate and incubated for 30 min. The phenol extraction was used to separate the proteins from the exopolysaccharides, DNA and other molecules that were also secreted in large amounts by *R. etli* CE3. Proteins were precipitated by the addition of 25 ml of -20°C methanol and incubation at -20°C for at least 12 h. Subsequently, proteins were pelleted by centrifugation (5000 × g at 4°C for 10 min). The precipitate was washed twice with 10 ml cold 80% (v/v) acetone, and dried at RT, and resuspended in 200 μl of solubilization buffer (7 M urea; 2% CHAPS; 1 mM DTT). For each experiment, we applied up to 500 μg of protein sample.

Tryptic digest and RP/LC prefractionation

After adjusting the pH to 8.0 with 100 mM Tris-HCl, the samples were reduced by adding 1 mM dithiothreitol and alkylated in 10 mM iodoacetamide. Proteins were then precipitated with 10% trichloroacetic acid, washed twice with 95% acetone and resolubilized in 100 mM Tris-HCl, pH 8.0. Digestion was performed overnight at 37°C with porcine modified trypsin (1/50, w/w, trypsin/sample). In order to simplify the complexity of the peptide mixture, tryptic digests were subjected to an off-line reverse-phase LC prefractionation of peptides on an Aquapore OD-300, 2.1 × 100 mm C18 column (Perkin-Elmer Corporation, Norwalk, USA). Digests were acidified with trifluoroacetic acid (TFA) and loaded onto the column previously equilibrated in 95% buffer A (0.1% TFA in water), 5% buffer B (0.08% TFA in acetonitrile). The chromatographic separation was developed at a flow rate of 100 μl/min as follows: 5-15% buffer B over 10 min, 15-35% buffer B over 65 min, 35-80% buffer B over 15 min, and a 10 min hold at 80% buffer B. Fractions of 0.5 ml were collected, dried by centrifugal evaporation and dissolved in 20 μl of 1% acetic acid prior to analysis by on-line LC-ESI-MS/MS.

LC-ESI-MS/MS

MS/MS analysis of each fraction obtained from the off-line separation step was carried out on a 3200 Q TRAP hybrid tandem mass spectrometer (Applied Biosystems/MDS Sciex, Concord, ON, Canada), equipped with a nano electro spray ion source (NanoSpray II). The instrument was coupled on-line to a nanoAcquity Ultra Performance LC system (Waters Corporations, Milford, MA). Samples were desalted by injection onto a Symmetry C18 UPLC trapping capillary column (180 μm × 20 mm, Waters Corporations) and washed with 0.1% formic acid in 100% MilliQ water at a flow rate of 15 μl/min. After 3 min, the trap column was switched in-line with the capillary analytical column. Peptides were separated on an ethylene bridged hybrid, C18 UPL column (75 μm × 100 mm, Waters Corporations) using a linear gradient of 2-70% acetonitrile, 0.1% formic acid over a 60 min period, at a flow rate of 0.25 μl/min. Spectra were acquired in automated mode using Information Dependent Acquisition (IDA). Precursor ions were selected in Q1 using the enhanced MS mode with a scan range of *m/z* 400-1500 and 4000 amu/s. Selected ions were subjected to an enhanced resolution scan at the low speed of 250 amu/s over a narrow (30 amu) mass range and then to an enhanced product ion scan (MS/MS). The precursor ions were fragmented by collision activated dissociation (CAD) in the Q2 collision cell using rolling collision energy. The fragment ions generated were captured and mass analyzed in the Q3 linear IT.

Bioinformatics analysis of the extracellular proteome of *R. etli* CE3

Data interpretation and protein identification were performed from the MS/MS spectra datasets using the MASCOT search algorithm (Matrix Science, London, UK available at <http://www.matrixscience.com>). Data were queried against a *R. etli* CE3 database (Retli_Marzo 2008, NC_007761.1). Trypsin was set as the specific digest reagent, the precursor mass tolerance and fragment ion tolerances equalled 0.5 and 0.8 Da, respectively, one missed cleavage site was allowed, and carbamidomethyl-cysteine as a fixed and methionine oxidation set as a variable modification. The protein identification reporting criteria included at least two MS/MS spectrum matched at the 95% level of confidence (Mowse score = 25 in the conditions used in this work) and the presence of a consecutive *y* and/or *b* ion series of three or more amino acids.

(Mowse score = $-10 \times \log_{10}(p)$, where *p* is the likelihood that the identification is a random event).

N-terminal signal-peptide prediction

The SignalP 2.0 software (<http://www.cbs.dtu.dk/services/SignalP-2.0/>) was used with the settings for Gram-negative organisms. Signal peptides were predicted by a Hidden Markov Model as well as by a neural networks method. Protein classification by cluster orthologous groups (COG) was performed on the NCBI server (<http://www.ncbi.nih.gov/COG>).

Results and discussion

A fundamental biological goal in bacteriology is the identification of gene expression patterns for all major functional classes during the exponential and stationary phases, and to relate these changes to important physiological events. To complement this knowledge, the goal must be to determine the specific role of different set of proteins secreted in the different growth phases. Towards this goal, we performed secretome analysis of *R. etli* batch cultures during the exponential-growth phase and early stationary phase with the purpose of identifying the extracellular proteome in these phases of growth. To do this we cultured a *R. etli* in Minimal Medium (MM) succinate-ammonium, because in the rizosphere *R. etli* lives in conditions similar to those in minimal medium; consequently growing *R. etli* in minimal medium may give a more realistic idea about the behaviour of this bacterium on its environment. In addition, the fermentative state of free living *R. etli* in MM may be closely related to the metabolism that these bacteria express during the infection process and nitrogen fixation in symbiosis [4,5].

Previous studies showed that *R. etli* under fermentative growth conditions reached mid-exponential phase,

and stationary phase at 6 h and 24 h respectively [4,5]. To reconfirm this, we cultured *R. etli* CE3 in MM, and demonstrated that our collection times represent both growth phases (data not show). Furthermore, the availability of nitrogen, succinate and oxygen, was measured throughout the incubation, to determine if some factors become limited and cause the culture to reach the stationary phase (data not shown), because any kind of restriction could alter protein synthesis, which would have an impact on the composition of the proteome. We concluded that the only factor limiting is oxygen, which becomes limiting after 10 h (this coincides with the culture entering pre-stationary phase). The other compounds (nitrogen and succinate) measured in the culture medium are in sufficient concentration in the period preceding sampling.

The extracellular proteins of *R. etli* CE3 of exponential or stationary growth phase were obtained from cell-free culture supernatant as indicated in Materials and Methods. A phenol extraction method was applied to separate the proteins from exopolysaccharides. Proteins derived from this purification method were analyzed by LC-MS/MS. Two independent experiments of each growth condition were run using 500 µg of protein.

By the criteria specified in Materials and Methods, we identified 327 and 280 proteins for the exponential and stationary phase, respectively; and a total of 192 and 191 proteins were found reproducibly in both replicates for each growth phases (Figure 1). Being the percentage of reproducibility for each pair of trials; 67% and 80%, (exponential phase) 82.3% and 79.9% (stationary phase) (Figure 1). Information about the identified proteins is given in the Additional file 1 and Additional file 2.

The proteins identified were grouped by COG (Cluster Orthologous Groups), according to NCBI. The proteins could be allocated to nineteen functional groups, the largest groups being translation, ribosomal structure and biogenesis, carbohydrate transport and metabolism,

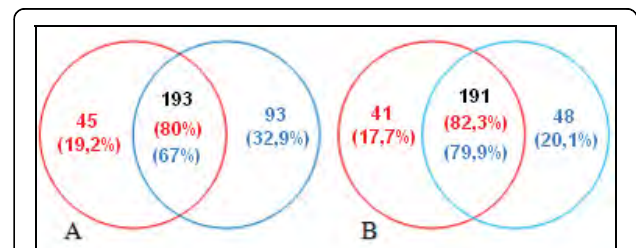


Figure 1 Common proteins found between two experiments and the number of identified proteins in each experiment.

Diagram A corresponds to the *R. etli* experiments in exponential phase. Diagram B corresponds to the *R. etli* experiments in stationary phase.

post-translational modification, protein folding, chaperons and unknown protein function.

Proteins secreted by *R. etli* CE3 in exponential phase

The proteins identified in the exponential growth phase were mainly proteins involved in bacterial physiology and survival. The most represented COGs were “translation, ribosomal structure and biogenesis” with 43 proteins identified, followed by the “energy production and conversion” that represented 19 proteins; “post-translational modification, folded proteins and chaperons” with 17 proteins; “metabolism and transport of carbohydrates” with 15 proteins; “amino acid transport and metabolism” 13 proteins; “cell wall, membrane, envelope biogenesis” with 10 proteins. The rest of the groups represented less than 5% of the total of proteins (Table 1).

From the identified proteins, those corresponding to defence mechanisms are; probable antibiotic resistance (streptomycin kinase) protein, putative 6-aminohexanoate-dimer hydrolase protein and probable efflux cation

transporter protein (Additional file 1). These proteins are essential since they are involved in the bacteria defence against other organisms [6].

ThuK (trehalosemaltose ABC transporter) and RHE_CH03989 (Probable sugar ABC transporter, ATP-binding protein) are two ABC proteins identified in exponential phase predicted to be involved in carbohydrate transport, which contained a conserved ATP-binding Walker A region (GXXGXGKT/S) [7], and were probably secreted by the type I protein secretion systems.

In this growth phase were also identified ClpX, an ATP-dependent Clp protease and heat shock protein. These proteins were also identified on the cell surface of *Corynebacterium glutamicum*, which might play a role in the degradation of misfolded proteins, and assist the refolding of proteins related to stress conditions [8]. The presence of proteins DegPch2, a serine protease DO-like protein, and Lon protein, an ATP-dependent protease LA protein, in addition to their participation in stress response, could suggest a putative role in the degradation of plant defensive barriers.

Table 1 Comparison of the extracellular proteins of *R. etli*; identified in the exponential growth phase (6 h) and the stationary growth phase (24 h)

COGs	prot 6 h (%)	prot 6 h (SP) (%)	prot 24 h (%)	prot 24 h (SP) (%)
Energy production and conversion (C)	9.8	1.0	3.1	-
Cell cycle control, cell division, chromosome partitioning (D)	0.5	-	0.5	-
Amino acid transport and metabolism (E)	6.7	1.0	7.3	4.7
Nucleotide transport and metabolism (F)	1.0	-	-	-
Carbohydrate transport and metabolism (G)	7.7	1.0	7.9	4.1
Coenzyme transport and metabolism (H)	1.5	-	0.5	-
Lipid transport and metabolism (I)	1.0	-	0.5	-
Translation, ribosomal structure and biogenesis (J)	22.3	-	6.3	-
Transcription (K)	3.1	-	3.1	-
Replication, recombination and repair (L)	3.6	-	5.2	0.5
Cell wall, membrane, envelope biogenesis (M)	5.1	1.5	7.3	3.7
Cell motility (N)	3.6	1.5	2.0	0.5
Posttranslational modification, protein turnover, chaperones (O)	8.8	1.0	6.8	1.0
Inorganic ion transport and metabolism (P)	3.1	0.5	5.7	3.7
Secondary metabolites biosynthesis, transport and catabolism (Q)	2.6	0.5	0.5	-
General function prediction only (R)	2.0	-	3.6	1.6
Function unknown (S)	3.6	1.5	28.8	11.0
Signal transduction mechanisms (T)	3.1	1.0	1.0	-
Intracellular trafficking, secretion, and vesicular transport (U)	4.1	0.5	1.5	1.6
Defense mechanisms (V)	1.5	-	1.0	-
Without COG (WC)	4.1	1.5	6.2	3.1
Total protein	192	25	191	68

The proteins were grouped in functional groups (COGs). The relative contribution to each COG of proteins with a potential signal peptide (SP) is indicated. Functional groups with a major abundance in proteins are shown in bold letters and numbers.

In our study, proteins found in the extracellular proteome in exponential growth phase, are also identified in the cytoplasmic fraction like SucB, SdhA, AtpA, AtpC, AtpD and AtpH proteins (data not shown). These proteins were reported also on the cell surface of *C. glutamicum*, and were related to stress response [8].

In agreement with our data, Kazemi-Pour et al [9], found many *Erwinia chrysanthemum* cytoplasmic proteins in the secretome. These proteins, identified also by us in the secretome of *R. etli*, included two major cytoplasmic proteins, the elongation factor EF-Tu and the chaperonin GroEL, two cytoplasmic enzymes, AcnB and RhaA, and periplasmic proteins involved in chemotaxis (Tsr) and transport (OppA). These proteins perform functions within the cell such as protein synthesis and carbon metabolism. Derived from their function and supported by the results of the Signal P analysis (see below), it is most likely that these proteins are not actively exported into the culture medium. Cell lysis could release proteins into the extracellular space, contaminating the extracellular proteome with cytosolic proteins. To assess if cytosolic proteins originating from cell lysis and contaminated the secretome, we compared the protein profiles of whole cell lysates with those of extracellular proteins using two-dimensional gel electrophoresis (Additional file 3). The protein composition of extracellular proteins was considerably different from that of whole cell samples. Moreover, a direct relationship between the relative spot volume in the cellular proteome and secretome was not found, suggesting that very little cell lysis occurred during growth and subsequent procedures. However, although these data suggest that large scale lysis does not occur during cultivation, we cannot totally rule out that the cytosolic proteins in the extracellular space are proteolysis resistant remains of lysed cells. Cell lysis during harvesting and processing of the sample can be ruled out, because this should release all of the most abundant cytosolic proteins into the extracellular space, which should then be detected in the extracellular proteome.

Additionally, in this phase of growth, we unexpectedly identified a great number of proteins belonging to the translation, ribosomal structure and biogenesis group. Together with elongation factor EF-Tu protein (TufB) which was also identified by Kazemi-Pour [9], we found; 16, 30S ribosomal proteins, 19, 50S ribosomal protein, a polyribonucleotide nucleotidyltransferase protein (Pnp), aspartyl-tRNA synthetase protein (AspS), elongation factor G protein (FusA2), glutamyl-tRNA amidotransferase subunit B protein (GatB), alanyl-tRNA synthetase protein (AlaS2) and glutamyl-tRNA amidotransferase subunit A protein (GatA2) (Additional file 1).

When we compared the secreted proteins obtained during the exponential phase versus stationary phase, we

observed that this group of proteins (detected in both experimental replicates of exponential phase growth) practically disappeared (Additional file 1 and 2), corroborating our conclusion that they are not mainly product of cell lysis. Therefore, cellular lysis during growth is not likely to be the sole reason for the appearance of cytosolic proteins in the extracellular proteome [1].

Chaperones and translation elongation factors were also found in secretomes of *Xanthomonas campestris* pv. *campestris* [10], *Corynebacterium efficiens* [11], *Bacillus anthracis* and *Bacillus thuringiensis* [12]. Our interpretation is that these proteins are involved in other cellular processes unrevealed until now, and these examples suggest that typical cytosolic proteins might have additional functions outside the cell.

Proteomic analysis of subcellular fractions of *Mycobacterium tuberculosis* revealed that some of the proteins, like Ef-Tu, could also be found in cell wall and membrane compartments [13]. In *Escherichia coli*, a similar result was reported [14] suggesting that these proteins are secreted through the outer membrane vesicles. Our preliminary results suggest that in *R. etli* CE3 similar export mechanisms are operating (data not shown), allowing the export of cytoplasmic, periplasmic and outer membrane proteins to the medium.

Proteins secreted by *R.etli* CE3 during stationary phase

Among the 191 proteins identified in the extracellular proteome in the stationary phase growth (Additional file 2), 55 proteins were without similarity to functionally characterized proteins previously (COG s, “function unknown”), which represents 39.9% of the total of proteins identified in this growth phase. In addition, 21 of the proteins identified in this COG had predicted N-terminal signal peptide, showing that there are many proteins with unknown functions in the extracellular proteome of *R. etli* CE3.

Other overrepresented COGs were “metabolism and carbohydrate transportation” with 15 proteins; “metabolism and amino acid transport” with 14 proteins; “cellular wall, membrane, biogenesis” with 14 proteins; “post-translational modification, protein folding, chaperons” with 13 proteins; and “translation, ribosomal structure and biogenesis” with 12 proteins (Additional file 2).

Among the proteins found in the secretome in this phase of growth, we identified a polysaccharidase protein (PlyA2). This protein could be crucial in the defence of *R. etli*, or in root colonization as was reported in *Rhizobium leguminosarum*, where PlyA showed the capacity to degrade exopolysaccharide (EPS) and carboxymethyl cellulose (CMC) [15]. We also identified the VirE3 protein (Additional file 2). In

Agrobacterium tumefaciens, it has been shown that VirE3 is related to the infection process [16]. Therefore, this protein in *R. etli* could be associated with the process of colonization of the plant.

Studies carried out with MAMP (microbe associated molecular patterns) and their perceptions by plants have been mainly performed with phytopathogenic bacteria [17]. These studies revealed that plants have developed perception systems for different bacterial MAMPs like flagellin, lipopolysaccharide (LPS), elongation factor Tu (EF-Tu), cold shock protein (CSP) or peptidoglycan (PGN) [17-20]. Several plant responses to MAMP molecules, occurring within minutes, include ion-flux across the plasma membrane, increased intracellular Ca²⁺ concentration, oxidative burst, MAP kinase activation and major transcriptional changes [17]. Interestingly, many of these responses have also been detected in epidermal cells of legume roots soon after application of Nod Factors [21]. In this study, we identified FlaCch1, FlaCe, FlaCch2, FlaCch3 and elongation factor Tu proteins in the extracellular proteome of *R. etli* CE3, suggesting that *R. etli* could also use these proteins for the initial recognition between the bacteria and the plant, as was reported previously in other bacterial models [18]. The release of flagellar proteins into the extracellular medium is commonly observed in bacterial cultures, because the flagellum is easily disrupted from the cell surface, however, in *C. jejuni* these proteins are secreted by the flagellar system and mutants in the *flaC* gene showed a significantly reduced level of invasion into HEp-2 cells [22].

The large amount of proteins (55) with unknown function identified under stationary growth condition is remarkable. The analysis and study of those proteins will allow the discovery of new functions in the bacteria. We carried out a search for domains using bioinformatics tools [<http://www.uniprot.org>], and found that several of the proteins have domains that could predict important physiological functions in symbiosis. 1) Proteins with domains of CLE (CLAVATA3/ESR-related), a class of plant peptides known to play roles as short-range signals in controlling the fate of stem cells; 2) peptidases; 3) adhesion proteins; 4) PKD, a domain first identified in the Polycystic kidney disease protein; 5) RTX toxins (RTX representing repeats in the structural toxin) which are important virulence factors produced by a wide range of Gram-negative bacteria; among others (Table 2). In general, these domains belong to enzymes that could be used by the bacteria in defence and adaptation to the environment [23,24]. However, we need additional experiments to determine their possible role in the *R. etli* secretome.

Comparison of the secretome of *R. etli* during exponential and stationary growth phases

The composition of the secretome of *R. etli* CE3 growing in either exponential or stationary phase revealed some similarities. Forty eight proteins were identified as being secreted in both, growth conditions, (Figure 2 and Additional file 4). These proteins could be important components of the *R. etli* CE3 secretome and play essential roles in this bacterium since they are present in both growth phases.

In the exponential phase 144 proteins were found as specific components of the *R. etli* secretome with the most represented COGs being "energy production and conversion", "amino acid transport and metabolism", "translation, ribosomal structure and biogenesis", "post-translational modification, protein turnover, chaperones" (Table 3). For the stationary phase 143 proteins were identified as specific, the most represented COGs being "amino acid transport and metabolism", "carbohydrate transport and metabolism" and "function unknown" (Table 3), which suggests that there are specific bacterial requirements in each growth phase.

The presence of proteins in the groups "production and energy conversion" and "metabolism and transportation of inorganic ions" could be related to the survival of the bacteria to adverse conditions and could contribute to diminish the levels of toxic substances, for instance superoxide dismutase. On the other hand, some proteins like malate dehydrogenase involved in the Krebs cycle, appear in most secretomes previously reported [1,25], suggesting that it could have other physiological role in the bacterium as part of the bacterial secretome.

The presence of superoxide dismutase (SodB) under both studied conditions, and catalase-peroxidase (KatG) in stationary phase, in the *R. etli* CE3 secretome, (Additional file 4), suggest that these proteins could participate in the neutralization of the oxidative burst, by which plants challenge infecting bacteria. SodB is responsible for the inactivation of O₂⁻ and H₂O₂ and is a virulence factor for some phytopathogenic bacteria [26]. In this way, the presence of this protein in the *R. etli* secretome could be crucial for the establishment of the symbiosis, in addition to its remarkable role in the maintenance of the symbiosis as an intracellular protein [26].

The high-molecular-weight heat shock proteins Hsp60 (GroEL) and Hsp70 (DnaK) were also identified, the Hsp60 in both growth phases and Hsp70 only in stationary phase. These proteins may function as chaperons in the type II and type IV secretion systems helping to fold other proteins accumulated in the acidic region, as reported previously in the *E. coli* EHEC EDL933 and EPEC E2348/69 extracellular proteomes [27]. For

Table 2 Functional domains present in some hypothetical proteins secreted by *R. etli* in stationary growth phase

ID from hypothetical proteins	Domain	Domain description found in Swissprot data base
gij21492958	IPB000209	Proteolytic enzymes that exploit serine in their catalytic activity are ubiquitous, being found in viruses, bacteria and eukaryotes, (serine-carboxyl peptidases).
gij86355844	IPR010662	The structure shows an alpha-beta hydrolase fold suggesting an enzymatic function for these proteins. The crystal structure from <i>B. subtilis</i> has been solved.
gij86356671	PDOC00169	Cytochrome c-type centres are also found in the active sites of many enzymes, including cytochrome cd1-nitrite reductase as the N-terminal haem c domain, in quinoprotein alcohol dehydrogenase as the C-terminal domain, in Quinohemoprotein amine dehydrogenase A chain as domains 1 and 2, and in the cytochrome bc1 complex as the cytochrome bc1 domain.
gij86357010	IPR010221 VCBS	<p>IPR006644 Dystroglycan-type cadherin-like In animals, cadherin domain-containing proteins are adhesion molecules that modulate a wide variety of processes including cell polarization and migration but they have also been identified in yeast and magnetotactic bacteria. Crystal structures have revealed that multiple cadherin domains form Ca²⁺-dependent rod-like structures with a conserved Ca²⁺-binding pocket at the domain-domain interface.</p> <p>IPR008009 Putative Ig This domain probably corresponds to a new superfamily in the immunoglobulin fold. The function of this domain is uncertain it may be involved in cell adhesion. In the Sushi repeat-containing protein (SrpX), this domain is found between two sushi repeats. PKD domains are also found in other proteins, usually in the extracellular parts of proteins involved in interactions with other proteins. For example, domains with a PKD-type fold are found in archaeal surface layer proteins that protect the cell from extreme environments.</p> <p>IPR005066 Molybdenum cofactor oxidoreductase, dimerisation The majority of molybdenum-containing enzymes utilise a molybdenum cofactor (MoCF or Moco) consisting of a Mo atom coordinated via a cisdithiolene moiety to molybdopterin (MPT). MoCF is ubiquitous in nature, and the pathway for MoCF biosynthesis is conserved in all three domains of life. MoCF-containing enzymes function as oxidoreductases in carbon, nitrogen, and sulphur metabolism.</p> <p>This domain of about 100 residues is found multiple (up to 35) copies in long proteins from several species of <i>Vibrio</i>, <i>Colwellia</i>, <i>Bradyrhizobium</i>, and <i>Shewanella</i> (hence the name VCBS) and in smaller copy numbers in proteins from several other bacteria. The large protein size and repeat copy numbers, species distribution, and suggested activities oouter membrane adhesin like protein.</p>
gij86357471	EC:2	<p>IPR013785 Aldolase-type TIM barrel This entry represents the TIM beta/alpha barrel found in aldolase and related proteins. This TIM barrel usually covers the entire protein structure. Proteins containing this TIM barrel domain include class I aldolases, class I DAHP synthases, class II fructose-bisphosphate aldolases (FBP aldolases), and 5-aminolaevulinic acid dehydratase (a hybrid of classes I and II aldolases)</p> <p>IPR006638 Elongator protein 3/MiaB/NifB This domain is found in MoaA, NifB, PqqE, coproporphyrinogen III oxidase, biotin synthase and MiaB families, and includes a representative in the eukaryotic elongator subunit, Elp-3. Some members of the family are methyltransferases</p> <p>IPR007197 Radical SAM Radical SAM proteins catalyze diverse reactions, including unusual methylations, isomerization, sulphur insertion, ring formation, anaerobic oxidation and protein radical formation. Evidence exists that these proteins generate a radical species by reductive cleavage of S-adenosylmethionine (SAM) through an unusual Fe-S centre.</p>
gij86357955	EC:4.2.3.5	Chorismate synthase catalyzes the last of the seven steps in the shikimate pathway which is used in prokaryotes, fungi and plants for the biosynthesis of aromatic amino acids. It catalyzes the 1,4-trans elimination of the phosphate group from 5-enolpyruvylshikimate-3-phosphate (EPSP) to form chorismate which can then be used in phenylalanine, tyrosine or tryptophan biosynthesis. Chorismate synthase requires the presence of a reduced flavin mononucleotide (FMNH ₂ or FADH ₂) for its activity. Chorismate synthase from various sources shows a high degree of sequence conservation. It is a protein of about 360 to 400 amino-acid residues).
gij86358739	GI:241205844	Integral membrane protein which is thought to regulate cation conductance. A variety of proteins belong to this family. These include the prohibitins, cytoplasmic anti-proliferative proteins and stomatin, an erythrocyte membrane protein. Bacterial HflC protein also belongs to this family.
gij86360288	PS00495	The apple domain has an N-terminal region that contains four tandem repeats of about 90 amino acids and a C-terminal catalytic domain. The 90 amino-acid repeated domain contains 6 conserved cysteines. It has been shown that three disulphide bonds link the first and sixth, second and fifth, and third and fourth cysteines. This entry contains apple-like domains, which are presented in Plasminogen, <i>Caenorhabditis elegans</i> hypothetical ORFs and the extracellular portion of plant S-locus glycoproteins and S-receptor kinases. The domain is predicted to possess protein-and/or carbohydrate-binding functions.
gij86360818	PF 06724	This region consists of two a pair of transmembrane helices and occurs three times in each of the family member proteins.

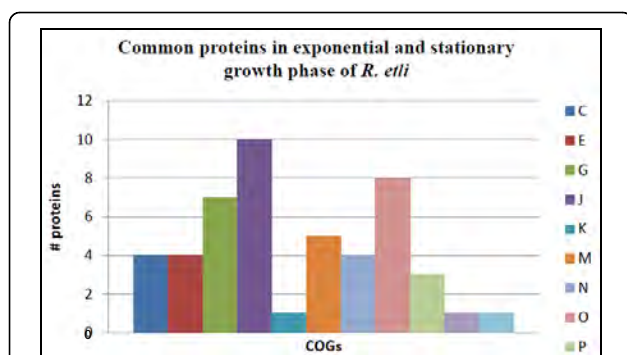


Figure 2 Proteins identified in both growth phases. Proteins were grouped into functional groups (COG). Energy production and conversion (C); Amino acid transport and metabolism (E); Carbohydrate transport and metabolism (G); Translation, ribosomal structure and biogenesis (J); Transcription (K); Cell wall, membrane, envelope biogenesis (M); Cell motility (N); Posttranslational modification, protein turnover, chaperones (O); Inorganic ion transport and metabolism (P); General function prediction only (R); Intracellular trafficking, secretion, and vesicular transport (U).

Legionella pneumophila and *Helicobacter pylori*, Hsp60 and Hsp70 were reported to be involved in colonization, attachment, and invasion [28]. In addition, previous studies have demonstrated the role of chaperones in secretion systems through chaperone-usher pathway [29]. Therefore, the presence of these proteins in the *R. etli* extracellular proteome could suggest that they take part in the colonization of the plant.

In stationary phase NodTch protein was identified, this protein was found to be essential for free-living *R. etli*. The authors proposed that NodTch may play a role in outer membrane stability or in chromosome segregation instead of, or in addition to, drug efflux [30].

We also identified flagellin-like proteins in both growth phases (FlaCch1, FlaCch2, FlaCch3 and FlaCe), and FlgD (flagellar basal-body rod protein) specifically in exponential phase, which are typical components of different bacterial secretomes [31-33]. These proteins are exported via a translocation complex involved in the assembly of flagella, which is related to the type III

Table 3 Comparison of specific proteins for each growth phase

COGs	prot 6 h (%)	prot 6 h (SP) (%)	prot24 h (%)	prot 24 h (SP) (%)
Energy production and conversion (C)	10.3	1.3	1.5	-
Cell cycle control, cell division, chromosome partitioning (D)	0.7	-	0.7	-
Amino acid transport and metabolism (E)	6.2	1.3	7.2	5.1
Nucleotide transport and metabolism (F)	1.3	-	-	-
Carbohydrate transport and metabolism (G)	5.5	-	5.8	5.8
Coenzyme transport and metabolism (H)	2.1	-	0.7	-
Lipid transport and metabolism (I)	1.3	-	0.7	-
Translation, ribosomal structure and biogenesis (J)	22.8	-	1.5	-
Transcription (K)	3.4	-	3.6	-
Replication, recombination and repair (L)	4.8	-	7.2	0.7
Cell wall, membrane, envelope biogenesis (M)	3.4	-	6.5	4.3
Cell motility (N)	2.1	1.3	-	-
Posttranslational modification, protein turnover, chaperones (O)	6.2	0.7	3.6	1.5
Inorganic ion transport and metabolism (P)	2.8	0.7	2.9	2.9
Secondary metabolites biosynthesis, transport and catabolism (Q)	3.4	0.7	0.7	-
General function prediction only (R)	2.1	-	4.3	2.1
Function unknown (S)	4.8	2.1	39.9	15.2
Signal transduction mechanisms (T)	4.1	1.3	1.5	-
Intracellular trafficking, secretion, and vesicular transport (U)	4.8	-	1.5	1.5
Defense mechanisms (V)	2.1	-	1.5	-
Without COG (WC)	5.5	2.1	8.7	4.3
Total protein	144	17	143	60

Namely exponential phase (6 h) and the stationary phase (24 h). The relative contribution to each COG of proteins with a potential signal peptide (SP) is indicated. Functional groups with a major abundance in proteins are shown in bold letters and numbers.

secretion system of Gram-negative bacteria [34]. However, in some cases, the transportation of flagellar proteins should not be considered as secretion in a general sense, because most of proteins to be transported are not intentionally exported into the media, but result from some leaks or spill-over during construction [35].

Two proteins identified as cold shock proteins (CspA4 and CspA5) were found specifically in the secretome of exponential phase. In tobacco cell cultures, it was suggested that cold shock proteins are a novel class of bacterial elicitors for which tobacco and other *Solanaceae* have evolved specific and sensitive chemoperception systems and can potentially serve as MAMPs [36]. Because lysis of bacteria under natural low titre conditions would probably not generate sufficient amounts of Csp to be detected by the plant, an alternative mechanism of protein release other than active secretion can be taken into consideration.

Bioinformatics analysis of the secreted proteins

The Signal P 3.0 Server program (<http://www.cbs.dtu.dk/services/SignalP-2.0/>), was used to scan the identified proteins for *N*-terminal signal peptides, indicative for type II and tat-mediated protein transport secretion system [37]. The first 60 amino acids from the *N*-termini were examined by the Signal P software, and only search results with a probability score higher than 0.98 were considered. This analysis revealed that 24.21% of the extracellular proteins identified in this study, (12.95% in the exponential phase and 35.60% in stationary growth phase) had putative secretion signal peptides and were probably secreted via a sec-dependent pathway (Table 1). Which is apparently more active in *R. etli* in the late stages of bacterial growth, in *X. campestris* was previously reported that the 53% of extracellular proteome had signal peptides in early stationary growth phase [1].

Since the software Signal P does not differentiate signal peptides for secretion system type II and Tat-secretion, it was necessary to do a manual search of all signal peptide sequences in order to find a twin arginine motif that is typical of Tat-secretion signal [37]. The twin-arginine motif could only be found in five of the putative secretion signals: Omp1, Omp2, MpA (outer membrane lipoprotein), PstS (phosphate ABC transporter, substrate-binding protein) and FabB (3-oxoacyl-(acyl-carrier-protein) synthase I protein) (data not show). These proteins are the only candidates to traverse the inner membrane in their folded conformation.

Extracellular proteins of *R. etli* with a putative signal peptides, identified in culture supernatant in exponen-

tial phase growth, include proteins for transport of carbohydrates (ChvE, AglE, and FrcB) and transport of amino acids (BraC1, BraC2, and DppAch1) (Additional file 1). In general, 30 proteins with signal peptide, identified in both growth phases, could be involved in transport processes. For stationary phase growth, the best represented group of proteins with signal peptides were proteins of unknown function (21 proteins). For some of these, we detected by bioinformatics analysis specific domains for proteases and peptidases (Table 1 and Additional file 2). These classes of enzymes are known to be secreted by pathogenic bacteria [38].

Conclusion

The results presented here render a more detailed picture of the *R. etli* CE3 extracellular proteome. A total of 384 proteins were identified by LC-MS/MS in two experimental replicates of exponential and stationary phases, and 607 proteins in total, which reveals an abundance of secreted proteins by *R. etli* in both phases of growth and represents the first report about the extracellular proteome in this bacterium. This study also provides a basis for comparative studies of the secretome in different growth conditions for *R. etli*. Our results suggest that the secretome of *R. etli* consists mainly of two groups of proteins. The actively secreted proteins are extracellular enzymes (mostly degradation enzymes) and proteins that bind nutrients and extracellular appendages. The secretion and synthesis of these proteins is apparently regulated by cellular phase growth. The second group includes proteins that have functions in the cytosol and are not actively secreted but may be released into the culture medium due to lysis of bacteria or by other, yet unknown mechanisms.

In addition, in the exponential phase of growth we found a large number of ribosomal proteins. The role that these proteins play in the extracellular proteome remains unknown. Our results also showed many proteins without assigned function mainly, but not exclusively, in stationary phase. Further studies are needed to know the function that these proteins play in the extracellular proteome, and the study of these proteins will clarify aspects of *R. etli* and its interaction with the environment, since the extracellular proteome must be considered an essential component of the bacterial cell, which plays an important role in at least three fundamental processes; protection against stresses, cell survival and colonization of the host.

Additional material

Additional file 1: Proteins secreted in exponential phase of growth in *R. etli*. description of all identified proteins in exponential growth phase in *R. etli*. In this table, we show the description of the proteins identified the score, the number of peptides, the molecular weight and isoelectric point.

Additional file 2: Proteins secreted in stationary phase of growth in *R. etli*. description of all identified proteins in stationary growth phase in *R. etli*. In this table, we show the description of the proteins identified the score, the number of peptides, the molecular weight and isoelectric point.

Additional file 3: Comparison of two-dimensional gels in exponential (6 h) and stationary (24 h) growth phase of *R. etli*. Comparison of 2D gels of secreted proteins and cytoplasm proteins in exponential and stationary growth phase of *R. etli*. In order to assess if proteins originated from cell lysis, and hence contaminated the secretome, we compared the protein profiles of whole cell lysates with those of extracellular proteins.

Additional file 4: Common proteins in both phases of growth of *R. etli*. Descriptions of all the common identified proteins are secreted in exponential and stationary growth phase in *R. etli*. In this table, we show the description of the proteins identified the score, the number of peptides, the molecular weight and isoelectric point.

Acknowledgements

We thank Michael F. Dunn, for useful scientific comments and to Alejandro Aguilar for his support in the Bioinformatics Analysis. Part of this work was supported by CONACyT grant 60641 and DGAPA-PAPIIT grant IN-216260. Niurka Meneses is a recipient of a PhD Studentship from the Universidad Nacional Autónoma de México.

Author details

¹Programa de Genómica Funcional de Procariones, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Apdo. Postal 565-A, Cuernavaca, Morelos, CP 62210, México. ²Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México, México, D.F., México.

Authors' contributions

NM carried of the experiments and prepared a first draft of the manuscript, GMH supplied methodological expertise and carried of some experiments, SE conceived the study, participated in its design and coordination, and wrote the final manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 28 April 2010 Accepted: 14 October 2010

Published: 14 October 2010

References

1. Watt SA, Patschkowski T, Niehaus K: **Comprehensive analysis of the extracellular proteins from *Xanthomonas campestris* pv. *campestris* B100.** *Proteomics* 2005, **5**:153-167.
2. Kambara K, Ardisson S, Kobayashi H, Saad MM, Schumpp O, Broughton WJ, Deakin WJ: **Rhizobia utilize pathogen-like effector proteins during symbiosis.** *Molecular Microbiology* 2009, **1**:192-106.
3. Gohar M, Gilois N, Graveline R, Garreau C, Sanchis V, Lereclus D: **A comparative study of *Bacillus cereus*, *Bacillus thuringiensis* and *Bacillus anthracis* extracellular proteomes.** *Proteomics* 2005, **5**:3696-3711.
4. Encarnación S, Dunn M, Willms K, Mora J: **Fermentative and aerobic metabolism in *Rhizobium etli*.** *J Bacteriol* 1995, **177**:3058-3066.
5. Encarnación S, Guzmán Y, Dunn MF, Hernández M, del Carmen Vargas M, Mora J: **Proteome analysis of aerobic and fermentative metabolism in *Rhizobium etli* CE3.** *Proteomics* 2003, **3**:1077-1085.
6. Baker S, Campbell JJ, Stabler R, Nguyen HVM, To DS, Nguyen DV, Farrar J: **Drug resistance in *Chromobacterium violaceum*.** *Genet Mol Res* 2004, **3**:134-147.
7. Hyde SC, Emsley P, Hartshorn MJ, Mimmack MM, Gileadi U, Pearce SR, Gallagher MP, Gill DR, Hubbard RE, Higgins CF: **Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport.** *Nature* 1990, **346**(6282):362-365.
8. Barriuso-Iglesias M, Schluesener D, Barreiro C, Poetsch A, Martín JF: **Response of the cytoplasmic and membrane proteome of *Corynebacterium glutamicum* ATCC 13032 to pH changes.** *BMC Microbiology* 2008, **8**:225.
9. Kazemi-Pour N, Condemine G, Hugouvieux-Cotte-Pattat N: **The secretome of the plant pathogenic bacterium *Erwinia chrysanthemi*.** *Proteomics* 2004, **4**:3177-3186.
10. Watt SA, Wilke A, Patschkowski T, Niehaus K: **Comprehensive analysis of the extracellular proteins from *Xanthomonas campestris* pv. *campestris* B100.** *Proteomics* 2005, **5**:153-167.
11. Hansmeier N, Chao TC, Kalinowski J, Pühler A, Tauch A: **Mapping and comprehensive analysis of the extracellular and cell surface proteome of the human pathogen *Corynebacterium diphtheriae*.** *Proteomics* 2006, **6**:2465-2476.
12. Gohar M, Gilois N, Graveline R, Garreau C, Sanchis V, Lereclus D: **A comparative study of *Bacillus cereus*, *Bacillus thuringiensis* and *Bacillus anthracis* extracellular proteomes.** *Proteomics* 2005, **5**:3696-3711.
13. Mawuenyega KG, Forst CV, Dobos KM, Belisle JT, Chen J, Bradbury EM, Bradbury AR, Chen X: ***Mycobacterium tuberculosis* functional network analysis by global subcellular protein profiling.** *Mol Biol Cell* 2005, **16**:396-404.
14. McBroom AJ, Kuehn MJ: **Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response.** *Molecular Microbiology* 2007, **63**:545-558.
15. Zorreguieta A, Finnie C, Downie JA: **Extracellular glycanases of *Rhizobium leguminosarum* are activated on the cell surface by an exopolysaccharide-related component.** *Journal of Bacteriology* 2000, **182**:1304-1312.
16. Benoit L, Manjusha V, Tzvi T, Vitaly C: **The VirE3 protein of *Agrobacterium mimics* a host cell function required for plant genetic transformation.** *The EMBO Journal* 2005, **24**:428-437.
17. Aslam SN, Erbs G, Morrissey KL, Newman M-A, Chinchilla D, Boller T, Molinaro A, Jackson W, Cooper RM: **Microbe-associated molecular pattern (MAMP) signatures, synergy, size and charge: influences on perception or mobility and host defence responses.** *Molecular Plant Pathology* 2009, **10**:375-387.
18. Gómez-Gómez L, Boller T: **FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*.** *Mol Cell* 2000, **5**:1003-1011.
19. Silipo A, Molinaro A, Sturiale L, Dow JM, Erbs G, Lanzetta R, Newman MA, Parrilli M: **The elicitation of plant innate immunity by lipooligosaccharide of *Xanthomonas campestris*.** *J Biol Chem* 2005, **280**(33):660-33, 668.
20. Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JDG, Boller T, Felix G: **Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation.** *Cell* 2006, **125**:749-760.
21. Soto MJ, Domínguez-Ferreras A, Pérez-Mendoza D, Sanjuán J, Olivares J: **Mutualism versus pathogenesis: the give-and-take in plant-bacteria interactions.** *Cellular Microbiology* 2009, **11**:381-388.
22. Song YC, Jin S, Louie H, Ng D, Lau R, Zhang Y, Weerasekera R, Al Rashid S, Ward LA, Der SD, et al: **FlaC, a protein of *Campylobacter jejuni* TGH9011 (ATCC43431) secreted through the flagellar apparatus, binds epithelial cells and influences cell invasion.** *Mol Microbiol* 2004, **53**:541-553.
23. Wlodawer A, Li M, Gustchina A, Oyama H, Dunn BM, Oda K: **Structural and enzymatic properties of the sedolisin family of serine-carboxyl peptidases.** *Acta Biochim Pol* 2003, **50**:81-102.
24. Okubo A, Li M, Ashida M, Oyama H, Gustchina A, Oda K, Dunn BM, Wlodawer A, Nakayama T: **Processing, catalytic activity and crystal structures of kumamolisin-As with an engineered active site.** *FEBS* 2006, **273**:2563-2576.

25. Fauvart M, Michiels J: **Rhizobial secreted proteins as determinants of host specificity in the rhizobium-legume symbiosis.** *FEMS Microbiology Letters* 2008, **1-9**.
26. Soto MJ, Sanjuán J, Olivares J: **Rhizobia and plant-pathogenic bacteria: common infection weapons.** *Microbiology* 2006, **152**:3167-3174.
27. Sandkvist M: **Type II Secretion and Pathogenesis.** *Infection and Immunity* 2001, **69**:3523-3535.
28. Bumann D, Aksu S, Wendland M, Janek K, Zimny-Arndt U, Sabarth N, Meyer TF, Jungblut PR: **Proteome analysis of secreted proteins of the gastric pathogen *Helicobacter pylori*.** *Infection and Immunity* 2002, **70**:3396-3403.
29. Waksman G, Hultgren SJ: **Structural biology of the chaperone-usher pathway of pilus biogenesis.** *Nature Reviews Microbiol* 2009, **7**:765-774.
30. Hernández-Mendoza A, Nava N, Santana O, Abreu-Goodger C, Tovar A, Quinto C: **Diminished redundancy of outer membrane factor proteins in rhizobiales: A *nodT* homolog is essential for free-living *Rhizobium etli*.** *Journal Molecular Microbiology Biotechnology* 2007, **13**:22-34.
31. Komoriya K, Shibano N, Higano T, Azuma N, Yamaguchi S, Aizawa S: **Flagellar proteins and type III-exported virulence factors are the predominant proteins secreted into the culture media of *Salmonella typhimurium*.** *Molecular Microbiology* 1999, **34**:767-779.
32. Süß C, Hempel J, Zehner S, Krause A, Patschkowski T, Göttfert M: **Identification of genistein-inducible and Type III-secreted proteins of *Bradyrhizobium japonicum*.** *Journal Biotechnology* 2006, **126**:69-77.
33. Nielsen H, Engelbrecht J, Brunak S, von Heijne G: **Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.** *Protein Engineering* 1997, **10**:1-6.
34. Blocker A, Komoriya K, Aizawa S: **Type III secretion systems and bacterial flagella: Insights into their function from structural similarities.** *Proc Natl Acad Sci USA* 2003, **100**:3027-3030.
35. Ikeda T, Oosawa K, Hotani H: **Self-assembly of the filament capping protein, FliD, of bacterial flagella into an annular structure.** *J Mol Biol* 1996, **259**:679-686.
36. Felix G, Boller T: **Molecular sensing of bacteria in plants. The highly conserved RNA-binding motif RNP-1 of bacterial cold shock proteins is recognized as an elicitor signal in tobacco.** *J Biol Chem* 2003, **278**(8):6201-6208.
37. Sandkvist M: **Type II Secretion and Pathogenesis.** *Infection and Immunity* 2001, **69**:3523-3535.
38. Antelmann H, Tjalsma H, Voigt B, Ohlmeier S, Bron S, van Dijk JM, Hecker M: **A proteomic view on genome-based signal peptide predictions.** *Genome Biology* 2001, **11**:1484-1502.

doi:10.1186/1477-5956-8-51

Cite this article as: Meneses *et al.*: The extracellular proteome of *Rhizobium etli* CE3 in exponential and stationary growth phase. *Proteome Science* 2010 **8**:51.

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

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

Submit your manuscript at
www.biomedcentral.com/submit



Proteínas identificadas en el secretoma de *R. etli*

AC	nombre	COGs	PU	PM	PI	PMSS	condiciones	descripción
B3PPT9	icd	C	9	45078.50	6.04	417.85	todas	Isocitrate dehydrogenase (NADP(+)) protein OS
Q2K3E9	mdh	C	5	33586.02	5.74	374.30	todas	Malate dehydrogenase OS
Q2K3F0	sucC	C	9	41856.26	5.08	421.90	todas	Succinyl-CoA ligase [ADP-forming] subunit beta OS
Q2K3H0	atpD	C	12	50537.33	4.91	839.57	todas	ATP synthase subunit beta OS
Q2K8W5	pdhB	C	3	45965.80	6.22	322.50	todas	Dihydrolipoamide acetyltransferase protein OS
Q2K8X4	gltA	C	7	48010.05	6.02	133.43	todas	Citrate synthase OS
Q2KE66	pckA	C	12	57562.22	5.71	803.73	todas	Phosphoenolpyruvate carboxykinase [ATP] OS
B3PR68	RHECIAT_CH0004294	E	27	27616.41	6.01	1778.36	todas	Probable amino acid ABC transporter, substrate-binding protein OS
B3PUX7	aatAch	E	7	43586.01	6.35	148.39	todas	Aspartate aminotransferase protein OS
B3Q1K9	RHECIAT_PA0000221	E	10	42611.17	5.74	286.82	todas	Probable branched-chain amino acid ABC transporter, substrate-binding protein OS
Q2JYN6	RHE_PF00410	E	10	35873.82	4.92	329.54	todas	Probable amino acid ABC transporter, substrate-binding protein OS
Q2K0L8	RHE_00228	E	12	29360.63	7.72	177.02	todas	Probable amino acid ABC transporter, substrate-binding protein OS
Q2K1K6	RHE_PC00160	E	21	28118.82	4.84	376.40	todas	Probable amino acid ABC transporter, substrate-binding protein OS
Q2K2U3	asd	E	4	37650.90	5.82	204.14	todas	Aspartate-B-semialdehyde dehydrogenase protein OS
Q2K336	RHE_CH04006	E	31	27647.42	5.76	2017.51	todas	Probable amino acid ABC transporter, substrate-binding protein OS
Q2K378	RHE_CH03963	E	16	28932.82	4.79	531.90	todas	Putative amino acid ABC transporter, substrate-binding protein OS
Q2K3C4	dcp	E	14	77224.36	5.32	450.59	todas	Peptidyl-dipeptidase Dcp protein OS
Q2K4N4	RHE_CH03445	E	12	41240.17	6.36	152.60	todas	Putative branched-chain amino acid ABC transporter, substrate-binding protein OS
Q2K4R4	potD	E	10	36900.25	5.11	238.17	todas	Spermidine/putrescine ABC transporter, substrate-binding protein OS
Q2K502	braC2	E	23	38902.07	5.03	2632.32	todas	Amino acid ABC transporter, substrate-binding protein OS
Q2K5M4	oppA	E	11	59021.86	5.71	380.65	todas	Oligopeptide ABC transporter, substrate-binding protein OS
Q2K5M7	braC1	E	13	37573.66	5.01	941.97	todas	Amino acid ABC transporter, substrate-binding protein OS
Q2K5N5	proX	E	9	34004.66	4.99	251.72	todas	Proline/glycine betaine ABC transporter, substrate-binding protein OS
Q2K5X1	aatAch	E	8	43527.85	5.96	239.30	todas	Aspartate aminotransferase protein OS
Q2K7J0	RHE_CH02418	E	19	27702.60	5.05	2456.86	todas	Probable amino acid ABC transporter, substrate-binding protein OS
Q2KA25	glyA	E	11	46598.02	6.44	525.54	todas	Serine hydroxymethyltransferase OS
Q2KCA6	potF	E	13	40481.00	5.06	653.54	todas	Spermidine/putrescine ABC transporter, substrate-binding protein OS
Q2KDI3	RHE_CH00280	E	7	44937.62	5.51	247.52	todas	Probable aminopeptidase protein OS
B3PT89	ilvC	EH	6	36748.07	5.88	203.99	todas	Ketol-acid reductoisomerase OS
Q2K902	aapJ	ET	23	37492.43	5.67	2797.14	todas	General L-amino acid ABC transporter, substrate-binding protein OS
Q2K4X9	purA	F	15	46493.84	5.47	192.91	todas	Adenylosuccinate synthetase OS
Q2KA61	ndk	F	5	15265.38	6.31	496.71	todas	Nucleoside diphosphate kinase OS
B3PYR0	eno	G	8	44781.50	4.78	254.56	todas	Enolase OS
B3Q501	RHECIAT_PC0000201	G	8	45916.04	4.99	141.39	todas	Putative sugar ABC transporter, substrate-binding protein OS
Q2JZJ7	RHE_PF00096	G	17	47282.89	6.40	782.13	todas	Probable sugar ABC transporter, substrate-binding protein OS
Q2JZK2	RHE_PF00091	G	12	47501.86	7.05	318.62	todas	Probable sugar ABC transporter, substrate-binding protein OS
Q2K0S9	RHE_00289	G	10	34503.62	6.65	292.74	todas	Probable sugar ABC transporter, substrate-binding protein OS
Q2K1C6	ugpBc	G	16	48382.60	6.18	440.84	todas	Sn-glycerol-3-phosphate ABC transporter, substrate-binding protein OS

Q2K414	tal	G	4	34863.09	5.62	182.36	todas	Transaldolase OS
Q2K433	xylA	G	7	48973.19	5.56	212.95	todas	Xylose isomerase OS
Q2K462	RHE_CH03619	G	7	40213.78	6.97	369.61	todas	Probable TRAP mannitol/C4-dicarboxylate transporter, substrate-binding protein OS
Q2K4D8	pykA	G	5	51516.36	6.49	109.05	todas	Pyruvate kinase OS
Q2K4I0	fbaB	G	20	36916.09	6.09	1310.14	todas	Fructose-bisphosphate aldolase OS
Q2K4I4	gap	G	12	35867.08	7.21	880.29	todas	Glyceraldehyde 3-phosphate dehydrogenase protein OS
Q2K5F5	chvE	G	22	37700.84	4.83	4932.96	todas	Xylose ABC transporter, substrate-binding protein OS
Q2K670	RHE_CH02898	G	12	45934.00	5.26	379.76	todas	Probable sugar ABC transporter, substrate-binding protein OS
Q2KCC4	aglE	G	20	49112.43	4.96	1368.71	todas	Alpha-glucoside ABC transporter, substrate-binding protein OS
Q2KCY8	pgi	G	15	58618.05	6.36	377.14	todas	Glucose-6-phosphate isomerase OS
Q2KD00	frcB	G	11	35204.28	5.33	875.29	todas	Fructose ABC transporter, substrate-binding protein OS
Q2K678	RHE_CH02890	H	14	38845.97	5.04	634.02	todas	Probable ABC transporter, substrate-binding protein OS
Q2K324	phbAch	I	10	40479.25	6.41	541.31	todas	Acetyl-CoA acetyltransferase (Beta-ketothiolase) protein OS
Q2K4J1	RHE_CH03489	I	3	17666.17	5.15	65.62	todas	Hypothetical conserved protein OS
Q2KE04	fabA	I	3	19014.75	5.55	117.95	todas	3-hydroxydecanoyl-(Acyl-carrier-protein) dehydratase protein OS
B3PV07	rplY	J	7	21985.09	5.55	83.56	todas	50S ribosomal protein L25 OS
B3PWT1	rplN	J	2	13418.94	10.63	216.94	todas	50S ribosomal protein L14 OS
Q2K379	rplS	J	3	19245.80	9.38	234.93	todas	50S ribosomal protein L19 OS
Q2K5U4	rplY	J	9	22071.13	5.66	131.55	todas	50S ribosomal protein L25 OS
Q2K8Y4	frr	J	9	20735.72	7.93	308.44	todas	Ribosome-recycling factor OS
Q2K8Y6	tsf	J	9	32200.79	5.14	559.96	todas	Elongation factor Ts OS
Q2K9K1	rplF	J	3	19323.41	9.83	139.01	todas	50S ribosomal protein L6 OS
Q2K9L9	fusA	J	9	77735.65	5.16	1076.41	todas	Elongation factor G OS
Q2K9M0	rpsG	J	2	17683.19	10.26	270.24	todas	30S ribosomal protein S7 OS
Q2K9M6	rplJ	J	5	18128.99	9.73	236.82	todas	50S ribosomal protein L10 OS
Q2KA05	RHE_CH01528	J	2	17575.87	9.61	184.82	todas	Hypothetical conserved protein OS
Q2KE13	rpsA	J	9	62356.69	5.11	455.00	todas	30S ribosomal protein S1 OS
Q2K9N2	tufI	JE	10	42637.75	5.33	1757.14	todas	Elongation factor Tu 1 OS
Q2K9J2	rpoA	K	3	37239.38	4.78	72.21	todas	DNA-directed RNA polymerase subunit alpha OS
Q2K9M9	nusG	K	9	20036.76	6.35	178.15	todas	Transcription antitermination protein nusG OS
B3PVY7	hrm	L	6	9367.57	9.70	395.28	todas	Bacterial nucleoid DNA-binding protein OS
Q2K8E4	ssb	L	3	18398.00	5.88	111.62	todas	Single-stranded DNA-binding protein OS
Q2KDE6	dnaN	L	9	40435.99	4.97	124.11	todas	DNA polymerase III, beta chain protein OS
B3PQ93	omp2	M	6	17466.68	9.45	965.52	todas	Outer membrane lipoprotein OS
B3PS05	ctpA	M	7	46974.54	5.22	101.79	todas	C-terminal processing peptidase protein OS
Q2JY83	RHE_PF00565	M	3	47610.89	5.09	70.30	todas	Putative sugar ABC transporter, substrate-binding protein OS
Q2K370	RHE_CH03971	M	9	24080.43	8.86	478.17	todas	Putative outer membrane lipoproteins carrier protein OS
Q2K4B9	exoN	M	7	32474.43	5.57	281.63	todas	UTP-glucose-1-phosphate uridylyltransferase protein OS
Q2K4L9	omp2	M	7	19057.54	9.71	1508.63	todas	Outer membrane lipoprotein OS
Q2K584	rkpK	M	10	47469.71	5.30	738.20	todas	UDP-glucose 6-dehydrogenase protein OS
Q2K699	RHE_CH02869	M	11	43342.22	5.48	657.77	todas	Probable membrane-bound lytic murein transglycosylase protein OS

Q2K6F8	ompA	M	6	22717.51	8.79	1021.59	todas	Outer membrane lipoprotein OS
Q2K7J1	RHE_CH02417	M	27	80501.39	5.19	1066.50	todas	Putative outer membrane protein, OmpA family OS
Q2KA52	ropB1	M	8	22335.50	4.63	8103.16	todas	Outer membrane protein OS
Q2KCY0	RHE_CH00485	M	8	31871.31	5.29	369.54	todas	Probable glycine betaine ABC transporter, substrate binding protein OS
B3PPM9	flaCch1	N	9	31096.33	4.59	2879.73	todas	Flagellin C protein OS
B3PPN0	flaCch2	N	13	31096.59	4.61	4395.78	todas	Flagellin C protein OS
Q2K133	flaCe	N	5	31288.40	4.13	988.84	todas	Flagellin C protein OS
Q2K8G8	RHE_CH02084	N	19	38814.96	6.10	1533.78	todas	Putative ABC transporter, substrate-binding protein OS
Q2KCE9	flaCch3	N	9	31554.19	4.49	2077.46	todas	Flagellin C protein OS
Q2KCF0	flaCch2	N	16	31100.58	4.56	5121.23	todas	Flagellin C protein OS
Q2KCF1	flaCch1	N	13	31079.27	4.70	8894.78	todas	Flagellin C protein OS
B3PU03	groELch2	O	17	57745.14	5.22	1510.02	todas	60 kDa chaperonin OS
B3Q351	groELc	O	10	57505.83	5.13	1348.63	todas	60 kDa chaperonin OS
Q2K237	msrAb	O	6	19120.40	5.72	147.40	todas	Peptide methionine sulfoxide reductase protein OS
Q2K3R8	ppiD2	O	8	32615.87	5.52	272.29	todas	Peptidyl prolyl cis-trans isomerase D signal peptide protein OS
Q2K3Z1	osmC	O	3	14663.69	5.75	174.84	todas	Osmotically inducible protein OS
Q2K8D7	ppiB	O	9	18533.96	5.93	646.64	todas	Peptidyl-prolyl cis-trans isomerase OS
Q2K8D8	ppiA	O	7	20485.41	7.88	1140.14	todas	Peptidyl-prolyl cis-trans isomerase OS
Q2K952	tig	O	6	54445.34	4.84	452.66	todas	Trigger factor OS
Q2K9A0	RHE_CH01791	O	9	23779.06	6.07	159.49	todas	Probable anti-oxidant protein, AhpCTSA family OS
Q2KBH6	RHE_CH00999	O	10	27828.77	5.27	243.63	todas	Putative thiol-disulfide oxidoreductase protein OS
Q2KBK7	RHE_CH00968	O	6	16869.36	5.11	138.64	todas	Probable peroxiredoxin protein OS
Q2KBZ6	groESch1	O	6	10561.13	5.85	253.16	todas	10 kDa chaperonin OS
Q2KBZ7	groL1	O	29	57718.41	5.07	3688.11	todas	60 kDa chaperonin 1 OS
Q2KDW6	dnaK	O	11	68350.22	4.92	447.16	todas	Chaperone protein dnaK OS
Q2K5Q0	clpP3	OU	4	21578.59	5.49	277.40	todas	ATP-dependent Clp protease proteolytic subunit 3 OS
B3PQ64	RHECIAT_CH0004147	P	6	35676.98	4.85	284.48	todas	Putative nitrate/sulfonate/bicarbonate ABC transporter, substrate-binding protein OS
Q2JZ27	RHE_PF00269	P	13	38353.00	5.62	778.92	todas	Probable iron ABC transporter, substrate-binding protein OS
Q2K0P9	RHE_00259	P	18	36984.78	6.35	1709.13	todas	Probable sulfate uptake ABC transporter, substrate-binding protein OS
Q2K2Y6	sfuA	P	21	36351.17	5.53	1121.44	todas	Iron(III) ABC transporter, substrate-binding protein OS
Q2K3H4	RHE_CH03866	P	9	35844.32	5.00	326.60	todas	Putative nitrate/sulfonate/bicarbonate ABC transporter, substrate-binding protein OS
Q2KAX8	sodB	P	8	22461.19	5.98	1320.43	todas	Superoxide dismutase OS
Q2KBK4	RHE_CH00971	P	13	31827.94	5.18	1614.96	todas	Probable zinc ABC transporter, substrate-binding protein OS
Q2KCV8	pstS	P	18	36036.95	5.22	1644.59	todas	Phosphate ABC transporter, substrate-binding protein OS
Q2K524	RHE_CH03298	Q	17	41312.87	5.23	251.25	todas	Putative uncharacterized protein OS
Q2K283	RHE_PB00160	R	11	30303.11	5.34	323.97	todas	Probable amino acid ABC transporter, substrate-binding protein, lipoprotein YaeC family OS
Q2K339	RHE_CH04003	R	3	20933.02	7.03	92.94	todas	Putative reductase protein OS
Q2K3R9	argJ	R	7	42957.83	5.07	146.74	todas	Glutamate N-acetyltransferase/amino-acid N-acetyltransferase protein OS
Q2K6P0	RHE_CH02726	R	12	25870.44	5.67	241.69	todas	Putative hydrolase protein OS
Q2K3B2	RHE_CH03929	S	16	32536.28	5.12	893.70	todas	Putative ABC transporter, substrate-binding protein OS
Q2K497	RHE_CH03584	S	4	33161.57	5.69	71.76	todas	Hypothetical conserved protein OS

Q2K4M0	RHE_CH03460	S	7	26338.02	5.59	499.85	todas	Hypothetical conserved protein OS
Q2KA26	RHE_CH01507	S	5	47549.99	6.99	229.60	todas	Putative peptidoglycan binding protein (Involved in cell wall degradation) OS
Q2KD62	RHE_CH00402	S	5	18477.70	5.57	153.33	todas	Hypothetical conserved protein OS
Q2K547	trcX	T	8	28938.22	4.91	233.31	todas	Two-component response regulator protein OS
Q2K6I9	RHE_CH02778	T	7	13173.08	9.19	189.95	todas	Hypothetical conserved protein OS
Q2K4L8	tolB	U	18	47833.94	7.75	1591.16	todas	Protein tolB OS
Q2KE96	secB	U	2	17591.93	4.92	215.02	todas	Protein-export protein secB OS
Q2JYD6	RHE_PF00511	WC	4	18954.65	5.49	112.44	todas	Hypothetical conserved protein OS
Q2K1V1	RHE_PB00025	WC	17	64082.11	5.48	730.68	todas	Probable sugar ABC transporter, substrate-binding protein OS
Q2K355	RHE_CH03987	WC	2	18129.71	8.80	103.10	todas	Hypothetical conserved protein OS
Q2K4H2	RHE_CH03508	WC	8	34259.35	6.08	166.78	todas	Hypothetical conserved protein OS
Q2K4P7	RHE_CH03432	WC	10	33991.20	6.44	410.74	todas	Hypothetical conserved protein OS
Q2K4Y6	RHE_CH03337	WC	6	20657.70	9.27	913.36	todas	Hypothetical conserved protein OS
Q2K723	RHE_CH02591	WC	2	16984.08	4.92	65.77	todas	Hypothetical conserved protein OS
Q2K8F2	RHE_CH02100	WC	4	23091.71	9.33	210.13	todas	Putative signal peptide protein OS
Q2KBJ5	RHE_CH00980	WC	5	22289.64	9.51	149.12	todas	Hypothetical conserved protein OS
Q2KBX4	RHE_CH00851	WC	6	20226.71	4.95	515.39	todas	Hypothetical conserved protein OS
B3PVD5	rrpP	WC	2	72195.22	5.86	12.18	6-8	Inorganic pyrophosphatase protein OS
Q2K1Q8	RHE_PC00213	G	3	36915.66	6.19	14.02	6-8	Probable C4-dicarboxylate transport system, substrate-binding protein OS
Q2KBU7	mntH	P	2	47717.57	8.70	23.32	6-8	Manganese transporter protein OS
B3PQY9	RHECIAT_CH0004214	WC	9	18101.71	9.49	830.90	5-8	Hypothetical conserved protein OS
Q2K111	RHE_00370	G	2	43716.94	5.30	61.20	5-8	Probable sugar uptake ABC transporter, substrate-binding protein OS
Q2K3P4	hbdA	I	2	31665.58	6.00	24.77	5-8	3-hydroxybutyryl-CoA dehydrogenase protein OS
Q2K5T4	apt1	F	3	19329.30	6.20	25.42	5-8	Adenine phosphoribosyltransferase 1 OS
Q2K6D9	senC	R	3	21545.67	6.14	25.89	5-8	Cytochrome-c oxidase assembly factor protein OS
Q2K7D3	RHE_CH02476	IQ	2	42255.08	5.76	21.66	5-8	Probable 3-oxoacyl-(Acyl-acyl) synthase II protein OS
Q2K8W2	lpdAchl	C	5	50956.19	8.07	139.88	5-8	Dihydrolipoyl dehydrogenase OS
Q2K967	serS	J	3	47580.11	5.76	28.14	5-8	Seryl-tRNA synthetase OS
Q2K9K2	rpsH	J	5	14358.51	10.00	163.74	5-8	30S ribosomal protein S8 OS
Q2KAU3	groESch2	O	7	11363.12	5.86	146.14	5-8	10 kDa chaperonin OS
Q2KBH9	RHE_CH00996	R	3	23244.34	5.44	27.27	5-8	Probable hydrolase protein OS
Q2KCF9	flgG	N	5	27882.17	4.61	23.03	5-8	Flagellar basal-body rod protein OS
Q2KCP0	murA	M	3	45656.53	6.52	19.15	5-8	UDP-N-acetylglucosamine 1-carboxyvinyltransferase OS
B3PWS8	rplP	J	2	15527.26	10.87	35.45	5-7	50S ribosomal protein L16 OS
Q2K0W1	RHE_00320	WC	2	16123.33	6.51	25.38	5-7	Hypothetical conserved protein OS
Q2K1F6	RHE_PC00110	T	3	11931.62	5.16	14.47	5-7	Putative DnaK suppressor protein OS
Q2KBM5	tldD	R	2	49608.93	5.42	24.62	5-7	Zn-dependent protease protein OS
Q2K078	RHE_00084	E	6	95352.09	5.56	52.70	5,7,8	Putative sarcosine dehydrogenase protein OS
Q2K0U1	RHE_00300	E	2	#####	5.68	18.74	5,7,8	Probable N-methylhydantoinase (ATP-hydrolyzing)/5-oxoprolinase protein OS
Q2K7H5	RHE_CH02433	I	2	46428.17	9.40	16.89	5,7,8	Probable 3-hydroxyisobutyrate dehydrogenase protein OS
Q2K2S8	RHE_CH04115	R	4	23606.94	5.40	13.22	5,6,8	Hypothetical conserved protein OS

Q2K5Z8	dctA	C	5	46110.75	7.69	12.36	5,6,8	C4-dicarboxylate transport protein OS
Q2K8X0	kdsA	M	5	29981.39	5.67	18.22	5,6,8	2-dehydro-3-deoxyphosphooc-tonate aldolase protein OS
Q2KA91	fabD	I	3	32162.02	5.12	33.43	5,6,8	[acyl-carrier protein] S-malonyltransferase protein OS
Q2KAB2	RHE_CH01421	WC	3	17673.86	6.18	10.37	5,6,8	Hypothetical conserved protein OS
Q2KBC9	RHE_CH01048	QR	7	27515.32	5.33	22.16	5,6,8	Probable 3-oxoacyl-[acyl-carrier protein] reductase protein OS
Q2KD99	grpE	O	8	22933.70	4.80	28.64	5,6,8	Protein grpE OS
B3Q2Y0	azuP	C	5	15761.51	9.30	36.64	4-8	Pseudoazurin blue-copper protein OS
Q2K4C7	hss	WC	4	54581.07	6.21	32.86	4-8	Homospermidine synthase protein OS
Q2K139	RHE_00400	WC	3	30924.07	6.15	15.62	4-7	Probable 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase protein OS
B3PZP4	RHECIAT_CH0002151	WC	3	23056.81	9.33	19.07	4-6,8	Putative invasion-associated protein OS
Q2K550	RHE_CH03272	Q	2	12925.48	9.29	32.27	4-6,8	Hypothetical conserved protein OS
Q2K6F7	flaCch5	N	2	35764.00	5.07	300.65	4-6,8	Flagellin C protein OS
Q2KD75	tme	C	5	82500.61	5.92	73.33	4-6,8	Malate dehydrogenase (Oxaloacetate-decarboxylating) protein OS
B3PPK2	cheY2	J	4	13936.37	9.52	39.01	4-6	Chemotaxis two-component response regulator protein OS
B3PU59	lppA	WC	2	14487.45	9.60	28.12	4-6	Outer membrane lipoprotein-related protein OS
Q2K7Y3	purB	F	3	48578.61	6.00	34.22	4-6	Adenylosuccinate lyase protein OS
Q2K813	gcvP	E	4	#####	5.94	24.35	4-6	Glycine dehydrogenase [decarboxylating] OS
Q2KBM3	ctaC	C	3	32552.74	5.63	14.75	4-6	Cytochrome c oxidase subunit 2 OS
Q2KBX3	RHE_CH00852	G	3	31922.94	4.94	23.33	4-6	Probable gluconolactonase protein (Calcium-binding regucalcin-like) OS
Q2K7T4	alaS	I	2	95245.88	5.35	78.74	4,7,8	Alanyl-tRNA synthetase OS
Q2K9P9	topA	R	4	96832.10	8.98	19.12	4,6,8	DNA topoisomerase OS
Q2K302	efp	J	4	20954.77	5.37	55.95	4,5,8	Elongation factor P OS
Q2KBA8	RHE_CH01070	WC	4	16404.53	9.48	14.97	4,5,8	Hypothetical conserved protein OS
Q2K6H7	ilvI	EH	5	64328.75	5.81	15.24	4,5,7	Acetolactate synthase OS
B3PPT3	rpsD	J	2	23450.91	9.99	127.54	3-8	30S ribosomal protein S4 OS
Q2K005	RHE_00010	R	5	#####	4.86	69.60	3-8	Hypothetical conserved protein OS
Q2K5Q5	RHE_CH03064	R	8	20994.68	9.08	509.21	3-8	Hypothetical conserved protein OS
B3PZQ9	ppiA	O	6	20567.55	7.88	436.62	3-7	Peptidyl-prolyl cis-trans isomerase OS
Q2K4C6	omp3	WC	2	12703.39	9.41	116.87	3-7	Outer membrane lipoprotein OS
Q2K6H1	hflC	O	3	35435.97	7.93	45.24	3-7	Hydrolase serine protease transmembrane subunit C protein OS
Q2K8A8	occT	E	3	29738.22	5.60	35.98	3-7	Octopine ABC transporter, substrate-binding protein OS
Q2KD36	argB	E	2	31174.21	5.35	44.40	3-7	Acetylglutamate kinase OS
B3PS43	flaCch6	N	2	35820.25	4.84	548.15	3-6,8	Flagellin C protein OS
B3PXS8	valS	J	4	#####	5.13	56.69	3-6,8	Valyl-tRNA synthetase protein OS
Q2K2S6	RHE_CH04117	WC	4	16623.55	8.97	71.95	3-6,8	Hypothetical conserved protein OS
Q2K4A2	RHE_CH03579	WC	4	16429.01	9.73	62.93	3-6,8	Putative uncharacterized protein OS
Q2K6D6	creA	S	6	18027.56	7.75	58.70	3-6,8	CreA protein OS
Q2K8Z5	RHE_CH01905	WC	2	13269.77	5.84	29.08	3-6,8	Hypothetical conserved protein OS
Q2KAX1	RHE_CH01210	G	3	32358.83	5.40	22.68	3-6,8	Probable sugar ABC transporter, substrate-binding protein OS
B3PS42	RHECIAT_CH0004405	WC	3	26029.45	5.91	16.07	3-6	Hypothetical conserved protein OS
B3PS98	fabI2	I	8	29257.39	5.70	34.07	3-6	Enoyl-[acyl-carrier-protein] reductase (NADH) protein OS

B3PSE8	RHECIAT_CH0001085	WC	3	13557.05	4.60	14.18	3-6	Hypothetical conserved protein OS
B3PWT8	rpsE	J	3	20694.63	10.76	101.14	3-6	30S ribosomal protein S5 OS
B3Q251	RHECIAT_PB0000035	T	10	29603.66	5.64	22.73	3-6	Putative universal stress protein OS
Q2JYF6	RHE_PF00490	R	2	24738.55	6.60	27.29	3-6	Hypothetical conserved protein OS
Q2JYI9	vbsS	Q	19	#####	5.60	114.36	3-6	Vicibactin biosynthesis non-ribosomal peptide synthase protein OS
Q2JZ62	RHE_PF00234	EP	3	72156.90	6.23	40.84	3-6	Probable oligopeptide ABC transporter, substrate-binding protein OS
Q2K2P4	traH	WC	4	21628.47	4.59	28.81	3-6	Conjugal transfer protein H OS
Q2K2Y0	rpmA	J	4	9464.72	10.54	62.76	3-6	50S ribosomal protein L27 OS
Q2K4H4	rpmE	J	3	8084.15	9.22	132.95	3-6	50S ribosomal protein L31 OS
Q2K613	RHE_CH02956	S	2	16624.08	7.87	51.18	3-6	Hypothetical conserved protein OS
Q2K7D8	RHE_CH02471	R	3	13324.77	4.94	10.47	3-6	Putative glyoxalase protein OS
Q2K7E1	gnd	G	9	52217.80	6.14	101.81	3-6	6-phosphogluconate dehydrogenase, decarboxylating OS
Q2K7M2	rapB2	WC	3	30233.80	4.34	51.92	3-6	Autoaggregation protein (Adhering protein) OS
Q2K7N5	nadE	H	2	61003.39	5.17	29.71	3-6	NAD(+) synthase (Glutamine-hydrolysing) protein OS
Q2K814	gcvH	E	2	12773.28	4.11	34.08	3-6	Glycine cleavage system H protein OS
Q2K914	gloA	E	5	16631.81	5.38	31.92	3-6	Lactoylglutathione lyase methylglyoxalase protein OS
Q2K9J4	rpsM	J	7	13788.14	11.10	66.92	3-6	30S ribosomal protein S13 OS
Q2K9L4	rplW	J	3	10543.30	10.06	65.18	3-6	50S ribosomal protein L23 OS
Q2K9Z5	RHE_CH01538	WC	4	26738.61	10.84	20.95	3-6	Hypothetical conserved protein OS
Q2KCE8	flaCch4	N	6	36257.55	4.53	94.70	3-6	Flagellin C protein OS
Q2KD27	RHE_CH00437	N	3	66502.73	8.62	25.23	3-6	Putative inner membrane transmembrane protein OS
Q2KDQ6	cpaD	WC	2	24746.50	6.11	77.36	3-6	Pilus assembly protein OS
Q2KDU2	gpmA	G	2	23206.40	5.97	57.24	3-6	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase OS
Q2KE53	RHE_CH00051	WC	4	34753.25	8.34	63.93	3-6	Hypothetical conserved protein OS
Q2KE87	RHE_CH00016	Q	4	29623.87	5.60	52.56	3-6	Probable bifunctional 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase/5-carboxymethyl-2-oxo-
Q2K163	tauA	P	8	34971.62	5.65	23.81	3-5,8	Probable taurine uptake ABC transporter, substrate-binding protein OS
Q2K5T3	fbcC	C	2	32020.05	5.18	18.63	3-5,8	Ubiquinol--cytochrome c reductase, cytochrome c1 subunit protein OS
Q2K7P5	RHE_CH02361	G	3	32706.65	5.98	17.19	3-5,8	Probable sugar ABC transporter, substrate-binding protein OS
Q2K9W7	argC	E	3	33012.86	5.58	67.87	3-5,8	N-acetyl-gamma-glutamyl-phosphate reductase OS
Q2JZM5	RHE_PF00068	G	3	36098.28	8.21	48.70	3-5,7	Probable C4-dicarboxylate ABC transporter, substrate-binding protein OS
B3PRW7	efp	J	3	21038.85	5.37	33.77	3-5	Elongation factor P OS
B3PYR5	RHECIAT_CH0002033	E	4	22865.16	5.34	21.38	3-5	Probable arylesterase protein OS
Q2K3G9	atpG	C	7	29324.82	8.35	41.09	3-5	ATP synthase gamma chain OS
Q2K4E9	purE	F	2	15123.38	5.11	22.27	3-5	Phosphoribosylaminoimidazole carboxylase catalytic subunit protein OS
Q2K4X4	mexE1	M	7	41595.30	7.86	25.37	3-5	Multidrug-efflux system secretion protein, HlyD family OS
Q2K741	RHE_CH02571	R	3	30673.78	6.22	17.21	3-5	Probable chloride peroxidase protein OS
Q2K7X5	purQ	F	3	23745.39	6.20	26.83	3-5	Phosphoribosylformylglycinamide synthase 1 OS
Q2K814	RHE_CH02068	G	2	45010.04	6.03	21.52	3-5	Probable sugar ABC transporter, substrate-binding protein OS
Q2K9T4	nuoB1	C	3	21293.67	8.60	23.40	3-5	NADH-quinone oxidoreductase subunit B 1 OS
Q2KA68	RHE_CH01465	E	9	28713.75	6.20	34.29	3-5	Probable amino acid ABC transporter, substrate-binding protein OS
Q2KD94	metK	H	6	44880.84	8.55	49.68	3-5	S-adenosylmethionine synthetase OS

B3PWA2	idhA	R	4	35085.17	5.93	32.18	3,7,8	Myo-inositol 2-dehydrogenase protein OS
B3PWS5	rpsS	J	2	10399.04	10.28	40.02	3,6,8	30S ribosomal protein S19 OS
Q2K340	pyc	C	7	#####	5.90	58.31	3,5-8	Pyruvate carboxylase protein OS
Q2JZD3	RHE_PF00162	ET	8	27572.73	5.94	30.96	3,5-7	Probable amino acid ABC transporter, substrate-binding protein OS
Q2K2Y2	RHE_CH04060	H	3	51844.23	5.30	24.91	3,5,7	Hypothetical conserved protein OS
B3PWI8	RHECIAT_CH0000053	WC	3	34572.07	8.65	43.53	3,5,6	Hypothetical conserved protein OS
B3PXN7	cpdB	F	7	71657.40	5.69	66.87	3,5,6	2',3'-cyclic-nucleotide 2'-phosphodiesterase protein OS
B3PZ95	rpsT	J	2	9791.27	12.01	57.03	3,5,6	30S ribosomal protein S20 OS
Q2K280	RHE_PB00156	Q	2	40593.15	5.39	30.04	3,5,6	Hypothetical conserved protein OS
Q2K5T2	fbcB	C	3	47899.41	9.13	22.20	3,5,6	Cytochrome b OS
Q2KC22	gsh	H	3	51187.07	5.39	31.82	3,5,6	Glutamate-cysteine ligase protein OS
Q2KCE4	motD	WC	3	48103.69	4.75	13.72	3,5,6	Chemotaxis motility protein OS
Q2KE03	fabB	IQ	5	43554.07	5.77	87.67	3,5,6	3-oxoacyl-(Acyl-carrier-protein) synthase I protein OS
Q2KE12	cmk	F	7	28316.51	8.96	22.58	3,5,6	Cytidylate kinase (Cytidine monophosphate kinase) protein OS
Q8KKV1	RHE_PD00022	WC	2	23769.01	5.73	25.35	3,5,6	Hypothetical conserved protein OS
B3PSX6	RHECIAT_CH0001161	WC	3	13062.88	5.82	17.33	3,4,8	Hypothetical conserved protein OS
Q2K013	RHE_00018	WC	3	11081.45	8.71	301.85	3,4,8	Hypothetical conserved protein OS
Q2K4B4	gltB	E	9	#####	5.68	50.08	3,4,8	Glutamate synthase large subunit protein OS
Q2K171	RHE_PC00028	E	2	36103.75	4.88	12.38	3,4,7,8	Probable polyamine ABC transporter, substrate-binding protein OS
Q2JYF9	RHE_PF00487	G	6	#####	5.80	68.14	3,4,6-8	Putative alpha-mannosidase protein OS
Q2K1L1	RHE_PC00167	P	3	34867.45	5.31	17.84	3,4,6,8	Probable aliphatic sulphonate ABC transporter, substrate-binding protein OS
Q2K373	RHE_CH03968	TK	3	25563.17	5.50	32.51	3,4,6	Putative two-component response regulator protein OS
Q2K4F7	cheY3	T	6	12889.92	6.23	58.23	3,4,6	Chemotaxis two-component response regulator protein OS
Q2K4M9	rapB3	WC	2	29898.63	3.97	48.09	3,4,6	Autoaggregation protein (Adhering protein) OS
Q2K5H8	RHE_CH03142	G	2	35093.66	5.55	13.17	3,4,6	Putative chitooligosaccharide deacetylase protein OS
Q2K807	RHE_CH02249	E	3	45585.03	5.94	26.93	3,4,6	Putative selenocysteine lyase protein OS
Q2KAB7	RHE_CH01416	S	2	68519.65	5.10	19.30	3,4,6	Putative peptidoglycan binding protein (LysM domain involved in cell wall degradation) OS
Q2KAR0	groL3	O	3	57575.25	5.35	15.04	3,4,6	60 kDa chaperonin 3 OS
B3PRY2	sfuA	P	12	36721.64	5.49	663.86	2-8	Iron(III) ABC transporter, substrate-binding protein OS
B3PUK3	greA	K	5	17461.55	4.94	98.53	2-8	Transcription elongation factor greA OS
Q2JYW6	groL4	O	12	57696.17	5.23	1119.44	2-8	60 kDa chaperonin 4 OS
Q2JZH2	htpG	O	9	70274.43	5.20	123.68	2-8	Chaperone protein htpG OS
Q2K113	RHE_00373	M	6	16580.24	5.68	167.32	2-8	Probable symbiotically induced surface protein OS
Q2K1D3	RHE_PC00088	N	7	38158.57	6.23	80.23	2-8	Hypothetical conserved protein OS
Q2K1E9	RHE_PC00103	G	8	59788.93	5.11	118.02	2-8	Probable sugar ABC transporter, substrate-binding protein OS
Q2K1G2	RHE_PC00118	E	16	29585.07	5.63	263.78	2-8	Probable amino acid ABC transporter, substrate-binding protein OS
Q2K249	RHE_PB00126	EP	6	57977.68	5.97	82.59	2-8	Probable oligopeptide ABC transporter, substrate-binding protein OS
Q2K2X2	modA	P	8	27353.28	5.37	197.25	2-8	Molybdenum ABC transporter, substrate-binding protein OS
Q2K350	dppAch3	E	27	60145.02	5.26	1138.36	2-8	Dipeptide ABC transporter, substrate-binding protein OS
Q2K3A4	tesA	E	4	18996.80	4.99	74.35	2-8	Acyl-CoA thioesterase protein OS
Q2K484	pgm	G	9	58669.04	5.24	289.99	2-8	Phosphoglucomutase protein OS

Q2K4D7	RHE_CH03543	F	5	17731.06	5.45	64.12	2-8	Hypothetical conserved membrane protein OS
Q2K4U9	RHE_CH03374	G	7	44367.30	5.01	90.70	2-8	Probable sugar ABC transporter, substrate-binding protein OS
Q2K525	RHE_CH03297	P	5	34133.87	5.06	64.49	2-8	Probable nitrate/sulfonate ABC transporter, substrate-binding protein OS
Q2K560	RHE_CH03262	R	7	12078.48	6.12	141.29	2-8	Hypothetical conserved protein OS
Q2K5D6	RHE_CH03185	WC	4	15402.44	5.35	72.71	2-8	Hypothetical conserved protein OS
Q2K5H9	RHE_CH03141	S	3	17952.32	5.25	53.75	2-8	Hypothetical conserved protein OS
Q2K6S8	RHE_CH02687	WC	2	16510.69	5.36	46.42	2-8	Hypothetical conserved protein OS
Q2K7E8	RHE_CH02461	G	15	43523.50	5.20	386.32	2-8	Probable sugar ABC transporter, substrate-binding protein OS
Q2K7L5	RHE_CH02391	S	3	14335.12	5.41	58.14	2-8	Hypothetical conserved protein OS
Q2K7W3	RHE_CH02293	E	19	29136.14	8.60	180.73	2-8	Probable amino acid ABC transporter, substrate-binding protein OS
Q2K7Y5	RHE_CH02271	WC	6	25783.47	8.47	103.71	2-8	Hypothetical conserved protein OS
Q2K817	RHE_CH02239	S	3	36478.52	5.73	90.85	2-8	Hypothetical conserved protein OS
Q2K8F3	RHE_CH02099	EP	9	67412.19	5.80	205.72	2-8	Probable peptide ABC transporter, substrate-binding protein OS
Q2K8J2	dppAch2	EP	5	58308.45	6.14	76.94	2-8	Dipeptide ABC transporter, substrate-binding protein OS
Q2K925	accBch	I	5	16345.78	4.79	92.92	2-8	Acetyl-CoA carboxylase, biotin carboxyl carrier protein OS
Q2K9C1	aldA	E	3	38756.46	6.23	42.61	2-8	Alanine dehydrogenase protein OS
Q2K9J8	rpmD	J	2	7844.25	11.51	42.72	2-8	50S ribosomal protein L30 OS
Q2K9K0	rplR	J	4	12743.60	11.15	165.04	2-8	50S ribosomal protein L18 OS
Q2KAF2	dapAch2	EM	4	31085.59	6.76	111.38	2-8	Dihydrodipicolinate synthase protein OS
Q2KBM4	RHE_CH00951	WC	5	27671.73	8.37	111.14	2-8	Hypothetical conserved protein OS
Q2KC31	guaB	F	2	52116.54	6.44	55.13	2-8	Inosine-5'-monophosphate dehydrogenase OS
Q2KC60	noeL1	M	7	41056.42	5.68	47.15	2-8	GDP-mannose 4,6-dehydratase protein OS
Q2KCK8	rbsBch1	G	5	34847.38	8.62	122.04	2-8	Ribose ABC transporter, substrate-binding protein OS
Q2KDP1	RHE_CH00221	EP	6	55801.79	5.32	69.28	2-8	Probable dipeptide ABC transporter, substrate-binding protein OS
Q2KDX8	cycM	C	2	19391.06	5.79	48.49	2-8	Cytochrome-c protein OS
B3PTX6	pfk	G	2	43343.21	5.83	42.60	2-6,8	6-phosphofructokinase protein OS
B3PXA6	RHECIAT_CH0003541	Q	11	41269.70	5.30	170.13	2-6,8	Putative uncharacterized protein OS
Q2K2Q6	trbJ	U	4	29270.63	5.01	121.57	2-6,8	Conjugal transfer protein J OS
Q2K4M6	serA	HE	11	56646.42	5.51	173.05	2-6,8	D-3-phosphoglycerate dehydrogenase protein OS
Q2K5T8	RHE_CH03031	R	2	47098.64	5.39	70.34	2-6,8	Putative GTP-binding protein OS
Q2K890	rapB1	WC	4	29201.09	4.09	179.13	2-6,8	Autoaggregation protein (Adhering protein) OS
Q2KA90	fabG2	IQR	4	26039.97	6.92	34.07	2-6,8	3-oxoacyl-(Acyl-carrier protein) reductase protein OS
Q2KBZ3	ileS	J	2	#####	5.70	87.36	2-6,8	Isoleucyl-tRNA synthetase protein OS
Q2K960	RHE_CH01833	S	4	13905.00	6.73	38.40	2-6	Hypothetical conserved protein OS
Q2K135	RHE_00394	WC	4	15461.24	6.97	145.25	2-5,8	Hypothetical conserved protein OS
Q2K6Z6	RHE_CH02618	C	3	14207.29	9.10	36.56	2-5,8	Probable cytochrome-c protein OS
Q2KC98	RHE_CH00723	WC	2	35765.43	4.57	145.44	2-5,8	Probable metalloproteinase protein OS
Q2KC13	lpxQ	P	2	24806.55	5.45	33.80	2-5,7,8	Lipid A oxidase (Involved in formation of 2-aminogluconate) protein OS
B3Q4F9	katG	P	8	79729.73	5.60	122.41	2-5	Catalase-peroxidase OS
Q2K8C6	afuA1	P	6	37268.39	6.85	29.64	2-4,8	Iron(III) ABC transporter, substrate-binding protein OS
B3Q2W2	RHECIAT_PB0000312	WC	2	15002.27	7.88	48.18	2-4,7,8	Hypothetical conserved protein OS

Q2K4C5	RHE_CH03555	F	17	69551.26	5.00	292.14	2-4,6-8	Probable 5'-nucleotidase protein OS
Q2K323	phbB	IQR	3	25578.09	6.52	81.30	2-4,6,8	Acetoacetyl-CoA reductase protein OS
Q2K7N7	aroG	E	6	38240.70	5.95	70.47	2-4,6,8	Phospho-2-dehydro-3-deoxyheptonate aldolase protein, Phe-sensitive OS
Q2K964	lppB	M	9	53996.77	9.73	273.61	2-4,6,8	Lipoprotein protein OS
Q2KA96	rplI	J	4	21069.66	4.87	116.01	2-4,6,8	50S ribosomal protein L9 OS
B3PXF2	RHECIAT_CH0000162	S	4	11424.47	5.45	58.86	2-4,6	Hypothetical conserved protein OS
Q2JYI7	vbsA	J	3	37821.86	6.53	38.16	2-4,6	N6-hydroxylysine O-acetyltransferase protein OS
Q2K0F6	RHE_00163	WC	2	13676.42	6.50	32.77	2-4,6	Hypothetical conserved protein OS
Q2K1J3	RHE_PC00149	WC	3	38040.24	5.73	25.42	2-4,6	Putative sugar ABC transporter, substrate-binding protein OS
Q2K320	mgpS	LKJ	3	#####	6.54	31.78	2-4,6	ATP-dependent DNA helicase protein OS
Q2K7V2	grlA	O	7	12367.04	4.81	48.00	2-4,6	Glutaredoxin protein OS
Q2K7W7	aroQ2	E	2	16814.09	6.54	16.33	2-4,6	3-dehydroquinone dehydratase protein OS
Q2K9J5	adk	F	5	23602.09	5.66	87.19	2-4,6	Adenylate kinase OS
Q2KCB6	zwf1	G	4	55301.00	6.26	17.09	2-4,6	Glucose-6-phosphate 1-dehydrogenase OS
Q2K5T5	RHE_CH03034	I	3	17549.96	8.69	18.63	2-4	Putative enoyl-CoA hydratase protein OS
Q2K773	RHE_CH02538	WC	2	20586.28	5.39	12.85	2-4	Hypothetical conserved protein OS
Q2KBM6	RHE_CH00949	KR	6	18055.68	5.72	24.29	2-4	Probable acetyltransferase protein OS
Q2KCD9	flgLch	N	8	37008.60	5.20	68.69	2-4	Flagellar hook-associated protein OS
Q2KD58	RHE_CH00406	J	2	23618.76	6.55	65.10	2-4	Probable sigma 54 modulation protein OS
B3PXD4	fabB	IQ	5	43490.96	5.90	119.69	2,6-8	3-oxoacyl-[acyl-carrier-protein] synthase protein OS
Q2K2D9	virB9a	U	2	32157.39	8.43	21.09	2,6-8	Transport secretion system IV, VirB9 protein OS
Q2K9H7	RHE_CH01714	I	3	37663.03	6.72	22.75	2,6,8	Putative esterase/lipase/thioesterase family protein OS
Q2K2T8	RHE_CH04105	E	5	#####	6.11	28.32	2,6,7	Putative NAD-specific glutamate dehydrogenase protein OS
Q2K9T3	nuoC	C	3	22839.93	4.91	22.32	2,6,7	NADH-quinone oxidoreductase subunit C OS
B3PVY4	clpP2	OU	4	23345.85	6.08	53.79	2,5-8	ATP-dependent Clp protease proteolytic subunit OS
Q2JYP2	etfAf	C	2	31489.87	4.98	69.35	2,5-8	Electron-transport flavoprotein, alpha-subunit OS
Q2K0F3	RHE_00162	WC	3	14454.96	9.77	40.09	2,5-8	Hypothetical conserved protein OS
Q2K0N6	RHE_00245	G	3	46367.02	5.05	48.81	2,5-8	Probable sugar ABC transporter, sugar-binding protein OS
Q2K274	RHE_PB00150	G	3	45387.14	4.94	43.06	2,5-8	Probable sugar ABC transporter, substrate-binding protein OS
Q2K315	RHE_CH04027	O	3	55547.93	7.82	46.80	2,5-8	Putative metalloprotease protein OS
Q2K4H1	RHE_CH03509	S	5	6890.71	9.30	87.38	2,5-8	Hypothetical conserved protein OS
Q2K587	pssN	M	7	43028.58	5.81	216.62	2,5-8	Exopolysaccharide export protein OS
Q2K665	RHE_CH02903	WC	4	17830.37	5.89	39.41	2,5-8	UPF0303 protein RHE_CH02903 OS
Q2K714	RHE_CH02600	WC	4	13585.56	6.56	67.04	2,5-8	Putative uncharacterized protein OS
Q2K789	RHE_CH02521	WC	4	10184.06	9.65	104.18	2,5-8	Hypothetical conserved protein OS
Q2K790	RHE_CH02520	WC	3	17091.76	7.88	63.90	2,5-8	Putative copper-containing oxidase protein OS
Q2K7M7	RHE_CH02379	S	3	19902.71	6.35	138.66	2,5-8	Hypothetical conserved protein OS
Q2K8G7	RHE_CH02085	G	8	34309.32	6.06	89.22	2,5-8	Probable sugar ABC transporter, substrate-binding protein OS
Q2K9B6	cysK2	E	3	36393.44	5.41	51.18	2,5-8	Cysteine synthase OS
Q2K9N4	RHE_CH01654	WC	2	24325.63	4.87	40.01	2,5-8	Putative metalloprotease inhibitor (Transcriptional regulator) protein OS
Q2KAC8	RHE_CH01405	WC	3	31942.20	5.10	43.14	2,5-8	Hypothetical conserved protein OS

Q2KB07	RHE_CH01174	O	4	14330.33	6.10	136.97	2,5-8	Probable organic hydroperoxide resistance (Osmotically inducible) protein OS
Q2KCN8	hisD	E	4	45179.51	5.08	42.32	2,5-8	Histidinol dehydrogenase OS
Q2KDE2	gshB	HJ	2	35369.54	5.96	24.50	2,5-7	Glutathione synthetase protein OS
Q2K9S4	nuoI1	C	4	18865.53	8.11	11.09	2,5,8	NADH-quinone oxidoreductase subunit I 1 OS
Q2K3B6	RHE_CH03925	QR	3	31368.80	5.92	45.86	2,5,6,8	Probable oxidoreductase protein OS
Q2K709	plyA1	WC	5	51532.49	4.56	112.80	2,5,6,8	Polysaccharidase protein OS
Q2JYH7	fhuA	P	10	79084.67	4.75	31.84	2,5,6	Ferrichrome-iron transporter, receptor protein OS
B3PPQ7	aglE	G	14	49074.44	4.94	773.82	2,4-8	Alpha-glucoside ABC transporter, substrate-binding protein OS
Q2K072	RHE_00077	ET	8	29920.05	4.86	74.90	2,4-8	Probable octopine ABC transporter, substrate-binding protein OS
Q2K1T3	uxuAb	G	3	43840.67	5.48	35.63	2,4-8	Mannonate dehydratase protein OS
Q2K2U2	RHE_CH04101	M	2	29000.66	9.62	59.76	2,4-8	Putative membrane-bound lytic murein transglycosylase protein OS
Q2K5E8	xylF	G	16	36024.93	5.33	238.60	2,4-8	Xylose ABC transporter, substrate-binding protein OS
Q2K5V7	RHE_CH03012	WC	4	22033.94	4.64	41.54	2,4-8	Hypothetical conserved protein OS
Q2K686	RHE_CH02882	S	14	49980.01	4.99	284.43	2,4-8	Hypothetical conserved protein OS
Q2K6A6	pfk	G	2	43524.40	5.78	54.39	2,4-8	6-phosphofructokinase protein OS
Q2K7Y6	r	G	3	23839.38	6.30	22.73	2,4-8	D-ribulose-5 phosphate 3-epimerase protein OS
Q2K8E9	glnB	E	5	12281.18	5.90	38.59	2,4-6	Nitrogen regulatory protein P-II OS
Q2KBY4	pmbA	R	15	47299.13	5.09	41.82	2,4-6	Modulator of DNA gyrase protein OS
Q8KL06	RHE_PD00077	WC	2	16073.31	9.58	28.29	2,4-6	Putative uncharacterized protein OS
B3PNL7	phoU	P	10	26142.69	5.21	17.34	2,4,8	Phosphate uptake transcriptional regulator protein OS
Q2K3N8	hpt	F	3	19835.92	5.60	18.92	2,4,8	Hypoxanthine phosphoribosyltransferase protein OS
B3PS54	rho	K	11	47016.19	6.01	60.15	2,4,6-8	Transcription termination factor protein OS
Q2K0D3	RHE_00139	S	2	18484.79	4.67	23.44	2,4,6-8	Hypothetical conserved protein OS
Q2K553	hmuT	P	8	32580.64	5.70	81.22	2,4,6-8	Hemin transport system, hemin-binding protein OS
Q2KAK3	RHE_CH01328	WC	4	19452.65	10.21	27.90	2,4,6-8	Hypothetical conserved protein OS
B3PW16	hslV	O	3	18699.45	6.75	73.93	2,4,6,8	ATP-dependent protease hslV OS
Q2K3G7	atpH	C	3	20015.07	7.99	61.79	2,4,6,8	ATP synthase subunit delta OS
Q2K9T7	btuE	O	2	19981.63	5.30	23.60	2,4,6,8	Glutathione peroxidase OS
B3PS16	plyA2	WC	3	51963.75	5.00	472.39	2,4,6	Polysaccharidase protein OS
Q2JYH6	RHE_PF00470	S	2	24865.32	9.45	12.27	2,4,6	Hypothetical conserved protein OS
Q2K939	gatA	G	5	52893.18	5.12	63.16	2,4,5,8	Glutamyl-tRNA(Gln) amidotransferase subunit A OS
Q2K9I4	degPch3	O	4	45507.59	5.24	27.62	2,3-5-8	Serine protease DO-like protein OS
Q2K447	RHE_CH03634	S	4	24434.66	4.89	25.57	2,3,8	Hypothetical conserved protein OS
Q2K5A4	pssJ	WC	2	30770.00	9.78	12.57	2,3,8	Galactosyltransferase protein OS
Q2K3V8	RHE_CH03723	C	4	55487.54	5.85	65.94	2,3,7,8	Putative aldehyde dehydrogenase protein OS
B3PTD3	cysD	EH	5	36470.45	6.43	19.16	2,3,7	Sulfate adenyltransferase subunit 2 OS
Q2KAW1	bla	V	4	31553.61	4.84	29.48	2,3,6-8	Beta-lactamase protein OS
Q2KAW7	cycF	WC	3	14956.97	8.55	35.90	2,3,6-8	Cytochrome-c-556 signal peptide protein OS
B3Q1A2	RHECIAT_PA0000112	M	5	76630.65	5.41	25.21	2,3,6,7	Putative C-terminal processing peptidase protein OS
B3PQ70	atpA	C	5	54572.55	6.95	88.44	2,3,6	ATP synthase subunit alpha OS
B3PRS2	cspA4	K	3	7431.38	9.30	70.24	2,3,6	Cold shock protein OS

Q2KE58	hisH	E	4	23359.80	6.06	11.62	2,3,6	Imidazoleglycerol phosphate synthase subunit (IGP synthase glutamine amidotransferase subunit)
Q2K226	RHE_PB00102	N	6	34579.44	5.55	63.40	2,3,5-8	Putative lipoprotein periplasmic protein OS
Q2K5F8	gguC	S	2	38316.33	5.98	28.81	2,3,5-8	GguC protein OS
Q2K6D0	RHE_CH02838	R	3	31978.19	5.68	229.12	2,3,5-8	Hypothetical conserved protein OS
Q2K9X9	RHE_CH01555	G	2	48146.00	5.52	71.70	2,3,5-8	Probable sugar ABC transporter, substrate-binding protein OS
Q2KCB7	pgl	G	8	24092.27	5.26	130.62	2,3,5-8	6-phosphogluconolactonase protein OS
Q2K3C5	typA	N	5	66695.63	5.41	65.58	2,3,5,8	GTP-binding elongation factor protein (Tyrosin phosphorylated protein) OS
Q2K083	agpA	EP	7	77207.32	5.82	143.84	2,3,5,6,8	Alpha-galactoside ABC transporter, substrate-binding protein OS
Q2K6T2	RHE_CH02683	P	5	35148.08	5.62	35.65	2,3,5,6	Probable ferrichrome ABC transporter, substrate-binding protein OS
Q2KBM8	RHE_CH00947	P	2	24353.87	4.93	22.15	2,3,5	Probable phosphate transporter (Pit) regulator protein OS
Q2K372	RHE_CH03969	T	5	16691.15	5.92	23.08	2,3,4	Putative cAMP-dependent kinase protein OS
Q2K397	acnA	C	6	96842.29	5.53	364.98	1-7	Aconitate hydratase (Citrate hydro-lyase) protein OS
Q2K5U8	RHE_CH03021	E	5	42726.87	5.61	82.60	1-7	Probable Xaa-Pro dipeptidase protein OS
Q2KDZ9	rpsO	J	4	10171.63	9.99	73.86	1-7	30S ribosomal protein S15 OS
B3PVV6	trxA1	OC	5	11391.11	4.65	220.27	1-6,8	Thioredoxin protein OS
B3PWT3	rpIE	J	3	20951.44	9.62	164.15	1-6,8	50S ribosomal protein L5 OS
Q2JYW7	groESf	O	6	11438.10	5.46	175.48	1-6,8	10 kDa chaperonin OS
Q2JZ85	thuE	G	8	45602.74	5.70	177.85	1-6,8	Trehalosemaltose ABC transporter, sugar-binding protein OS
Q2K437	RHE_CH03644	G	3	34088.00	5.19	26.83	1-6,8	Probable sugar ABC transporter, substrate-binding protein OS
Q2K4I1	pgk	G	7	43493.60	4.98	130.60	1-6,8	Phosphoglycerate kinase OS
Q2K6N6	ilvE2	EH	3	39616.31	6.04	49.48	1-6,8	Branched-chain-amino-acid aminotransferase OS
Q2K805	sufC	R	4	27407.47	5.17	118.84	1-6,8	ABC transporter involved in Fe-S cluster assembly, ATP-binding protein OS
Q2K8Y0	omp1	M	13	84637.63	4.89	539.61	1-6,8	Outer membrane lipoprotein OS
Q2K9C5	cpdB	F	7	71705.32	5.54	173.68	1-6,8	2',3'-cyclic-nucleotide 2'-phosphodiesterase protein OS
Q2K9K7	rpsQ	J	2	9088.56	10.07	40.48	1-6,8	30S ribosomal protein S17 OS
Q2KAN8	degPch2	O	14	54953.25	5.06	137.96	1-6,8	Serine protease DO-like protein OS
Q2KAU5	RHE_CH01236	S	2	24391.40	5.26	47.53	1-6,8	Hypothetical conserved protein OS
Q2KCD6	flgD	N	4	14760.19	4.12	258.80	1-6,8	Flagellar basal-body rod protein OS
Q2KDD4	dut	F	4	16284.53	5.17	47.55	1-6,8	Deoxyuridine 5'-triphosphate nucleotidohydrolase OS
Q2JYI5	vbsL	E	6	43384.60	6.47	200.28	1-6	Cystathionine gamma-synthase protein OS
Q2KAV5	ptrB	E	9	78053.84	5.13	209.37	1-6	Protease II protein OS
B3PRE2	RHECIAT_CH0000940	WC	4	20159.62	4.97	359.36	1-5	Hypothetical conserved protein OS
Q2K7F3	RHE_CH02456	WC	5	28065.66	5.78	123.33	1-4,6-8	Hypothetical conserved protein OS
Q2K9R4	proS	J	3	49449.37	5.70	60.35	1-4,6,7	Prolyl-tRNA synthetase OS
Q2JZ78	RHE_PF00217	G	3	33832.81	5.48	29.34	1-4,6	Probable D-ribose ABC transporter, substrate-binding protein OS
B3PP82	lysC	E	2	45303.78	5.50	81.16	1-4	Aspartokinase OS
Q2K636	ropB2	M	2	21977.14	4.15	30.94	1-4	Outer membrane protein OS
Q2K8T9	RHE_CH01961	WC	2	18744.41	6.84	18.88	1-4	Hypothetical conserved protein OS
Q2K0V6	RHE_00318	R	3	19088.62	9.00	27.57	1-3,7,8	Hypothetical conserved protein OS
Q2KD46	dapD	E	2	30570.88	5.99	27.60	1-3,6,8	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase OS
Q2JYD7	azuPf	C	5	15639.27	7.86	64.25	1-3,5-8	Pseudoazurin blue-copper protein OS

Q2K1Q3	rapAc	WC	3	25377.87	4.38	92.33	1-3,5-8	Autoaggregation protein OS
Q2K3F8	lpdAch2	C	9	48617.03	5.98	157.86	1-3,5-8	Dihydrolipoamide dehydrogenase protein OS
Q2K3Y6	RHE_CH03695	G	5	34679.59	5.30	58.68	1-3,5-8	Putative sugar ABC transporter, substrate-binding protein OS
Q2K473	RHE_CH03608	R	6	31572.97	8.85	405.12	1-3,5-8	Probable O-linked GlcNAc transferase protein OS
Q2K708	plyA2	WC	22	51787.62	4.98	3662.63	1-3,5-8	Polysaccharidase protein OS
Q2K7H2	ropAch2	WC	4	36637.03	3.96	262.52	1-3,5-8	Porin outer membrane protein OS
Q2K8W6	pdhA2	C	5	49347.67	4.74	414.83	1-3,5-8	Pyruvate dehydrogenase beta subunit protein OS
Q2K8X8	fabZ	I	2	17106.83	6.51	67.46	1-3,5-8	(3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase OS
Q2K9J7	rplO	J	2	16250.79	10.59	169.71	1-3,5-8	50S ribosomal protein L15 OS
Q2K9M8	rplK	J	2	15118.82	9.80	161.49	1-3,5-8	50S ribosomal protein L11 OS
Q2K9W3	rplM	J	10	17286.04	10.12	148.58	1-3,5-8	50S ribosomal protein L13 OS
Q2KA93	rpsF	J	3	17541.78	9.47	155.46	1-3,5-8	30S ribosomal protein S6 OS
Q2KCJ1	RHE_CH00629	TK	3	16616.85	4.68	38.15	1-3,5-8	Putative two-component response regulator protein OS
Q2K220	RHE_PB00095	O	5	32439.74	6.00	33.06	1-3,5-7	Probable glutathione S-transferase protein OS
Q2K4M5	serC	HE	13	42454.52	5.51	94.83	1-3,5-7	Phosphoserine aminotransferase protein OS
Q2K9M3	rpoC	K	23	#####	6.48	339.70	1-3,5,6,8	DNA-directed RNA polymerase subunit beta' OS
Q2KB13	RHE_CH01168	S	2	26803.52	9.27	72.29	1-3,5	Hypothetical conserved protein OS
Q2K619	sigA	K	3	77201.30	5.02	36.06	1-3	RNA polymerase sigma factor OS
Q2KDX6	RHE_CH00135	G	3	35739.49	4.96	28.04	1-3	Probable sugar kinase protein OS
Q2JYH5	RHE_PF00471	M	4	19210.78	7.66	36.12	1,6,7	Probable outer membrane lipoprotein OS
Q2K2V0	leuB	CE	8	39711.80	5.10	158.48	1,5-8	3-isopropylmalate dehydrogenase OS
Q2K325	aniA	S	2	21973.07	5.80	73.73	1,5-8	Carbon flux regulator protein OS
Q2K578	noeJ	M	5	53084.71	6.44	33.63	1,5-8	Mannose-1-phosphate guanylyltransferase (GDP) protein OS
Q2K5R7	RHE_CH03052	R	2	36097.56	5.30	40.69	1,5-8	Putative methanol dehydrogenase regulator MoxR-like protein OS
Q2K5U2	RHE_CH03027	E	3	26434.77	7.63	48.89	1,5-8	Putative amino acid ABC transporter, substrate-binding protein OS
Q2K6R4	RHE_CH02701	Q	4	25015.49	5.69	48.22	1,5-8	Probable fumarylpyruvate hydrolase protein OS
Q2K7T1	RHE_CH02325	Q	6	32597.36	4.92	26.18	1,5-8	Hypothetical conserved protein OS
Q2KA82	RHE_CH01451	O	2	34354.23	9.63	24.87	1,5-8	Probable peptidyl-prolyl cis-trans isomerase protein OS
Q2KB45	RHE_CH01134	WC	2	14050.82	9.51	20.23	1,5-8	Hypothetical conserved protein OS
Q2KBS8	RHE_CH00897	WC	17	#####	4.65	65.73	1,5-8	Hypothetical conserved protein OS
Q2KBU1	aroA	E	4	44376.85	5.14	44.13	1,5-8	3-phosphoshikimate 1-carboxyvinyltransferase OS
B3PXU8	yajC	U	3	11670.78	10.07	21.26	1,5-7	Preprotein translocase protein subunit YajC OS
Q2K5U5	RHE_CH03024	WC	5	55213.53	4.72	54.90	1,5-7	Hypothetical conserved protein OS
Q2K8E2	gyrA	L	6	#####	5.02	42.46	1,5-7	DNA gyrase subunit A protein OS
Q2KDD9	cysK1	E	9	33593.64	5.93	30.21	1,5-7	Cysteine synthase protein OS
Q2K3Z9	smoEch	G	3	47510.26	5.88	11.92	1,5,7	Sorbitol/mannitol ABC transporter, substrate-binding protein OS
Q2K8H4	fpr	C	2	29947.36	5.82	50.92	1,5,7	Ferredoxin-NADP+ reductase protein OS
B3PWH9	hisF	E	3	27645.44	4.93	12.55	1,5,6,8	Imidazole glycerol phosphate synthase subunit hisF OS
Q2K6Q3	znuA	P	6	37596.43	5.14	37.34	1,5,6,8	Zinc uptake ABC transporter, substrate-binding protein OS
Q2K7M3	fumC	C	5	49150.11	5.81	36.73	1,5,6,8	Fumarate hydratase class II protein OS
Q2K823	RHE_CH02233	WC	2	16048.61	9.13	19.81	1,5,6,8	Hypothetical conserved protein OS

Q2K9L6	rplC	J	3	22522.81	10.17	91.42	1,5,6,8	50S ribosomal protein L3 OS
Q2KAU2	groL2	O	15	57692.04	5.22	1349.19	1,5,6,8	60 kDa chaperonin 2 OS
Q2KDG7	guaA	F	2	57348.50	5.75	26.53	1,5,6,8	GMP synthase [glutamine-hydrolyzing] OS
B3PP95	secA	U	7	#####	5.16	67.57	1,5,6	Protein translocase subunit secA OS
B3PU13	msrB1	O	2	15375.07	5.26	27.85	1,5,6	Peptide-methionine (R)-S-oxide reductase protein OS
B3PYX1	ftsH	O	3	70123.08	5.80	18.12	1,5,6	Cell division metalloproteinase protein OS
Q2JYP1	etfBf	C	5	26613.93	7.75	56.98	1,5,6	Electron transport flavoprotein, beta subunit OS
Q2JYS2	RHE_PF00374	G	2	49158.10	4.92	24.92	1,5,6	Putative sugar ABC transporter, substrate-binding protein OS
Q2JYS8	RHE_PF00368	R	2	21748.91	6.19	11.88	1,5,6	Hypothetical conserved protein OS
Q2JZ89	thuAf	WC	2	29185.99	5.95	13.53	1,5,6	Trehalosemaltose utilization protein OS
Q2K4L5	tolR	U	2	15836.43	9.39	24.15	1,5,6	Protein transporter protein OS
Q2K5V5	csaA	R	2	12387.40	5.38	15.00	1,5,6	Methionine--tRNA ligase protein OS
Q2K6F1	RHE_CH02817	T	18	67938.22	5.87	59.17	1,5,6	Putative serine protein kinase protein OS
Q2K6L4	ugpBch1	G	2	40160.35	6.67	11.74	1,5,6	sn-glycerol-3-phosphate ABC transporter, substrate-binding protein OS
Q2K7M1	fumB	C	3	54919.86	6.26	17.36	1,5,6	Fumarate hydratase class I protein OS
Q2K9T1	nuoD1	C	2	44819.46	6.17	7.37	1,5,6	NADH-quinone oxidoreductase subunit D 1 OS
Q2KAP3	lppA	M	5	14446.41	9.78	48.18	1,5,6	Outer membrane lipoprotein-related protein OS
Q2KBN9	RHE_CH00936	WC	4	9374.72	4.47	17.84	1,5,6	Hypothetical conserved protein OS
Q2KE85	RHE_CH00018	WC	4	20198.04	10.28	18.47	1,5,6	Hypothetical conserved protein OS
B3Q350	groESc	O	4	11362.00	5.70	72.04	1,4-8	10 kDa chaperonin OS
Q2K4F3	thiEch	H	3	26408.16	6.09	29.38	1,4-8	Thiamine-phosphate pyrophosphorylase protein OS
Q2K589	pssP	D	4	84127.83	5.37	159.96	1,4-7	Exopolysaccharide polymerization protein OS
B3PW93	ilvD2	EG	2	62599.81	6.01	23.87	1,4-6,8	Dihydroxy-acid dehydratase protein OS
Q2K375	RHE_CH03966	R	2	33115.31	6.84	63.53	1,4-6,8	Probable oxidoreductase protein, aldo/keto reductase family OS
Q8KL91	RHE_PD00192	P	4	22908.29	5.55	62.08	1,4-6,8	Carbonic anhydrase OS
B3PQF6	dppAch1	E	23	58461.33	5.14	2860.51	1,4-6	Dipeptide ABC transporter, substrate-binding protein OS
B3PZ64	cysK1	E	8	33583.66	5.93	30.24	1,4-6	Cysteine synthase protein OS
Q2JYP6	RHE_PF00400	S	2	20190.17	7.03	16.71	1,4-6	Hypothetical conserved protein OS
Q2K0X8	thiM	H	5	26338.05	4.91	25.63	1,4-6	Hydroxyethylthiazole kinase OS
Q2K9M7	rplA	J	7	24122.09	9.76	101.45	1,4,6	50S ribosomal protein L1 OS
Q2KDZ2	rimP	S	4	22246.27	4.65	12.59	1,4,6	Ribosome maturation factor rimP OS
B3PNQ0	RHECIAT_CH0000622	M	5	21912.68	9.49	27.75	1,4,5,8	Hypothetical conserved protein OS
Q2KBF1	RHE_CH01025	WC	4	14388.39	9.40	32.69	1,4,5,6	Hypothetical conserved protein OS
B3PWU4	rpsK	J	3	13892.93	10.95	116.34	1,3-8	30S ribosomal protein S11 OS
B3Q3H5	RHECIAT_PC0000655	WC	3	41533.40	4.68	79.48	1,3-8	Hypothetical conserved protein OS
Q2JYL4	RHE_PF00432	WC	4	41481.34	4.52	113.64	1,3-8	Hypothetical conserved protein OS
Q2K2W2	bfr	P	7	19459.97	4.95	84.91	1,3-8	Bacterioferritin OS
Q2K2Y7	RHE_CH04055	C	6	54165.88	5.94	70.15	1,3-8	Probable aldehyde dehydrogenase protein OS
Q2K3C7	ppa	C	2	19870.97	5.63	145.84	1,3-8	Pyrophosphate phospho-hydrolase protein OS
Q2K3E5	sdhA	C	6	67051.73	5.89	67.56	1,3-8	Succinate dehydrogenase, flavoprotein subunit protein OS
Q2K4I5	tkt	G	3	69757.18	6.13	152.93	1,3-8	Transketolase protein OS

Q2K695	RHE_CH02873	R	6	60971.19	5.46	81.91	1,3-8	Probable ABC transporter, ATP-binding protein OS
Q2K7D4	RHE_CH02475	IQ	2	44565.30	5.38	39.70	1,3-8	Probable 3-oxoacyl-(Acyl-carrier-protein) synthase II protein OS
Q2K803	sufB	R	3	54482.81	5.59	56.84	1,3-8	ABC transporter involved in Fe-S cluster assembly, permease protein OS
Q2K8Y8	RHE_CH01912	WC	2	30565.16	5.07	63.91	1,3-8	Hypothetical conserved protein OS
Q2K9L0	rpsC	J	2	27625.74	9.96	143.13	1,3-8	30S ribosomal protein S3 OS
Q2K9L5	rplD	J	3	22406.94	10.10	50.61	1,3-8	50S ribosomal protein L4 OS
Q2KAI4	ropAch1	WC	4	36547.99	3.98	267.66	1,3-8	Porin outer membrane protein OS
Q2KCE0	flgKch	N	2	52144.76	4.50	121.42	1,3-8	Flagellar hook-associated protein OS
Q2KCE1	flgEch	N	4	44366.29	4.68	253.28	1,3-8	Flagellar hook protein OS
Q2KCV1	argD1	E	6	42387.53	5.98	77.84	1,3-8	Diaminobutyrate--pyruvate aminotransferase protein OS
Q2KDH1	RHE_CH00293	R	7	18823.34	8.52	165.45	1,3-8	Putative phosphatidylethanolamine-binding protein OS
Q2KE48	actR	TK	2	21340.12	6.86	19.56	1,3-8	Two-component response regulator protein OS
Q2K5C9	RHE_CH03192	G	2	34882.03	5.03	32.60	1,3-7	Probable sugar ABC transporter, substrate-binding protein OS
Q2K5W3	RHE_CH03006	WC	5	16877.86	8.90	95.10	1,3-7	Hypothetical conserved protein OS
Q2K926	accCch	I	2	48886.65	5.69	63.48	1,3-7	Acetyl-CoA carboxylase, biotin carboxylase protein OS
Q2K933	gatB	J	3	54993.50	4.97	87.26	1,3-7	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B OS
Q2K975	argS	J	3	64686.62	5.55	84.21	1,3-7	Arginyl-tRNA synthetase OS
Q2K9M4	rpoB	K	30	#####	5.03	445.16	1,3-7	DNA-directed RNA polymerase subunit beta OS
Q2K9W4	rpsI	J	5	16667.32	11.06	139.17	1,3-7	30S ribosomal protein S9 OS
Q2KDZ8	RHE_CH00113	R	9	30991.50	5.64	121.74	1,3-7	Putative hydrolase protein OS
B3PXH6	RHECIAT_CH0000186	WC	3	27225.07	9.51	63.93	1,3-6,8	Hypothetical conserved protein OS
Q2JZT8	katG	P	8	79870.91	5.78	259.29	1,3-6,8	Catalase-peroxidase OS
Q2K2S7	leuS	J	4	97999.95	5.36	69.09	1,3-6,8	Leucyl-tRNA synthetase OS
Q2K4Y0	RHE_CH03343	WC	8	88915.64	4.46	62.31	1,3-6,8	Hypothetical conserved protein OS
Q2KDW4	RHE_CH00147	WC	4	34589.50	9.83	72.31	1,3-6,8	Hypothetical conserved protein OS
Q2KE00	pnp	J	4	77482.57	5.22	229.93	1,3-6,8	Polyribonucleotide nucleotidyltransferase OS
B3PRD1	pmbA	R	2	47170.19	5.40	26.60	1,3-6	Modulator of DNA gyrase protein OS
Q2K077	RHE_00082	WC	6	19856.68	7.73	76.48	1,3-6	Putative uncharacterized protein OS
Q2K3F3	sucB	C	2	43235.69	5.87	112.46	1,3-6	Dihydrolipoamide succinyltransferase subunit of 2-oxoglutarate dehydrogenase complex prot
Q2K3I7	rocD	E	3	42952.92	5.56	33.06	1,3-6	Ornithine aminotransferase protein OS
Q2K603	carB	EF	6	#####	5.18	46.23	1,3-6	Carbamoyl-phosphate synthase large chain (Ammonia chain) protein OS
Q2K684	RHE_CH02884	C	2	44737.19	5.60	34.16	1,3-6	Probable alcohol dehydrogenase protein OS
Q2K7R0	RHE_CH02346	S	5	21518.65	9.41	46.38	1,3-6	Putative ABC transporter protein OS
Q2K806	sufD	R	2	45295.08	4.87	43.45	1,3-6	ABC transporter involved in Fe-S cluster assembly, permease protein OS
Q2K9L3	rplB	J	2	30364.81	10.83	55.38	1,3-6	50S ribosomal protein L2 OS
Q2KBZ2	RHE_CH00833	WC	2	22986.65	8.70	82.87	1,3-6	Hypothetical conserved protein OS
Q2KDR1	RHE_CH00201	WC	2	10830.22	6.04	31.87	1,3-6	Hypothetical conserved protein OS
Q2KDK1	rplT	J	2	15061.52	10.67	29.41	1,3-5-7	50S ribosomal protein L20 OS
B3PYJ4	RHECIAT_CH0000351	WC	3	10585.05	6.58	29.99	1,3-5,8	Hypothetical conserved protein OS
Q2KC97	rzcA	Q	4	#####	3.43	701.33	1,3-5,7,8	Rhizobiocin/RTX toxin and hemolysin-type calcium binding protein OS
Q2KE55	hslV	O	6	18777.65	6.75	90.69	1,3-5,7,8	ATP-dependent protease hslV OS

Q2K7H1	RHE_CH02438	K	6	18779.78	6.61	45.51	1,3-5	Probable transcriptional regulator protein, AsnC family OS
Q2KBV1	RHE_CH00874	R	4	66449.32	5.27	31.11	1,3-5	Hypothetical conserved protein OS
Q8KIL5	RHE_PD00087	G	4	42130.18	5.34	29.80	1,3-5	Ribose ABC transporter, substrate-binding protein OS
Q8KLF9	RHE_PD00276	S	2	15003.26	6.83	53.08	1,3-5	Hypothetical conserved protein OS
B3PR16	rpsP	J	5	13597.51	10.27	83.80	1,3,6,7	30S ribosomal protein S16 OS
Q2K2U0	cynT	P	4	23675.75	5.26	29.14	1,3,6	Carbonic anhydrase OS
Q2K7F4	RHE_CH02455	R	5	40579.48	5.61	37.07	1,3,6	Probable oxidoreductase protein OS
Q2K9S6	nuoG1	C	5	74216.68	5.60	27.20	1,3,6	NADH-quinone oxidoreductase OS
B3PT30	RHECIAT_CH0002824	WC	2	15999.10	5.43	50.82	1,3,5-8	Hypothetical conserved protein OS
Q2JYD8	fixL	T	13	68226.79	5.88	106.58	1,3,5-8	Sensor protein OS
Q2JZA9	RHE_PF00186	ET	5	29689.20	5.86	48.80	1,3,5-8	Probable amino acid ABC transporter, substrate-binding protein OS
Q2K308	galM	G	4	37651.15	5.61	40.03	1,3,5-8	Aldose-1-epimerase protein OS
Q2K456	RHE_CH03625	R	3	41008.05	5.42	41.99	1,3,5-8	Putative cobalamin synthesis protein, P47K family OS
Q2K6Q1	RHE_CH02714	E	3	13657.55	6.09	48.18	1,3,5-8	Putative glyoxalase/bleomycin resistance-like protein OS
Q2KDG0	RHE_CH00304	P	7	45315.63	4.57	139.41	1,3,5-8	Hypothetical conserved protein OS
Q2KDS5	RHE_CH00187	N	34	34821.36	5.76	2500.28	1,3,5-8	Probable outer membrane lipoprotein OS
Q2KDT6	RHE_CH00175	EP	13	68503.37	5.82	111.15	1,3,5-8	Probable peptide ABC transporter, substrate-binding protein OS
Q2KDZ3	nusA	K	8	59626.26	4.60	95.08	1,3,5-8	N-utilization substance protein A OS
Q2KDZ5	infB	J	7	99078.15	8.18	75.67	1,3,5-8	Translation initiation factor IF-2 OS
B3PUT8	RHECIAT_CH0001509	C	6	38491.62	6.20	34.20	1,3,5-7	Putative oxidoreductase protein OS
Q2JZD7	RHE_PF00158	R	12	23984.32	5.41	177.51	1,3,5-7	Hypothetical conserved protein OS
Q2K3F2	sucA	C	3	#####	6.28	177.99	1,3,5-7	Oxoglutarate dehydrogenase E1 subunit protein OS
Q2K3R7	secA	U	10	#####	5.18	116.20	1,3,5-7	Protein translocase subunit secA OS
Q2K488	glgP	G	3	91876.62	6.19	42.34	1,3,5-7	Phosphorylase OS
Q2K4R6	nodTch	MU	15	50288.88	7.78	152.42	1,3,5-7	Nodulation protein (Outer membrane efflux protein) OS
Q2K9U5	lon	O	3	88934.55	5.93	46.51	1,3,5-7	ATP-dependent protease La OS
Q2KD83	nifUch1	O	5	20389.07	4.75	51.52	1,3,5-7	Nitrogen fixation protein OS
Q2KE90	gyrB	L	5	88766.60	5.55	49.15	1,3,5-7	DNA gyrase subunit B OS
Q2K3X2	RHE_CH03709	G	10	44558.84	5.58	147.74	1,3,5,8	Probable sugar ABC transporter, substrate-binding protein OS
Q2K4D1	pepF	E	3	78339.60	6.01	48.37	1,3,5,8	Oligoendopeptidase F protein OS
Q2K450	RHE_CH03631	WC	2	23707.11	5.09	28.80	1,3,5,6	Hypothetical conserved protein OS
Q2KA07	nusB	K	4	17678.18	5.28	51.15	1,3,5,6	N utilization substance protein B homolog OS
Q2K5Q2	leuA2	E	19	63446.77	5.34	114.52	1,3,5	2-isopropylmalate synthase protein OS
Q2K6H0	hflK	O	3	40048.89	5.45	64.83	1,3,5	Hydrolase serine protease transmembrane subunit K protein OS
Q2KCN5	infA	J	5	8381.85	9.77	34.85	1,3,5	Translation initiation factor IF-1 OS
Q2KDJ9	pheT	R	3	86532.66	5.28	15.27	1,3,5	Phenylalanyl-tRNA synthetase, beta chain protein OS
Q2K446	eda	G	4	21703.22	5.30	35.35	1,3,4,8	4-hydroxy-2-oxoglutarate aldolase/2-dehydro-3-deoxyphosphogluconate aldolase protein OS
Q2K6B6	murE	M	17	53087.04	5.48	50.55	1,3,4,6,7	UDP-N-acetylmuramoylalanyl-D-glutamate--2, 6-diamin opimelate ligase protein OS
Q2K424	lysS	J	4	56738.64	5.43	25.04	1,3,4,6	Lysyl-tRNA synthetase OS
Q2K4K5	RHE_CH03475	S	3	26732.79	4.90	65.22	1,3,4,6	UPF0082 protein RHE_CH03475 OS
B3PQ18	rbsBch2	G	6	32224.06	6.84	23.71	1,3,4	Ribose ABC transporter, substrate-binding protein OS

B3PYR2	pdhA1	C	3	37997.93	6.81	170.46	1,3,4	Pyruvate dehydrogenase (Acetyl-transferring) protein, alpha subunit OS
Q2JZW0	RHE_00429	WC	3	15147.42	9.37	16.05	1,3,4	Hypothetical conserved protein OS
Q2K2Q3	cinRa	K	4	29402.20	5.54	25.63	1,3,4	Autoinducer transcriptional regulator protein OS
Q2K3M5	RHE_CH03813	S	2	15555.37	5.55	10.03	1,3,4	Hypothetical conserved protein OS
Q2K7T3	recA	L	7	38862.42	5.45	25.33	1,3,4	Protein recA OS
Q2K8D3	RHE_CH02119	WC	3	15823.03	6.72	9.31	1,3,4	Hypothetical conserved protein OS
Q2K984	pcm1	O	2	23261.87	5.34	19.47	1,3,4	Protein-L-isoaspartate O-methyltransferase protein OS
Q2K9N7	RHE_CH01650	WC	2	18135.49	4.59	42.29	1,3,4	Hypothetical conserved protein OS
Q2KA45	purM	F	4	36689.18	4.99	24.90	1,3,4	Phosphoribosylformylglycinamide cyclo-ligase OS
Q2KBV8	atpF	J	2	18033.75	6.86	15.93	1,3,4	ATP synthase subunit b OS
Q2KCS2	RHE_CH00546	S	2	17188.70	7.76	21.73	1,3,4	Hypothetical conserved protein OS
Q2KDA3	RHE_CH00361	F	2	22955.75	5.01	30.94	1,3,4	Probable nucleoside-triphosphatase HAM1-like protein OS
Q2JYX5	RHE_PF00321	G	2	46000.39	6.02	38.50	1,2-6	Putative sugar ABC transporter, substrate-binding protein OS
Q2K9X3	RHE_CH01561	M	2	20659.73	9.48	136.25	1,2-5-8	Probable polysaccharide export system protein OS
B3PS93	ribA	H	2	40055.45	5.80	11.32	1,2,8	GTP cyclohydrolase II protein (Riboflavin biosynthesis) OS
B3PPN1	flaCch3	N	9	31508.16	4.48	1522.97	1,2,5-8	Flagellin C protein OS
B3PWU6	rplQ	J	5	15374.83	10.80	101.13	1,2,5-8	50S ribosomal protein L17 OS
Q2K6Y0	RHE_CH02634	WC	2	54023.26	3.98	187.36	1,2,5-8	Putative RTX toxin hemolysin-type protein OS
Q2K881	ppiD1	O	7	69890.97	4.92	80.48	1,2,5-8	Peptidyl prolyl cis-trans isomerase D signal peptide protein OS
Q2K8Y5	pyrH	F	5	25168.00	6.21	59.68	1,2,5-8	Uridylate kinase OS
Q2K8Y7	rpsB	J	15	28246.23	6.46	173.63	1,2,5-8	30S ribosomal protein S2 OS
Q2K908	RHE_CH01892	Q	7	71750.76	4.50	620.73	1,2,5-8	Putative RTX toxin hemolysin-type calcium-binding protein OS
Q2K982	valS	J	6	#####	5.13	80.78	1,2,5-8	Valyl-tRNA synthetase (Valine-tRNA ligase) protein OS
Q2KA22	ribE	H	6	21837.61	5.14	42.46	1,2,5-8	Riboflavin synthase, alpha chain protein OS
Q2KC63	RHE_CH00759	WC	5	21549.75	5.23	46.93	1,2,5-8	Putative uncharacterized protein OS
Q2K9S9	nuoE1	C	2	41437.11	6.93	65.14	1,2,5-7	NADH-ubiquinone oxidoreductase chain E protein OS
Q2K766	RHE_CH02546	WC	4	47531.76	8.42	36.94	1,2,5,7	Hypothetical conserved protein OS
B3PUA2	RHECIAT_CH0003034	S	9	49816.79	4.91	137.37	1,2,5,6,8	Hypothetical conserved protein OS
Q2KAT2	rosR	WC	4	15652.84	7.83	27.94	1,2,5,6	Transcriptional regulator protein-nodulation competitiveness determinant OS
Q2JZP5	RHE_PF00047	G	8	46670.93	4.78	146.33	1,2,4-8	Probable sucrose ABC transporter, substrate-binding protein OS
Q2K1U0	RHE_PB00014	WC	3	17546.85	6.25	95.08	1,2,4-8	Hypothetical conserved protein OS
Q2K929	RHE_CH01866	WC	4	15977.40	7.78	64.66	1,2,4-8	Hypothetical conserved protein OS
Q2KCU5	metZ	E	10	42611.57	6.18	238.83	1,2,4-8	O-succinylhomoserine (Thiol)-lyase protein OS
Q2K3F1	sucD	C	3	31138.90	7.66	128.40	1,2,4-6	Succinyl-CoA ligase [ADP-forming] subunit alpha OS
B3PZA6	hsp	O	2	15632.93	6.61	42.21	1,2,4,6	Molecular chaperone small heat shock protein OS
Q2KB32	gstch6	O	2	22728.01	6.20	23.81	1,2,4,6	Glutathione S-transferase protein OS
B3Q0C4	RHECIAT_CH0000563	G	12	34226.57	6.65	281.40	1,2,4	Probable sugar ABC transporter, substrate-binding protein OS
Q2KA88	fabF1	IQ	3	44103.92	5.44	26.56	1, 5, 6	3-oxoacyl-(Acyl-carrier protein) synthase II protein OS
B3PR90	gshA	H	4	51179.06	5.57	5.79	8	Gamma-glutamylcysteine synthetase protein OS
B3PS36	purH	F	4	57111.41	6.11	6.83	8	Bifunctional purine biosynthesis protein purH OS
B3PS94	aroC	E	2	39216.34	6.67	1.89	8	Chorismate synthase OS

B3PTS0	flaCch5	N	5	35789.99	5.00	37.66	8	Flagellin C protein OS
B3PTU4	RHECIAT_CH0002982	S	2	17914.36	6.73	4.60	8	Hypothetical conserved protein, CreA family OS
B3PUY1	RHECIAT_CH0003158	G	2	41086.19	5.11	6.73	8	Putative glucose dehydrogenase B protein OS
B3PVV2	trpB	E	2	43564.65	5.57	14.67	8	Tryptophan synthase beta chain OS
B3PWH4	pckA	C	16	57525.16	5.63	101.21	8	Phosphoenolpyruvate carboxykinase [ATP] OS
B3PWR8	fusA	J	38	77793.73	5.16	479.31	8	Elongation factor G OS
B3PWU3	rpsM	J	3	13758.11	11.10	11.06	8	30S ribosomal protein S13 OS
B3PXH2	dnaJ	O	3	40790.86	8.28	3.32	8	Chaperone protein dnaJ OS
B3PZP6	glnA2	E	2	51933.14	5.38	25.56	8	Glutamine synthetase OS
Q2K0J8	glgXe	G	4	80292.83	5.09	6.07	8	Glycosyl hydrolase (Glycogen debranching) protein OS
Q2K694	RHE_CH02874	WC	4	25714.99	7.62	6.48	8	Probable enterobacter ribonuclease protein OS
Q2K7S1	folB	H	3	13504.40	5.10	9.52	8	Dihydroneopterin aldolase protein OS
Q2K8V0	RHE_CH01950	C	5	36760.60	6.19	4.90	8	Probable voltage-gated potassium channel beta subunit protein OS
Q2KDX5	RHE_CH00136	WC	4	19374.36	9.37	7.31	8	Hypothetical conserved protein OS
Q2K520	ureE	O	3	25773.80	6.53	7.01	7.8	Urease accessory protein OS
B3PQR8	RHECIAT_CH0002539	K	6	18793.81	6.62	26.40	7	Probable transcriptional regulator protein, AsnC family OS
B3PUG9	RHECIAT_CH0003102	WC	3	12684.67	11.31	9.95	7	Hypothetical conserved protein OS
B3PYX3	RHECIAT_CH0003699	S	6	34938.02	6.05	52.07	7	Hypothetical conserved protein OS
Q2K4A7	RHE_CH03573	G	2	29492.46	5.38	18.44	7	Probable monophosphatase protein OS
B3PX44	rkpK	M	14	47701.99	5.25	228.25	6.8	UDP-glucose 6-dehydrogenase protein OS
Q2JZ57	dppAf	E	3	57803.69	5.89	16.34	6.8	Dipeptide ABC transporter, substrate-binding protein OS
Q2K6B2	mraZ	S	3	23177.85	9.33	6.91	6.8	MraZ protein OS
Q2K917	thrA	E	6	46923.61	5.62	11.42	6.8	Homoserine dehydrogenase OS
Q2KE21	uxuB	G	8	52692.06	5.61	14.45	6.8	D-mannionate oxidoreductase protein OS
B3Q1H8	RHECIAT_PA0000190	G	3	46292.06	5.25	26.17	6.7	Putative sugar ABC transporter, substrate-binding protein OS
Q2K9K5	rplX	J	2	11209.10	10.51	48.73	6.7	50S ribosomal protein L24 OS
B3PQ87	sucA	C	11	#####	6.26	75.41	6	Oxoglutarate dehydrogenase E1 protein OS
B3PW58	rplL	J	2	12849.75	4.89	36.83	6	50S ribosomal protein L7/L12 OS
B3PWS1	rplC	J	3	22464.77	10.25	10.04	6	50S ribosomal protein L3 OS
B3PX11	RHECIAT_CH0001839	F	2	#####	5.74	3.86	6	Putative ribonucleotide-diphosphate reductase protein OS
B3PXK3	gpmA	G	11	23153.34	6.16	32.08	6	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase OS
B3PY12	RHECIAT_CH0003587	WC	2	89065.86	4.41	2.75	6	Hypothetical conserved protein OS
Q2JYZ9	RHE_PF00297	G	2	46705.98	5.48	1.76	6	Probable sugar ABC transporter, substrate-binding protein OS
Q2K0H7	RHE_00184	G	2	52046.60	5.03	5.23	6	Putative sugar ABC transporter, ATP-binding protein OS
Q2K3M9	tyrC	E	4	33597.26	5.71	3.41	6	Prephenate dehydrogenase protein OS
Q2K4C8	RHE_CH03552	WC	2	10467.02	6.83	16.05	6	Hypothetical conserved protein OS
Q2K7R7	RHE_CH02339	T	3	13164.22	7.74	5.66	6	Hypothetical conserved protein OS
Q2K808	RHE_CH02248	R	3	13797.65	4.03	8.83	6	Hypothetical conserved protein OS
Q2K851	RHE_CH02204	R	2	15695.21	9.10	4.90	6	Putative inosine-5'-monophosphate dehydrogenase protein OS
Q2KAF4	RHE_CH01379	R	4	32705.31	5.53	8.63	6	Putative hydrolase protein OS
Q2KAP5	feuP	TK	7	24914.42	5.63	9.67	6	Two-component response regulator protein OS

Q2KCR6	RHE_CH00552	H	8	20625.79	5.34	24.83	6	Putative pyrimidine reductase protein (Dihydrofolate reductase family) OS
Q2KDF2	RHE_CH00312	MG	5	34643.90	9.72	9.51	6	Probable NADH dehydrogenase/NADH dehydrogenase (Ubiquinone) protein OS
Q2KDG6	pfs	F	3	22774.08	4.60	4.11	6	Methylthioadenosine nucleosidase/adenosylhomocysteine nucleosidase protein OS
Q2KE56	hisB	E	4	21802.61	6.50	5.82	6	Imidazoleglycerol-phosphate dehydratase OS
B3PYC7	RHECIAT_CH0000284	K	4	25641.18	5.18	9.37	5.8	Probable transcriptional activator regulator protein, TenA family OS
Q2K460	RHE_CH03621	CHR	2	34217.26	5.61	8.01	5.8	Probable 2-hydroxyacid dehydrogenase protein OS
Q2K667	metH	E	3	#####	5.05	16.94	5.8	5-methyltetrahydrofolate--homocysteine S-methyltransferase protein OS
Q2KD55	rbsK	G	3	31914.60	4.90	30.63	5.7	Ribose kinase protein OS
B3PPX7	dme	C	4	83657.14	5.64	29.74	5.6	Malate dehydrogenase (Oxaloacetate decarboxylating) (NADP+) protein OS
B3PUZ0	RHECIAT_CH0003167	TK	6	26023.91	5.60	25.76	5.6	Probable two-component response regulator protein OS
B3PWW2	RHECIAT_CH0001790	I	3	37530.83	7.08	7.82	5.6	Putative esterase/lipase/thioesterase family protein OS
Q2K187	RHE_PC00042	EM	4	33103.02	5.44	15.79	5.6	Probable dihydrodipicolinate synthetase protein OS
Q2K2M2	RHE_PA00136	WC	2	16839.83	6.56	6.34	5.6	Hypothetical conserved protein OS
Q2K318	fdxA	C	2	12608.28	4.35	13.99	5.6	Ferredoxin III protein OS
Q2K3D2	RHE_CH03908	WC	5	18387.92	5.14	16.94	5.6	Hypothetical conserved protein OS
Q2K3G0	RHE_CH03880	WC	2	38707.03	5.55	10.71	5.6	Hypothetical conserved protein OS
Q2K504	RHE_CH03318	I	13	72275.33	6.38	37.15	5.6	Putative phosphatase protein OS
Q2K5A7	pssM	WC	4	33993.74	8.59	8.08	5.6	Ketal pyruvate transferase protein OS
Q2K5L6	glnII	E	6	38794.71	5.45	13.94	5.6	Glutamine synthetase OS
Q2K5W7	RHE_CH03002	G	3	40908.96	5.22	13.46	5.6	Putative glucose dehydrogenase B protein OS
Q2K6K7	ndhch	C	2	45609.69	9.16	11.58	5.6	NADH dehydrogenase protein OS
Q2K7M4	RHE_CH02382	S	2	68779.62	5.67	33.24	5.6	Hypothetical conserved protein OS
Q2K7Z4	bcp	O	2	16916.33	6.90	11.45	5.6	Bacterioferritin comigratory protein OS
Q2K8X9	lpxD	M	2	36608.92	6.51	10.35	5.6	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase OS
Q2KAY6	RHE_CH01195	H	2	29593.96	5.57	12.85	5.6	Probable NH(3)-dependent NAD(+)synthetase protein OS
Q2KBU6	sohB	OU	2	31004.16	9.94	4.12	5.6	Protease protein OS
Q2KC23	RHE_CH00802	WC	7	31224.64	5.10	13.95	5.6	Hypothetical conserved protein OS
Q2KD86	fur	P	2	16644.85	5.76	10.86	5.6	Ferric uptake regulation protein OS
Q2KDA2	RHE_CH00362	E	2	15063.04	5.01	10.61	5.6	Probable lactoylglutathione lyase protein OS
Q2KEA0	RHE_CH00002	D	4	21652.92	5.81	7.61	5.6	Maf-like protein RHE_CH00002 OS
B3PP51	purB	F	2	48550.55	6.00	15.59	5	Adenylosuccinate lyase protein OS
B3PP56	purC1	F	2	29038.67	5.93	2.38	5	Phosphoribosylaminoimidazolesuccinocarboxamide synthase protein OS
B3PPC5	RHECIAT_CH0004070	T	3	13088.07	4.80	2.72	5	Probable two-component response regulator protein OS
B3PTY2	lyc	M	2	30644.37	8.27	3.57	5	Glycoside hydrolase protein OS
B3PU93	RHECIAT_CH0001417	M	2	37808.16	9.02	3.76	5	Probable glycosyltransferase protein OS
B3PV66	proX	E	8	34015.68	4.99	63.06	5	Proline/glycine betaine ABC transporter, substrate-binding protein OS
B3PVE6	RHECIAT_CH0001609	WC	3	38798.28	9.84	2.30	5	Hypothetical conserved protein OS
B3PW00	nuoD1	C	2	44755.42	6.17	3.47	5	NADH-quinone oxidoreductase subunit D 1 OS
B3PW94	RHECIAT_CH0003383	C	6	50986.29	6.07	18.35	5	Probable aldehyde dehydrogenase protein OS
B3PWU2	adk	F	3	23634.13	5.72	13.61	5	Adenylate kinase OS
B3PXR2	RHECIAT_CH0001879	O	8	23751.05	6.43	44.35	5	Putative peroxidase protein OS

B3PZV7	gltD	ER	3	52876.01	5.95	20.64	5	Glutamate synthase (NADPH) protein, small subunit OS
B3Q174	RHECIAT_PA0000084	E	2	29910.05	4.86	2.18	5	Probable octopine ABC transporter, substrate-binding protein OS
B3Q181	RHECIAT_PA0000091	E	10	95203.91	5.50	19.58	5	Putative sarcosine dehydrogenase protein OS
Q2JYC8	hemNf	H	2	48800.45	5.59	2.43	5	Coproporphyrinogen III oxidase protein OS
Q2K0G7	RHE_00174	R	8	36234.33	5.87	10.17	5	Putative alpha/beta hydrolase protein OS
Q2K0X4	thiD	H	2	27945.02	5.42	19.21	5	Phosphomethylpyrimidine kinase protein OS
Q2K3A7	RHE_CH03934	H	4	11261.76	6.84	12.95	5	Putative pterin-4-alpha-carbinolamine dehydratase OS
Q2K487	glgB1	G	3	83272.02	5.91	3.03	5	1,4-alpha-glucan-branching enzyme 1 OS
Q2K4J2	RHE_CH03488	G	2	36468.35	5.76	1.74	5	Probable 6-phosphogluconate dehydrogenase protein OS
Q2K4J5	atvA	U	4	48925.36	4.49	7.11	5	Acid tolerance and virulence protein OS
Q2K5I1	allA	F	2	18452.86	5.37	6.34	5	Ureidoglycolate hydrolase OS
Q2K5S3	hemF	H	2	34265.44	5.48	4.67	5	Coproporphyrinogen-III oxidase, aerobic OS
Q2K844	RHE_CH02211	M	2	54839.54	9.95	4.41	5	Probable serine-type D-Ala-D-Ala carboxypeptidase protein OS
Q2K861	RHE_CH02194	C	2	21708.07	5.36	7.58	5	Putative nitroreductase protein OS
Q2K928	RHE_CH01867	WC	2	35202.38	7.81	1.84	5	Hypothetical conserved protein OS
Q2K9M5	rpL	J	3	12849.75	4.89	40.18	5	50S ribosomal protein L7/L12 OS
Q2KAI8	iolC	G	6	70331.48	5.38	4.66	5	Myo-inositol catabolism, sugar kinase protein OS
Q2KAM2	RHE_CH01309	S	2	22512.77	9.84	2.14	5	Hypothetical conserved protein OS
Q2KBA0	RHE_CH01078	S	3	17019.40	6.41	8.53	5	Hypothetical conserved protein OS
Q2KD16	RHE_CH00449	C	6	32050.54	6.84	9.60	5	Probable oxidoreductase protein OS
Q2KDN0	RHE_CH00232	L	4	39795.44	5.28	4.50	5	Hypothetical conserved protein OS
Q2KDS0	deoD	F	2	27990.38	5.70	1.83	5	Purine-nucleoside phosphorylase protein OS
Q2KDT9	glk	G	4	36426.78	6.08	11.29	5	Glucokinase OS
Q2KE60	hisA	E	5	26307.20	5.72	8.58	5	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxami
Q8KLI3	RHE_PD00310	WC	2	21800.24	8.33	2.70	5	Acyltransferase nodulation protein A OS
Q2K6B7	murF	M	10	49755.07	6.03	24.84	4.8	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase OS
Q2K871	pyrG	F	2	59982.41	5.80	4.91	4.8	CTP synthase OS
Q2K9G8	RHE_CH01723	G	10	44782.01	5.70	87.64	4.8	Probable sugar ABC transporter, substrate-binding protein OS
B3PPR0	aglA	G	3	62440.05	4.89	8.21	4.6	Alpha-glucosidase protein OS
Q2K1V9	RHE_PB00032	EP	7	59453.87	5.41	20.97	4.6	Putative oligopeptide ABC transporter, substrate-binding protein OS
Q2K2U8	RHE_CH04095	GEPR	2	66807.34	9.11	14.13	4.6	Putative transporter, permease protein OS
Q2K361	RHE_CH03980	WC	5	24345.98	7.83	18.16	4.6	Hypothetical conserved protein OS
Q2K857	RHE_CH02198	WC	2	16067.97	4.82	5.59	4.6	Hypothetical conserved protein OS
Q2KA35	RHE_CH01498	WC	2	41844.36	4.46	16.21	4.6	Hypothetical conserved protein OS
Q2KBT9	purD	F	2	44969.29	5.11	17.47	4.6	Phosphoribosylamine-glycine ligase protein OS
Q2KCK5	RHE_CH00615	WC	2	42204.26	6.03	12.92	4.6	Probable polysaccharidase protein OS
Q2KDF0	pyrF	F	2	24632.53	5.67	13.45	4.6	Orotidine 5'-phosphate decarboxylase OS
B3PY05	RHECIAT_CH0003580	WC	3	22875.28	5.13	9.35	4.5	Hypothetical conserved protein OS
Q2K1F3	adhA2	R	5	36311.84	5.55	25.12	4.5	Alcohol dehydrogenase protein OS
Q2K3D4	RHE_CH03906	N	7	41483.75	4.42	13.00	4.5	Hypothetical conserved protein OS
Q2K3T1	prfA	J	7	39991.12	5.20	13.75	4.5	Peptide chain release factor 1 OS

Q2K4Y9	ctrA	TK	3	25986.05	5.47	8.44	4.5	Two-component response regulator protein controls chromosomal replication initiation OS
Q2K5L1	RHE_CH03109	WC	3	24564.88	5.81	15.81	4.5	Hypothetical conserved protein OS
Q2K6U8	RHE_CH02667	T	3	14775.97	4.38	19.10	4.5	Putative two-component response regulator protein OS
Q2K777	arsC	P	2	15356.68	5.35	11.81	4.5	Arsenate reductase protein OS
Q2K869	tpiA	G	2	26952.66	5.29	13.70	4.5	Triosephosphate isomerase 1 OS
Q2KB00	RHE_CH01181	R	7	56422.29	5.65	10.33	4.5	Putative ATP-binding protein OS
B3PPC1	argH	E	3	51047.06	5.84	6.71	4	Argininosuccinate lyase OS
B3PPE3	rpmB	J	2	10959.77	11.37	28.62	4	50S ribosomal protein L28 OS
B3PQY3	RHECIAT_CH0004208	QR	2	31320.61	5.47	12.01	4	Probable oxidoreductase protein OS
B3PTP2	ilvH	E	2	20804.97	6.45	15.31	4	Acetolactate synthase III protein, small subunit OS
B3PTV3	ftsZ2	D	14	61317.99	5.15	68.84	4	Cell division protein ftsZ OS
B3PTZ3	RHECIAT_CH0003031	J	3	12143.88	5.73	10.06	4	Putative translation initiation inhibitor protein OS
B3PUU5	gmk	F	6	25219.49	6.01	10.83	4	Guanylate kinase protein OS
B3PVD7	RHECIAT_CH0001600	WC	2	19949.07	6.61	2.73	4	Hypothetical conserved protein OS
B3PW56	rplA	J	4	24149.12	9.76	23.61	4	50S ribosomal protein L1 OS
B3PW60	rpoC	K	33	#####	6.57	174.73	4	DNA-directed RNA polymerase subunit beta' OS
B3PW62	rpsL	J	2	13987.42	11.62	14.19	4	30S ribosomal protein S12 OS
B3PXD3	fabA	I	3	18937.85	6.90	12.90	4	3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase protein OS
B3PXD5	fabI1	I	2	28609.65	6.30	10.55	4	Enoyl-[acyl-carrier-protein] reductase (NADH) protein OS
B3PY86	RHECIAT_CH0003668	WC	6	34088.26	6.11	140.89	4	Hypothetical conserved protein OS
B3PYR7	lpdA1	C	2	51485.74	8.07	12.98	4	Dihydrolipoyl dehydrogenase OS
B3PYT4	RHECIAT_CH0002053	C	4	36748.68	7.06	4.39	4	Probable aldo/keto reductase protein (Voltage-gated potassium channel beta subunit) OS
B3PZT1	pykA	G	5	51414.25	6.49	56.61	4	Pyruvate kinase OS
B3Q0A7	frk	G	3	32572.30	5.12	6.68	4	Fructokinase protein OS
B3Q0J4	metS	J	2	57583.04	5.69	11.73	4	Methionyl-tRNA synthetase protein OS
B3Q126	RHECIAT_PA0000036	T	4	16991.79	6.30	18.41	4	Putative transcriptional regulator protein OS
B3Q290	RHECIAT_PB0000083	WC	4	#####	5.12	6.42	4	Hypothetical conserved protein OS
Q2JYC5	RHE_PF00522	O	2	20025.86	4.87	6.15	4	Hypothetical conserved protein OS
Q2JYC7	RHE_PF00520	WC	2	10595.84	5.54	3.29	4	Hypothetical conserved protein OS
Q2JYI2	fhuD	P	6	29997.65	5.59	15.85	4	Ferrichrome-iron ABC transporter, substrate-binding protein OS
Q2JYR2	putA	C	2	#####	6.37	3.05	4	Proline 1-pyrroline-5-carboxylate dehydrogenase/proline dehydrogenase protein OS
Q2JYR4	RHE_PF00382	G	4	33343.99	5.31	2.82	4	Putative epimerase protein OS
Q2JZ41	glnH	E	3	29386.31	4.85	8.95	4	Glutamine ABC transporter, substrate binding protein OS
Q2JZP2	cscA	G	5	61968.06	5.36	4.05	4	Beta-fructofuranosidase protein OS
Q2K095	RHE_00100	G	7	54100.86	5.64	12.39	4	Probable sugar ABC transporter, substrate-binding protein OS
Q2K0L4	RHE_00224	WC	5	17296.76	9.29	4.98	4	Hypothetical conserved protein OS
Q2K0L5	RHE_00223	S	4	20750.41	4.64	10.46	4	Hypothetical conserved protein OS
Q2K0P5	pT	E	4	44621.41	4.97	5.46	4	Aminopolypeptidase protein OS
Q2K0S2	RHE_00283	WC	2	26664.11	5.76	7.45	4	Hypothetical conserved protein OS
Q2K0X6	thiEe	H	3	21746.10	5.09	12.10	4	Thiamine-phosphate pyrophosphorylase OS
Q2K191	RHE_PC00045	H	2	25978.66	5.98	3.01	4	Hypothetical conserved protein OS

Q2K1I1	RHE_PC00135	EP	2	57342.19	6.77	7.57	4	Probable peptide ABC transporter, substrate-binding-protein OS
Q2K1T9	RHE_PB00010	G	6	30864.01	5.55	10.35	4	Putative aldose 1-epimerase protein OS
Q2K245	RHE_PB00122	G	4	36939.95	5.12	8.29	4	Probable sugar ABC transporter, substrate-binding protein OS
Q2K264	RHE_PB00139	G	7	49297.63	5.07	8.41	4	Putative sugar ABC transporter, substrate-binding protein OS
Q2K269	RHE_PB00141	WC	2	35797.04	9.59	10.95	4	Hypothetical conserved protein OS
Q2K2Q4	trbG	U	4	29132.34	9.51	11.61	4	Conjugal transfer protein G OS
Q2K2Q5	trbH	WC	3	14660.78	7.82	6.54	4	Conjugal transfer protein H OS
Q2K2Q8	trbF	WC	5	24776.21	9.49	22.52	4	Conjugal transfer protein F OS
Q2K2R5	traIa	TQ	3	22408.87	6.59	7.80	4	Autoinducer synthesis protein I OS
Q2K2S4	parB	K	3	32643.15	6.05	7.75	4	Chromosome partitioning protein B OS
Q2K2T6	purH	F	2	57150.37	6.01	27.23	4	Bifunctional purine biosynthesis protein purH OS
Q2K351	RHE_CH03991	E	3	37764.81	5.04	4.94	4	Putative membrane dipeptidase protein OS
Q2K388	ftsY	U	2	55632.12	4.54	3.13	4	Cell division protein OS
Q2K3K6	dacF2	M	2	38333.82	8.95	3.08	4	D-alanyl-D-alanine carboxypeptidase protein OS
Q2K3S6	RHE_CH03761	S	2	30173.26	5.78	2.96	4	Putative amidohydrolase protein OS
Q2K441	bglSch	G	2	51121.67	5.35	3.71	4	Beta-glucosidase protein OS
Q2K4C3	RHE_CH03557	O	2	37346.99	6.09	3.03	4	Putative membrane protease subunit protein OS
Q2K4D5	RHE_CH03545	S	6	11728.27	7.81	9.33	4	Hypothetical conserved protein OS
Q2K4E1	RHE_CH03539	R	2	15821.93	5.41	1.86	4	Hypothetical conserved protein OS
Q2K4G6	suhBch	G	7	28507.78	6.67	5.20	4	Myo-inositol-1(Or 4)-monophosphatase protein OS
Q2K4G7	RHE_CH03513	WC	2	39690.54	8.88	3.32	4	Hypothetical conserved protein OS
Q2K4W3	mexE2	M	6	42565.08	5.21	6.31	4	Multidrug-efflux system secretion protein, HlyD family OS
Q2K4Y7	RHE_CH03336	WC	8	22770.17	5.14	31.94	4	Hypothetical conserved protein OS
Q2K4Y8	RHE_CH03335	T	2	13228.47	7.73	8.23	4	Probable two-component response regulator protein OS
Q2K582	RHE_CH03240	R	4	21668.69	6.23	5.30	4	Probable esterase/lipase/thioesterase protein OS
Q2K5A0	pssF	M	2	32000.35	6.76	3.57	4	Glycosyltransferase protein OS
Q2K5J7	bdhA	IQR	2	27263.44	7.86	5.83	4	D-beta-hydroxybutyrate dehydrogenase protein OS
Q2K5P8	queF	R	2	17302.67	5.83	6.85	4	NADPH-dependent 7-cyano-7-deazaguanine reductase OS
Q2K5S6	RHE_CH03043	J	2	16798.26	6.41	2.97	4	Probable tRNA/rRNA methyltransferase protein OS
Q2K5T1	fbcF	C	4	20510.42	6.30	19.26	4	Ubiquinol-cytochrome c reductase iron-sulfur subunit OS
Q2K5T6	RHE_CH03033	I	5	18529.26	5.14	17.85	4	Probable enoyl-CoA hydratase protein OS
Q2K5V2	RHE_CH03017	S	6	9481.73	5.18	19.33	4	Hypothetical conserved protein OS
Q2K5V4	proC	E	6	27824.34	9.10	8.69	4	Pyrroline-5-carboxylate reductase protein OS
Q2K634	RHE_CH02935	S	3	26161.81	5.23	4.17	4	Hypothetical conserved membrane protein OS
Q2K680	RHE_CH02888	WC	6	25528.86	5.34	10.27	4	Putative N-formylglutamate amidohydrolase protein OS
Q2K6C3	murB	M	8	35317.49	6.27	30.53	4	UDP-N-acetylenolpyruvoylglucosamine reductase OS
Q2K6H2	degPch4	O	2	58830.14	4.70	25.06	4	Serine protease DO-like protein OS
Q2K6K8	adh	R	13	37481.84	5.57	20.71	4	Alcohol dehydrogenase (NADP+) protein OS
Q2K6P1	afuA3	H	3	34355.36	5.30	8.75	4	Iron(III) ABC transporter, substrate-binding protein OS
Q2K6Q4	znuC	P	3	33857.06	6.46	2.24	4	Zinc import ATP-binding protein znuC OS
Q2K6R8	RHE_CH02697	K	3	21020.17	6.51	2.99	4	Probable transcriptional regulator protein, TetR family OS

Q2K6S7	ribH2	H	2	15689.97	6.36	11.35	4	Riboflavin synthase, beta chain protein OS
Q2K6W3	RHE_CH02652	WC	5	20452.38	5.10	14.94	4	Hypothetical conserved protein OS
Q2K6W5	RHE_CH02650	T	5	12666.72	6.97	6.11	4	Hypothetical conserved protein OS
Q2K759	RHE_CH02553	MG	5	31664.70	6.54	8.68	4	Hypothetical conserved protein OS
Q2K779	RHE_CH02532	T	5	18944.70	8.48	10.03	4	Putative uncharacterized protein OS
Q2K7B6	RHE_CH02493	T	4	43852.03	5.45	2.70	4	Probable adenylate cyclase protein OS
Q2K7B8	cobP	H	2	18530.45	6.20	12.96	4	Bifunctional cobinamide kinase/cobinamide phosphoguanlyltransferase OS
Q2K7B9	cobW	R	4	37941.94	4.69	4.72	4	Cobalamin biosynthesis protein OS
Q2K7C4	iumH2	F	2	32496.26	5.18	2.06	4	Inosine-uridine preferring nucleoside hydrolase protein OS
Q2K7E5	ddlA	M	9	38040.50	5.79	8.66	4	D-alanine-D-alanine ligase protein OS
Q2K7G6	cobT	H	3	34848.12	5.69	6.40	4	Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase OS
Q2K7J5	RHE_CH02413	TK	4	27925.24	6.24	3.20	4	Probable two-component response regulator protein OS
Q2K7N6	RHE_CH02370	N	5	30612.92	4.70	4.77	4	Putative periplasmic component of the Tol biopolymer transport system OS
Q2K7S2	folK	H	2	19030.71	5.07	4.87	4	7, 8-dihydro-6-hydroxymethylpterin-pyrophosphokinase protein OS
Q2K7U8	RHE_CH02308	G	6	29735.96	5.04	8.31	4	Putative inositol monophosphatase protein OS
Q2K7X0	purL	F	5	79485.27	4.88	55.94	4	Phosphoribosylformylglycinamide synthase 2 OS
Q2K7X1	RHE_CH02285	WC	2	20645.02	4.91	13.69	4	Hypothetical conserved protein OS
Q2K7Z0	RHE_CH02266	LR	3	13411.06	4.60	5.62	4	Putative hydrolase protein, MutT/nudix family OS
Q2K822	RHE_CH02234	WC	2	29030.97	5.07	2.04	4	Hypothetical conserved protein OS
Q2K855	nifUch2	C	5	15983.30	6.41	9.12	4	Nitrogen fixation protein OS
Q2K896	metS	J	12	57607.02	5.61	24.76	4	Methionyl-tRNA synthetase protein OS
Q2K8F7	mfd	LK	3	#####	6.14	4.35	4	Transcription-repair coupling factor (TRCF) protein OS
Q2K8W7	pdhA1	C	6	37981.93	6.81	47.10	4	Pyruvate dehydrogenase alpha subunit protein OS
Q2K911	RHE_CH01888	WC	4	11693.32	9.79	6.02	4	Probable NADH-ubiquinone oxidoreductase protein OS
Q2K945	map1	J	4	30434.61	6.14	3.31	4	Methionine aminopeptidase OS
Q2K966	surE	R	2	27960.81	5.13	3.58	4	5'-nucleotidase surE OS
Q2K974	RHE_CH01819	WC	3	#####	4.28	8.57	4	Hypothetical conserved protein OS
Q2K977	RHE_CH01816	S	2	11346.70	4.35	6.82	4	Hypothetical conserved protein OS
Q2K981	RHE_CH01812	I	5	44349.40	4.77	7.23	4	Probable outer membrane protein OS
Q2K9A4	rbsBch2	G	2	39973.39	6.32	4.40	4	Ribose ABC transporter, substrate-binding protein OS
Q2K9D7	gpt	F	2	18115.88	5.89	12.83	4	Xanthine phosphoribosyltransferase OS
Q2K9E2	RHE_CH01749	WC	2	18164.35	5.42	2.21	4	Hypothetical conserved protein OS
Q2K9F4	RHE_CH01737	WC	2	29580.82	7.78	7.39	4	Hypothetical conserved protein OS
Q2K9J6	secY	N	2	47625.26	9.91	3.06	4	Protein-export translocase protein OS
Q2K9K9	rplP	J	2	15527.26	10.87	27.70	4	50S ribosomal protein L16 OS
Q2K9M1	rpsL	J	3	14015.43	11.70	20.48	4	30S ribosomal protein S12 OS
Q2K9P6	RHE_CH01641	F	2	45481.96	5.47	3.41	4	Probable dihydroorotase protein OS
Q2K9V1	RHE_CH01583	R	2	12862.50	5.16	5.58	4	Hypothetical conserved protein OS
Q2KA04	RHE_CH01529	WC	2	16285.71	4.76	5.40	4	Hypothetical conserved protein OS
Q2KA09	RHE_CH01524	N	2	12947.31	8.38	3.65	4	Hypothetical conserved protein OS
Q2KA18	RHE_CH01515	M	2	21322.06	5.63	4.75	4	Probable dTDP-rhamnose-3,5-epimerase protein OS

Q2KA29	hemB	H	4	36676.99	5.19	5.58	4	Delta-aminolevulinic acid dehydratase OS
Q2KA57	moaE	H	5	16505.86	5.69	14.55	4	Molybdopterin converting factor subunit 2 protein OS
Q2KA75	gal	R	3	33714.43	6.05	5.76	4	D-galactose-1-dehydrogenase protein OS
Q2KAB3	pipch1	R	2	36293.13	5.96	5.68	4	Proline iminopeptidase OS
Q2KAD7	RHE_CH01396	R	3	31543.90	6.23	7.37	4	Hypothetical conserved protein OS
Q2KAD8	RHE_CH01395	C	7	35835.73	5.52	7.16	4	Putative oxidoreductase protein OS
Q2KAI9	iolD	E	5	65542.01	5.92	5.72	4	Myo-inositol catabolism, acetolactate synthase protein OS
Q2KAK0	hisCch1	E	4	39371.07	6.11	5.53	4	Histidinol-phosphate aminotransferase protein OS
Q2KAU8	RHE_CH01233	T	6	14968.43	7.90	13.91	4	Probable universal stress protein OS
Q2KAV4	RHE_CH01227	WC	4	14611.65	8.50	20.40	4	Hypothetical conserved protein OS
Q2KAW5	RHE_CH01216	O	4	23602.82	5.69	3.74	4	Probable glutathione S-transferase protein OS
Q2KB03	RHE_CH01178	R	3	28128.14	4.83	12.50	4	Putative glyoxalase protein OS
Q2KB15	RHE_CH01166	WC	2	29138.59	5.60	1.79	4	Hypothetical conserved protein OS
Q2KB47	cysH	EH	2	26800.37	5.16	7.19	4	Phosphoadenosine-phosphosulfate reductase protein OS
Q2KB49	cysN	P	3	53730.02	6.07	7.90	4	Sulfate adenylate transferase subunit 1 protein (Cysteine biosynthesis protein) OS
Q2KB71	RHE_CH01108	QR	2	33688.24	5.61	3.20	4	Putative uncharacterized protein OS
Q2KBB4	RHE_CH01064	P	2	25050.70	5.55	6.28	4	Carbonic anhydrase OS
Q2KBI9	RHE_CH00986	R	5	24482.33	5.39	9.30	4	Probable hydrolase phosphatase protein OS
Q2KBK3	RHE_CH00972	P	2	23889.32	4.90	4.63	4	Probable molybdate ABC transporter, substrate-binding protein OS
Q2KBN4	pdxH	H	4	23683.83	9.21	4.20	4	Pyridoxine/pyridoxamine 5'-phosphate oxidase OS
Q2KBN5	RHE_CH00940	WC	3	16201.30	8.52	10.41	4	Hypothetical conserved protein OS
Q2KBT4	RHE_CH00891	R	2	83151.81	6.89	2.60	4	Probable transcriptional accessory protein (RNA binding S1 domain) OS
Q2KBV3	moaB	H	6	19987.75	6.32	11.56	4	Molybdenum cofactor biosynthesis protein OS
Q2KBX1	dgoK	WC	5	31258.52	5.51	6.82	4	2-dehydro-3-deoxygalactonokinase protein OS
Q2KBY0	RHE_CH00845	R	3	25234.58	4.69	3.13	4	Probable phosphoglycolate phosphatase protein OS
Q2KBY9	RHE_CH00836	L	4	20362.20	6.97	5.99	4	Putative methylase protein OS
Q2KC04	hisS	J	3	60611.77	5.52	9.27	4	Histidyl-tRNA synthetase protein OS
Q2KC12	RHE_CH00813	S	2	25074.01	10.11	3.00	4	Hypothetical conserved protein OS
Q2KC34	RHE_CH00791	E	2	13747.27	6.49	1.64	4	Putative lactoylglutathione lyase protein OS
Q2KC59	nolK	MG	2	34183.98	6.00	5.17	4	GDP-L-fucose synthase protein OS
Q2KC62	RHE_CH00760	M	3	33131.48	6.59	3.49	4	Probable nucleoside-diphosphate-sugar epimerase protein OS
Q2KCG5	flgF	N	6	26679.82	4.90	16.89	4	Flagellar basal-body rod protein OS
Q2KCJ5	RHE_CH00625	T	3	15912.20	4.63	7.96	4	Probable two-component response regulator protein OS
Q2KCL1	appA	E	6	57030.20	8.47	6.98	4	Oligopeptide ABC transporter, substrate-binding protein OS
Q2KCN9	RHE_CH00580	WC	6	15717.06	8.89	8.77	4	Hypothetical conserved protein OS
Q2KCP7	RHE_CH00571	I	4	37579.87	5.54	3.35	4	Probable enoyl-CoA hydratase protein OS
Q2KCS6	RHE_CH00542	S	2	17867.36	9.88	4.57	4	Hypothetical conserved protein OS
Q2KCS9	RHE_CH00539	S	3	12713.56	5.34	3.47	4	Hypothetical conserved protein OS
Q2KD11	RHE_CH00454	WC	5	13151.77	5.20	6.99	4	Hypothetical conserved protein OS
Q2KD43	RHE_CH00421	R	2	26796.49	5.32	3.87	4	Putative hydrolase protein OS
Q2KD90	RHE_CH00374	R	3	19226.83	4.51	5.89	4	Putative metalloprotease RHE_CH00374 OS

Q2KDL5	lepA	M	2	67338.53	5.57	4.75	4	GTP-binding protein lepA OS
Q2KE82	trpB	E	8	43537.67	5.74	23.12	4	Tryptophan synthase beta chain OS
Q2KE89	RHE_CH00013	C	2	22263.02	6.53	10.59	4	Probable oxidoreductase protein OS
B3PYX5	tolB	U	27	47984.26	8.50	622.98	3.8	Protein tolB OS
Q2K579	noeK	G	9	49981.87	5.29	23.09	3.8	Phosphomannomutase protein OS
Q2K5J9	RHE_CH03121	WC	3	13587.49	4.87	10.57	3.8	Hypothetical conserved protein OS
Q2KA36	aspS	J	5	67705.13	5.47	44.70	3.7	Aspartyl-tRNA synthetase protein OS
B3PVA2	aspS	J	3	67598.90	5.40	5.73	3.6	Aspartyl-tRNA synthetase protein OS
Q2JZJ9	dapAf1	EM	2	31341.91	5.59	13.97	3.6	Dihydrodipicolinate synthase protein OS
Q2K2Q0	traR	K	6	26866.51	9.48	24.21	3.6	Conjugal transfer transcriptional regulator protein, LuxR family OS
Q2K7Y0	RHE_CH02276	R	4	20248.82	6.65	9.80	3.6	Hypothetical conserved protein OS
Q2K829	adhCch	C	4	39860.86	6.31	13.40	3.6	Formaldehyde dehydrogenase (Glutathione)/alcohol dehydrogenase protein OS
Q2K8F0	glnA2	E	3	51973.18	5.38	92.82	3.6	Glutamine synthetase OS
B3PYK5	gloA	E	2	16573.73	5.38	7.47	3.5	Lactoylglutathione lyase methylglyoxalase protein OS
Q2K4S0	purC2	F	2	35735.57	5.34	27.70	3.5	Phosphoribosylaminoimidazole-succinocarboxamide synthase OS
Q2K689	RHE_CH02879	J	2	12129.85	5.73	49.24	3.5	Putative translation initiation inhibitor protein OS
Q2K6C8	ftsZ2	D	15	58141.20	5.08	86.53	3.5	Cell division protein ftsZ OS
Q2K973	RHE_CH01820	G	2	38183.69	5.02	8.65	3.5	Probable beta-N-acetylhexosaminidase protein OS
Q2KBT5	glyS	J	2	77046.01	5.72	41.66	3.5	Glycyl-tRNA synthetase beta subunit OS
B3PPJ7	cheY1	J	2	13469.44	5.25	13.16	3.4	Chemotaxis two-component response regulator protein OS
B3PQZ6	RHECIAT_CH0004221	R	5	24850.27	5.59	7.42	3.4	Putative ABC transporter, ATP-binding protein (Involved in lysophospholipase L1 biosynthesis)
B3PR36	lonD	H	6	24624.33	4.75	10.46	3.4	ATP-dependent protease La protein OS
B3PR87	RHECIAT_CH0004313	K	3	21310.19	6.78	23.42	3.4	Probable transcriptional regulator protein, CarD family OS
B3PSD9	RHECIAT_CH0001076	WC	3	22394.73	9.48	12.43	3.4	Hypothetical conserved protein OS
B3PT79	ilvE2	EH	6	39650.31	5.76	13.16	3.4	Branched-chain-amino-acid aminotransferase OS
B3PUT9	fabD	I	2	31989.86	5.12	18.95	3.4	[acyl-carrier-protein] S-malonyltransferase protein OS
B3PWG9	RHECIAT_CH0000034	G	3	14162.28	5.68	10.00	3.4	Probable protein-N(Pi)-phosphohistidine-sugar phosphotransferase protein (Enzyme II of the
B3PW17	hslU	O	2	48074.79	5.56	28.62	3.4	ATP-dependent hsl protease ATP-binding subunit hslU OS
B3PWT5	rpsH	J	4	14358.51	10.00	88.52	3.4	30S ribosomal protein S8 OS
B3PY14	RHECIAT_CH0000341	P	4	45334.71	4.70	42.49	3.4	Hypothetical conserved protein OS
B3PZM8	RHECIAT_CH0002135	G	5	34292.34	6.38	16.95	3.4	Probable sugar ABC transporter, substrate-binding protein OS
Q2JYB3	RHE_PF00534	C	8	53596.06	6.00	20.10	3.4	Probable aldehyde dehydrogenase protein OS
Q2JYL0	dadA	E	4	45695.67	8.87	9.30	3.4	D-amino-acid dehydrogenase, small subunit protein OS
Q2JZY3	RHE_00448	WC	2	28878.87	5.72	16.99	3.4	Putative sugar epimerase protein OS
Q2K0F4	RHE_00161	WC	2	17339.52	4.82	5.90	3.4	Hypothetical conserved protein OS
Q2K101	RHE_00362	Q	5	45302.70	6.31	10.58	3.4	Hypothetical conserved protein OS
Q2K1F7	RHE_PC00112	R	2	17417.88	5.80	4.36	3.4	Hypothetical conserved protein OS
Q2K202	thiEb	H	7	22488.04	5.79	21.73	3.4	Thiamine-phosphate pyrophosphorylase OS
Q2K2P3	traCa	J	3	10120.21	8.01	10.76	3.4	Conjugal transfer protein C OS
Q2K2V2	leuD	E	4	22183.27	5.13	34.87	3.4	3-isopropylmalate dehydratase small subunit OS
Q2K2V3	RHE_CH04090	S	2	15453.60	5.45	8.49	3.4	Hypothetical conserved protein OS

Q2K306	glmS2	M	5	35328.29	5.73	6.73	3.4	Glucosamine--fructose-6-phosphate aminotransferase protein OS
Q2K3B7	argG	E	5	45212.73	5.93	40.38	3.4	Argininosuccinate synthase OS
Q2K3D6	hemC	H	5	33406.96	6.01	7.57	3.4	Porphobilinogen deaminase protein OS
Q2K3D9	RHE_CH03901	S	4	10546.18	5.61	9.83	3.4	Hypothetical conserved protein OS
Q2K3G8	atpA	C	21	54698.66	6.51	211.33	3.4	ATP synthase subunit alpha OS
Q2K3H1	atpC	C	2	15049.24	5.30	20.99	3.4	ATP synthase epsilon chain OS
Q2K3M2	rpmB	J	3	10860.64	11.37	55.22	3.4	50S ribosomal protein L28 OS
Q2K3M6	gloB	R	10	28559.47	5.51	24.10	3.4	Hydroxyacylglutathione hydrolase OS
Q2K3M8	hisCch2	E	9	39034.79	6.34	25.02	3.4	Histidinol-phosphate aminotransferase protein OS
Q2K3N9	RHE_CH03799	T	6	13102.09	4.80	12.30	3.4	Probable two-component response regulator protein OS
Q2K3X6	lacZ2	G	2	83712.56	5.49	18.52	3.4	Beta-D-galactosidase protein OS
Q2K3Z3	RHE_CH03688	Q	2	40167.34	5.27	5.06	3.4	Putative multidrug efflux transporter protein OS
Q2K425	gltX	J	4	54726.14	6.16	21.24	3.4	Glutamyl-tRNA synthetase OS
Q2K4H0	RHE_CH03510	WC	2	12430.12	4.96	17.25	3.4	Hypothetical conserved protein OS
Q2K4L4	tolQ	U	2	25733.85	9.43	7.92	3.4	Protein transporter protein OS
Q2K4L6	tolA	M	2	40598.64	5.14	16.90	3.4	Probable signal peptide protein OS
Q2K4Z1	RHE_CH03332	R	2	13539.37	7.87	3.54	3.4	Putative acetyltransferase protein OS
Q2K523	RHE_CH03299	Q	3	#####	5.32	21.00	3.4	Probable acriflavin resistance protein (Transporter protein) OS
Q2K5F7	gguB	G	2	42521.56	9.79	8.13	3.4	Xylose ABC transporter, permease protein OS
Q2K5F9	RHE_CH03161	C	8	51112.41	6.30	39.26	3.4	Probable aldehyde dehydrogenase protein OS
Q2K5P4	tr	EH	5	80647.70	5.81	40.77	3.4	Anthranilate synthase includes: glutamine amidotransferase protein OS
Q2K5W2	RHE_CH03007	QR	4	30498.46	5.41	6.46	3.4	Probable short-chain dehydrogenase protein OS
Q2K6B0	RHE_CH02858	WC	2	41254.61	5.75	6.21	3.4	Probable sarcosine oxidase protein OS
Q2K6J8	pdxJ	H	8	26997.76	5.29	12.62	3.4	Pyridoxal phosphate biosynthetic (Pyridoxine 5'-phosphate synthase) protein OS
Q2K6Y8	ftsZ1	D	3	34830.67	4.99	20.00	3.4	Cell division protein ftsZ OS
Q2K7D1	acpXL	IQ	2	13968.97	5.07	17.41	3.4	Acyl carrier protein OS
Q2K7J8	RHE_CH02410	O	2	22778.26	6.86	6.22	3.4	Probable maleylpyruvate isomerase protein OS
Q2K7M6	rpiA	G	5	24138.72	5.03	26.03	3.4	Ribose-5-phosphate isomerase A OS
Q2K7N1	aroF	E	3	50909.59	6.13	8.52	3.4	Aldolase DAHP synthetase protein OS
Q2K7Q1	dme	C	7	83660.10	5.57	30.89	3.4	Malate dehydrogenase (Oxaloacetate decarboxylating) (NADP+) protein OS
Q2K7R6	RHE_CH02340	N	2	#####	4.89	48.36	3.4	Hypothetical conserved protein OS
Q2K7S3	RHE_CH02333	WC	2	18985.88	9.20	10.15	3.4	Hypothetical conserved protein OS
Q2K809	RHE_CH02247	S	3	13695.54	5.04	15.65	3.4	Hypothetical conserved protein OS
Q2K880	trpD	E	8	34921.04	5.55	12.86	3.4	Anthranilate phosphoribosyltransferase OS
Q2K8Q2	RHE_CH01999	WC	4	18519.45	6.60	24.34	3.4	Putative uncharacterized protein OS
Q2K8Z0	RHE_CH01910	C	3	26602.00	4.68	18.11	3.4	Probable glycerophosphoryl diester phosphodiesterase protein OS
Q2K8Z2	RHE_CH01908	FGR	4	15371.62	6.34	33.74	3.4	Putative hydrolase protein, HIT family OS
Q2K8Z8	RHE_CH01902	R	4	27830.78	5.43	7.38	3.4	Probable hydrolase protein OS
Q2K912	RHE_CH01886	S	2	14569.81	9.05	13.64	3.4	Hypothetical conserved protein OS
Q2K997	ttuD	G	2	43851.86	5.53	12.06	3.4	Hydroxypyruvate reductase protein OS
Q2K9B8	sseA	P	10	31230.15	5.44	25.05	3.4	Thiosulfate sulfurtransferase protein OS

Q2K9D1	RHE_CH01760	WC	2	11755.14	5.77	10.70	3.4	Hypothetical conserved protein OS
Q2K9D6	RHE_CH01755	T	2	29778.95	5.21	24.38	3.4	Putative universal stress protein OS
Q2K9J9	rpsE	J	2	20694.63	10.76	62.89	3.4	30S ribosomal protein S5 OS
Q2K9P0	mrcA	M	2	88351.48	6.05	7.82	3.4	Penicillin binding peptidoglycan synthetase protein OS
Q2K9U0	RHE_CH01594	S	4	16916.07	7.11	9.38	3.4	Hypothetical conserved protein OS
Q2K9U6	clpX	O	10	46936.87	5.45	21.03	3.4	ATP-dependent Clp protease ATP-binding subunit clpX OS
Q2K9Z8	RHE_CH01535	T	3	12437.27	4.47	15.16	3.4	Putative two-component response regulator protein OS
Q2KA77	pA	E	2	52154.86	6.58	7.06	3.4	Probable cytosol aminopeptidase OS
Q2KA81	RHE_CH01452	M	3	85990.71	4.98	24.34	3.4	Putative organic solvent tolerance protein OstA OS
Q2KA89	acpP	IQ	2	8368.37	4.00	7.71	3.4	Acyl carrier protein OS
Q2KAN3	pN	E	3	97603.93	5.29	43.51	3.4	Aminopeptidase N protein OS
Q2KAQ1	RHE_CH01280	S	3	18221.50	7.90	8.72	3.4	Hypothetical conserved protein OS
Q2KAQ8	RHE_CH01273	O	6	32041.47	4.57	9.23	3.4	Probable serine protease protein OS
Q2KBA3	RHE_CH01075	S	5	17654.51	5.31	10.69	3.4	Hypothetical conserved protein OS
Q2KBL3	thrB	R	4	35842.26	5.29	10.88	3.4	Homoserine kinase OS
Q2KBW6	RHE_CH00859	D	10	41145.70	6.32	45.93	3.4	Probable ATP-binding protein (Mrp protein-like) OS
Q2KC86	dppAchl	EP	21	58461.33	5.14	1483.09	3.4	Dipeptide ABC transporter, substrate-binding protein OS
Q2KC91	iolA	C	2	53528.58	6.06	12.56	3.4	Methylmalonate-semialdehyde dehydrogenase (Acylating) protein OS
Q2KC99	RHE_CH00721	N	5	16001.25	5.95	19.40	3.4	Hypothetical conserved protein OS
Q2KCI9	secD1	U	2	90727.37	8.72	12.41	3.4	Protein-export translocase protein OS
Q2KCN4	RHE_CH00585	D	3	22639.25	9.71	11.86	3.4	Maf-like protein RHE_CH00585 OS
Q2KCQ0	aspB1	E	3	42126.13	5.22	6.82	3.4	Aspartate transaminase protein OS
Q2KCX3	RHE_CH00492	G	11	34255.61	7.74	204.14	3.4	Probable sugar ABC transporter, substrate-binding protein OS
Q2KCY6	RHE_CH00479	WC	3	14814.30	4.59	11.15	3.4	Hypothetical conserved protein OS
Q2KD40	RHE_CH00424	G	3	26579.50	5.70	7.39	3.4	Putative carboxyphosphoenolpyruvate phosphonmutase protein OS
Q2KD47	da	E	2	42212.70	5.31	39.49	3.4	Succinyl-diaminopimelate desuccinylase OS
Q2KD57	ptsNch	GT	2	16666.16	5.78	36.97	3.4	Protein-N(Pi)-phosphohistidine-sugar phosphotransferase protein (Enzyme II of the phosphotransferase system) OS
Q2KDJ8	RHE_CH00265	WC	4	13318.16	5.23	13.89	3.4	Hypothetical conserved protein OS
Q2KE54	hslU	O	5	48040.79	5.55	35.28	3.4	ATP-dependent hsl protease ATP-binding subunit hslU OS
Q2KE67	chvI	TK	2	28046.12	5.77	32.14	3.4	Two-component response regulator protein OS
Q2KE81	trpA	E	6	29503.08	5.76	22.20	3.4	Tryptophan synthase alpha chain OS
Q2KE94	RHE_CH00008	WC	2	25821.92	4.93	10.25	3.4	Putative translocase transmembrane protein OS
B3PP45	RHECIAT_CH0002377	WC	2	11838.58	4.96	5.61	3	Putative regulator of competence-specific genes protein, TfoX family OS
B3PS29	RHECIAT_CH0004392	M	7	28989.67	9.62	24.20	3	Putative membrane-bound lytic murein transglycosylase protein OS
B3PTA3	adh	R	2	37174.45	5.48	3.36	3	Alcohol dehydrogenase (NADP+) protein OS
B3PTS8	RHECIAT_CH0002966	T	9	67877.09	5.82	18.12	3	Putative serine protein kinase protein OS
B3PTV0	comL	R	2	31918.14	5.68	2.63	3	Competence lipoprotein protein OS
B3PUD5	RHECIAT_CH0003067	WC	4	14218.25	5.10	32.10	3	Hypothetical conserved protein OS
B3PUM0	iolE	G	4	33516.03	5.26	12.17	3	Myo-inositol catabolism protein OS
B3PVW2	ahcY	H	6	50695.97	5.50	52.89	3	Adenosylhomocysteinase OS
B3PXA5	RHECIAT_CH0003540	P	3	35358.55	5.21	2.81	3	Probable nitrate/sulfonate ABC transporter, substrate-binding protein OS

B3PX2	RHECIAT_CH0001952	WC	4	16002.45	8.62	15.91	3	Hypothetical conserved protein OS
B3PY7	serA	E	4	56598.33	5.51	6.81	3	D-3-phosphoglycerate dehydrogenase protein OS
B3PZ32	suhBch	G	5	28429.65	6.61	3.78	3	Myo-inositol-1(Or 4)-monophosphatase protein OS
B3PZV3	exoN	M	12	33208.29	5.73	74.97	3	UTP--glucose-1-phosphate uridylyltransferase protein OS
B3PZW5	RHECIAT_CH0003827	G	2	28137.85	5.23	4.66	3	Probable inositol monophosphatase protein OS
B3PZX8	ropAch3	WC	3	36836.53	4.19	18.79	3	Porin outer membrane protein OS
B3Q0S3	gltX	J	2	54210.52	6.15	6.21	3	Glutamyl-tRNA synthetase OS
B3Q1J2	RHECIAT_PA0000204	P	11	37017.80	5.75	415.37	3	Probable sulfate uptake ABC transporter, substrate-binding protein OS
B3Q5B9	RHECIAT_PC0000320	EP	2	59405.84	5.41	13.57	3	Putative oligopeptide ABC transporter, substrate-binding protein OS
B3Q5L4	RHECIAT_PC0000417	R	2	20697.63	8.97	2.00	3	Putative NADPH:quinone reductase protein OS
Q2JZ68	RHE_PF00228	R	4	20677.64	8.97	7.67	3	Putative NADPH:quinone reductase protein OS
Q2JZ72	RHE_PF00224	T	2	13416.30	5.86	4.25	3	Putative two-component response regulator protein OS
Q2JZU2	minC	D	3	25620.72	5.61	6.59	3	Cell division inhibitor MinC protein OS
Q2K063	hutUe	E	2	61048.44	6.14	4.49	3	Urocanate hydratase protein OS
Q2K0J6	RHE_00207	QR	5	27741.48	7.89	11.09	3	Probable dehydrogenase/reductase protein OS
Q2K0W3	RHE_00322	J	2	47688.66	5.35	2.17	3	Probable glutamyl-tRNA amidotransferase subunit A protein OS
Q2K1H6	RHE_PC00130	EP	4	61384.69	6.82	5.59	3	Probable peptide ABC transporter, substrate-binding protein OS
Q2K206	thiC	H	3	67120.48	5.66	12.12	3	Phosphomethylpyrimidine synthase OS
Q2K208	thiG	H	3	27178.55	6.08	3.95	3	Thiazole biosynthesis protein OS
Q2K219	RHE_PB00096	T	7	20812.11	5.53	14.52	3	Putative universal stress protein OS
Q2K2A1	RHE_PA00012	WC	2	19342.00	6.40	8.36	3	Putative uncharacterized protein OS
Q2K2A6	RHE_PA00017	WC	3	10396.76	6.58	4.29	3	Hypothetical conserved protein OS
Q2K2Q2	trbI	N	6	48067.14	6.14	11.89	3	Conjugal transfer protein I OS
Q2K2R0	trbB	U	3	35237.33	5.73	20.86	3	Conjugal transfer protein B OS
Q2K376	leuC	E	12	50846.77	5.78	32.89	3	3-isopropylmalate dehydratase large subunit OS
Q2K382	rpsP	J	2	13526.43	10.27	34.96	3	30S ribosomal protein S16 OS
Q2K3I6	cheW5	J	5	17175.80	4.97	6.87	3	Chemotaxis signal transduction protein OS
Q2K3J5	RHE_CH03844	G	4	44001.91	5.01	5.80	3	Putative sugar ABC transporter, substrate-binding protein OS
Q2K3K5	hipO	R	2	41578.07	5.54	2.40	3	Hyppurate hydrolase protein OS
Q2K3L0	accA	I	4	34583.27	6.36	6.25	3	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha OS
Q2K3L3	aroB	E	2	40276.01	6.09	5.30	3	3-dehydroquinate synthase OS
Q2K3P2	argH	E	2	50978.90	5.83	8.25	3	Argininosuccinate lyase OS
Q2K3T4	clpB	O	10	95856.24	5.58	49.84	3	ATP-dependent Clp protease, ATP-binding subunit protein OS
Q2K3W6	nrdF	F	5	36744.43	4.76	7.49	3	Ribonucleoside-diphosphate reductase, beta subunit protein OS
Q2K3W7	cyaG2	T	3	45235.82	5.62	4.36	3	Adenylate cyclase protein OS
Q2K432	xylBch	G	4	50602.33	5.73	8.23	3	Xylulokinase protein OS
Q2K458	RHE_CH03623	K	2	20647.77	9.54	2.14	3	Probable transcriptional regulator protein, MarR family OS
Q2K471	RHE_CH03610	WC	2	16270.93	9.12	3.94	3	Hypothetical conserved protein OS
Q2K4L7	RHE_CH03463	WC	4	16216.76	9.37	8.32	3	Putative signal peptide protein OS
Q2K4Z0	RHE_CH03333	WC	4	13593.48	6.63	5.23	3	Hypothetical conserved protein OS
Q2K5L8	kdgK	G	2	31370.22	4.95	8.70	3	2-dehydro-3-deoxygluconokinase protein OS

Q2K5P9	RHE_CH03070	S	4	27210.15	4.32	7.99	3	Hypothetical conserved protein OS
Q2K626	RHE_CH02943	S	5	14545.34	6.28	7.94	3	Hypothetical conserved protein OS
Q2K656	RHE_CH02912	WC	8	14275.36	5.10	41.32	3	Hypothetical conserved protein OS
Q2K683	RHE_CH02885	C	3	50245.35	5.20	4.61	3	Probable aldehyde dehydrogenase protein OS
Q2K688	RHE_CH02880	S	2	12425.02	5.24	22.72	3	Hypothetical conserved protein OS
Q2K697	metF	E	2	32498.69	5.92	7.79	3	Methylenetetrahydrofolate reductase OS
Q2K6B9	murD	M	7	49002.70	6.12	15.59	3	UDP-N-acetylmuramoylalanine--D-glutamate ligase OS
Q2K6R3	RHE_CH02702	R	3	18528.30	6.28	6.66	3	Probable acetyltransferase protein OS
Q2K702	RHE_CH02612	E	2	20904.66	4.83	1.79	3	Hypothetical conserved protein OS
Q2K733	RHE_CH02579	H	4	23124.25	6.23	6.54	3	Probable dihydrofolate reductase protein OS
Q2K7D2	RHE_CH02477	I	3	17000.11	5.38	10.02	3	Probable hydroxymyristoyl-[acyl carrier] dehydratase transmembrane protein OS
Q2K7J7	RHE_CH02411	O	2	14595.51	6.10	2.46	3	Probable stress-induced OsmC-like protein OS
Q2K7S6	RHE_CH02330	T	3	17703.10	4.66	2.85	3	Putative phosphohistidine phosphatase protein OS
Q2K7T5	RHE_CH02321	K	5	28562.83	8.97	7.89	3	Probable transcriptional regulator protein, AraC family OS
Q2K7T6	gstch8	O	3	23168.57	7.69	5.91	3	Glutathione S-transferase protein OS
Q2K7Z8	tyrS	J	2	46225.66	5.99	5.42	3	Tyrosyl-tRNA synthetase protein OS
Q2K815	gcvT	E	4	38168.51	5.53	5.27	3	Aminomethyltransferase OS
Q2K859	thrS	J	2	74962.36	5.88	9.62	3	Threonyl-tRNA synthetase OS
Q2K860	RHE_CH02195	WC	2	38312.32	4.72	3.03	3	Hypothetical conserved protein OS
Q2K8D9	coaD	M	3	17314.10	6.83	4.77	3	Phosphopantetheine adenylyltransferase OS
Q2K8T8	RHE_CH01962	HC	2	38963.80	7.09	4.07	3	Probable 2-octaprenyl-6-methoxyphenol hydroxylase protein OS
Q2K9I7	ilvD	EG	6	65216.19	5.85	22.76	3	Dihydroxy-acid dehydratase OS
Q2K9I8	yedY	R	3	35119.17	9.05	2.49	3	Sulfoxide reductase catalytic subunit yedY OS
Q2K9L1	rplV	J	4	14137.39	10.45	18.90	3	50S ribosomal protein L22 OS
Q2K9Q5	ceiR1	T	3	13932.18	5.05	4.28	3	Two-component response regulator protein OS
Q2K9T8	RHE_CH01596	WC	4	13647.47	6.06	14.73	3	Hypothetical conserved protein OS
Q2KAB6	RHE_CH01417	R	2	22425.85	5.69	4.52	3	Hypothetical conserved protein OS
Q2KAC6	RHE_CH01407	F	2	20663.28	5.88	1.84	3	Putative ADP-ribose pyrophosphatase protein, mutT/nudiX family OS
Q2KAC9	RHE_CH01404	R	3	29640.71	6.92	6.00	3	Putative aminomethyltransferase protein (Glycine cleavage) OS
Q2KAF0	RHE_CH01383	S	2	21990.06	5.31	3.46	3	Hypothetical conserved protein OS
Q2KAR5	RHE_CH01266	E	2	13834.52	4.72	3.63	3	Probable lactoylglutathione lyase protein OS
Q2KAU6	RHE_CH01235	M	3	19133.63	6.66	3.42	3	Probable phosphinothricin acetyltransferase (Antibiotic resistance) protein OS
Q2KBV0	ispB	H	3	37285.56	5.00	9.33	3	Octaprenyl-diphosphate synthase protein OS
Q2KBX0	fabG1	IQR	2	27176.17	6.44	2.06	3	3-oxoacyl-(Acyl carrier protein) reductase protein OS
Q2KBZ5	RHE_CH00830	G	10	30928.34	5.28	10.52	3	Putative hydrolase, haloacid dehalogenase-like family protein OS
Q2KC00	gstch3	O	4	26328.19	6.75	5.22	3	Glutathione S-transferase protein OS
Q2KC64	asnB	E	3	68445.65	5.69	4.40	3	Asparagine synthetase (Glutamine-hydrolyzing) protein OS
Q2KCJ2	RHE_CH00628	K	5	25963.58	6.40	6.66	3	Putative transcriptional regulator protein, LuxR family OS
Q2KCK0	RHE_CH00620	O	2	15252.38	5.84	7.95	3	Putative membrane-anchored protein (Osmotically inducible protein) OS
Q2KCT4	RHE_CH00534	WC	3	15036.50	5.01	5.22	3	Hypothetical conserved protein OS
Q2KD65	ihfB	L	3	11190.57	6.74	7.99	3	Integration host factor subunit beta OS

Q2KDA8	rpsT	I	4	27660.95	6.36	6.92	3	Enoyl-CoA hydratase protein OS
Q2KDC1	RHE_CH00343	QR	2	25944.23	5.42	1.79	3	Probable oxidoreductase protein OS
Q2KDF5	gstch2	O	3	26644.46	8.61	12.43	3	Glutathione S-transferase protein OS
Q2KDR6	upp	F	3	23744.60	6.66	7.89	3	Uracil phosphoribosyltransferase protein OS
Q2KDW3	da1	E	3	49872.53	5.12	16.76	3	Succinyl-diaminopimelate desuccinylase protein OS
Q2KDY4	RHE_CH00127	FGR	3	15215.50	6.74	5.40	3	Hypothetical conserved protein OS
Q2KDZ6	rbfA	J	2	15460.54	6.75	4.43	3	Ribosome-binding factor A OS
Q2KE28	RHE_CH00082	G	4	46270.54	8.36	9.03	3	Probable sugar ABC transporter, substrate-binding protein OS
Q2KE51	glcB	C	5	79898.98	5.67	72.09	3	Malate synthase G OS
Q2KE72	ahcY	H	8	50727.03	5.52	100.51	3	Adenosylhomocysteinase OS
Q8KIL7	RHE_PD00091	C	3	52811.89	5.47	2.87	3	Succinate-semialdehyde dehydrogenase [NAD(P)+] protein OS
Q8KLE5	RHE_PD00258	WC	2	21766.21	6.83	5.85	3	Polysaccharide inhibition protein OS
Q2K4X5	mexF1	V	4	#####	5.37	13.92	2.8	Multidrug-efflux system, acriflavin resistance protein OS
Q2KBV2	is	I	6	30623.31	5.59	9.60	2.8	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase OS
B3PZX6	RHECIAT_CH0003838	T	2	13471.63	5.18	36.25	2.6	Probable two-component response regulator protein OS
Q2JZ23	RHE_PF00273	E	2	43316.71	5.80	5.43	2.6	Probable aspartate aminotransferase protein OS
Q2K3V7	RHE_CH03724	S	3	14031.94	5.56	14.39	2.6	Hypothetical conserved protein OS
B3PWS3	rplW	J	3	10543.30	10.06	38.70	2.5	50S ribosomal protein L23 OS
Q2K328	RHE_CH04014	O	2	24238.86	5.98	7.07	2.4	Glutathione S-transferase protein OS
Q2K6H3	serB	E	4	30901.48	4.68	15.26	2.4	Phosphoserine phosphatase protein OS
Q2KD82	RHE_CH00382	T	5	17556.10	5.17	10.95	2.4	Putative universal stress protein OS
Q2KD98	iunH1	F	2	34500.91	5.37	12.93	2.4	Inosine-uridine preferring nucleoside hydrolase protein OS
Q2K6B3	mraW	M	5	35868.03	9.18	13.42	2.3	S-adenosyl-L-methionine-dependent methyltransferase mraW OS
B3PQC7	pgl	G	3	24074.25	5.13	2.56	2	6-phosphogluconolactonase protein OS
B3PRF1	RHECIAT_CH0000949	D	3	41762.32	6.56	10.90	2	Hypothetical conserved protein OS
B3PV47	leuA2	E	2	63396.65	5.40	14.71	2	2-isopropylmalate synthase protein OS
B3PZ10	tkt	G	6	69762.35	6.08	17.55	2	Transketolase protein OS
B3PZ53	RHECIAT_CH0003779	WC	3	6603.62	8.98	5.19	2	Hypothetical conserved protein OS
B3PZB0	metK	H	6	44956.85	8.09	31.77	2	S-adenosylmethionine synthetase OS
Q2K0Z4	RHE_00353	S	2	14808.99	5.87	5.95	2	Putative carboxymuconolactone decarboxylase protein OS
Q2K2Q9	trbE	N	11	91904.65	5.80	12.88	2	Conjugal transfer protein E OS
Q2K4V8	RHE_CH03365	QR	3	27517.53	5.33	3.27	2	Probable short-chain alcohol dehydrogenase protein OS
Q2K4X7	RHE_CH03346	KR	3	30581.84	6.86	2.34	2	Putative acetyltransferase protein OS
Q2K6C1	murG	M	3	39283.51	9.47	2.10	2	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol
Q2K6N7	RHE_CH02729	KE	3	43895.28	5.69	3.78	2	Probable aminotransferase protein OS
Q2K7X4	RHE_CH02282	WC	2	13730.86	10.40	7.42	2	Hypothetical conserved protein OS
Q2K7Z1	rhIE1	LKJ	2	58574.51	9.36	1.94	2	ATP dependent RNA helicase protein OS
Q2K879	trpC	E	4	28960.24	5.14	5.20	2	Indole-3-glycerol phosphate synthase OS
Q2K8H2	qor	CR	2	34950.70	6.40	5.03	2	Quinone oxidoreductase protein OS
Q2K8U3	RHE_CH01957	FJ	2	21995.92	5.37	2.24	2	Probable deoxycytidylate deaminase family protein OS
Q2K924	aroQ1	E	2	15815.10	6.49	5.59	2	3-dehydroquinate dehydratase protein OS

Q2K930	RHE_CH01865	WC	2	15126.98	9.74	8.29	2	Putative NADH-ubiquinone oxidoreductase chain b17.2 like protein OS
Q2K983	RHE_CH01810	WC	3	30778.42	4.56	8.16	2	Hypothetical conserved protein OS
Q2K9C8	RHE_CH01763	K	2	20867.45	5.42	2.14	2	Probable transcriptional regulator protein, TetR family OS
Q2KA46	purN	F	2	21558.83	5.32	2.60	2	Phosphoribosylglycinamide formyltransferase protein OS
Q2KDH9	RHE_CH00284	P	2	13260.27	6.29	5.92	2	Probable arsenate reductase protein OS
B3PWS6	rplV	J	3	14137.39	10.45	22.68	1.8	50S ribosomal protein L22 OS
Q2K1F4	RHE_PC00109	T	2	29476.19	6.00	16.57	1.7	Putative universal stress UspA protein OS
B3PPV0	dksA	T	3	17137.15	5.52	14.11	1.6	DnaK suppressor protein OS
Q2K5U6	prsAch	FE	5	33550.42	5.98	9.49	1.6	Ribose-phosphate pyrophosphokinase OS
Q2K5U7	RHE_CH03022	WC	4	28232.45	6.84	9.57	1.6	Hypothetical conserved protein OS
Q2KA00	ihfA	L	3	12519.57	10.56	15.83	1.6	Integration host factor subunit alpha OS
Q2KDF1	RHE_CH00313	WC	3	10629.10	6.58	21.02	1.6	Hypothetical conserved protein OS
B3PPS1	purL	F	3	79526.26	4.83	13.03	1.5	Phosphoribosylformylglycinamide synthase 2 OS
Q2JZW7	RHE_00431	QR	4	25099.74	5.91	7.69	1.5	Probable oxidoreductase protein, short-chain-dehydrogenase (SDR) superfamily OS
Q2K666	RHE_CH02902	G	2	46837.49	5.55	6.53	1.5	Hypothetical conserved protein OS
Q2K3B0	RHE_CH03931	WC	10	18030.63	9.49	518.36	1.4	Hypothetical conserved protein OS
Q2K7R8	fdx	C	3	11671.18	4.49	13.14	1.4	Ferredoxin, 2Fe-2S FDII electron transport iron-sulfur protein OS
Q2K7X7	purS	F	2	8399.79	5.68	20.62	1.4	Phosphoribosylformylglycinamide synthase protein OS
Q2K9F7	RHE_CH01734	TK	3	22157.35	6.21	16.17	1.4	Probable two-component response regulator protein OS
Q2KA86	RHE_CH01447	S	5	32321.93	5.37	9.33	1.4	Hypothetical conserved protein OS
B3PRS6	RHECIAT_CH0002684	R	2	30513.54	6.12	9.20	1.3	Probable chloride peroxidase protein OS
B3PTQ6	RHECIAT_CH0002944	WC	2	18938.45	5.32	12.53	1.3	Hypothetical conserved protein OS
B3Q2V6	RHECIAT_PB0000305	WC	2	26296.83	9.57	7.53	1.3	Hypothetical conserved protein OS
Q2K107	RHE_00366	Q	3	32084.30	6.24	15.15	1.3	Probable dienelactone hydrolase protein OS
Q2K383	tyrA	E	3	12054.74	6.41	28.72	1.3	Chorismate mutase protein OS
Q2K3F6	RHE_CH03884	QR	4	25867.40	5.32	7.41	1.3	Putative oxidoreductase protein OS
Q2K4V1	ugpBch2	G	4	49095.11	5.67	29.64	1.3	sn-glycerol 3-phosphate ABC transporter, substrate-binding protein OS
Q2K5V0	RHE_CH03019	S	2	39753.40	5.02	7.77	1.3	Hypothetical conserved protein OS
Q2K5Z0	nerA	C	2	43689.84	6.42	25.31	1.3	Glycerol trinitrate reductase protein OS
Q2K660	RHE_CH02908	E	2	47186.20	8.36	6.07	1.3	Probable spermidine/putrescine ABC transporter, substrate-binding protein OS
Q2K6D4	RHE_CH02834	KR	4	19769.43	6.93	13.50	1.3	Probable acetyltransferase protein OS
Q2K8Y9	RHE_CH01911	J	4	15877.40	5.86	14.46	1.3	Putative translation initiation inhibitor protein OS
Q2K9V8	RHE_CH01576	R	2	15131.45	5.75	35.24	1.3	Hypothetical conserved protein OS
Q2KDW0	polA	L	2	#####	5.29	22.23	1.3	DNA-directed DNA polymerase, DNA polymerase I protein OS
Q2KDX9	kdsB	M	2	27451.48	5.37	21.28	1.3	3-deoxy-manno-octulosonate cytidyltransferase OS
B3PYU7	hfq	R	4	9044.49	9.03	28.84	1.2	Protein hfq OS
B3PZK1	dppAch2	E	5	58210.36	6.78	47.69	1.2	Dipeptide ABC transporter, substrate-binding protein OS
Q2K3P1	lysA	E	3	45783.34	5.63	7.63	1.2	Diaminopimelate decarboxylase OS
Q2K9H8	RHE_CH01713	WC	3	20324.15	5.74	16.45	1.2	Putative nucleoside 2-deoxyribosyltransferase protein OS
Q2KBT2	RHE_CH00893	G	6	20841.49	6.92	18.30	1.2	Probable phosphoglycerate mutase protein OS
Q2KBX2	kdgA	G	2	21768.38	5.33	48.56	1.2	4-hydroxy-2-oxoglutarate aldolase/2-dehydro-3-deoxyphosphogluconate aldolase protein OS

Q2KDA1	rph	J	3	26029.66	6.84	5.38	1.2	Ribonuclease PH OS
Q2KDL4	RHE_CH00249	K	5	25751.39	5.23	11.90	1.2	Probable transcriptional activator (Regulator) protein, TenA family OS
Q2KDR9	deoC	F	9	26916.94	5.06	24.18	1.2	2-deoxyribose-5-phosphate aldolase protein OS
B3PNT7	maf	D	3	22625.22	9.71	5.25	1	Septum formation protein OS
B3PPD6	hisCch2	E	2	39442.21	6.12	5.20	1	Histidinol-phosphate aminotransferase protein OS
B3PQ90	sucC	C	19	41844.24	5.08	81.54	1	Succinyl-CoA ligase [ADP-forming] subunit beta OS
B3PQE8	RHECIAT_CH0000805	N	3	15972.21	6.62	12.02	1	Hypothetical conserved protein OS
B3PQU3	RHECIAT_CH0002564	G	6	43539.50	5.20	18.31	1	Probable sugar ABC transporter, substrate-binding protein OS
B3PR22	leuC	E	4	50856.72	5.62	5.77	1	3-isopropylmalate dehydratase large subunit OS
B3PR47	mocBch2	G	3	33032.58	4.79	6.75	1	Rhizopine ABC transporter, substrate-binding protein OS
B3PU22	rosR	WC	4	15638.81	7.83	7.54	1	Transcriptional regulator protein-nodulation competitiveness determinant OS
B3PV44	RHECIAT_CH0003221	R	3	20724.36	9.30	9.71	1	Hypothetical conserved protein OS
B3PVU7	RHECIAT_CH0000017	Q	3	29486.68	5.78	16.66	1	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase/5-carboxymethyl-2-oxo-hex-3-ene-1, 7-dioate
B3PW05	nuoG1	C	3	74312.79	5.53	3.82	1	NADH-quinone oxidoreductase OS
B3PW39	RHECIAT_CH0001717	J	2	#####	4.94	6.31	1	Probable ribonuclease protein OS
B3PWT2	rplX	J	2	11209.10	10.51	36.03	1	50S ribosomal protein L24 OS
B3PXE6	rimP	S	2	22277.34	4.66	2.44	1	Ribosome maturation factor rimP OS
B3PY41	ugpB2	G	5	49003.10	5.83	10.19	1	sn-glycerol 3-phosphate ABC transporter, substrate-binding protein OS
B3PYN2	clpA	O	5	90836.58	5.96	12.85	1	ATP-dependent Clp protease protein, ATP-binding subunit OS
B3Q1T4	RHECIAT_PA0000296	G	4	43626.82	5.29	19.89	1	Probable sugar uptake ABC transporter, substrate-binding protein OS
B3Q2B9	rbsBb	G	4	42130.18	5.34	11.58	1	Ribose ABC transporter, substrate-binding protein OS
Q2JYH3	bglSf	G	6	88423.42	5.16	10.22	1	Beta-glucosidase protein OS
Q2K303	RHE_CH04039	WC	2	15745.91	5.48	3.75	1	Hypothetical conserved protein OS
Q2K345	ndvB	G	11	#####	5.91	21.96	1	Beta-1,2-glucan production associated transmembrane protein OS
Q2K354	mocBch2	G	8	33112.67	4.85	18.15	1	Rhizopine ABC transporter, substrate-binding protein OS
Q2K357	galE3	M	2	35898.23	5.64	3.83	1	UDP-glucose 4-epimerase protein OS
Q2K367	glnK	E	2	12178.12	7.97	16.98	1	Nitrogen regulatory protein P-II OS
Q2K5A5	pssK	GM	5	38823.39	8.47	4.44	1	Exopolysaccharide polymerization protein OS
Q2K7Y8	RHE_CH02268	R	5	21238.10	7.05	11.62	1	Probable 4-hydroxyphenylacetate 3-monooxygenase small chain protein OS
Q2K864	RHE_CH02191	WC	3	14663.76	6.08	2.92	1	Hypothetical conserved protein OS
Q2K8F6	RHE_CH02096	O	3	17086.21	5.95	5.30	1	Putative redox protein OS
Q2K8U1	RHE_CH01959	WC	2	11392.91	4.76	8.55	1	Hypothetical conserved protein OS
Q2K9Y3	RHE_CH01551	K	3	10276.47	4.59	14.11	1	Hypothetical conserved protein OS
Q2KBA7	RHE_CH01071	JE	2	12182.66	5.05	3.02	1	Hypothetical conserved protein OS
Q2KBG1	RHE_CH01015	WC	2	26365.10	5.51	3.05	1	Putative uncharacterized protein OS
Q2KBJ7	RHE_CH00978	R	2	47743.24	5.08	7.66	1	Probable processing protease protein OS
Q2KBN0	RHE_CH00945	WC	4	12158.56	5.22	4.38	1	Hypothetical conserved protein OS
Q2KBY3	RHE_CH00842	G	4	29293.10	4.99	6.32	1	Probable inositol monophosphatase protein OS
Q2KC94	RHE_CH00727	WC	5	15663.83	5.08	7.79	1	Hypothetical conserved protein OS
Q2KCQ3	RHE_CH00565	K	4	22319.58	6.16	3.73	1	Probable transcriptional regulator protein, TetR family OS
Q2KD80	RHE_CH00384	E	2	14343.22	5.85	1.65	1	Putative glyoxalase protein OS

Q2KDD3	prfC	J	3	58661.93	5.40	2.83	1	Peptide chain release factor 3 protein OS
Q2KDI2	gstch1	O	2	25470.12	5.25	1.81	1	Glutathione S-transferase protein OS
Q2KDJ3	RHE_CH00270	KR	4	16828.98	5.35	4.55	1	Probable acetyltransferase protein OS
Q2KDJ5	RHE_CH00268	WC	6	18262.82	5.56	12.23	1	Hypothetical conserved protein OS
Q2KDK4	RHE_CH00259	R	3	29111.35	5.68	6.25	1	Putative hydrolase protein OS
Q2KDV5	RHE_CH00156	WC	4	28110.17	8.86	7.82	1	Hypothetical conserved protein OS
Q2KDX1	moeB	H	3	27255.29	6.34	3.34	1	Molybdopterin biosynthesis protein OS
Q2KE50	RHE_CH00053	S	2	24994.56	9.54	7.11	1	Hypothetical conserved protein OS

Número	Condición
1	OMV fase estacionaria
2	OMV fase exponencial
3	periplasma fase exponencial
4	periplasma fase estacionaria
5	sobrenadante con naringenina
6	sobrenadante sin naringenina
7	OMV con naringenina
8	OMV sin naringenina

PU	péptidos únicos
PM	peso molecular
PI	punto isoelectrico
PMSS	sumatoria de la calificación de los péptidos identificados (expresión relativa)

COG	Grupos funcionales
C	Energy production and conversion
D	Cell cycle control, cell division, chromosome partitioning
E	Amino acid transport and metabolism
F	Nucleotide transport and metabolism
G	Carbohydrate transport and metabolism
H	Coenzyme transport and metabolism
I	Lipid transport and metabolism
J	Translation, ribosomal structure and biogenesis
K	Transcription
L	Replication, recombination and repair
M	Cell wall/membrane/envelope biogenesis
N	Cell motility
O	Posttranslational modification, protein turnover, chaperones
P	Inorganic ion transport and metabolism
Q	Secondary metabolites biosynthesis, transport and catabolism
R	General function prediction only
T	Signal transduction mechanisms
U	Intracellular trafficking, secretion, and vesicular transport
V	Defense mechanisms
WC	without COG
S	Function unknown