



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**LA RECONSTRUCCIÓN DEL GENOMA
ANCESTRAL DE PROTEOBACTERIAS:
UN CONTROL METODOLÓGICO PARA EL
ESTUDIO DEL ÚLTIMO ANCESTRO COMÚN
(LUCA)**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGA

P R E S E N T A:

CORAL CRUZ GONZÁLEZ LUNA



**DIRECTOR DE TESIS:
DR. ARTURO CARLOS II BECERRA BRACHO
2015**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Hoja de Datos del Jurado

1. Datos del alumno

Cruz González
Luna
Coral
56 06 06 30
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
407088216

2. Datos del Tutor

Doctor
Arturo Carlos II
Becerra
Bracho

3. Datos del Sinodal 1

Doctora
Claudia Andrea
Segal
Kischinevzky

4. Datos del Sinodal 2

Doctor
Luis David
Alcaraz
Peraza

5. Datos del Sinodal 3

Maestra en Ciencias
Fabiola
Ramírez
Corona

6. Datos del Sinodal 4

Maestro en Ciencias
Ricardo
Hernández
Morales

7. Datos del Trabajo Escrito

La reconstrucción del genoma ancestral de Proteobacterias:
Un control metodológico para el estudio del último ancestro común (LCA)
99 pp
2015

A mis abuelas, Elsa y Jose, por ser los pilares de mi vida

A mis papás, Gabriel y Gabriela, con todo el amor

Agradecimientos

A mi tutor, Arturo Becerra, por la confianza que me dio para realizar este trabajo y para continuar con mis estudios.

Al Dr. Lazcano y a todos mis compañeros del laboratorio de Origen de la Vida. No imaginaba que existiera un lugar con una camaradería como la que he encontrado con ustedes, gracias por compartir su conocimiento y sobretodo por su calidad humana, Mario, Pepón, Beto, Luciana, Ricardo, Wolfie, Claudia (ambas), Héctor, Sara, Edwin, Isra, Caro, Rodrigo, Alex, Pablo: son la onda, me siento afortunada de formar parte del grupo de trabajo.

A mi familia, pues esta tesis no hubiera sido posible sin todo el cariño que me han brindado. A mi hermano, el mejor amigo que me dio la vida. A mi abuela Jose y a mis tías Ceci y Laura, por todas las risas. A mis primos Oscar y en especial a Fabián, por su apoyo incondicional. A mi familia Cholulteca, Elsa, Alina y Darío.

A mis amigos de la infancia, Cris, Pablo, Serch, Roy y Panca, quienes siguen caminando a mi lado, gracias por estar en las buenas y en las malas.

A toda la chilanga banda que me ha acompañado en las distintas etapas de este largo proceso, en particular a Julia, Rafin, Sandra, Sara, Carito, Aura y a Dodis, por los buenos momentos que pasamos juntos.

Resumen

Sin lugar a dudas, uno de los grandes retos de la biología evolutiva es descubrir la naturaleza del último ancestro común (LCA por sus siglas en inglés). Éste se define como el ancestro más reciente que comparten todos los organismos modernos y probablemente vivió hace más de 3,500 millones de años.

En principio, se pueden inferir sus características identificando las proteínas compartidas entre los dominios *Bacteria*, *Archaea* y *Eukarya*, pero la realidad es que existe una gran controversia sobre los rasgos que podría tener el LCA. En particular, entre las diferentes reconstrucciones que se han hecho, no hay un consenso sobre cuál sería el contenido proteico del ancestro.

Con el objetivo de identificar el alcance y los límites del análisis de la estructura primaria de las proteínas para la reconstrucción de ancestros, en el presente trabajo se aplicó dicha metodología para un grupo de bacterias bien conocido: las proteobacterias.

Los catálogos ancestrales de los diferentes subgrupos de proteobacterias tuvieron entre 69 y 598 proteínas, lo que sugiere que el método recupera entre el 1.7% y el 25% del contenido proteico del ancestro. Además no todas las categorías funcionales estuvieron igualmente conservadas, siendo la más evidente la de traducción, biogénesis y estructura del ribosoma. Los factores que influyeron en el tamaño de las reconstrucciones ancestrales son la pérdida de genes, la transferencia horizontal, la diversidad metabólica y fisiológica de las especies que conforman la muestra, y sobretodo la distancia filogenética entre éstas. Aunque el método subestima cuantitativa y cualitativamente el contenido de los ancestros de las proteobacterias, se puede concluir que éstos ya contaban con una diversidad funcional equivalente a la de las especies modernas.

Los resultados muestran que a mayor distancia filogenética entre las especies de la muestra, el catálogo ancestral se hace menor, es decir, hay menos proteínas conservadas. Dado que el LCA representa el nodo más alejado del árbol universal, se concluye que el método planteado resulta insuficiente para su reconstrucción.

Palabras clave: LCA, Proteobacteria, ancestro común de proteobacterias, comparación de estructura primaria de las proteínas, catálogo ancestral, núcleo genómico.

Índice

Agradecimientos	4
Resumen	5
Índice	6
Capítulo 1. Introducción	8
1.1 El último ancestro común	8
1.2 Temporalidad relativa del LCA	9
1.3 Metodologías para reconstruir al LCA	10
1.4 Polémicas sobre las características del LCA	13
1.5 Genómica del LCA	15
1.6 Planteamiento del problema y objetivos	16
Capítulo 2. Metodología	18
Capítulo 3. Resultados	22
3.1 Tamaño de la muestra	22
3.2 Tamaño del genoma	22
3.3 Análisis de similitud y tamaño de los catálogos ancestrales	23
3.4 Clasificación funcional de las proteínas	24
3.4.1. Proteobacteria alfa subgrupo Rhizobiales	24
3.4.2. Proteobacteria alfa no Rhizobiales	25
3.4.3. Proteobacterias beta subgrupo Burkholderiales	26
3.4.4. Proteobacterias beta no Burkholderiales	27
3.4.5. Proteobacterias gamma subgrupo Enterobacteriales	28
3.4.6. Proteobacterias gamma no Enterobacteriales	29
3.4.7. Proteobacterias delta	31
3.4.8. Proteobacterias epsilon	32
3.5 Comparación entre las reconstrucciones ancestrales	33
Capítulo 4. Biología de Proteobacteria	35
4.1 Generalidades de las proteobacterias	35
4.2 Descripción de la muestra	38
4.2.1. Proteobacterias alfa	38

4.2.1.1.	Proteobacterias alfa subgrupo Rhizobiales	38
4.2.1.2.	Proteobacterias alfa no Rhizobiales	43
4.2.2.	Proteobacterias beta	46
4.2.2.1.	Proteobacterias beta subgrupo Burkholderiales	47
4.2.2.2.	Proteobacterias beta no Burkholderiales	50
4.2.3.	Proteobacterias gamma	52
4.2.3.1.	Proteobacterias gamma subgrupo Enterobacteriales	53
4.2.3.2.	Proteobacterias gamma no Enterobacteriales	55
4.2.4.	Proteobacterias delta	58
4.2.5.	Proteobacterias epsilon	60
Capítulo 5.	Discusión	62
5.1	Sesgo inicial de la muestra	62
5.2	Análisis cuantitativo	62
5.3	Factores que afectaron las reconstrucciones ancestrales	63
5.4	Alcances y limitaciones del método	69
5.5	Perspectivas	70
Conclusiones	71
Referencias	72
Anexos	88
Anexo A.	Listado de proteobacterias alfa, orden Rhizobiales	88
Anexo B.	Listado de proteobacterias alfa no Rhizobiales	89
Anexo C.	Listado de proteobacterias beta, orden Burkholderiales	91
Anexo D.	Listado de proteobacterias beta no Burkholderiales	93
Anexo E.	Listado de proteobacterias gamma, orden Enterobacteriales	94
Anexo F.	Listado de proteobacterias gamma no Enterobacteriales	95
Anexo G.	Listado de proteobacterias delta	97
Anexo H.	Listado de proteobacterias epsilon	99

Capítulo 1. Introducción

1.1 El último ancestro común

Aunque hoy sabemos que todos los seres vivos están genéticamente relacionados, fue Charles Darwin (1859) quien planteó por primera vez la idea de ancestría común. En la parte final de *El origen de las especies* se refiere al ancestro común:

"... hay también que admitir que todos los seres que en todo tiempo han vivido sobre la Tierra pueden haber descendido de alguna forma primordial..."

A pesar de esto, el estudio moderno de la naturaleza del LCA inició hasta la segunda mitad del siglo XX, gracias a la aparición de las secuencias moleculares, las cuales hicieron posible la construcción de árboles filogenéticos de un gran número de organismos.

Fue Woese quien utilizó por primera vez el RNA ribosomal como marcador molecular y propuso un árbol universal sin raíz (Fig. 1), que se divide en los tres grandes grupos que hoy llamamos dominios: *Bacteria*, *Archaea* y *Eukarya* (Woese, 1987; Woese *et al.* 1990). A partir de dicho trabajo quedó en evidencia la existencia de un ancestro común a los tres dominios y sobre todo se hizo patente la posibilidad de reconstruirlo.

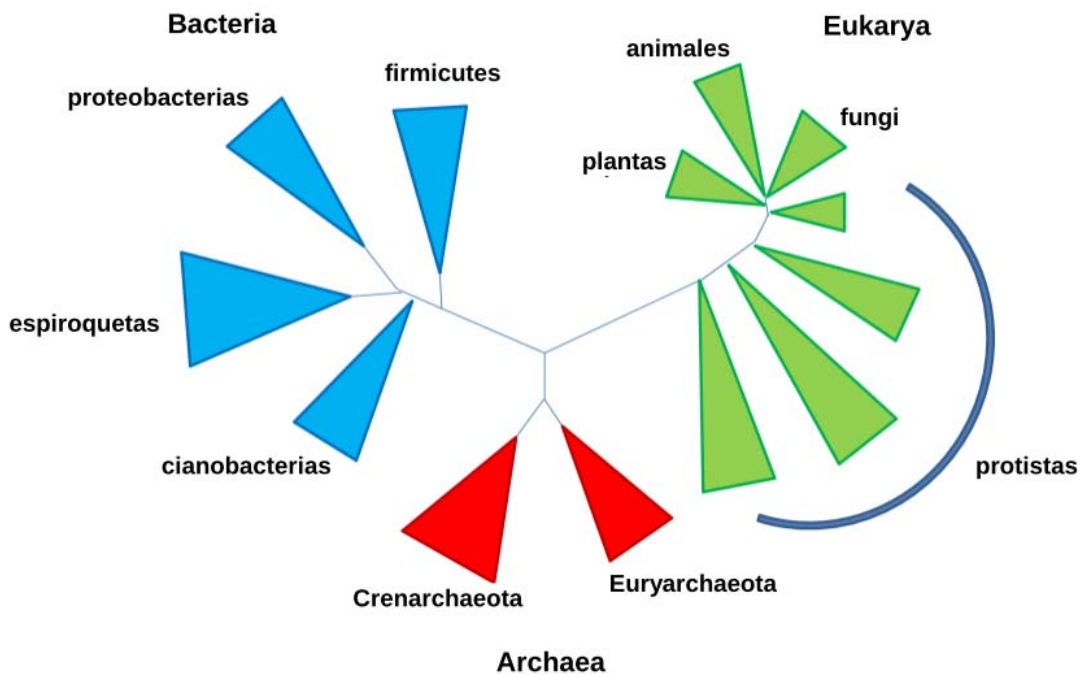


Figura 1. Árbol de la vida construido con el 16S/18S rRNA, tomado de Delaye & Becerra (2012).

1.2 Temporalidad relativa del LCA

El LCA, entendido como el ancestro más reciente a partir del cual evolucionaron todos los seres vivos que habitan la Tierra, no representa la etapa más temprana de la vida, sino la fase de evolución biológica previa a la diversificación de los dominios celulares (Fig. 2). Es decir que el LCA ya es en sí un producto del proceso evolutivo que inició con el origen de la vida.

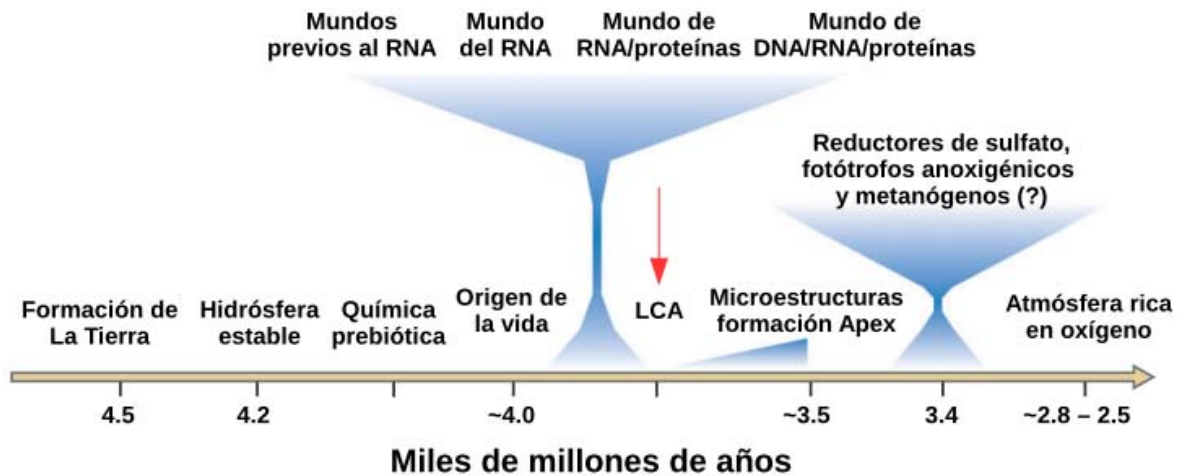


Figura 2. Línea del tiempo que muestra los eventos importantes en la evolución temprana de la vida, modificada de Becerra *et al.* 2007.

Si el LCA no es un descendiente directo de la sopa primitiva, entonces le anteceden otras etapas de organización biológica. En ese sentido se ha propuesto que la etapa previa a las células con DNA, RNA y proteínas fue un mundo de RNA, en el que dicha molécula tendría un papel preponderante en las células, al mostrar tanto actividad catalítica como informativa (Joyce, 2002). Más aún, se ha sugerido la existencia de mundos previos al del RNA en el que habría fuerzas evolutivas actuando sobre diversas moléculas informativas y autocatalíticas (Joyce, 2002). Sin embargo, la naturaleza de dichos polímeros no se puede deducir con los métodos filogenéticos o de la genómica comparada, ya que éstos tienen como límite el mundo del RNA (Poole, 2002; Becerra *et al.* 2007).

La evidencia fósil más antigua data de hace 3,500 Millones de años (ma), sugiriendo que la vida ya existía para ese momento, y que había comunidades de procariontes con diversidad metabólica, tales como fotótrofos anoxigénicos, reductores de sulfato y quizás arqueas metanogénicas (Canfield, 2006). Podemos afirmar entonces que el LCA vivió hace

más de 3,500 ma, periodo para el que no se tiene evidencia geológica que aporte información sobre las condiciones ambientales en las que se habría desarrollado, pues la veracidad de posibles fósiles más antiguos a dicha fecha ha sido fuertemente debatida (Schopf, 2006; Brasier *et al.* 2006; Canfield, 2006).

1.3 Metodologías para reconstruir al LCA

A pesar de la falta de fósiles morfológicos es posible reconstruir al LCA con métodos de cladística molecular y de genómica comparada, o una combinación de ambos, gracias a que permiten inferir sus características aplicando el principio de parsimonia.

Desde el punto de vista cladístico, los caracteres plesiomórficos de un árbol filogenético corresponden a los rasgos primitivos, es decir, al fenotipo del ancestro. En ese sentido se necesita una referencia que bien puede ser el árbol universal trifurcado construido a partir de la comparación de la subunidad pequeña del RNA ribosomal, debido a que el patrón se repite si se utilizan otras proteínas como polimerasas, subunidades de la ATPasa, factores de elongación o proteínas ribosomales (Cicarelli *et al.* 2006), lo que demuestra que muy temprano en la evolución celular la vida se separó en tres linajes (Delaye & Becerra, 2012).

La cuestión es establecer el orden de aparición de los linajes, para lo cual es necesario colocarle raíz al árbol universal. Sin embargo enraizar el árbol es un tema controversial, y las propuestas van desde *Bacteria* como el grupo plesiomórfico (Iwabe *et al.* 1989; Valas & Bourne 2009), *Archaea* (Di Giulio 2007a) e incluso *Eukarya* (Forterre & Philippe 1999). Lo anterior es importante porque de la posición de la raíz dependerán las inferencias que se puedan hacer sobre la naturaleza del LCA.

Estudios recientes indican que el origen de *Eukarya* se encuentra en un linaje de *Archaea*, lo que implicaría la existencia de sólo dos dominios, apareciendo el tercero como resultado de la fusión de una arquea y una bacteria (Williams *et al.* 2013). De hecho en el registro geológico aparecen microfósiles de posibles eucariontes hace 1800 ma (Knoll *et al.* 2006) y de manera indiscutible fósiles de algas rojas entre 1200 y 720 ma (Butterfield 2000), mucho tiempo después que los primeros fósiles de procariontes. Además mediante análisis moleculares se ha fechado al ancestro de los eucariontes entre 1900 y 1700 ma (Parfrey *et al.* 2011) e incluso se ha planteado que fue un organismo aerobio (Di Giulio 2007b), todo lo

cual sugiere la aparición posterior de los eucariontes. De ser este el caso el dominio *Eukarya* no sería un descendiente directo del LCA como algunos autores lo han sugerido (Forterre & Philippe 1999; Jeffares & Poole, 2000; Glansdorff *et al.* 2008).

Otra manera de reconstruir al LCA consiste en comparar las secuencias de genes o de proteínas de un conjunto de organismos para encontrar aquellas que son universales, las cuales se supone que debieron estar presentes en el ancestro (Fig. 3). Con este procedimiento se puede elaborar un catálogo genético del LCA que contenga a los genes o a las proteínas homólogas, a partir del cual se pueden inferir las funciones mínimas presentes en el ancestro.

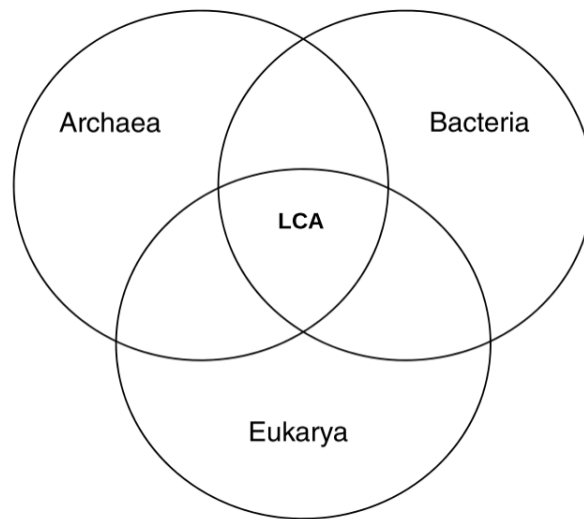


Figura 3. El catálogo o complemento genético del LCA se puede inferir a partir de los genes o de las proteínas compartidas. Modificado de Delaye & Becerra (2012).

Si bien la lógica para reconstruir al LCA es relativamente simple, las diversas propuestas sobre su naturaleza han levantado una gran polémica. Lo anterior se ve reflejado en la diversidad de nombres que se le han dado, como por ejemplo: *progenote* (Woese & Fox 1977a), *cenancestro* (Fitch & Upper 1987), último ancestro universal común (Kyrpides *et al.* 1999), último ancestro celular universal (Forterre & Philippe 1999), ancestro universal (Doolittle 2000), última comunidad común (Line 2002), ancestro común más reciente (Zhaxybayeva & Gogarten 2004) o *urancestro* (Kim & Caetano-Anollés, 2011).

En parte, la reconstrucción del LCA se ha dificultado por dos procesos que pueden obscurecer la señal filogenética:

- Los genes se pueden transferir entre organismos de distintas especies, proceso que se conoce como transferencia horizontal o lateral de genes (HGT, por sus siglas en inglés). En la naturaleza esto es posible gracias a diversos mecanismos como la conjugación, la transformación o la transducción bacteriana, y a la presencia de endonucleasas de restricción y ligasas capaces de cortar el DNA y unirlo de nuevo. Si las tasas de HGT han sido elevadas en el pasado, el contenido genético del ancestro se estaría sobrestimando.
- Algunos genes que pudieron haber estado presentes en el ancestro ya no son universales debido a que se han perdido en algunos linajes celulares. Si las tasas de pérdida de genes fueron muy altas, entonces se subestimaría la cantidad de genes del ancestro. Por lo tanto la precisión de las reconstrucciones depende de la intensidad relativa de ambos procesos en el pasado (Delaye & Becerra 2012).

Doolittle (2000) sugirió que la transferencia horizontal fue masiva durante la evolución temprana y planteó que la base de la filogenia universal es reticulada. Esta idea es afín a las de otros autores que proponen que el nodo de la filogenia universal corresponde a una población de ancestros precelulares, metabólicamente complementarios, que se beneficiaban de eventos masivos y multidireccionales de HGT (Kandler 1994; Woese 1998; Koonin & Martin 2005).

De ser cierto, reconstruir al LCA sería imposible. Ahora bien, si la HGT existe en la naturaleza, asumir que fue muy intensa cuando no se conoce la tasa de ocurrencia no es una postura razonable. De hecho, algunos autores han reportado que la HGT no afecta al momento de construir árboles filogenéticos (Penny *et al.* 2001; Sicheritz-Pontén & Andersson 2001) y que interpretar una topología distinta a la filogenia del rRNA como HGT *a priori* puede ser inexacto (Kurland 2000). Otra postura sería entonces considerar que la historia de la vida no ha estado dominada completamente por la transferencia horizontal (Tekaia *et al.* 1999), de ser así es posible recuperar los rasgos universales de las células para reconstruir al LCA (Poole, 2002).

1.4 Polémicas sobre las características del LCA

El primer planteamiento sobre la naturaleza del ancestro común a todos los seres vivos fue de Woese y Fox (1977a, 1977b), quienes lo llamaron *progenote*, una entidad primitiva y rudimentaria que a diferencia de las células modernas no tendría una separación clara entre genotipo y fenotipo. Para Woese (1987), los tres dominios evolucionarían de manera independiente de esta entidad hipotética, lo que explicaría las diferencias entre las maquinarias de transcripción y traducción de *Bacteria*, *Archaea* y *Eukarya*.

Si bien es razonable suponer que antes de la organización celular hubo sistemas de menor complejidad, la propuesta del progenote rápidamente encontró oposición. Fitch y Upper (1987) concluyeron que el ancestro común o *cenancestro* (*cen* del Griego *kainos* que significa reciente y *koinos* que significa común) era complejo y ya poseía un código genético equivalente al de las células contemporáneas. Utilizando enfoques diferentes otros grupos de trabajo han coincidido en que el ancestro era un organismo celular, parecido a las bacterias modernas en cuanto a su complejidad (Gogarten *et al.* 1989; Iwabe *et al.* 1989; Lazcano *et al.* 1992; Lazcano 1995).

Aunado a lo anterior, la universalidad del código genético, de los rasgos esenciales de la replicación y la expresión de los genes, de algunas reacciones anabólicas básicas, así como la producción de energía mediante ATPasas asociadas a membranas, son características lo suficientemente complejas como para suponer que evolucionaron de manera independiente (Becerra, *et al.* 2007). Por lo tanto, la conclusión de que el LCA ya era complejo parece ineludible, y los rasgos comunes a todas las células son un parámetro mínimo de las funciones presentes en el ancestro. Más aún, las variaciones entre los dominios se pueden explicar como el resultado de eventos divergentes a partir del ancestro (Delaye *et al.* 2001, Becerra *et al.* 2007).

Otro debate importante es la temperatura a la que se habría desarrollado el LCA. El descubrimiento de procariontes en ambientes extremos y su posición aparentemente basal en las filogenias universales llevó a proponer que el LCA era un organismo hipertermófilo (Di Giulio 2003, Stetter 2006). Cabe mencionar que los grupos de bacterias con dicho hábitat se han reclasificado, además de que hay evidencia que muestra que hubo un intercambio de genes importante entre arqueas y bacterias termófilas (Makarova & Koonin 2003, Becerra *et al.* 2007). Por lo tanto, es probable que el LCA haya sido mesófilo como propusieron Galtier y

sus colegas (1999), quienes con un modelo de evolución molecular dedujeron que el rRNA del ancestro sería más parecido al de los mesófilos por su contenido de G-C, que al de los hipertermófilos quienes poseen más enlaces A-U.

Una controversia más tiene que ver con la naturaleza del genoma del LCA. Por ejemplo, Woese (1987) sugirió que el *progenote* tendría RNA como material genético. Asimismo, la aparente falta de homología entre las polimerasas de los diferentes dominios llevó a sugerir no sólo un ancestro con RNA, sino también un origen polifilético del DNA (Leipe *et al.* 1999, Koonin & Martin 2005).

De manera fascinante, se ha encontrado que a nivel de estructura terciaria las polimerasas comparten un dominio homólogo, el cual cataliza la formación del enlace fosfodiéster (Steitz, 1999). Esto indica que la polimerasa ancestral ya presentaba dicho dominio y que probablemente actuaba como replicasa y transcriptasa en el mundo del RNA y proteínas (Delaye *et al.* 2001), lo que implica que antes de la divergencia ya existía un mecanismo moderno de transcripción (Becerra *et al.* 2007).

Ciertamente los genomas de RNA existieron durante etapas tempranas de la evolución celular (Lazcano *et al.* 1998b), pero hay buenas razones para sugerir que el DNA se estableció antes de la diversificación. En primer lugar es poco probable que la reducción biológica de los ribonucleótidos se haya originado en múltiples ocasiones, ya que es una reacción termodinámicamente desfavorable. Además, se ha demostrado que las ribonucleótido reductasas de *Archaea*, *Bacteria* y *Eukarya* tienen un origen monofilético (Stubbe *et al.* 2001). Otro argumento es que la presencia de proteínas de gran tamaño necesita de un sistema genético con alta fidelidad, lo cual no sucede con genomas de RNA (Lazcano *et al.* 1992). Aunado a todo lo anterior, es un hecho que todas las células tienen genomas de DNA, por lo que la conclusión más parsimoniosa es que el ancestro tenía como material genético ácido desoxirribonucleico (Becerra *et al.* 2007).

Por lo tanto los componentes genéticos del LCA aparecieron antes de la diversificación y su genoma se formó gracias a procesos de transferencia vertical, duplicaciones de genes, pérdida y la transferencia horizontal de genes (Becerra *et al.* 1997; Gogarten *et al.* 2002; Mirkin *et al.* 2003; Becerra *et al.* 2007). El objetivo de la genómica del LCA es descubrir qué genes o proteínas estaban presentes en el ancestro.

1.5 Genómica del LCA

De forma pionera, Mushegian & Koonin (1996) compararon los primeros dos genomas completamente secuenciados y concluyeron que el ancestro tenía 256 genes. El problema de dicho trabajo fue que los organismos involucrados eran parásitos y que además representaban solo al dominio *Bacteria* (Becerra *et al.* 1997). Tiempo después Kyrpides y su grupo de trabajo (1999) hicieron la primera reconstrucción utilizando genomas de los tres dominios, lo que los llevó a proponer que el LCA tendría 324 proteínas, y que sería un organismo similar a los procariontes en cuanto a su complejidad.

Al identificar los COG's (*Clusters of Orthologous Genes*, Tatusov *et al.* 2003) más conservados entre los dominios celulares, Harris y sus colegas (2003) arguyeron que el LCA tendría 80 COG's, con lo que infirieron que el ancestro presentaba transcripción, replicación y que tenía funciones asociadas a membranas. El mismo año Mirkin (2003) concluyó que el LCA tendría 600 genes, y argumentó que éstos eran suficientes para que un organismo funcione. Contrariamente, Koonin (2003) al encontrar solamente 63 genes ubicuos durante la comparación de 100 genomas, consideró que el ancestro sería muy simple, con pocos genes y sin replicación moderna ni genoma de DNA.

Al fijarse en la conservación de proteínas a nivel de superfamilias, Yang y sus colegas (2005) encontraron 49 plegamientos conservados y sugieren que el ancestro contaba con un inventario genético sofisticado. Por otro lado, Delaye y su grupo de trabajo (2005) al detectar proteínas homólogas entre secuencias, lograron identificar 115 dominios que apoyan la noción de un LCA genéticamente complejo.

Con una búsqueda de los motivos proteínicos más conservados, Sobolevsky & Trifonov (2006) concluyeron que hay 20 motivos universales y que el ancestro no parece diferir mucho de los organismos modernos. Ouzounis y su equipo (2006) buscaron secuencias homólogas y encontraron aproximadamente 1000 genes universales, apoyando la noción de un LCA complejo. Al rastrear dominios ancestrales, Ranea y colegas (2006) detectaron un total de 140 dominios, pudiendo afirmar que muchos de los rasgos modernos ya estaban presentes en el LCA.

Tuller y sus colegas (2010) propusieron un catálogo de 784 proteínas ancestrales y sugieren que el LCA sería parecido a las bacterias en lo que respecta a su naturaleza y también al tamaño del genoma. Por otro lado, Kim & Caetano-Anollés (2011) obtuvieron un

conjunto de 70 a 152 dominios estructurales para el ancestro, lo que refleja capacidades metabólicas avanzadas, pero al no encontrar funciones como la transcripción o la síntesis de desoxirribonucleótidos, concluyen que el LCA sería parecido a los progenotes pero con algunas funciones modernas ya presentes. Finalmente Kannan y su equipo (2013) calcularon que el ancestro tendría de 517 a 597 genes, pero el modelo probabilístico que desarrollaron arrojó diferentes proteínas en cada análisis, por lo que discutieron los pros y contras de su metodología y no las posibles características del LCA.

Las diferentes reconstrucciones coinciden en algunos aspectos importantes: que el LCA era biológicamente complejo y que tenía un genoma de DNA. Además de que la alta conservación de proteínas relacionadas con el RNA refleja la existencia del mundo de RNA (Becerra *et al.* 2007). Sin embargo, uno de los principales problemas es que los genes o proteínas conservados no son los mismos en los diferentes análisis y en general las rutas metabólicas no aparecen completas, no pudiendo concluir sobre la fuente de energía, carbono, nitrógeno o el tipo de membrana del ancestro.

Los trabajos anteriormente mencionados también han sido útiles para detectar los factores que afectan a las reconstrucciones ancestrales, como lo son sesgos en las bases de datos, los niveles de HGT, la tasa de pérdida de genes, la variación en las tasas de sustitución de las diferentes proteínas y en particular las limitaciones de los métodos aplicados (Becerra *et al.* 2007). Sobre este último punto hace falta explorar con mayor detalle si la comparación de secuencias de proteínas es un método robusto para reconstruir al LCA.

1.6 Planteamiento del problema y objetivos

Sin lugar a dudas, descubrir la naturaleza del LCA es uno de los grandes retos de la biología moderna (Morange 2011). Conocer sus características es importante para entender los procesos evolutivos que dieron origen a las células modernas, así como las entidades que le precedieron.

Actualmente se puede inferir su contenido genético o proteínico gracias a métodos de cladística molecular y de genómica comparada. En particular, los trabajos que han reconstruido el catálogo de proteínas del ancestro sugieren que era un organismo con genoma de DNA, biológicamente complejo y semejante a los procariontes actuales. Sin embargo, existe una gran disimilitud en cuanto al número y la identidad de proteínas que

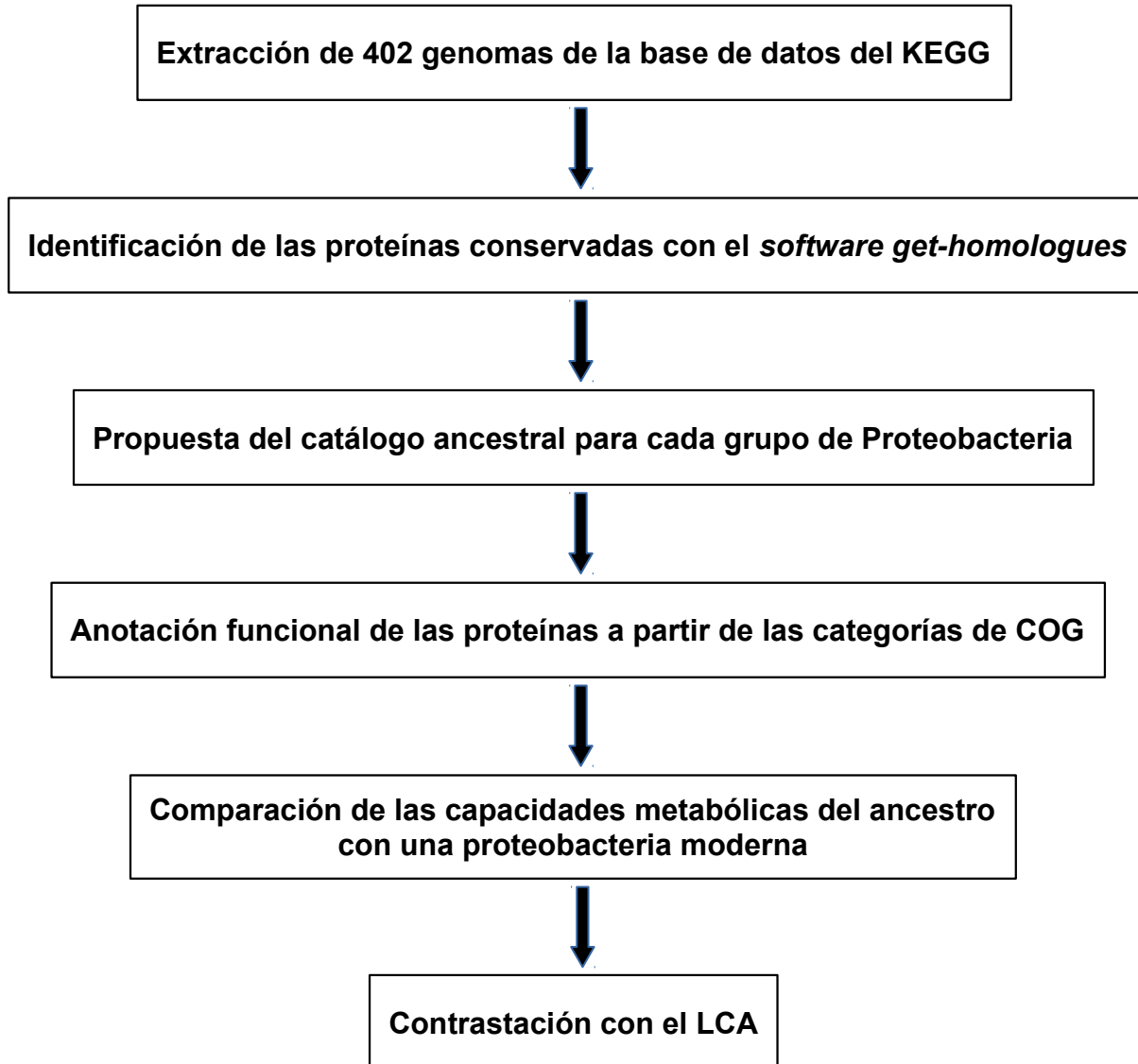
componen al LCA. Entre los diversos factores responsables de dichas diferencias se incluyen las limitaciones de la metodología aplicada en cada caso.

Tomando en consideración lo anterior, el presente trabajo tiene como objetivo reconstruir el catálogo genético de las proteobacterias, grupo cuya biología se conoce relativamente bien, además de que se cuenta con un gran número de genomas completamente secuenciados.

La reconstrucción del catálogo de proteínas ancestrales de las proteobacterias fungirá como un control metodológico para el caso del LCA, ya que permitirá identificar el alcance y los límites de comparar la estructura primaria de proteínas. Las preguntas de investigación de la tesis fueron ¿cuál es el contenido genético del ancestro común de cada clase de Proteobacteria? y ¿cuáles eran sus capacidades metabólicas?

Capítulo 2. Metodología

La metodología que se utilizó en este trabajo se resume en el siguiente esquema:



Extracción de secuencias

Se analizaron un total de 402 genomas, para los cuales se descargaron sus secuencias de aminoácidos respectivas de la base de datos KEGG (*Kyoto Encyclopedia of Genes and Genomes*, Kanehisa *et al.* 2006), en formato FASTA (terminación .faa).

Identificación de las proteínas conservadas

Para identificar las proteínas conservadas, se realizaron análisis de similitud entre los genomas de las diferentes especies de Proteobacterias con el *software get-homologues* (Contreras-Moreira & Vinuesa 2013), el cual detecta similitud vía BLASTP y posteriormente agrupa las posibles secuencias ortólogas utilizando el algoritmo BDBH (*bidirectional best hit*), y tomando como referencia el genoma de menor tamaño (secuencia *query*).

Los parámetros que se establecieron para el BLASTP fueron una cobertura del 75% y un valor de corte de $e < 1 \times 10^{-5}$, los cuales permiten detectar secuencias homólogas. En un BLASTP la cobertura o *query coverage* es el porcentaje de la secuencia *query* que se sobrelapa con la secuencia de la base datos (conocida como secuencia *subject*).

Cabe mencionar que con el *get-homologues* se puede detectar el *núcleo genómico* estricto y el *núcleo laxo* (*soft core*). Las proteínas presentes en el 100% de los genomas analizados representan el *núcleo genómico* estricto, mientras que el *núcleo laxo* son aquellas proteínas presentes en todos los individuos de la muestra, menos uno (n-1). Más aún, el *software* nos permitió relajar un poco más el criterio para incluir a las proteínas compartidas a partir del 80% de los organismos de la muestra.

Propuesta del catálogo ancestral

El análisis de similitud arroja una matriz de presencia/ausencia de las diferentes proteínas contra las distintas especies. En dicho archivo las proteínas venían en términos de la secuencia usada como referencia, por ejemplo la proteína HI0013 perteneciente a *Haemophilus influenzae*. Para poder identificarlas fue necesario descargar el genoma anotado de la especie que fungió como secuencia *query* (en formato .ptt), el cual se obtuvo de la página del NCBI (<ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/>) y posteriormente utilizar un programa de Perl que extrae de manera automatizada la identidad de la proteína, así como su clave COG. Lo anterior permitió construir una base de datos de las proteínas conservadas, denominado por nosotros como catálogo ancestral.

Debido a la cantidad de información y al poder computacional, los grupos más numerosos de proteobacterias alfa, proteobacterias beta y proteobacterias gamma, se redujeron numéricamente seleccionando sólo un representante por especie. Además cada uno de los grupos mencionados de proteobacterias se dividieron en dos conjuntos que se

estudiaron por separado, lo que hizo un total de ocho análisis para el presente trabajo. El criterio que se utilizó para formar los diferentes grupos se describe a continuación:

- Alfa: un grupo contiene especies del orden Rhizobiales y el otro conjunto al resto de las especies (no Rhizobiales).
- Beta: el primer grupo contiene solamente representantes del orden Burkholderiales y el segundo a bacterias que no pertenecen a dicho linaje (no Burkholderiales).
- Gamma: se separaron entre bacterias del orden Enterobacteriales y el resto de las especies (no Enterobacteriales).

Anotación funcional de las proteínas

Para clasificar funcionalmente a las proteínas se utilizaron sus claves COG, y las categorías propuestas en dicha base de datos (<http://www.ncbi.nlm.nih.gov/COG/>), las cuales forman cuatro grupos mayores según su función (Tabla 1):

Tabla 1. Categorías funcionales de los COG, tomado de: <ftp://ftp.ncbi.nih.gov/pub/COG/COG/fun.txt>.

ALMACENAMIENTO Y PROCESAMIENTO DE INFORMACIÓN

[J] Traducción, biogénesis y estructura del ribosoma

[A] Modificación y procesamiento del RNA

[K] Transcripción

[L] Replicación, reparación y recombinación

[B] Estructura y dinámica de la cromatina

PROCESOS CELULARES Y SEÑALIZACIÓN

[D] Control del ciclo celular, división celular y segregación de cromosomas

[Y] Estructura nuclear

[V] Mecanismos de defensa

[T] Mecanismos de transducción de señales

[M] Biogénesis de las membranas/pared celular

[N] Movilidad celular

[Z] Citoesqueleto

[W] Estructuras extracelulares

[U] Tráfico intracelular, transporte de vesículas y secreción

[O] Modificación postraduccional, chaperonas

METABOLISMO

[C] Conversión y producción de energía

[G] Metabolismo y transporte de carbohidratos

[E] Metabolismo y transporte de aminoácidos

[F] Metabolismo y transporte de nucleótidos

[H] Metabolismo y transporte de coenzimas
[I] Metabolismo y transporte de lípidos
[P] Metabolismo y transporte de iones inorgánicos
[Q] Biosíntesis, transporte y catabolismo de metabolitos secundarios

POBREMENTE CARACTERIZADAS

[R] Predicción general de la función
[S] Función desconocida

Comparación con las proteobacterias modernas y contrastación con el LCA

El tamaño del genoma se refiere al número total de proteínas que contiene. Para analizar la variación en dicho rasgo, se calculó el promedio y la desviación estándar correspondiente a cada grupo de proteobacterias.

Asimismo, el tamaño promedio se contrastó con el número de proteínas contenidas en los distintos catálogos ancestrales. Como el tamaño de cada reconstrucción representó una fracción del tamaño promedio de la muestra correspondiente, consideramos que dicho porcentaje refleja la *resolución* del análisis. Con los datos anteriores se elaboró la tabla 3 (ver resultados).

Para poder comparar las categorías funcionales de las distintas reconstrucciones, se relativizó el número de proteínas presente en cada una y posteriormente se realizó un *heatmap* utilizando la paquetería de R (R Core Team, 2014).

Capítulo 3. Resultados

3.1 Tamaño de la muestra

El tamaño de la muestra fue un parámetro que tuvo variación en el estudio. El grupo más numeroso fue el de las proteobacterias gamma, en particular el conjunto de las no Enterobacteriales, mientras que el conjunto más pequeño correspondió a las proteobacterias epsilon, con una diferencia de 58 genomas menos (tabla 2).

Grupo	Tamaño de la muestra
1. Proteobacteria alfa subgrupo Rhizobiales	51
2. Proteobacteria alfa subgrupo no Rhizobiales	61
3. Proteobacteria beta subgrupo Burkholderiales	63
4. Proteobacteria beta subgrupo no Burkholderiales	27
5. Proteobacteria gamma subgrupo Enterobacteriales	61
6. Proteobacteria gamma subgrupo no Enterobacteriales	71
7. Proteobacteria delta	55
8. Proteobacteria epsilon	13

Las diferencias en el número de genomas que se incluyeron en cada grupo, se debió a la disponibilidad de secuencias en la base de datos del KEGG, ya que no están igualmente representadas las cinco clases de Proteobacteria.

3.2 Tamaño del genoma

Hubo una variación importante en el tamaño de los genomas considerados en la muestra (fig. 4). De hecho se ha reportado que las proteobacterias, comparadas con otros grupos de procariontes, exhiben la mayor variación en dicho rasgo (Islas *et al.* 2004). El genoma con el menor número de proteínas tuvo solamente 1,354, y pertenece a *Candidatus Pelagibacter ubique*, la cual es una especie de vida libre de la clase alfa. En el otro extremo se sitúa la especie *Sorangium cellulosum*, la cual pertenece a las delta proteobacterias y posee 10,400 proteínas. Es decir que hay 9046 proteínas de diferencia entre el genoma menor y el mayor.

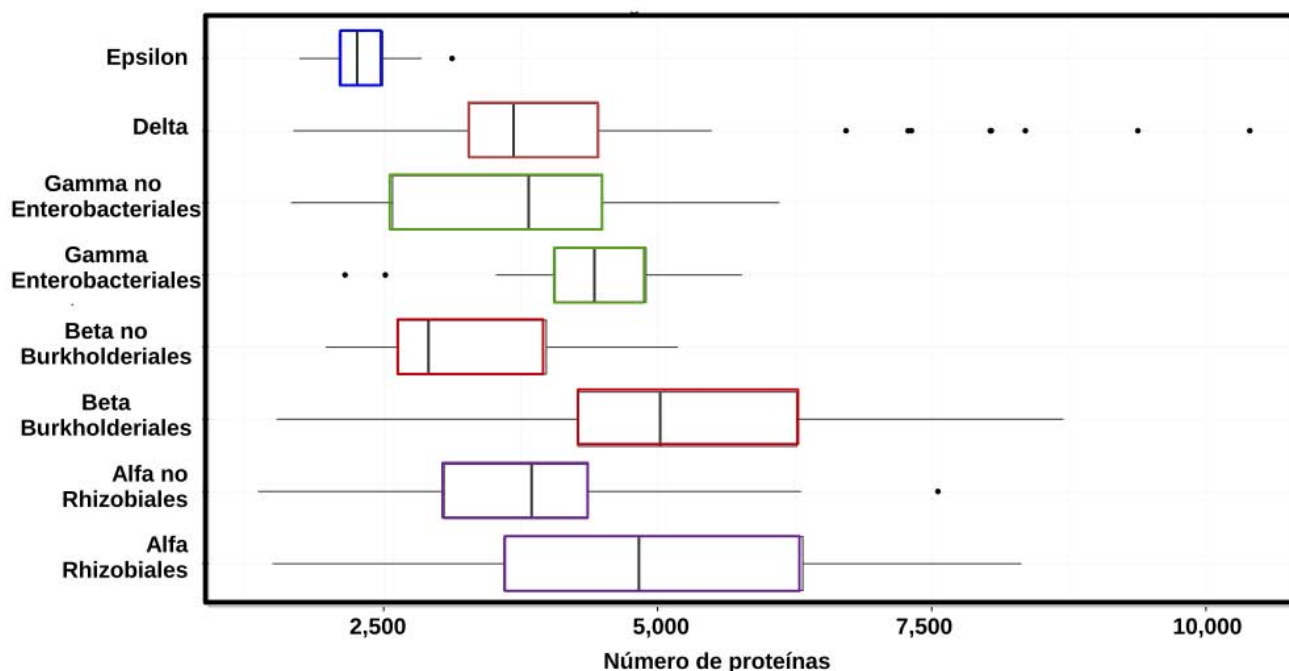


Figura 4. Variación en el tamaño del proteoma entre los diferentes subgrupos de proteobacterias.

3.3 Análisis de similitud y tamaño de los catálogos ancestrales

En la tabla 3 se observa el tamaño de la muestra, así como el promedio, la desviación estándar y los mínimos y máximos del tamaño del genoma de las proteobacterias modernas. Asimismo se muestran los resultados de los análisis de similitud, en función del número de proteínas presentes en el *núcleo* genómico estricto y en el *núcleo* genómico laxo. Además se reporta la resolución calculada para cada análisis.

Tabla 3. Comparación entre el tamaño del genoma de las proteobacterias modernas y las reconstrucciones ancestrales de los distintos grupos. La resolución se calculó como el porcentaje del tamaño promedio que representa el *núcleo* estricto. DE = desviación estándar.

Grupo	Tamaño de la muestra	Tamaño del genoma				Core			Resolución
		Promedio	DE	Min	Max	Estricto	Soft	Total	
Alfaproteobacteria Rhizobiales	51	4962.19	1691.88	1488	8317	382	127	509	7%
Alfaproteobacteria no Rhizobiales	61	3712.04	1269.37	1354	7557	69	65	134	1.8%
Betaproteobacteria Burkholderiales	63	5218.65	1540.74	1524	8702	261	191	452	5%
Betaproteobacteria no Burkholderiales	27	3187.4	885.86	1972	5188	410	138	548	12%
Gammaproteobacteria Enterobacteriales	61	4384.37	635.16	2148	5770	585	119	704	13%
Gammaproteobacteria no Enterobacteriales	71	3748.97	1199.32	1657	6108	173	139	312	4%
Deltaproteobacteria	55	4302.8	1810.18	1677	10400	97	57	154	2%
Epsilonproteobacteria	13	2330.61	401.815	1730	3126	598	84	682	25%

El *núcleo* de mayor tamaño fue el de las proteobacterias epsilon, seguido por las proteobacterias gamma subgrupo Enterobacteriales, mientras que los más pequeños fueron los de las proteobacteria alfa no Rhizobiales y las proteobacterias delta.

En ese sentido, para los grupos que presentaron más proteínas conservadas podemos decir que el análisis tuvo una mayor resolución, mientras que aquellos cuyos catálogos ancestrales fueron pequeños, entonces la resolución fue baja. Aún así la máxima resolución que se alcanzó fue del 25 %.

3.4 Clasificación funcional de las proteínas

A continuación se muestran las gráficas de las categorías funcionales con el número de proteínas que se encontraron en cada una. Debido a que es una gran cantidad de información se presentan los resultados por grupo.

3.4.1 Proteobacteria alfa subgrupo Rhizobiales

La categoría con más proteínas compartidas entre las especies de Rhizobiales fue la de traducción, biogénesis y estructura del ribosoma, seguida por la categoría de replicación, recombinación y reparación, ambas son funciones esenciales de todos los seres vivos. Es interesante observar que el núcleo estricto tiene proteínas que pertenecen a 16 categorías diferentes, lo que indica un ancestro equipado con una diversidad de funciones (fig. 5).

De acuerdo a la reconstrucción, el ancestro ya contaba con distintas proteínas que participan en la producción de energía, en la biosíntesis y metabolismo de aminoácidos, nucleótidos, coenzimas, carbohidratos y lípidos. Además contaba con proteínas para la modificación postraduccional de las cadenas de polipéptidos (fig. 5).

El ancestro poseía también proteínas que participan en la biogénesis de la membrana y de la pared celular así como en el transporte interno de vesículas. La transcripción es otra función que parece estar conservada en el ancestro, además de la transducción de señales y la división celular (fig. 5).

Finalmente también aparecieron conservadas proteínas clasificadas en las categorías de transporte de iones y mecanismos de defensa (fig. 5). Cabe mencionar que se observó un componente importante de proteínas que no están bien caracterizadas, muchas de las cuales se incorporan a la reconstrucción ancestral cuando se relaja el criterio del *núcleo* estricto.

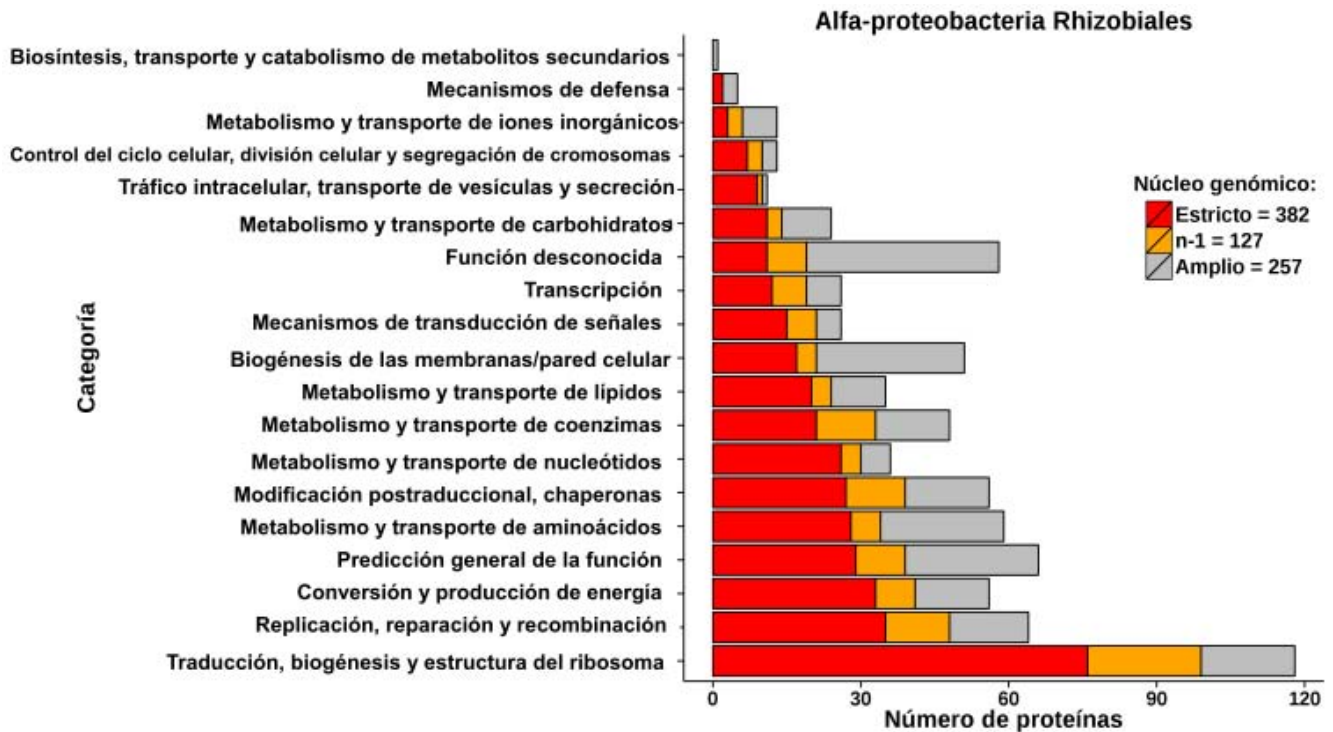


Figura 5. Clasificación funcional de las proteínas del catálogo ancestral de las proteobacteria alfa, subgrupo Rhizobiales.

3.4.2 Proteobacteria alfa no Rhizobiales

A diferencia del grupo anterior, la reconstrucción ancestral de las proteobacterias alfa subgrupo no Rhizobiales, resultó sorprendentemente pequeña. Es notable la falta de conservación en la mayoría de las categorías. Aún así la mejor representada fue la de traducción, biogénesis y estructura del ribosoma, seguida por la de transcripción y por la de metabolismo y transporte de coenzimas (fig. 6).

Otras categorías funcionales presentes en el ancestro fueron las de modificación postraduccional de proteínas, biogénesis de la membrana, producción de energía, metabolismo de lípidos y transporte de membranas, metabolismo de aminoácidos y nucleótidos y finalmente la relacionada con el ciclo celular. Aunque las categorías mencionadas aparecieron en la reconstrucción, hay que enfatizar que no tuvieron más de 10 proteínas cada una, además de que hay un buen número de proteínas cuya función se desconoce (fig. 6).

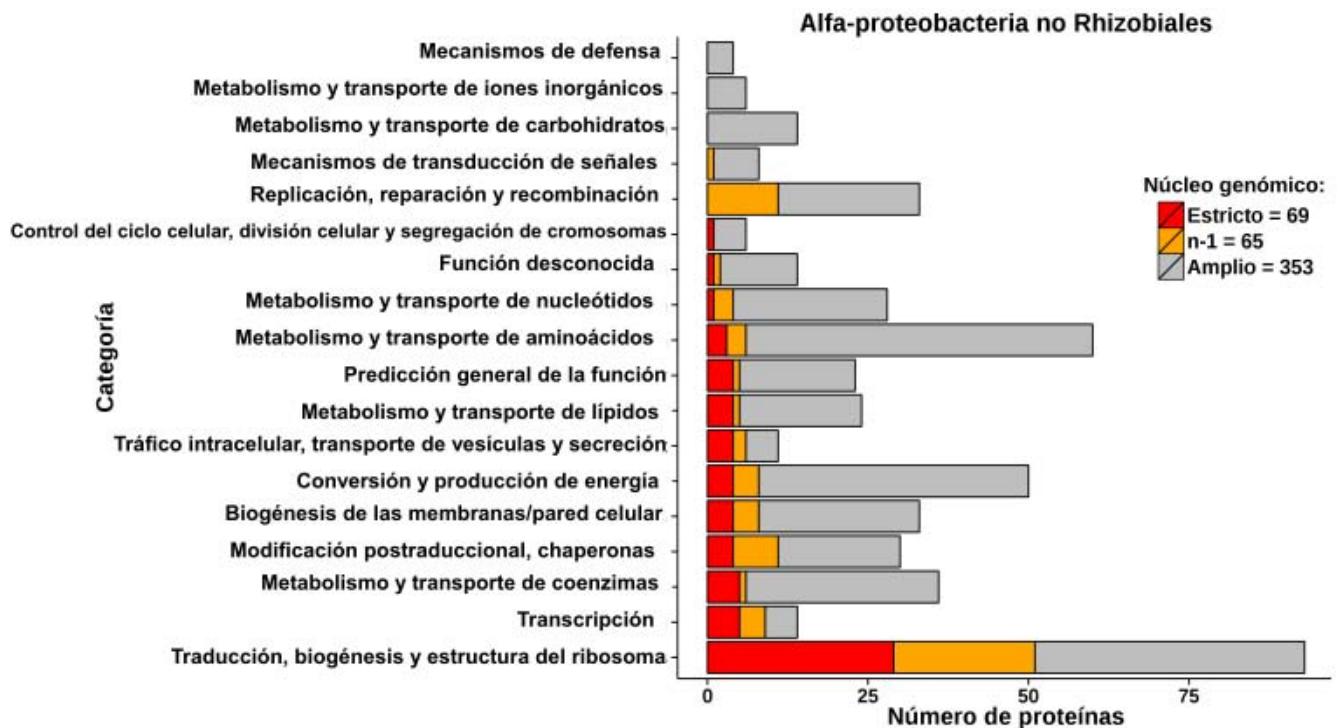


Figura 6. Clasificación funcional de las proteínas del catálogo ancestral de las proteobacterias alfa, no Rhizobiales.

3.4.3 Proteobacterias beta subgrupo Burkholderiales

En la reconstrucción ancestral de bacterias pertenecientes al subgrupo Burkholderiales, las categorías con más proteínas conservadas fueron las de traducción, biogénesis y estructura del ribosoma, y la de replicación, recombinación y reparación. Como ya se mencionó, ambas son funciones esenciales para la vida. Asimismo hubo 18 categorías funcionales presentes en el *núcleo* estricto, lo que sugiere un ancestro que contaba con diversas funciones (fig. 7).

Dicha entidad tenía proteínas necesarias para la producción de energía, también presentaba la capacidad de metabolizar coenzimas, aminoácidos, nucleótidos, carbohidratos y lípidos. Asimismo podía sintetizar su membrana y su pared celular, presentaba modificación postraduccional y también transcripción. Además tenía proteínas que participan en el tráfico intracelular de vesículas, en el transporte de iones, en el ciclo celular, en la transducción de señales y en la biosíntesis y catabolismo de metabolitos secundarios. Hubo también un componente importante de proteínas cuya función se desconoce (fig. 7).

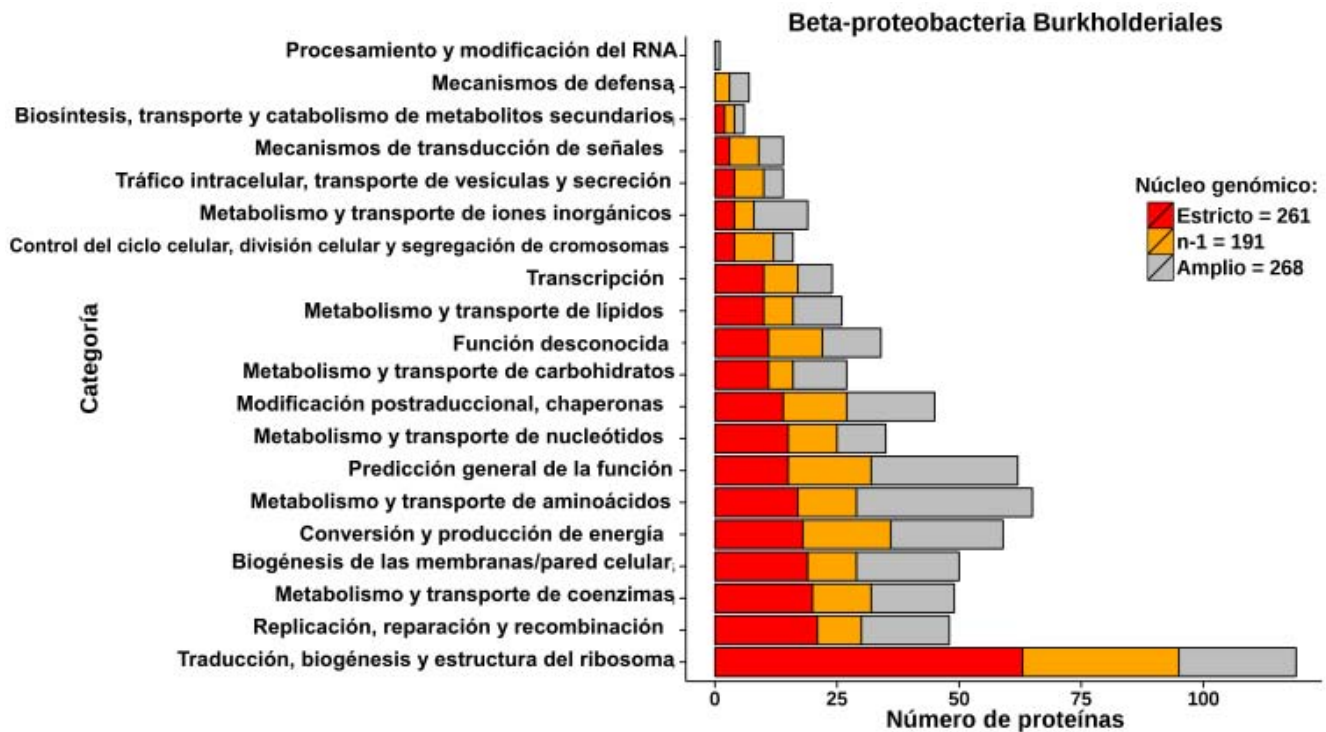


Figura 7. Clasificación funcional de las proteínas del catálogo ancestral de las proteobacterias beta, subgrupo Burkholderiales.

3.4.4 Proteobacterias beta no Burkholderiales

Como es de esperar, el ancestro del grupo de las proteobacterias beta cuenta con las funciones esenciales de la vida, pues la categoría mejor representada fue la de traducción, biogénesis y estructura del ribosoma. En total encontramos 20 categorías funcionales representadas, lo que sugiere que era funcionalmente complejo (fig. 8).

El metabolismo y transporte de aminoácidos y de coenzimas, así como la replicación, recombinación y reparación, son funciones presentes en el ancestro. También contaba con biogénesis de la membrana y de la pared celular, además de tener proteínas para la producción de energía, la modificación postraduccional de cadenas de aminoácidos, para el metabolismo de nucleótidos y lípidos, así como para la transcripción (fig. 8).

Asimismo, el metabolismo de carbohidratos, el tráfico intracelular de vesículas, la transducción de señales y algunas proteínas que participan en el ciclo celular, ya estaban presentes en el ancestro. Con menor número de proteínas encontramos las funciones de biosíntesis de metabolitos secundarios, mecanismos de defensa, movilidad celular, y

finalmente el transporte de iones. Una vez más las categorías que agrupan a proteínas con función desconocida estuvieron bien representadas (fig. 8).

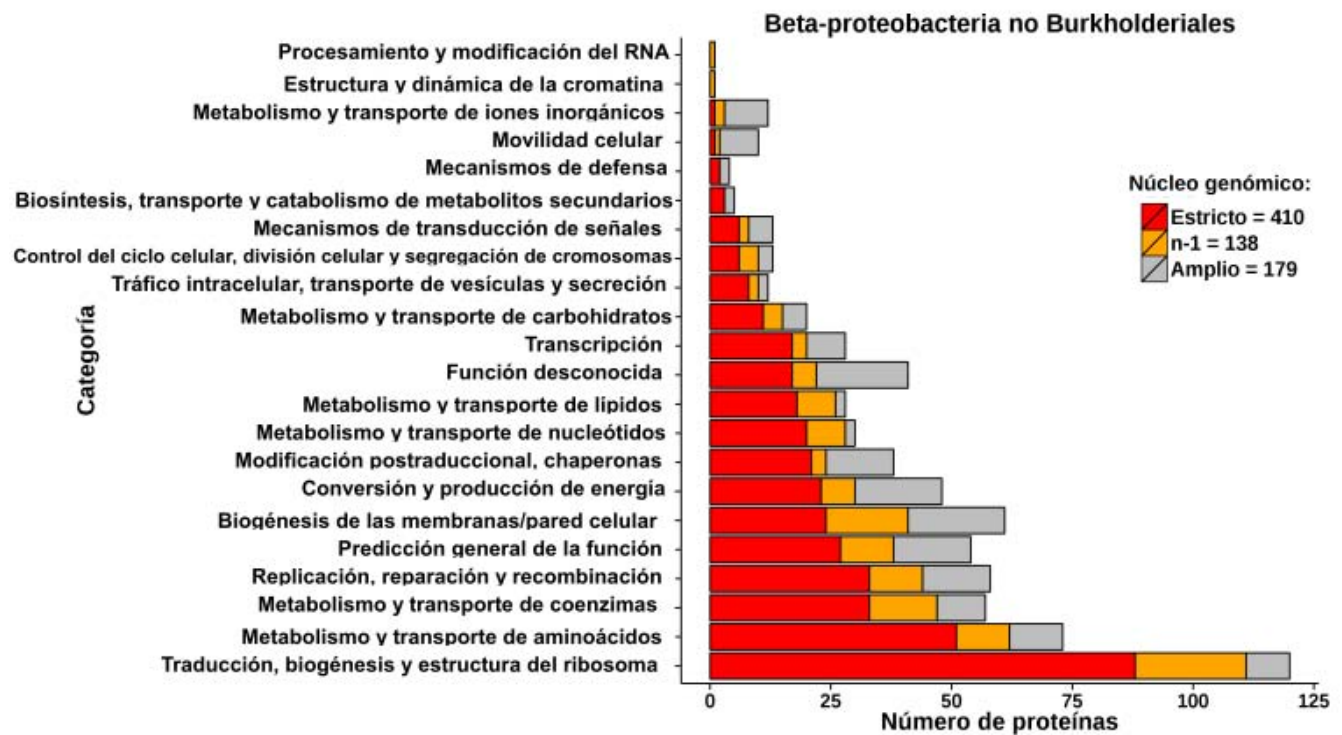


Figura 8. Clasificación funcional de las proteínas del catálogo ancestral de las proteobacterias beta, no Burkholderiales.

3.4.5 Proteobacterias gamma subgrupo Enterobacteriales

El catálogo ancestral de las proteobacterias gamma, subgrupo Enterobacteriales, fue de los más numerosos. Además hubo 18 categorías representadas en el *núcleo* estricto, lo que sugiere un ancestro con complejidad funcional. También sobresale el hecho de que a muchas de las proteínas del *núcleo* estricto no se les ha podido asignar una función (fig. 9).

Una vez más la categoría con mayor número de proteínas conservadas fue la de traducción, biogénesis y estructura del ribosoma, cuestiones esenciales para todo ser vivo. De forma interesante, la siguiente categoría se relaciona con la formación de las membranas y las paredes celulares, lo cual nos sugiere que dichas estructuras han jugado un papel importante en la evolución del grupo. El metabolismo y transporte de coenzimas así como la

replicación, reparación y recombinación, fueron también de las funciones más conservadas (fig. 9).

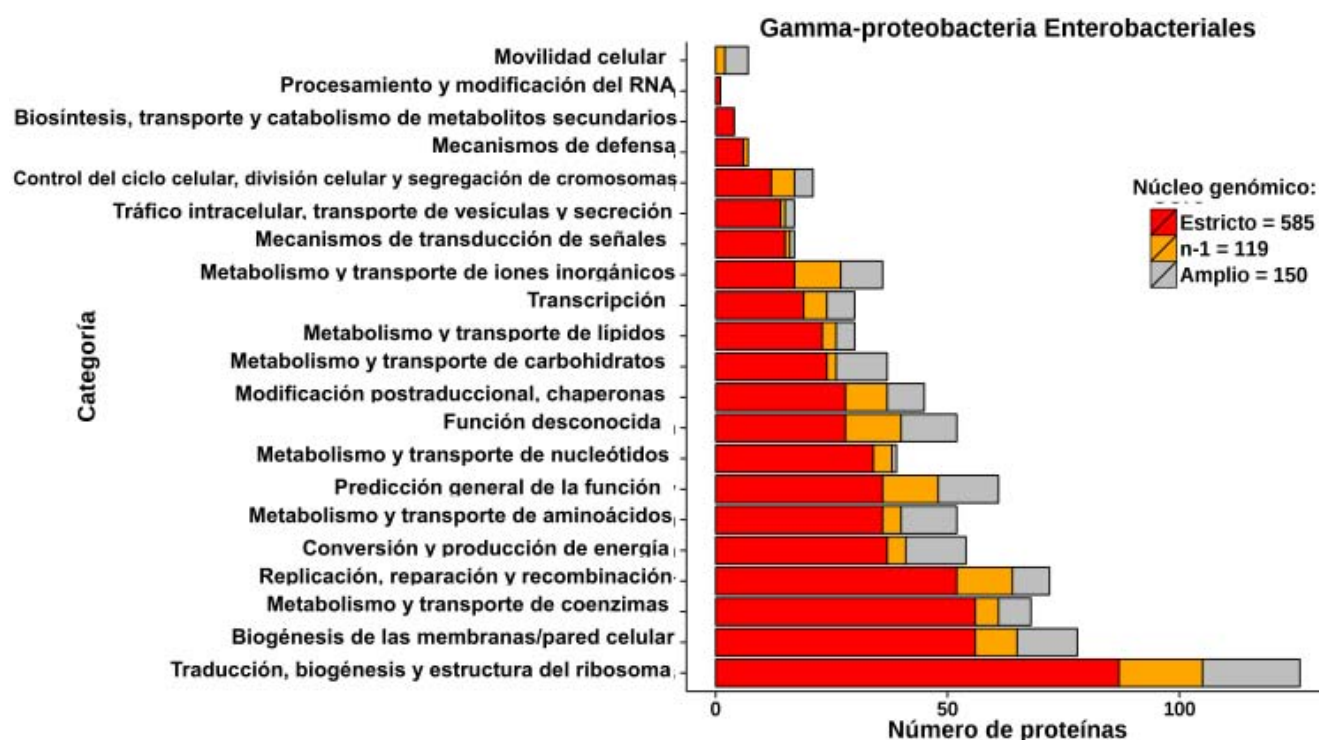


Figura 9. Clasificación funcional de las proteínas del catálogo ancestral de las proteobacterias gamma, subgrupo Enterobacteriales.

El ancestro contaba además con proteínas útiles para la producción de energía, para el metabolismo de aminoácidos, de nucleótidos, de carbohidratos y de lípidos. Tenía proteínas útiles para la modificación postraduccional de polipéptidos, contaba con transcripción, podía transportar iones, presentaba transducción de señales y tráfico intracelular de vesículas y control del ciclo celular. Finalmente presentaba proteínas que se clasifican como mecanismos de defensa y otras que participan en la síntesis y catabolismo de metabolitos secundarios (fig. 9).

3.4.6 Proteobacterias gamma no Enterobacteriales

Dentro de este grupo se mantuvo el patrón de que la categoría más conservada fue la de traducción, biogénesis y estructura del ribosoma, además aunque la reconstrucción

recuperó 19 categorías en el *núcleo* estricto, sobresale el hecho de que la mayoría presentan pocas proteínas conservadas (fig. 10).

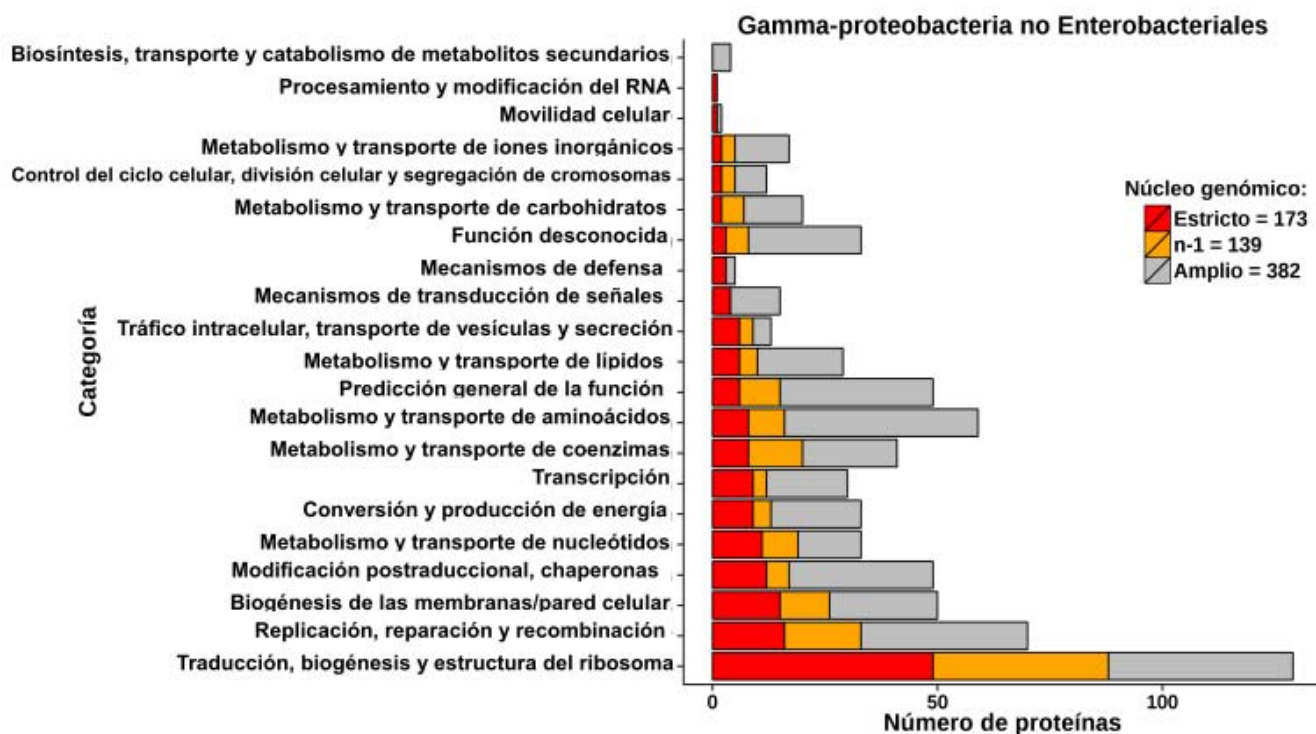


Figura 10. Clasificación funcional de las proteínas del catálogo ancestral de las proteobacterias gamma, no Enterobacteriales.

Aún así en el *núcleo* estricto se presentaron diversas funciones como la replicación, reparación y recombinación, la biogénesis de la membrana y la pared celular, la modificación postraduccional de proteínas, el metabolismo de nucleótidos, coenzimas, aminoácidos y lípidos, así como proteínas cuya función no se ha determinado (fig. 10).

Asimismo contaba con proteínas para producir su energía, realizaba transcripción, movimiento intracelular de vesículas, transducción de señales, además de que contaba con mecanismos de defensa. Las categorías menos conservadas fueron el metabolismo de carbohidratos, aquellas proteínas relacionadas con el ciclo celular, el transporte de iones y la movilidad celular (fig. 10).

3.4.7 Proteobacterias delta

La reconstrucción para este grupo resultó en un ancestro con las funciones esenciales de un ser vivo, empezando por la traducción, biogénesis y estructura del ribosoma. Las siguientes categorías fueron el metabolismo y transporte de nucleótidos, coenzimas y aminoácidos. También contaba con replicación, reparación y recombinación, así como con síntesis de membranas y pared celular. Cabe mencionar que fue un *núcleo* pequeño, pues a excepción de la primera categoría, las 14 restantes tuvieron pocas proteínas conservadas (fig. 11).

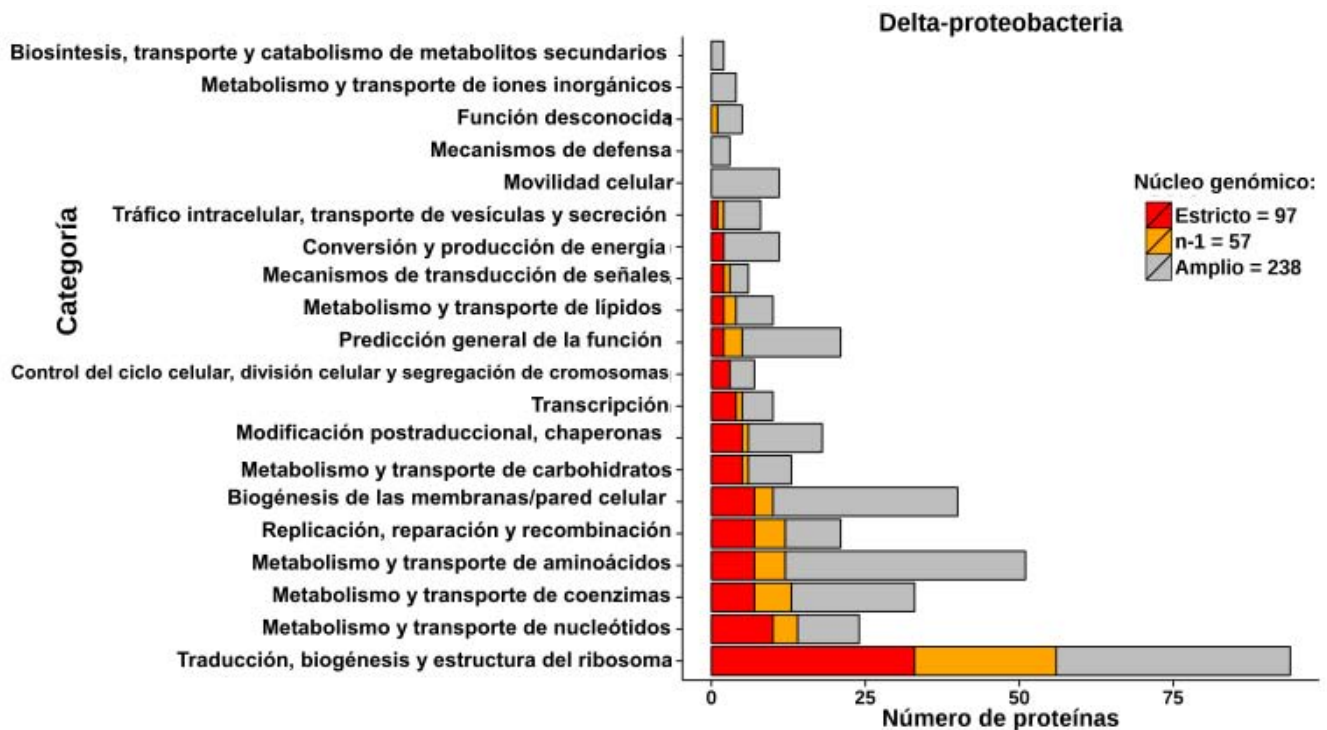


Figura 11. Clasificación funcional de las proteínas del catálogo ancestral de las proteobacterias delta.

Asimismo el ancestro contaba con metabolismo de carbohidratos, modificación postraduccional, transcripción, proteínas que participan en el ciclo celular, metabolismo de lípidos y transducción de señales. De manera notable no aparecieron en el *núcleo* estricto las proteínas relacionadas con la producción de energía, con el tráfico intracelular de vesículas, con el metabolismo y transporte de iones, así como con los mecanismos de defensa y la movilidad celular (fig. 11).

3.4.8 Proteobacterias epsilon

Fue el grupo de las epsilon la reconstrucción con más proteínas conservadas de todas. Además de la traducción, biogénesis y estructura del ribosoma, hubo otras 17 categorías bien representadas en el *núcleo* estricto. Hay que mencionar que un componente importante fueron proteínas cuya función no se ha identificado (fig. 12).

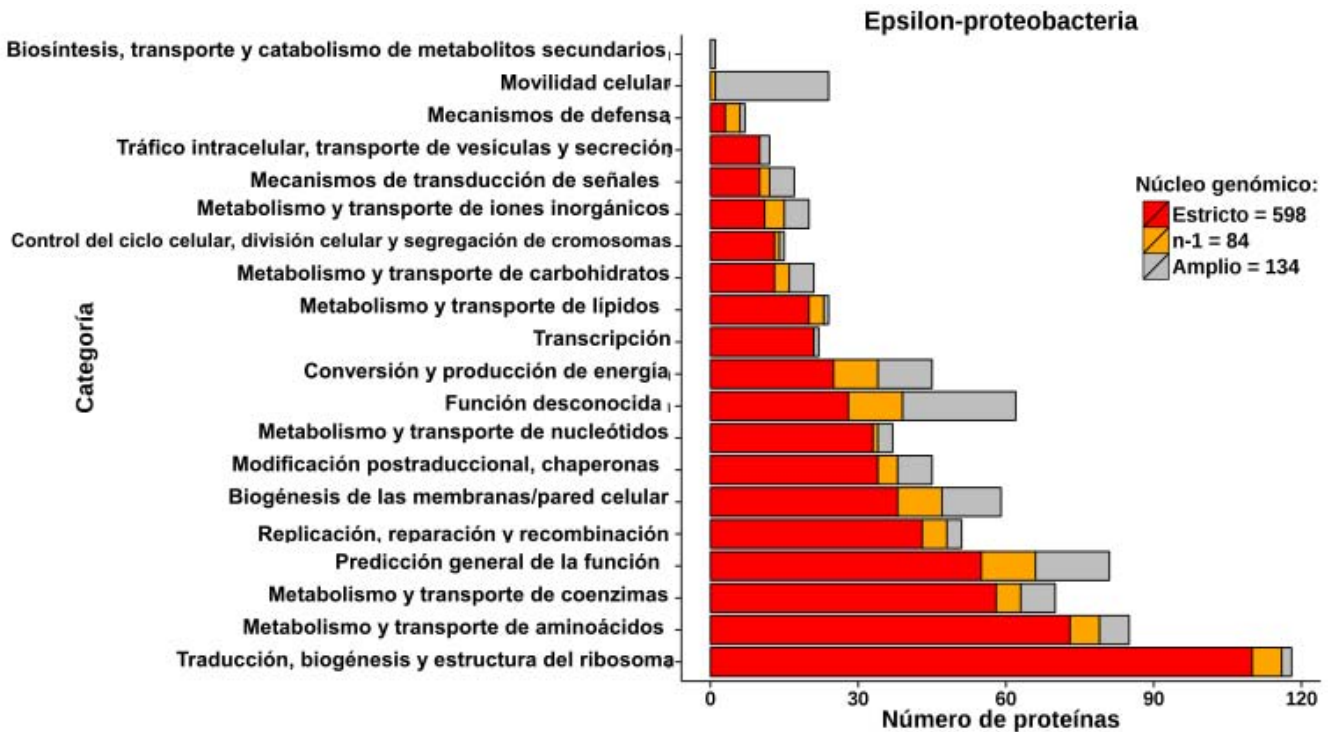


Figura 12. Clasificación funcional de las proteínas del catálogo ancestral de las proteobacterias epsilon.

El ancestro contaba con las funciones de metabolismo y transporte de aminoácidos y coenzimas, con replicación, reparación y recombinación, podía sintetizar sus membranas y la pared celular. Presentaba proteínas responsables de la modificación postraduccional de cadenas de aminoácidos, así como metabolismo de nucleótidos (fig. 12).

Además tenía proteínas asociadas a la producción de energía, a la transcripción, al metabolismo de lípidos y carbohidratos, al control del ciclo celular, al transporte de iones, a la transducción de señales, al tráfico de vesículas y finalmente a mecanismos de defensa.

3.5 Comparación entre las reconstrucciones ancestrales

De manera sobresaliente, en todas las reconstrucciones la categoría con más proteínas conservadas fue la de traducción, estructura y biogénesis del ribosoma, y en ese sentido fue única, pues no hubo ninguna otra que presentara un patrón de conservación tan claro para todos los grupos (fig. 13).

Por lo que respecta a las demás categorías, en la figura 13 se observan dos grandes tendencias. Los colores verde azules y naranjas representan funciones que el método sí detectó y que por lo tanto contienen proteínas con cierto grado de conservación en la estructura primaria. Estas son:

- Metabolismo y transporte de aminoácidos
- Replicación, recombinación y reparación
- Metabolismo y transporte de nucleótidos
- Metabolismo y transporte de coenzimas
- Biogénesis de la membrana, la envoltura y la pared celular
- Predicción general de la función
- Conversión y producción de energía
- Modificación postraducciona de proteínas

El segundo grupo son categorías que en general tuvieron muy pocas proteínas conservadas, o incluso ninguna, representadas por el color rojo en la figura 13. Entre ellas se encuentran:

- Metabolismo y transporte de iones
- Mecanismos de defensa
- Biosíntesis y catabolismo de metabolitos secundarios
- Movilidad celular
- Procesamiento y modificación del RNA
- Tráfico intracelular, transporte de vesículas y secreción
- Transcripción
- Metabolismo y transporte de lípidos
- Proteínas de función desconocida

- Metabolismo y transporte de carbohidratos
- Control del ciclo celular, división celular y segregación de cromosomas
- Mecanismos de transducción de señales

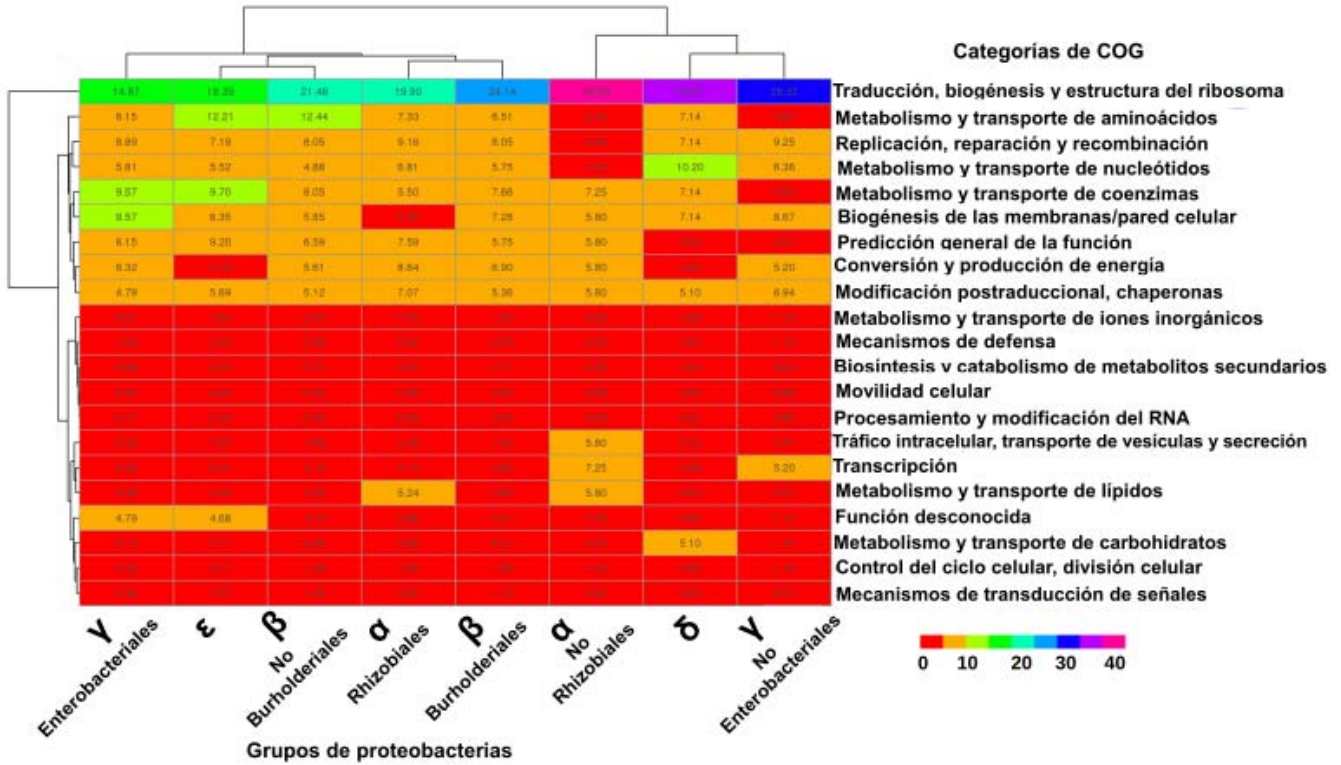


Figura 13. Comparación de las frecuencias relativas de las distintas categorías presentes en el catálogo ancestral de cada grupo de proteobacterias.

Capítulo 4. Biología de Proteobacteria

4.1 Generalidades de las proteobacterias

Actualmente el conjunto de bacterias Gram-negativas conocido como Proteobacteria -del Griego *Proteus*, dios del mar que podía tomar distintas formas (Stackebrandt *et al.* 1988)- constituye uno de los linajes más diversos dentro del dominio Bacteria, no solo por la gran cantidad de especies, pues se conocen más de 2000 contenidas en al menos 500 géneros, sino también por su notable variedad morfológica, metabólica, ecológica y reproductiva (Willey *et al.* 2008).

Como grupo se define únicamente en términos filogenéticos. En un principio se le denominó “bacterias púrpuras” (Woese 1987), debido a que contiene a bacterias fotosintéticas del azufre (PS, por sus siglas en inglés) y no del azufre (PNS, por sus siglas en inglés). Sin embargo a medida que se secuenciaron los genomas de más especies se hizo evidente que hay una enorme diversidad fisiológica (Tabla 4).

Tabla 4. Diversidad de grupos fenotípicos dentro de Proteobacteria. Tomada de Kersters *et al.* 2006.

Fenotipo	Clase
Bacterias aerobias/microaerofílicas, móviles y con forma helicoidal o vibrioide	$\alpha, \beta, \gamma, \delta, \epsilon$
Bacilos o cocos aerobios/microaerofílicos	$\alpha, \beta, \gamma, \delta, \epsilon$
Bacilos anaerobios facultativos	$\alpha, \beta, \gamma, \delta, \epsilon$
Bacilos anaerobios obligados curvados o helicoidales	β, γ, δ
Bacterias simbioses o parásitas de vertebrados e invertebrados	α, γ
Bacterias fototróficas anoxigénicas	α, β, γ
Bacterias quimiolitotróficas aerobias y géneros asociados	β, γ, ϵ
Bacterias no fotosintéticas con gemación o apéndices	α, γ
Bacterias deslizantes no fotosintéticas no frutescentes	β, γ
Bacterias curvas no móviles o raramente móviles	α, ϵ
Bacterias disimilatorias reductoras del azufre o del sulfato	δ
Bacterias con vaina	β
Bacterias deslizantes frutescentes	δ

Los análisis con rRNA e incluso con otras proteínas como marcadores moleculares, han demostrado la existencia de cinco grupos (figuras 14 y 15), actualmente tratados como

clases: Alfa, Beta, Gamma, Delta y Epsilon (Garrity 2001). Recientemente se ha propuesto la existencia de una sexta clase llamada proteobacteria Zeta, la cual contiene a bacterias microaerófilas oxidantes de hierro que se han encontrado en el fondo marino (Emerson *et al.* 2007).

Gupta (2000) dedujo la relación evolutiva utilizando diversas proteínas y propuso el siguiente orden de aparición de los grupos: Epsilon y Delta en primer lugar, después las Alfa y finalmente las Gamma y las Beta (Fig. 14). De hecho estas últimas se encuentran muy cercanamente relacionadas, al grado que se ha sugerido que no son clases separadas y que las Beta representan un subgrupo dentro de las Gamma (Willey *et al.* 2008).

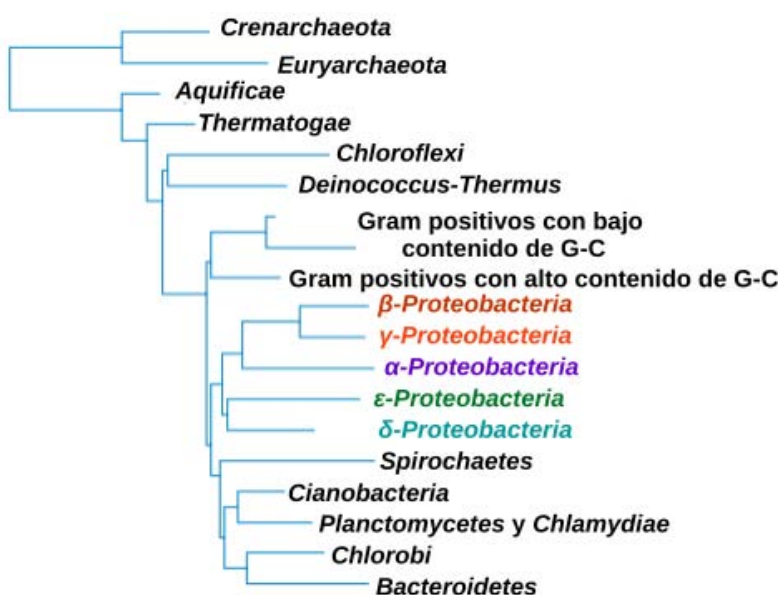


Figura 14. Relaciones filogenéticas de las Proteobacterias con el resto de procariontes.

Tomado de Willey *et al.* 2008.

Es uno de los grupos más estudiados debido a que algunos miembros son patógenos del hombre o de plantas y animales de interés agrícola y ganadero, razón por la cual es uno de los grupos bacterianos mejor conocidos. Tal es el caso de los géneros *Bartonella*, *Brucella*, *Rickettsia*, *Bordetella*, *Burkholderia*, *Neisseria*, *Escherichia*, *Haemophilus*, *Pseudomonas*, *Salmonella*, *Shigella*, *Vibrio*, *Yersinia*, *Campylobacter*, *Helicobacter*, por mencionar algunos.

Otras especies de proteobacterias se utilizan en la industria como por ejemplo *Acetobacter* en la producción de vinagre, *Agrobacterium* en la ingeniería genética, *Gluconobacter* para la obtención de vitamina C o *Zymomonas* en el caso del etanol. Hay especies que producen bioplásticos o incluso hidrógeno gaseoso. Incluso se utilizan en la biorremediación de suelos y cuerpos de agua (Kersters *et al.* 2006).

Las proteobacterias son organismos ubicuos: habitan en todos los ambientes posibles de la Tierra (Willey *et al.* 2008). Por mencionar algunos ejemplos, se ha reportado que son el grupo bacteriano más abundante en el suelo (Spain *et al.* 2009), que las bacterias fotosintéticas anoxigénicas forman al menos el 11% de la comunidad microbiana de la superficie del océano (Kolber *et al.* 2001), o que son de suma importancia en las ventilas hidrotermales y en otros ambientes extremos (Campbell *et al.* 2006).

Ecológicamente las proteobacterias son muy relevantes, pues participan en los ciclos de los elementos más importantes como el carbono, el azufre o el nitrógeno, entre otros (Kolber *et al.* 2001; Kowalchuk & Stephen 2001, Campbell *et al.* 2006). Además pueden encontrarse como colonizadores, productores, depredadores, desintegradores o en interacción biológica formando simbiosis o parasitismo. Asimismo, se han reportado parasitando hospederos que van desde mamíferos, aves, reptiles, peces, insectos, arácnidos, moluscos y hasta los gusanos que habitan en las ventilas hidrotermales (Kersters *et al.* 2006), lo que sugiere que muy probablemente todos los animales tienen un parásito proteobacteriano.

Cabe decir que la mayoría de las especies de proteobacterias descritas hasta ahora se pueden cultivar en el laboratorio, pero sin duda representan una pequeña fracción del total de especies que se sospecha existen en la Tierra. Lo anterior implica que apenas empezamos a conocer la enorme diversidad de este grupo de bacterias Gram-negativas.

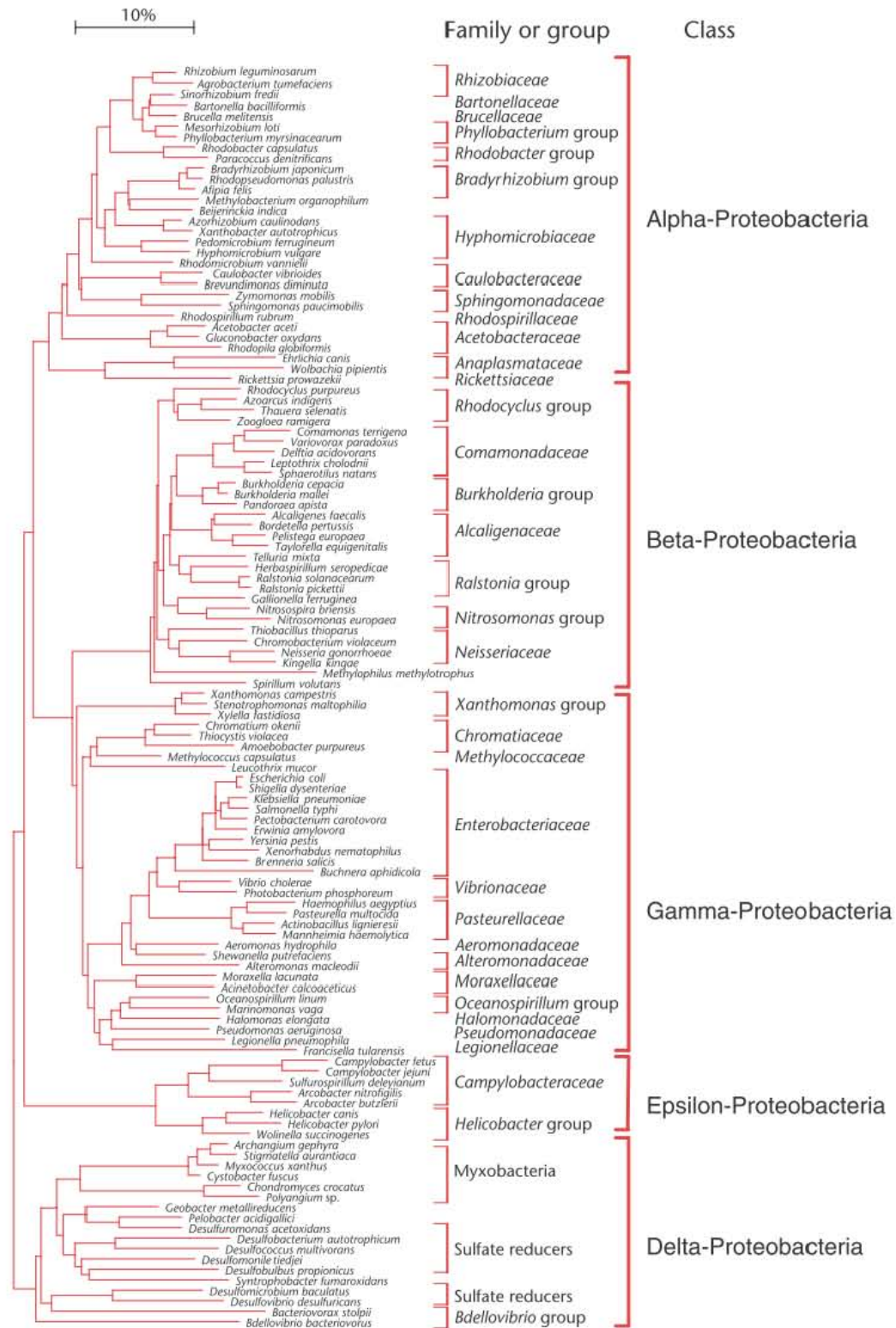


Figura 15. Árbol filogenético de Proteobacteria construido con el 16S rRNA. Tomado de Kersters *et al.* 2006.

4.2 Descripción de la muestra

A continuación se describirán brevemente las características generales de las cinco clases con énfasis en los géneros muestreados.

4.2.1 Proteobacterias alfa

Para este grupo se reconstruyeron dos ancestros, el primero utilizando exclusivamente especies del orden Rhizobiales, mientras que el segundo se construyó con representantes de siete órdenes distintos: Rhodobacterales, Rhodospirillales, Sphingomonadales, Caulobacterales, Rickettsiales, Magnetococcales, Parvularculales. En general, las proteobacteria alfa son morfológica y metabólicamente muy diversas. En ellas se incluyen miembros fotótrofos, quimiolitótrofos, quimiorganótrofos, heterótrofos, acidófilos y metilótrofos (Kerstens *et al.* 2006). Algunos son patógenos importantes de plantas o animales, incluyendo al humano; otros son parásitos obligados y causan enfermedades que se transmiten por picadura de insectos o garrapatas. También pertenecen a este conjunto las bacterias fijadoras de nitrógeno, tanto las de vida libre, como las que forman simbiosis con la raíz de las plantas. Representan el grupo bacteriano más abundante gracias a que sus miembros poseen adaptaciones para vivir en todos los ambientes posibles (Ettema & Andersson, 2009).

4.2.1.1 Proteobacterias alfa subgrupo Rhizobiales

Se analizaron un total de 51 especies (Anexo A) que se distribuyen en 11 familias: Rhizobiaceae (10), Bradyrhizobiaceae (9), Brucellaceae (8), Methylobacteriaceae (5), Phyllobacteriaceae (5), Hyphomicrobiaceae (4), Bartonellaceae (3), Xanthobacteraceae (3), Beijerinckiaceae (2), Methylocystaceae (1), Rhodobiaceae (1) (Fig. 16).

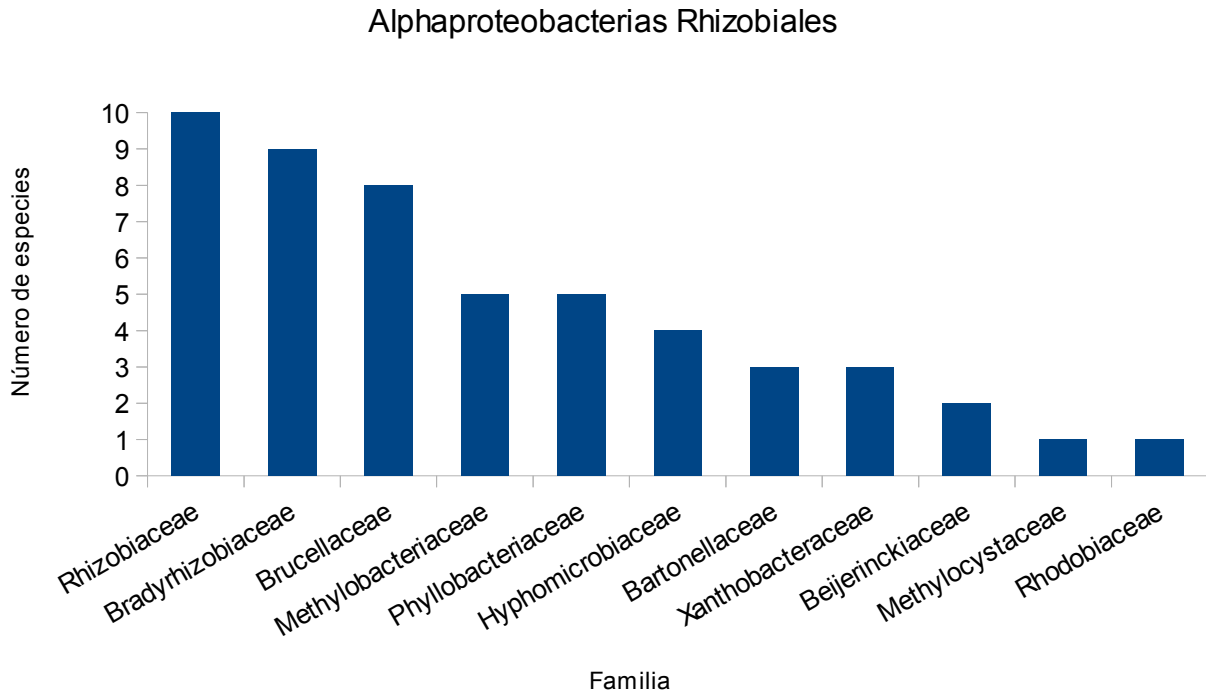


Figura 16. Familias de Rhizobiales representadas en la muestra.

Las familias Rhizobiaceae, Bartonellaceae, Brucellaceae y Phyllobacteriaceae se agrupan en la filogenia (Fig. 17) y comparten el rasgo de que todas interactúan con eucariontes (Kerstens *et al.* 2006).

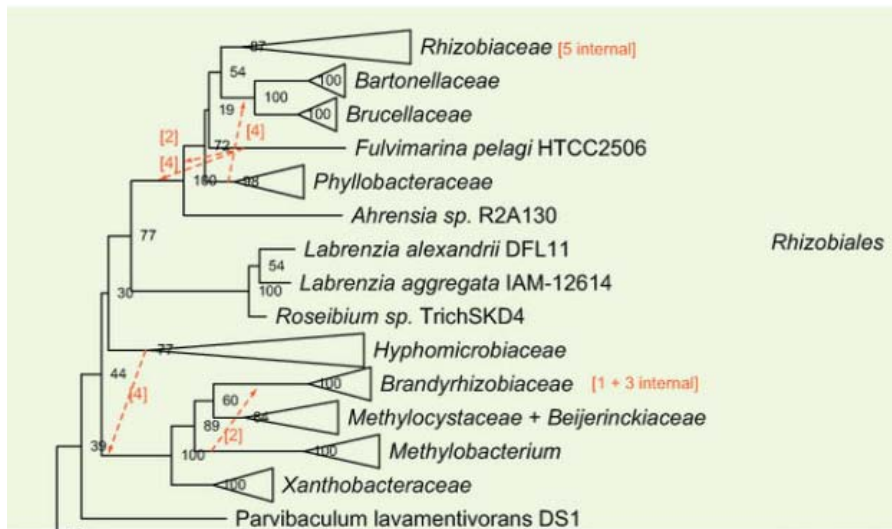


Figura 17. Segmento de la filogenia de proteobacterias alfa donde se observan los grupos pertenecientes a Rhizobiales. Tomado de Ferla *et al.* 2013.

Dentro de Rhizobiaceae, las bacterias del género *Agrobacterium* son comunes en el suelo, y también se encuentran asociadas a plantas; invaden los tejidos vegetales causando enfermedades tumorosas. Los géneros *Rhizobium* y *Sinorhizobium* contienen bacterias fijadoras de nitrógeno, móviles, capaces de inducir la formación de nódulos y vivir en simbiosis con la raíz de las plantas (generalmente leguminosas) (Kerstens *et al.* 2006).

Las bacterias de la familia Bartonellaceae, género *Bartonella*, son patógenas intracelulares que se caracterizan por infectar los glóbulos rojos de diversos vertebrados. Utilizan artrópodos como vectores y son agentes causantes de diversas enfermedades en el humano (Kerstens *et al.* 2006).

En la familia Brucellaceae hubo representantes de los géneros *Brucella* y *Ochrobactrum*; las primeras son patógenas intracelulares facultativas que infectan a mamíferos a nivel mundial y en el humano son responsables de una enfermedad llamada brucelosis. Las segundas viven en el suelo, son aerobias, quimiorganótrofas y en raras ocasiones llegan a infectar a humanos (Kerstens *et al.* 2006).

De la familia Phyllobacteriaceae hay bacterias fijadoras de nitrógeno del género *Mesorhizobium*, las cuales viven asociadas a las raíces de las plantas y son capaces de formar nódulos. Por otro lado las del género *Chelativorans* también forman nódulos pero en las hojas de las plantas y a diferencia de las anteriores no son fijadoras de nitrógeno (Kerstens *et al.* 2006).

Cercanamente emparentada con las Phyllobacteriaceae se encuentra la familia Hyphomicrobiaceae (Fig. 17), con los géneros *Hyphomicrobium*, *Pelagibacterium* y *Rhodomicrobium*. Las especies de *Hyphomicrobium* están bien adaptadas a condiciones oligotróficas en ambientes dulceacuícolas, su citoplasma se extiende y forma estructuras llamadas próstecas. Además su ciclo de vida es dimórfico, consta de una célula sésil con una hifa alargada, la cual mediante gemación forma una célula flagelada y móvil (Poindexter 2006). Son quimiorganótrofas y usan preferentemente con compuestos de un carbono como el metanol. Las del género *Rhodomicrobium* son bacterias móviles, fotoheterótrofas, anoxigénicas, acuáticas, también con próstecas, y se reproducen mediante gemación. En su interior poseen pliegues de la membrana llamados lamelas, además de presentar bacterioclorofila a. Por otro lado el género *Pelagibacterium* incluye bacterias acuáticas, aerobias, móviles y quimioheterotróficas (Kerstens *et al.* 2006).

Otro subgrupo importante es la familia Bradyrhizobiaceae, presente con cuatro géneros. *Bradyrhizobium* son bacterias del suelo, fijadoras de nitrógeno, que forman nódulos en las raíces de diversas especies de plantas en todo el mundo, tienen hábito intracelular y se caracterizan por crecer lentamente. *Nitrobacter* es un género sobresaliente de bacterias nitrificantes porque a diferencia del resto puede crecer heterotróficamente usando acetato o piruvato como fuente de carbono y de energía, además de que en condiciones anaeróbicas puede crecer mediante reducción disimilatoria del nitrato (Kersters *et al.* 2006).

El género *Oligotropha* contiene bacterias con respiración aerobia, aunque algunas desnitrifican, utilizan como fuente de energía el monóxido o el dióxido de carbono y también el hidrógeno, no forman nódulos en plantas ni fijan nitrógeno. En el género *Rhodopseudomonas* hay bacterias fotosintéticas no del azufre (PNSB), metabólicamente versátiles, son de agua dulce y aunque preferentemente crecen de manera fotoheterotrófica, también pueden ser fotoautótrofas usando hidrógeno, tiosulfato o sulfuro como donadores de electrones. Las PNSB pertenecientes a las proteobacterias alfa se encuentran ampliamente distribuidas en la naturaleza, viven en el mar, en el agua dulce y también en ambientes extremos (Kersters *et al.* 2006).

En la familia Methylocystaceae hubo un representante del género *Methylocystis*, como su nombre lo indica, poseen un metabolismo en el que utilizan metano y metanol como su fuente de energía, además son bacterias que pueden formar exosporas (Kersters *et al.* 2006).

La familia Methylobacteriaceae está relacionada con las familias Beijerinckiaceae y Xanthobacteraceae. Contiene bacterias del género *Methylobacterium*, las cuales utilizan preferentemente compuestos de un carbono, son aerobias, el oxígeno es el aceptor terminal de electrones en la respiración, son quimiorganotróficas y facultativamente metilotróficas. Se mueven con un flagelo y se encuentran ampliamente distribuidas en la naturaleza. La presencia de pigmentos sugiere un origen común con las bacterias fototróficas (Kersters *et al.* 2006).

Dentro de Beijerinckiaceae hubo especies de los géneros *Beijerinckia* y *Methylocella*. Las primeras son bacterias fototróficas, fijadoras de nitrógeno y facultativamente metilotróficas, las segundas son aerobias y metanotróficas (Kersters *et al.* 2006).

En Xanthobacteraceae encontramos tres géneros. *Azorhizobium* son bacterias fijadoras de nitrógeno que viven en simbiosis con el tallo de plantas, particularmente leguminosas; *Starkeya* fue la primera especie reportada que es facultativamente heterotrófica, capaz de crecer quimiolitotrofiticamente con compuestos reducidos del azufre (tiosulfato) como fuente de energía, o heterotróficamente en un medio orgánico, son aerobias y ubicuas en el suelo; *Xanthobacter* se encuentran también en el suelo, pero pueden estar asociadas a raíces o incluso ser acuáticas, son capaces de fijar nitrógeno, y también de crecer autotróficamente si en la atmósfera hay hidrógeno, oxígeno y dióxido de carbono, por lo que pertenecen al grupo de bacterias oxidantes del hidrógeno. Por si fuera poco también son metilotróficas, crecen heterotróficamente bajo condiciones aerobias o microaerófilas con compuestos orgánicos como ácidos, alcoholes o algunos carbohidratos como fuente de energía y de carbón (Kerstens *et al.* 2006).

Finalmente el género *Parvibaculum* de la familia Rhodobiaceae corresponde a bacterias heterótrofas capaces de catabolizar el alquilbenceno sulfonato, un tensoactivo muy utilizado presente en detergentes y productos de limpieza (Kerstens *et al.* 2006).

4.2.1.2 Proteobacterias alfa no Rhizobiales

Se analizaron un total de 61 especies pertenecientes a los órdenes Rhodobacterales (18), Rhodospirillales (16), Sphingomonadales (10), Caulobacterales (6), Rickettsiales (3), Magnetococcales (1), Parvularculales (1), cuatro especies del clado SAR11 y dos no clasificadas (Fig. 18) (Anexo B).

El orden mejor representado, Rhodobacterales, es un linaje heterogéneo. En la muestra hubo especies de los géneros *Hirschia* e *Hyphomonas*, bacterias quimioheterotróficas, aerobias, acuáticas, con una hifa alargada y un ciclo de vida dimórfico, que se reproducen por gemación. El género *Maricaulis* también son bacterias con una hifa alargada, marinas (halofílicas), se desarrollan en ambientes oligotróficos y toleran un amplio rango de temperatura y de oxígeno (Kerstens *et al.* 2006).

Dinoroseobacter y *Jannaschia* son géneros de bacterias fototróficas anoxigénicas aerobias, poseen bacterioclorofila a, son heterotróficas y marinas, las primeras se han aislado de dinoflagelados. *Ketogulonicigenium* son bacterias utilizadas en la producción industrial de vitamina C, gracias a que pueden oxidar L-sorbose en un intermediario de la

síntesis de dicha sustancia, se encuentran comúnmente en el suelo y son anaerobias facultativas. *Octadecabacter* son bacterias heterotróficas, psicrófilas, que se han aislado del mar helado en los polos ártico y antártico. *Paracoccus* representa bacterias muy versátiles, aerobias y anaerobias, conocidas primero por su habilidad para desnitrificar. Son capaces de crecer quimioorganotróficamente con compuestos orgánicos como fuente de carbono pero también quimiolitotóticamente con hidrógeno gaseoso o compuestos reducidos del azufre como fuente de energía (Kersters *et al.* 2006).

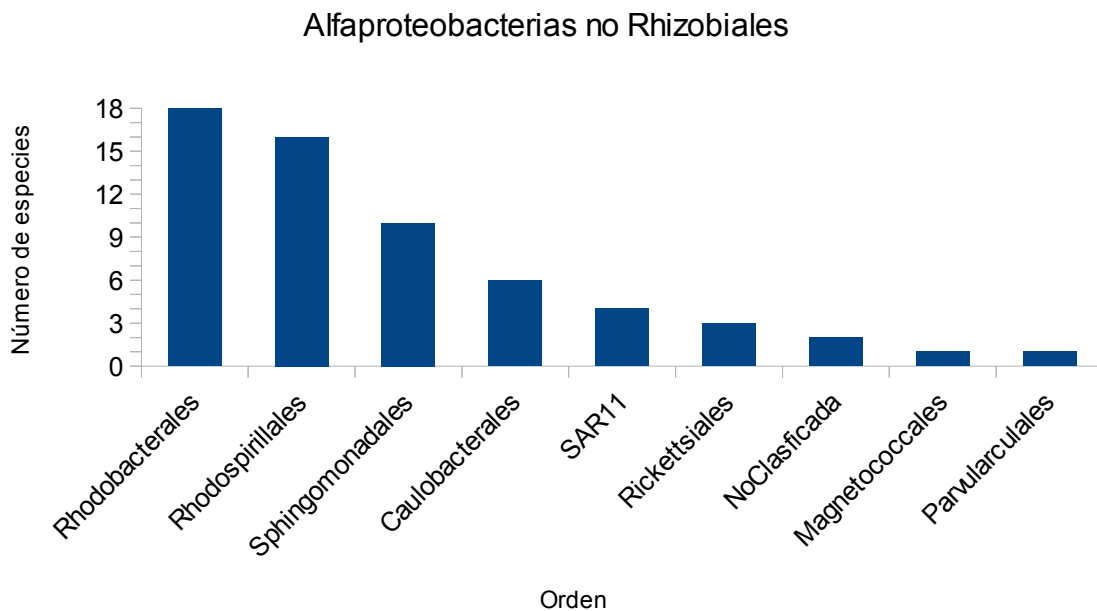


Figura 18. Grupos de proteobacterias alfa no Rhizobiales representados en la muestra.

Phaeobacter, son bacterias heterotróficas, marinas, conocidas por su capacidad para producir antibióticos. *Pseudovibrio* es un género sorprendentemente versátil, son bacterias que se han aislado de invertebrados marinos y pueden oxidar un amplio rango de compuestos orgánicos e inorgánicos para conseguir su energía, carbono y nitrógeno, tanto en condiciones aerobias como anaerobias. *Rhodobacter* es el género tipo, son bacterias púrpuras no del azufre (PNS), acuáticas, que se caracterizan por presentar pigmentos carotenoides y esferoidenos, pero también por su versatilidad: son aerobias y anaerobias, pueden fijar nitrógeno, presentar litotrofia e incluso se coordinan con otras bacterias. *Roseobacter* son bacterias fototróficas aerobias y marinas. Finalmente el género *Ruegeria*

contiene bacterias marinas, aerobias, que no presentan bacterioclorofila ni aparato fotosintético. Cercanamente emparentado se encuentra el orden Parvularculales con una especie del género *Parvularcula*, bacteria marina, aerobia y quimioheterótrofa (Kerstens *et al.* 2006).

El siguiente grupo importante fue el de Caulobacterales, organismos quimiorganótrofos y aerobios, comúnmente acuáticos, que se unen a distintas superficies con ayuda de una estructura con forma de tallo. Presentan un ciclo de vida dimórfico, con una célula sésil y otra móvil. En la muestra se presentan los géneros *Asticcacaulis*, *Brevundimonas*, *Caulobacter* y *Phenylobacterium*. Este último género representa bacterias aerobias, de suelo, capaces de crecer con compuestos artificiales como cloridazón, antipirina y piramidón, los cuales se utilizan en herbicidas comerciales (Kerstens *et al.* 2006).

Relacionado con el grupo anterior (ver Fig. 5), el orden Sphingomonadales contiene bacterias quimiorganótrofas, aerobias, que presentan glicoesfingolípidos en sus membranas. *Erythrobacter* es un género de bacterias que contienen bacterioclorofilas, fotosintetizan, pero de manera interesante no son capaces de crecer fototróficamente en condiciones anaerobias. *Novosphingobium*, *Sphingobium*, *Sphingomonas* y *Sphingopyxis* son géneros de bacterias capaces de degradar compuestos orgánicos muy diversos, se encuentran en el suelo, en cuerpos de agua e incluso se han encontrado en muestras clínicas. El género *Zymomonas* son bacterias presentes en el vino de palma y en el pulque, donde fermentan azúcares en etanol (Kerstens *et al.* 2006).

En el orden Rhodospirillales encontramos a los géneros *Acetobacter*, *Acidiphilium*, *Gluconacetobacter*, *Gluconobacter* y *Granulibacter*, los cuales se caracterizan por ser bacterias acidófilas, estrictamente aerobias ya que utilizan el oxígeno como aceptor de electrones en la respiración, y heterotróficas. Se encuentran en bebidas alcohólicas, frutas, flores y otros materiales vegetales. *Acetobacter* se conoce por su habilidad para formar ácido acético (vinagre) a partir de etanol. *Gluconacetobacter* contiene algunas especies capaces de fijar nitrógeno y de vivir en asociación con plantas como la caña o el café (Kerstens *et al.* 2006).

Otro subgrupo corresponde a Rhodospirillales. Dentro hubo representantes del género *Azospirillum*, bacterias de un característico color rojo, capaces de fijar nitrógeno y muy abundantes en el suelo y en la rizósfera de muchas plantas de interés agrícola.

Endolissoclinum es un género de bacterias simbiotes de tunicados. Como su nombre lo indica las bacterias pertenecientes a *Magnetospirillum* tienen la capacidad de orientarse con el campo magnético, son acuáticas, móviles, con forma de espiral, micraerofilicas y oxidan el hierro. El género *Tistrella* son bacterias aerobias y acuáticas, de interés comercial por su capacidad para formar poliésteres (polihidroxicanoatos). Finalmente el género *Rhodospirillum* contiene bacterias púrpuras no del azufre, con forma de espiral y pueden fotosintetizar (Kersters *et al.* 2006).

El orden Rickettsiales se encuentra cercanamente emparentado con los grupos que contienen al ancestro de la mitocondria. También formó parte de la muestra con los géneros *Orientia* y *Rickettsia*. Ambos contienen bacterias que son parásitas intracelulares obligadas y aerobias, son causantes de enfermedades como el tifus y se transmiten por la picadura de ácaros, cabe decir que no sólo infectan al humano sino también a otros hospederos e incluso se ha reportado que algunas cepas no causan ningún daño (Kersters *et al.* 2006).

Se presentaron cuatro especies clasificadas como SAR11, por ejemplo candidatos de los géneros *Pelagibacter* y *Puniceispirillum*. Son bacterias heterótrofas, ubicuas y muy abundantes en los océanos, oligotróficas, que dependen de carbono orgánico y del nitrógeno disuelto en el agua (Morris *et al.* 2002).

Finalmente el orden Magnetococcales se coloca en la base de las proteobacterias alfa y presentó el género *Magnetococcus*, bacterias capaces de formar cristales magnéticos (magnetosomas) para orientarse, son marinas y versátiles, pues crecen quimiolitotótróficamente con tiosulfato o sulfuro, o también quimiorganotótróficamente con acetato (Bazylinski *et al.* 2013).

Cabe decir que en la muestra hubo dos especies no clasificadas: *Micavibrio* y *Polymorphum*. Las primeras son bacterias depredadoras que exhiben un comportamiento tipo "vampiro", pues se unen a otras bacterias Gram negativas y las succionan hasta matarlas. Las segundas son de interés ecológico pues se encontraron en sitios contaminados con aceite y son capaces de degradarlo (Li *et al.* 2011).

4.2.2 Proteobacterias beta

Al igual que con las proteobacterias alfa, para las proteobacterias beta se reconstruyeron dos ancestros: uno con miembros del orden Burkholderiales, y el otro con el

resto de las especies contenidas en la muestra. En general son un grupo metabólica, morfológica y ecológicamente muy heterogéneo. Contiene bacterias fotosintetizadoras no del azufre (PNSB), varios quimiolitótrofos, algunos metilótrofos, un gran número de quimiorganótrofos, algunas bacterias que fijan nitrógeno, y patógenos importantes, tanto de plantas como de animales, incluyendo al humano. Algunas especies son de interés biotecnológico por su capacidad para degradar compuestos (Kersters *et al.* 2006).

4.2.2.1 Proteobacterias beta subgrupo Burkholderiales

Se analizaron un total de 63 especies pertenecientes a 25 géneros, repartidos en cuatro familias (Anexo C): Burkholderiaceae (29), Comamonadaceae (17), Alcaligenaceae (10), Oxalobacteraceae (4) y tres especies no clasificadas (Fig. 19).

En la familia Burkholderiaceae se consideraron los géneros *Burkholderia*, *Ralstonia*, *Cupriavidus* y *Polynucleobacter*. El primero es un género con organismos aerobios, que causan enfermedades a humanos, animales o plantas, además de tener especies útiles para la biorremediación o para promover el crecimiento vegetal (Woods & Sokol 2006). *B. cepacia* es una especie sobresaliente, ya que puede utilizar más de 100 compuestos orgánicos para obtener su energía, se conoce bien porque infecta de manera oportunista a humanos. También hay representantes capaces de fijar nitrógeno que incluso pueden formar nódulos en las raíces de las plantas (Willey & Sherwood, 2008).

El género *Ralstonia* contiene bacterias antes consideradas especies de *Pseudomonas*, son comunes en el suelo, algunas especies son patógenas de plantas de interés agrícola, también se presentan patógenos oportunistas del humano. Ciertas especies se requieren por la industria debido a que degradan compuestos xenobióticos, a que son resistentes a metales pesados o porque producen bioplásticos (Kersters *et al.* 2006). El género *Cupriavidus* contiene organismos aerobios, con flagelos peritricios y que pueden ser quimiheterótrofos o quimiolitótrofos, son comunes en el suelo y de interés debido a que algunas especies tienen resistencia a metales pesados (Vandamme & Coenye 2004). El género *Polynucleobacter* corresponde a bacterias inicialmente descritas como simbiotes de un ciliado, pero que ahora se sabe son de vida libre y ubicuas en ambientes de agua dulce, pueden formar hasta el 11% del bacterioplancton. Son quimiorganótrofas y utilizan sustratos

de bajo peso molecular derivados de la fotoxidación de la materia orgánica presente en el agua (Hahn *et al.* 2012).

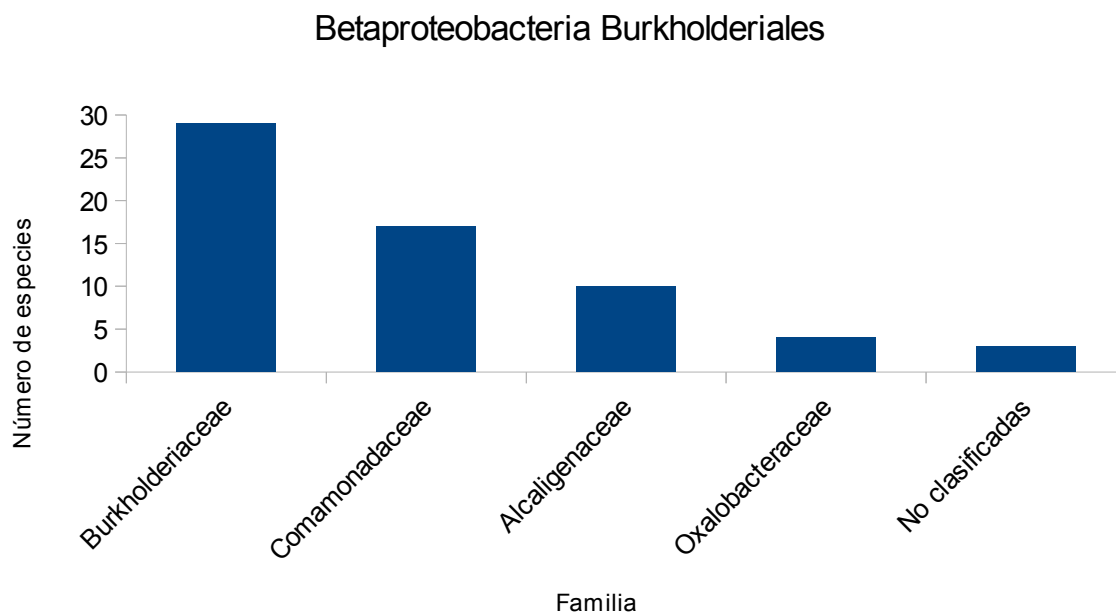


Figura 19. Familias representadas en la muestra de las proteobacteria beta Burkholderiales.

Emparentada con la anterior se encuentra la familia Oxalobacteraceae con los géneros *Herbaspirillum*, *Herminiimonas*, *Minibacterium* y *Collimonas*. Las bacterias del género *Herbaspirillum* son diazótroficas, es decir, fijan nitrógeno y se encuentran colonizando el interior de raíces, tallos u hojas de gramíneas (Kerstens *et al.* 2006). El género *Herminiimonas* contiene bacterias aerobias, organotróficas, inicialmente aisladas de manantiales de agua mineral. Actualmente se han descrito más especies de hábitats diversos, por ejemplo se han encontrado en sitios muy contaminados con metales pesados, donde pueden oxidar arsénico. *Minibacterium* contiene una especie recién descrita, de tamaño notablemente pequeño, es acuática, oligotrófica, resistente a metales pesados y a antibióticos (Audic *et al.* 2007). Las bacterias del género *Collimonas* son aerobias, se caracterizan porque se pueden alimentar de hifas gracias a que digieren la quitina, entre otros compuestos orgánicos (De Boer *et al.* 2004).

Otro subgrupo importante relacionado con los anteriores es la familia Comamonadaceae, la cual presentó 11 géneros. Las bacterias de *Comamonas* y *Delftia* son

ubicuas en el suelo y en cuerpos de agua, incluso en ambientes contaminados, además se han aislado de muestras clínicas y de hospitales, así como de la sangre de caballo o conejo. Son aerobias, móviles, y quimiorganótrofas (Willems & De Vos 2006). Además *Comamonas* y *Delftia* degradan compuestos orgánicos aromáticos, por lo que son de interés para la biodegradación tóxica en sitios contaminados. En particular la especie *D. acidovorans* es famosa porque produce nano partículas de oro (Kerstens *et al.* 2006).

El género *Acidovorax* contiene bacterias comunes en suelos y en cuerpos de agua, son quimiorganótrofas y aerobias, sin embargo, pueden crecer litoautotróficamente con hidrógeno como fuente de energía (Kerstens *et al.* 2006). El género *Alicyclophilus* contiene una bacteria facultativamente desnitrificante, su nombre se debe a su capacidad de utilizar ciclohexanol como fuente de carbono y nitrato como aceptor de electrones (Mechichi *et al.* 2003). El género *Verminiphrobacter* comprende bacterias flageladas, heterotróficas, que viven en simbiosis con el nefridio de una especie de lombriz (Pinel *et al.* 2008). El género *Polaromonas* contiene bacterias heterotróficas, psicrófilas y marinas.

Rhodoferax y *Rubrivivax* son bacterias fotosintéticas no del azufre (PNSB) comunes en cuerpos de agua. En particular *Rhodoferax ferrireducens* es una especie psicrotolerante que reduce azúcares, pero de manera interesante, no transfiere electrones al oxígeno sino al hierro, lo que genera una corriente eléctrica (Kerstens *et al.* 2006).

Variovorax se llama así porque puede crecer con diversos compuestos orgánicos (quimiorganótrofo), pero también es litoautótrofo ya que puede usar gas hidrógeno como fuente de energía (al igual que *Rhodoferax*), son organismos aerobios (Kerstens *et al.* 2006).

El género *Ramlibacter* contiene bacterias aerobias, quimiorganótrofas, que se encuentran en el suelo. Su ciclo de vida es único; presentan células esféricas, muy resistentes, llamadas quistes, las cuales se dividen por fisión binaria y presentan otras células con forma de bastón, móviles, que no se pueden dividir pero sirven para colonizar otros ambientes y luego pasar al estado de quiste (Kerstens *et al.* 2006). El género *Thiomonas* contiene bacterias antes consideradas *Thiobacillus*, capaces de crecer en ambientes ácidos con compuestos reducidos del azufre (Robertson & Kuenen, 2006).

Otro subgrupo emparentado con los anteriores es la familia Alcaligenaceae, presente con cinco géneros. *Achromobacter* es un género de bacterias aerobias, potencialmente patógenas, encontradas en muestras clínicas y de hospitales, pero también en suelos y

cuerpos de agua, que además bajo condiciones anaerobias desnitrifican, ya que reducen óxidos de nitrógeno a nitrógeno gaseoso (Kersters *et al.* 2006). El género *Bordetella* contiene bacterias aerobias, que se han adaptado a vivir en asociación con otros organismos, ya sea de manera patógena o en comensalismo, por lo que enferman comúnmente a mamíferos y aves en las vías respiratorias; una especie es responsable de la tosferina en el humano (Kersters *et al.* 2006).

El género *Advenella* contiene bacterias aerobias, heterótrofas, comunes en muestras clínicas. Dentro de *Pusillimonas* encontramos miembros de suelo o acuáticos, aerobios y en particular la especie de la muestra puede degradar diésel (Cao *et al.* 2011). El género *Taylorella* son especies microaerofílicas aisladas del tracto urinario de caballos y de burros, sólo en el primer caso tienen efectos patógenos (Hébert *et al.* 2012).

Finalmente hubo tres especies no clasificadas. *Leptothrix* son bacterias dulceacuícolas, oligotróficas, que crecen de manera filamentosa, son quimioheterótrofas y forman a su alrededor una cápsula de hierro y manganeso (Spring, 2006). Las bacterias del género *Methylibium* son metilótrofas y aerobias. La especie presente en la muestra es facultativamente metilótrofa ya que utiliza metanol como fuente de carbono, pero puede crecer heterotróficamente con otros compuestos orgánicos, y es de interés porque puede degradar aditivos presentes en la gasolina (Nakatsu *et al.* 2006). Hubo una especie designada tentativamente como *Thiomonas*, recuperada de un drenaje, la cual vive en ambientes ligeramente ácidos y puede oxidar compuestos de azufre.

4.2.2.2 Proteobacterias beta no Burkholderiales

De las 27 especies analizadas (Anexo D) solamente hay dos sin clasificar y el resto se distribuyen en los órdenes Neisseriales (6), Rhodocyclales (6), Methylophilales (5), Nitrosomonadales (5), Gallionellales (2) e Hydrogenophilales (1) (Fig. 20).

En el grupo Methylophilales estuvieron representados los géneros metanótrofos de las proteobacterias beta: *Methylobacillus*, *Methylotenera* y *Methylovorus*. Este grupo se encuentra relacionado con las Nitrosomonadales, en el cual se presentaron los géneros *Nitrosomonas* y *Nitrospira*, entre los que encontramos bacterias aerobias, quimioautótrofas capaces de oxidar amoníaco (NH₃), las primeras tienen forma de bacilo y las segundas de espiral (Kersters *et al.* 2006).

Betaproteobacterias no Burkholderiales

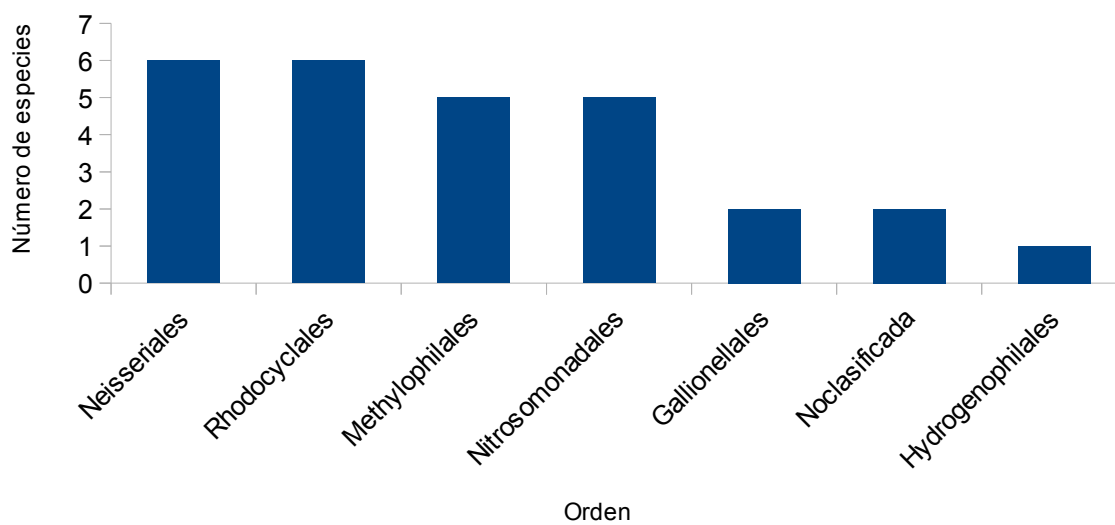


Figura 20. Órdenes representados en la muestra de proteobacteria beta no Burkholderiales.

En el orden Gallionellales encontramos bacterias de los géneros *Gallionella* y *Sideroxydans*. Las primeras son acuáticas, quimiolitótrofas, capaces de oxidar hierro, que viven en ambientes microaerófilos. Por su metabolismo "fijan" el hierro y lo transforman a hidróxido férrico (Halbach *et al.* 2001). De manera semejante el género *Sideroxydans* contiene miembros dulceacuícolas y oxidantes del hierro, que se han aislado de cuerpos de agua contaminados (Kerstens *et al.* 2006).

El único género dentro de Hydrogenophilales fue *Thiobacillus*, bacterias despigmentadas, quimiolitótrofas, que oxidan compuestos de azufre. La especie de la muestra es capaz de desnitrificar (Kerstens *et al.* 2006).

Dentro de Rhodocyclales se presentaron cinco géneros. Las bacterias de *Aromatoleum* se caracterizan por degradar una variedad de compuestos orgánicos en condiciones aerobias y anaerobias, por lo que son de interés en la biodegradación de contaminantes, además de que son facultativamente desnitrificadoras. Las bacterias del género *Azoarcus* son diplobacilos, móviles, se reconocen especies asociadas a raíces de plantas, capaces de fijar nitrógeno y especies de suelo capaces de degradar compuestos aromáticos en condiciones desnitrificantes (Kerstens *et al.* 2006). El género *Thauera* son bacterias acuáticas o de suelos húmedos, aisladas de plantas de tratamiento y reconocidas

porque pueden degradar compuestos aromáticos en condiciones anóxicas, además se reportan como organismos desnitrificantes (Liu *et al.* 2013). Los géneros *Dechloromonas* y *Dechlorosoma* son bacterias capaces de reducir el perclorato, sustancia común en herbicidas y en la industria de las municiones, algunas especies fijan nitrógeno.

Cercanamente relacionado al grupo Rhodocyclales (ver Fig. 5) se encuentra el orden Neisseriales, presente con cuatro géneros. Las bacterias de *Chromobacterium* contienen el pigmento morado violaceína, son facultativamente anaerobias y quimiorganótrofas, ya que fermentan carbohidratos. Pueden reducir nitratos y nitritos a amoníaco. Son organismos saprofiticos de suelo y cuerpos de agua, además de ser patógenos oportunistas del humano (Kersters *et al.* 2006). *Laribacter* contiene bacterias patógenas, encontradas en peces de agua dulce, incapaces de utilizar carbohidratos como fuente de energía pero sí como fuente de carbono (Woo *et al.* 2009). En el género *Pseudogulbenkiania* encontramos bacterias desnitrificantes de suelo. Finalmente en *Neisseria* encontramos diplococos quimiorganótrofos, comensales de las membranas mucosas de mamíferos, hay patógenos importantes como los causantes de gonorrea o meningitis (Kersters *et al.* 2006).

De las dos especies sin clasificar una no tiene nombre asignado y la otra fue del género *Accumulibacter*, las cuales son bacterias aisladas de aguas tratadas de desecho, de interés por su capacidad para remover el fósforo del medio y acumularlo en su interior como polifosfato (Seviour *et al.* 2003).

4.2.3 Proteobacterias gamma

Al igual que en los grupos anteriores la muestra se dividió en dos, reconstruyendo al ancestro de las Enterobacteriales, y al ancestro del resto de las especies. Como ya se mencionó, las proteobacterias gamma son un grupo monofilético (que de hecho incluye a las proteobacterias beta). Son la clase más numerosa, quizá por los miembros conocidos que son patógenos tanto del humano como de otros animales y de plantas. El grupo incluye especies púrpuras del azufre (PSB), muchas quimiorganótrofas y algunas especies quimiolitótrofas, en particular oxidantes del azufre o del hierro (Kersters *et al.* 2006).

4.2.3.1 Proteobacterias gamma subgrupo Enterobacteriales

Las 61 especies analizadas pertenecen a 24 géneros, todos de la familia Enterobacteriaceae, además de una especie no clasificada (Anexo E). En general son bacilos móviles, facultativamente anerobios, capaces de catabolizar azúcares y de producir ácidos por la fermentación de la glucosa (Francino *et al.* 2006).

Las bacterias del género *Escherichia* son quimiorganótrofas y habitan el interior del tracto intestinal de animales de sangre caliente, algunas especies son comensales pero otras son patógenas. Son fáciles de reproducir en laboratorio, lo que ha hecho de *E. coli* la especie bacteriana probablemente mejor conocida, se puede hallar también en suelos o en cuerpos de agua contaminados con materia fecal (Welch 2006). *Salmonella* es el género hermano de *Escherichia* y al igual que *Shigella*, son bacterias fermentadoras (no pueden fermentar lactosa) y habitan el tracto digestivo de animales. Algunas especies de *Salmonella* son responsables de la fiebre tifoidea, paratifoidea, salmonelosis o de infecciones producidas por el consumo de alimentos contaminados, en el caso de *Shigella* son el agente causante de disentería o shigellosis (Kersters *et al.* 2006).

El género *Citrobacter*, cercanamente emparentado con los géneros anteriores, contiene bacterias capaces de fermentar compuestos orgánicos, como su nombre lo indica el citrato es la fuente de carbono. Son comensales del tracto digestivo de animales y se conocen por su patogenicidad en humanos, en el ambiente se dispersan junto con la contaminación fecal (Borenshtein & Schauer 2006).

Las bacterias pertenecientes a *Enterobacter* poseen fermentación láctica y son comunes en el suelo y en cuerpos de agua, curiosamente antes del uso de antibióticos se encontraban raramente como patógenos, pero en la actualidad son comunes en infecciones de vías respiratorias y urinarias (Grimont & Grimont 2006). *Cronobacter*, antes *Enterobacter*, son bacterias conocidas por ser patógenas oportunistas de neonatos (Kersters *et al.* 2006). *Rahnella* son bacterias capaces de fermentar compuestos orgánicos y también de oxidar fenilalanina desaminada. Se encuentran ampliamente distribuidas en el ambiente, sobre todo en sitios húmedos, sobresale su hallazgo en el intestino de un mamut de 12 000 años de edad. Son comunes en los alimentos y pueden causar de manera oportunista enfermedades en el humano (Janda 2006).

Klebsiella contiene bacterias con cápsula, fermentativas, que utilizan como sustrato compuestos orgánicos. Las especies se diferencian por las enfermedades que provocan, sin embargo, se han reportado en el suelo, en cuerpos de agua y en partes de plantas (raíces, hojas, frutos), además de que algunas cepas pueden fijar nitrógeno (Brisse *et al.* 2006), algunas especies se reclasificaron como *Raoultella*.

Las especies de *Serratia* producen un pigmento que les da un color rojizo, son comunes en muestras clínicas pero también en plantas, en el suelo, en cuerpos de agua, en otros animales vertebrados y hasta en insectos. Es patógeno oportunista del humano (Grimont & Grimont 2006b).

Sodalis es un género de endosimbiontes de la mosca tsetse *Glossina mositans*, los cuales habitan en diversos tejidos del hospedero (Janda 2006). *Xenorhabdus* y *Photorhabdus* son simbiosis de nemátodos entomopatógenos (Janda 2006). *Hamiltonella* son bacterias endosimbiontes de áfidos y de homópteros, reportados como mutualistas nutricionales, pues proveen al hospedero de nutrientes importantes (Su *et al.* 2014). *Shimwellia* son bacterias que se han aislado del tracto digestivo de cucarachas y no son patógenas del humano.

Los géneros *Pantoea*, *Pectobacterium* y *Erwinia* contienen bacterias asociadas a plantas, por lo general patógenas, que provocan en sus hospederos necrosis, maceración del tejido vegetal, oclusión de tejidos vasculares o tumores (Kado 2006). El género *Dickeya*, antes *Pectobacterium*, contiene bacterias patógenas de plantas, por lo general herbáceas.

Yersinia contiene varias especies, tres de ellas son conocidas por causar enfermedades al humano y otros animales, en particular una es el agente etiológico de la peste. *Edwardsiella* son bacterias dulceacuícolas o encontradas al interior de animales como anfibios, reptiles, aves o más comúnmente en peces. Son patógenos oportunistas del humano (Abbot & Janda 2006).

Los géneros *Morganella*, *Proteus* y *Providencia* son bacilos móviles, con flagelos peritricios, que comparten la capacidad de oxidar fenilalanina desaminada y de hidrolizar urea (a excepción de unas especies de *Providencia*), algunas especies son patógenas oportunistas del humano, en particular en pacientes con cateter urinario (Manos & Belas 2006). Las especies de *Proteus* se encuentran ampliamente distribuidas en el ambiente, así como en el tracto digestivo de animales, en cultivo forman unas células nadadoras

características. Las bacterias *Morganella* son comensales del tracto digestivo y *Providencia* se han aislado de muestras clínicas, aunque se han detectado también en fetos de ovinos o heces de pingüino (Manos & Belas 2006).

4.2.3.2 Proteobacterias gamma no Enterobacteriales

Se analizaron un total de 71 especies (Anexo F) pertenecientes a los órdenes Pseudomonadales (21), Alteromonadales (17), Pasteurellales (14), Xanthomonadales (10) y Vibrionales (9) (Fig. 21).

El orden *Vibrionales* contiene bacilos anaerobios facultativos, ampliamente distribuidos en ambientes marinos, que además son parte de la microbiota intestinal de moluscos, camarones y peces. Se conocen especies importantes por su patogenicidad. En la muestra se presentaron los géneros *Vibrio*, *Aliivibrio* y *Photobacterium*. En el género *Vibrio* hay especies que provocan infecciones intestinales o extraintestinales como *V. cholerae*, agente causante del cólera y presente a nivel mundial. *Vibrio* y *Photobacterium* son bacterias versátiles ya que pueden utilizar diversos compuestos orgánicos como fuente de energía y de carbono (Kerstens *et al.* 2006). El género *Photobacterium* contiene bacterias bioluminiscentes que pueden ser de vida libre o encontrarse en simbiosis con invertebrados o con peces pelágicos. El género *Aliivibrio* contiene especies marinas que antes se clasificaban como *Vibrio* o *Photobacterium* (Kerstens *et al.* 2006).

El grupo Alteromonadales contiene aerobios estrictos, quimiorganótrofos y marinos, en la muestra se presentaron los géneros *Shewanella*, *Colwellia*, *Idiomarina* y *Pseudoalteromonas*. El género *Shewanella* son bacterias que habitan un amplio rango de ambientes, desde comida podrida, cuerpos de agua dulce, el mar, sus sedimentos, e incluso algunas especies se asocian a plantas, animales o son patógenas del humano. Se caracterizan por utilizar casi cualquier aceptor de electrones en la respiración anaerobia. Algunos miembros son psicrófilos, barófilos o presentan bioluminiscencia (Kerstens *et al.* 2006). Sus fuentes de carbono son limitadas y generalmente son productos de la fermentación, por lo que viven en sintropía con comunidades de fermentadores que les proveen de nutrientes. Además, bajo ciertas condiciones pueden reducir óxidos de hierro y manganeso, además de polisulfatos y compuestos nitrogenados (Nealson & Scott 2006).

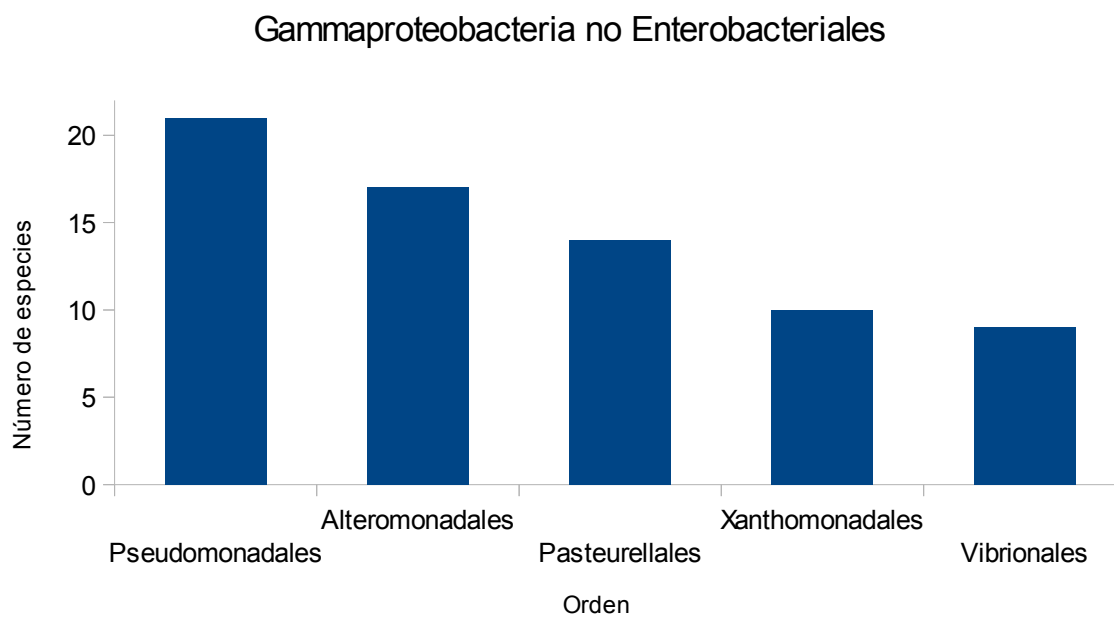


Figura 21. Gammaproteobacterias no Enterobacteriales consideradas en la muestra.

En el mismo grupo el género *Colwellia* contiene bacterias psicrófilas y barófilas, mientras que *Idiomarina* son bacterias encontradas en el fondo del océano Pacífico y obviamente también soportan una gran presión. *Pseudoalteromonas* son bacterias marinas, aerobias, móviles y quimiorganótrofas (Mikhailov *et al.* 2006).

Otro subgrupo importante fue el de Pseudomonadales con siete géneros representados en la muestra. Los géneros *Acinetobacter*, *Moraxella* y *Psychrobacter* forman un subconjunto (Kerstens *et al.* 2006). *Acinetobacter* son bacterias aerobias, no fermentativas, de vida libre, saprófitas, comunes en suelo, cuerpos de agua, drenajes, alimentos, en la piel de las personas sanas e incluso como patógenos oportunistas del humano (Towner 2006). Las bacterias *Psychrobacter* ocurren en el mar, son estrictamente aerobias, quimioheterótrofas, no móviles, adaptadas al frío y osmotolerantes. Se encuentran en ambientes polares pero también en muestras clínicas, pues se reportan algunos patógenos del humano y de otros animales como peces (Bowman 2006). *Moraxella* es un género de bacterias comunes en animales o humanos, con especies de importancia por ser patógenas del humano, gatos, perros, caballos, ovejas, cerdos, cabras, conejos, entre otros, no poseen pigmentos y crecen idealmente en condiciones aerobias (Pays 2006).

El género *Azotobacter* contiene bacterias de vida libre, aerobias, heterótrofas, que se caracterizan por fijar nitrógeno, son comunes en el suelo, en cuerpos de agua y en raíces y hojas de plantas (Becking 2006). *Cellvibrio* y *Pseudomonas* son bacterias aerobias y quimiorganótrofas, que utilizan diversos compuestos orgánicos como sustrato. *Cellvibrio* son capaces de degradar celulosa. *Pseudomonas* son especies ubicuas en ecosistemas acuáticos y en el suelo, además de que algunas son patógenas de plantas y animales, incluyendo al humano. Es un género conocido por su versatilidad metabólica, pues también pueden usar compuestos inorgánicos o tóxicos como sustrato, además de que presentan resistencia a detergentes, desinfectantes, metales pesados, antibióticos o solventes orgánicos (Moore *et al.* 2006). Se utilizan para purificar el agua de desechos o para limpiar derrames de aceite (Kersters *et al.* 2006).

El siguiente grupo fue Pasteurellales con siete géneros, algunas especies son de importancia por ser agentes etiológicos de enfermedades en vertebrados. *Pasteurella* son bacterias consideradas invasoras oportunistas -aunque la especie presente en la muestra se considera patógena- pero de manera normal habitan las mucosas de mamíferos y de aves (Christensen & Bisgaard 2006). *Mannheimia*, *Bibersteinia* y *Gallibacterium* son especies antes consideradas *Pasteurella*. *Haemophilus* son cocobacilos pleomórficos, algunas especies son patógenas de animales mientras que otras son comensales (Kersters *et al.* 2006).

También estuvieron presentes otros géneros: *Actinobacillus* son bacterias pequeñas, pleomórficas y facultativamente anaerobias que se encuentran como parásitos o comensales específicos en el tracto respiratorio, digestivo o genital de animales domésticos y primates (Macinnes & Lally 2006). *Aggregatibacter* son bacilos o cocobacilos facultativamente anaerobios, capnofílicos y heterótrofos, pueden reducir nitrato y fermentar algunos azúcares. Son parte de la flora bucal del humano y pueden ocasionar enfermedades de manera oportunista (Norksov-Lauristein & Kilian 2006).

Finalmente el orden Xanthomonadales estuvo presente con seis géneros. *Xanthomonas* son bacterias patógenas de plantas, con pigmentos amarillos y de utilidad porque producen goma xantana, material usado en la industria alimentaria, farmacéutica, agrícola y petrolera. También *Frateuria* son fitopatógenas y son fenotípicamente parecidas a las proteobacterias alfa del ácido acético (Swings 2006). *Xylella* toma su nombre por ser un

patógeno de especies importantes, como los cítricos o la vid, que habita en el xilema de la planta (Kersters *et al.* 2006).

Pseudoxanthomonas son bacilos aerobios ampliamente distribuidos a nivel mundial. Además tienen la capacidad de reducir nitrito en óxido de dinitrógeno, crecen en un medio alcalino y las colonias se observan de color amarillo (Thierry *et al.* 2004). *Rhodanobacter* contiene bacterias aerobias, quimorganótrofas, de vida libre y acuáticas, la especie de la muestra es capaz de desnitrificar (Nalin *et al.* 1999). *Stenotrophomonas* son bacterias aerobias y heterótrofas, que viven asociadas a plantas, además de que se utilizan para degradar compuestos contaminantes (Palleroni & Bradbury 1993).

4.2.4 Proteobacterias delta

Todas las proteobacterias delta son quimiorganótrofas: utilizan compuestos orgánicos como su fuente de energía y de carbono. Se dividen en dos grandes grupos: bacterias anaeróbicas reductoras de compuestos del azufre y bacterias de hábito depredador como los bdellovibrios y las mixobacterias. Estas últimas exhiben los ciclos de vida notablemente complejos dentro de los procariontes. Notablemente el grupo no contiene miembros fotosintéticos o patógenos de animales (Willey *et al.* 2008).

La muestra que se analizó consistió de 55 genomas pertenecientes a 8 órdenes distintos (Anexo G). Los grupos que contienen a bacterias reductoras del azufre son: Desulfovibrionales (14), Desulfuromonadales (11), Desulfobacterales (8), Syntrophobacterales (4), Desulfarculales (1) y Desulfurellales (1). También estuvieron presentes los Bdellovibrionales (4) y los Myxococcales (12) (Fig. 22).

Las bacterias capaces de reducir compuestos inorgánicos oxidados del azufre utilizan sulfato SO_4^{-2} o azufre elemental S^0 (aunque puede ser otro compuesto inorgánico oxidado del azufre), como aceptores terminales de electrones, y lo transforman en sulfuro de hidrógeno H_2S durante la oxidación anaeróbica de compuestos orgánicos. Son un grupo heterogéneo pues en su conjunto pueden oxidar completa o parcialmente una gran diversidad de compuestos orgánicos pequeños (Kersters *et al.* 2006).

Son organismos muy importantes dentro del ciclo del azufre y participan en las cadenas alimenticias microbianas responsables de la degradación de biopolímeros

complejos hasta CO_2 (Schulze & Mooney, 1994). Se encuentran ampliamente distribuidos en hábitats terrestres y acuáticos con condiciones anóxicas, las cuales por lo general son producto de la descomposición microbiana de la materia orgánica. Algunas especies impactan negativamente a la industria porque generan corrosión en materiales como el hierro (Willey *et al.* 2008). De manera interesante se consideran una de las formas de vida bacteriana más antiguas, su actividad dejó evidencia en rocas de hace más de 3,000 ma (Schulze & Mooney, 1994).

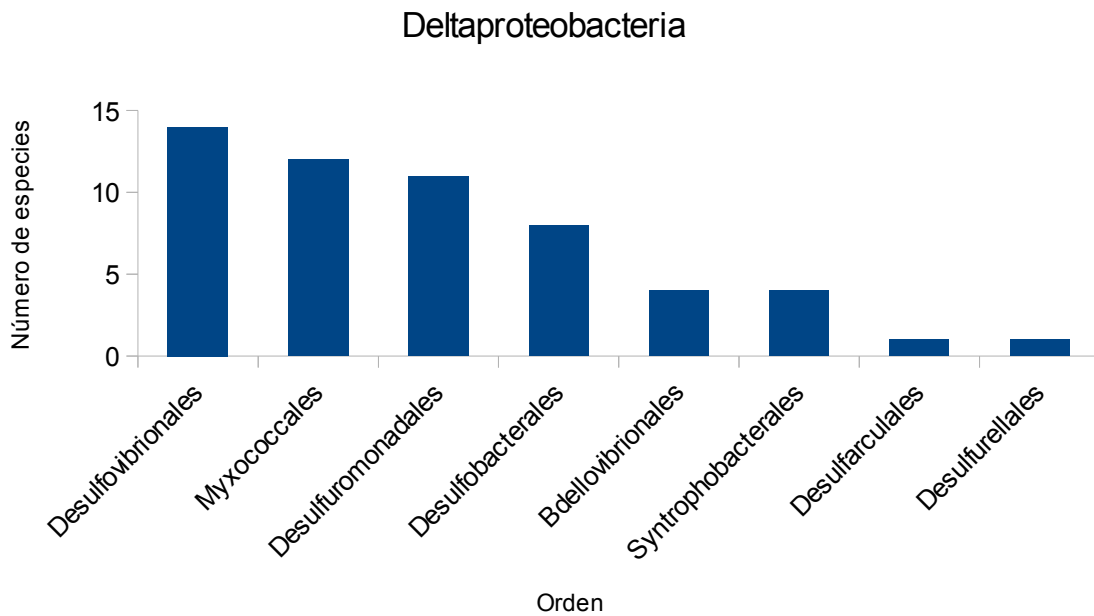


Figura 22. Órdenes de proteobacterias delta representados en el análisis.

En Desulfovibrionales hubo representantes de los géneros *Desulfohalobium*, *Desulfomicrobium* y *Desulfovibrio*. En Desulfuromonadales hubo los géneros *Geobacter* y *Pelobacter*, mientras que *Desulfatibacillum*, *Candidatus Desulfococcus*, *Desulfobacterium*, *Desulfobacula*, *Desulfobulbus*, *Desulfocapsa*, *Desulfotalea* y *Desulfurivibrio* pertenecen al grupo Desulfobacterales.

En el orden Syntrophobacterales hubo cuatro géneros: *Desulfobacca*, *Desulfomonile*, *Syntrophus* y *Syntrophobacter*. En Desulfarculales solamente se tuvo el género *Desulfarculus* así como en Desulfurellales sólo se presentó el género *Hippea*.

El grupo de los bdellovibrios contiene a bacterias pequeñas, bacilos con forma de coma y un flagelo grueso, que depredan a otras bacterias. Son organismos aerobios que presentan un ciclo de vida bifásico; alternan entre una fase depredadora donde no crecen, con otra fase reproductiva que se da dentro de una célula hospedera (Kersters *et al.* 2006). En esta última los bdellovibrios penetran al espacio periplásmico de células Gram negativas, crecen y se reproducen digiriendo los nutrientes de la célula hospedera, la cual termina muriendo. Se han aislado de muestras de suelo así como de agua dulce, marina e incluso de aguas residuales (Willey *et al.* 2008). Los géneros *Bacteriovorax* y *Bdellovibrio* estuvieron presentes en la muestra.

Las mixobacterias tienen forma de bacilos, son aerobias, se mueven deslizándose y se alimentan provocando la lisis a bacterias y a levaduras. Utilizan aminoácidos como su fuente de energía, de carbono y de nitrógeno. Se caracterizan por tener un ciclo de vida complejo porque en condiciones de pocos nutrientes forman cuerpos fructíferos vía procesos de señalización celular. Dichos cuerpos pueden ser de color amarillo, naranja o rojo debido a la presencia de pigmentos carotenoides. En su interior algunas células se desarrollan en estructuras de resistencia llamadas mixosporas, las cuales pueden resistir a la desecación y a periodos largos de tiempo. En condiciones favorables, el cuerpo fructífero y las esporas pueden formar rápidamente una colonia, la cual secreta enzimas hidrolíticas para digerir a sus presas más efectivamente que si lo hiciera una bacteria sola. Aunado a la complejidad en el comportamiento, son bacterias con genomas grandes. Son ubicuas y habitan los suelos en todo el mundo, aunque son más comunes en ambientes tropicales (Willey *et al.* 2008). En la muestra se presentaron los géneros *Anaeromyxobacter*, *Stigmatella*, *Haliangium*, *Corallocooccus*, *Myxococcus* y *Sorangium*.

4.2.5 Proteobacterias epsilon

En esta clase no se conocen miembros fotosintéticos, los géneros más famosos son enteropatógenos del hombre y de otros animales, aunque en los últimos tiempos ha crecido el interés por las especies oxidantes del azufre que se han asociado al grupo, algunas de las cuales son simbiotes de invertebrados marinos. En general son microaerófilicos o anaerobios, quimiorganótrofos o quimiolitótrofos, con forma de espiral o curvos y móviles (Kersters *et al.* 2006). La mayoría de las 13 especies analizadas pertenecen al orden

Campylobacterales (10), mientras que hay una especie del orden Nautiliales y dos sin clasificar (Anexo H).

Dentro de las Campylobacterales encontramos cinco géneros. En *Sulfurimonas* hay bacterias quimiolitótrofas que pueden utilizar compuestos reducidos del azufre o hidrógeno molecular como donadores de electrones, y nitrato o nitrito como aceptores, además del anhídrido carbónico como la fuente de carbono (Inagaki *et al.* 2003). Son comunes en sedimentos costeros o en ventilas hidrotermales. *Sulfuricurvum* son bacterias anaerobias facultativas capaces de oxidar compuestos del azufre que, se han encontrado en sitios donde hay petróleo (Kodama & Watanabe 2004). *Sulfurospirillum* son bacterias heterótrofas, reductoras del azufre y microaerofílicas, comunes en ambientes acuáticos, ya sea de agua dulce o marinos. Se han descrito especies capaces de utilizar arseniato o seleniato como aceptores de electrones (Kersters *et al.* 2006).

Las bacterias del género *Arcobacter* son abundantes en cuerpos de agua, drenajes, sitios con aceite o petróleo y en ambientes salinos. Algunas especies oxidan azufre mientras que otras pueden fijar nitrógeno. Se asocian con abortos en diversos animales de interés ganadero como vacas o borregos (Kersters *et al.* 2006). También son bacterias oxidantes del azufre las del género *Sulfurovum*, las cuales tienen forma de huevo y se han encontrado en los sedimentos marinos (Inagaki *et al.* 2004).

Los géneros *Nitratifactor* y *Nitratiruptor* contienen bacterias reductoras de nitrato, que habitan el fondo marino, son anaerobias a microaerofílicas, capaces de utilizar hidrógeno como donador de electrones, y oxígeno o nitrato como aceptor, por lo que se consideran quimiolitótrofas (Nakagawa *et al.* 2005).

Finalmente en el grupo Nautiliales hubo un representante del género *Nautilia*, bacterias encontradas en ventilas hidrotermales, ya sea formando tapetes bacterianos o en asociación simbiótica con la superficie de gusanos poliquetos. Son quimiolitótrofas; utilizan hidrógeno o formiato como donador de electrones y sulfuro como aceptor. Además pueden fijar carbono, azufre y nitrógeno (Smith *et al.* 2008).

Capítulo 5. Discusión

5.1 Sesgo inicial de la muestra

Los ancestros de las distintas clases de proteobacterias, se reconstruyeron comparando la estructura primaria de las proteínas presentes en un total de 402 genomas completamente secuenciados. Una cuestión que se hizo evidente al inicio del trabajo, fue el sesgo en la base de datos KEGG, ya que claramente la mayoría de los proyectos de secuenciación están dirigidos a organismos que poseen alguna utilidad, ya sea en el campo de la medicina, de la agricultura, de la ganadería, de la industria o en la biorremediación.

La consecuencia fue que hubo grupos muy bien representados, como en el caso de las alfa, las beta y las gamma proteobacterias, los cuales se tuvieron que dividir en dos para analizarse. En cambio, las clases delta y epsilon no tuvieron tantas especies. Esto resulta un tanto desafortunado, pues desde el punto de vista evolutivo, los grupos que se han sugerido como los más antiguos (Gupta 2000) siguen siendo los menos conocidos.

5.2 Análisis cuantitativo

Una vez que se obtuvieron las secuencias, fue posible comparar la variación en el tamaño del genoma dentro de las proteobacterias. Como se puede observar en la figura 5 y en la tabla 3, resulta que el tamaño es un rasgo extremadamente plástico: el genoma más pequeño tiene 1,354 proteínas y el mayor tiene 10,400. Ahora bien, los catálogos ancestrales que reconstruimos para los diferentes grupos, van desde 69 hasta 598 proteínas. El hecho de que no hay proteobacterias modernas con genomas tan pequeños, sugiere que las reconstrucciones ancestrales subestimaron el contenido proteico del ancestro.

Uno de los resultados importantes del trabajo tiene que ver con el parámetro que llamamos *resolución* del análisis, el cual nos indica qué tanto se parecen los ancestros a las especies actuales de proteobacterias. Como ya se mencionó, los catálogos ancestrales tuvieron un tamaño considerablemente menor. De hecho pudimos calcular que la resolución tuvo una variación del 1.7% al 25%. Esto implica que el ancestro tenía más proteínas que las que detectamos al comparar la estructura primaria.

5.3 Factores que afectaron las reconstrucciones ancestrales

En general, los factores que afectaron las reconstrucciones ancestrales fueron la transferencia horizontal de genes, la reducción o pérdida de genes, la diversidad fisiológica dentro del grupo, la biología del organismo de la secuencia *query*, en particular su estilo de vida, y la distancia filogenética entre las especies de la muestra.

La reconstrucción del ancestro de las proteobacterias alfa, subgrupo Rhizobiales, resultó en 382 proteínas conservadas en el *núcleo* estricto, las cuales pertenecen a 18 categorías funcionales. En este grupo la secuencia *query* fue la especie *Bartonella henselae*, la cual es un patógeno intracelular del humano. Las especies de este género, por lo general están adaptadas a sus huéspedes, aunque se ha reportado que *B. henselae* puede utilizar diversos hospederos. A nivel de genoma, se ha documentado una tendencia a perder genes (Alsmark *et al.* 2004), así como procesos de transferencia horizontal que los hacen parásitos muy exitosos (Berglund & Frank, 2012).

Por otro lado, las especies contenidas en la muestra presentan una notable diversidad fisiológica: se incluyen bacterias fotoheterótrofas, quimiolitótrofas, quimiorganótrofas, metilótrofas, fijadoras de N, aerobias, anaerobias, oxidantes del hidrógeno, especies que interactúan con eucariontes, como los patógenos intracelulares de plantas y animales, y especies de vida libre que en general son organismos ubicuos.

De manera interesante, el catálogo ancestral se hace mucho más numeroso conforme se relaja el criterio del *núcleo* estricto (barras de color amarillo y gris en la fig. 5), y son posiblemente aquellas proteínas que no encontramos en la secuencia *query*, debido a que las ha perdido por el hábito parásito. Aún así es una de las reconstrucciones más numerosas, posiblemente porque la muestra contiene organismos cercanamente emparentados, pues todos pertenecen al mismo orden.

La reconstrucción ancestral para las proteobacterias alfa que no pertenecen al orden Rhizobiales, resultó en tan sólo 69 proteínas en el núcleo estricto, contenidas en 13 categorías funcionales, lo que a primera vista parece indicar un ancestro con diversidad funcional, aunque claramente es un catálogo muy reducido numéricamente. La secuencia *query* fue una especie sumamente interesante, *Pelagibacter ubique*, la cual toma su nombre por el hecho de que constituye del 20 al 50% del bacteriopláncton oceánico (Morris *et al.* 2002). Es una especie perteneciente al clado SAR11: bacterias marinas, cosmopolitas y

difícilmente cultivables. Contiene el genoma más pequeño para ser de vida libre, pero curiosamente no hay evidencia de pérdida de genes, por el contrario, no ha descartado ningún metabolismo básico. Se ha propuesto que este patrón se debe a dos fuerzas evolutivas antagónicas: debe fabricar lo que necesita para vivir de forma independiente en el mar, pero tener un genoma grande es costoso energéticamente (Giovannoni *et al.* 2005). El hecho de que este organismo de vida libre no tenga un genoma más pequeño, apoya la noción de que los catálogos ancestrales están subestimados numéricamente.

Ahora bien la muestra de las alfa no Rhizobiales es muy diversa fisiológicamente, contiene especies aerobias, anaerobias o microaerofílicas, quimioheterotróficas, heterotróficas, quimiorganótrofas, quimiolitótrofas (oxidantes de hierro, hidrógeno o compuestos del azufre), bacterias púrpuras no del azufre, parásitos intracelulares, especies que pueden fijar nitrógeno o desnitrificar, e incluso con ciclo de vida dimórfico o acidófilas.

Además comparado con el grupo anterior, las especies guardan una mayor distancia filogenética entre ellas, lo que implica que la reconstrucción ancestral, al buscar las proteínas conservadas entre especies lejanas, se remonta más atrás en el tiempo. Es interesante notar que al relajar el criterio del *núcleo* estricto hay categorías que se vuelven muy numerosas, como la replicación, reparación y recombinación, el metabolismo de aminoácidos, la producción y conversión de energía, o el metabolismo de transporte de coenzimas, lo que podría indicar adquisiciones novedosas que les han permitido diversificarse en los estilos de vida tan variados que muestran (barras de color amarillo y gris en la fig. 6).

En el caso de la reconstrucción para las proteobacterias beta, subgrupo Burkholderiales, se encontraron 261 proteínas en el *núcleo* estricto, contenidas en 18 categorías distintas. La secuencia *query* de esta reconstrucción fue la especie *Taylorella asinigenitalis*, una bacteria que habita el tracto genital de burros. Se ha reportado que posee un genoma reducido, y además que tiene unos requerimientos nutricionales muy estrictos, lo que indica una especialización del nicho, pues se reporta que no posee vías catabólicas para sobrevivir en condiciones diferentes (Hébert *et al.* 2012).

Al igual que las alfa, las proteobacterias beta exhiben una gran diversidad fisiológica. Pueden ser aerobias o anaerobias, la mayoría son quimioheterótrofas pero también encontramos quimiolitótrofas con compuestos del azufre, quimiorganótrofas, litoautótrofas, bacterias desnitrificantes o fijadoras de nitrógeno y también bacterias PNSB. Todas

pertenecen a un mismo orden, lo que sugiere que están más cercanamente emparentadas entre ellas que con el resto de las especies.

Además del genoma *query* reducido, se puede ver que subestimamos al ancestro porque al relajar el criterio del núcleo estricto, se incorporan más de 400 proteínas en todas las categorías funcionales, lo que podría indicar las proteínas que ha perdido *Taylorella* o también aquellas que han permitido a los otros linajes diversificar sus capacidades metabólicas.

Para el grupo de las beta no Burkholderiales, el catálogo ancestral estricto presentó 410 proteínas conservadas, repartidas en 19 categorías funcionales. La especie *Neisseria lactamica* fue la secuencia *query*, un organismo heterótrofo. A dicho género pertenecen algunos patógenos importantes, aunque *N. lactamica* se ha reportado como un comensal estricto. Si bien lo anterior sugiere que estaríamos subestimando el contenido del ancestro, se ha reportado que las especies de *Neisseria* presentan transferencia horizontal de genes (Alber *et al.* 2001; Bennett *et al.* 2005). Además al relajar el criterio en la reconstrucción, no se incorporan más proteínas de las presentes en el *núcleo* estricto (ver color amarillo y gris de figura 8).

Dentro de la muestra de proteobacterias beta no Burkholderiales, encontramos especies quimiorganótrofas, quimiolitótrofas capaces de oxidar hidrógeno, amoníaco, fierro o compuestos del azufre, algunos metanótrofos, especies aerobias o anaerobias y bacterias que fijan nitrógeno así como otras que desnitrifican, así que claramente son un grupo con una diversidad fisiológica importante. Asimismo se encuentran menos emparentadas entre ellas, comparadas con las especies pertenecientes al orden Burkholderiales.

La reconstrucción ancestral de las proteobacterias gamma, subgrupo Enterobateriales, mostró 585 proteínas conservadas en el *núcleo* estricto, repartidas en 19 categorías. Esta fue una de las reconstrucciones más numerosas de todo el trabajo, cabe decir que todas las especies de la muestra pertenecen a la misma familia.

La secuencia *query* fue la especie *Hamiltonella defensa*, un endosimbionte de áfidos, que se presume evolucionó a partir de una ancestro patógeno. En la literatura se reporta que posee una reducción en su genoma, pero que todavía retiene genes y vías para una variedad de procesos y estructuras, que ya no se observan en otras especies como las del género *Buchnera* (Degnan *et al.* 2009). Además la relación que establece con su hospedero es un

mutualismo nutricional, lo que implica que tiene que intercambiar sustancias (Degnan *et al.* 2009). Al relajar el criterio del *núcleo* estricto, no se incorporan más proteínas que las observadas inicialmente, lo que sugiere que no estamos subestimando tanto el contenido del ancestro, comparado con el caso de las proteobacterias beta Burkholderiales (ver colores amarillo y gris en figuras 7 y 9).

Este grupo es mucho menos diverso que los discutidos anteriormente, pues todas las especies son quimiorganótrofas (algunas son fermentadoras) y la mayoría presentan el hábito típico de las Enterobacteriales, aunque también encontramos especies de vida libre en suelos y cuerpos de agua, algunas otras asociadas a plantas e incluso fijadoras de nitrógeno. Sobresale que la categoría relacionada con la pared y la membrana celular está en segundo lugar, lo que sugiere la importancia de dicho rasgo en la aparición del hábito Enterobacterial.

En el caso de las gamma proteobacteria no Enterobacteriales, la reconstrucción ancestral arrojó un total de 173 proteínas en el *núcleo* estricto, contenidas en 18 categorías, aunque el *núcleo* estricto se ve bastante reducido. El parásito obligado *Haemophilus influenzae* resultó ser la secuencia *query*, especie para la cual se ha reportado una reducción importante en su genoma (Tatusov *et al.* 1996). Al relajar el criterio del *núcleo* estricto se incorporan más de 400 proteínas, muchas de las cuales no aparecen en la reconstrucción ancestral, pues *H. influenzae* ha perdido una gran cantidad de éstas. Aunado a esto, las especies de la muestra se encuentran lejanamente emparentadas entre ellas.

Este grupo contiene bacterias más diversas que las gamma Enterobacteriales pero menos que los grupos alfa y beta. Se presentan organismos aerobios y anaerobios, quimioheterótrofos y quimiorganótrofos, algunos pocos viven asociados a plantas y animales, además de que hay miembros patógenos. Sólo una especie de la muestra es capaz de desnitrificar.

Otra reconstrucción con pocas proteínas en el núcleo estricto fue para el grupo de las proteobacterias delta. Se observaron 97 proteínas conservadas repartidas en 15 categorías funcionales. La secuencia *query* fue una especie llamada *Hippea maritima*, la cual es anaerobia obligada, vive en las ventilas hidrotermales y es reductora de sulfato (Miroshnichenko *et al.* 1999). La muestra contiene poca diversidad fisiológica, pues todas las bacterias son quimiorganótrofas.

Debido a que son uno de los grupos basales en el árbol de proteobacterias (fig. 15), esta reconstrucción se remonta más atrás en el tiempo que la de los grupos alfa, beta y gamma. Es interesante ver que al relajar el criterio del núcleo estricto se agregan muchas proteínas de las categorías: traducción, estructura y biogénesis del ribosoma, el metabolismo y transporte de coenzimas y de aminoácidos, además de proteínas relacionadas con la membrana y la pared celular, todo lo cual parece necesario para la diversificación que se observa en las proteobacterias. Resalta el hecho de que la categoría de producción de energía, la cual normalmente aparece en los primeros lugares, estuvo casi ausente del *núcleo* estricto. Esto sugiere que mientras más atrás nos vamos en el tiempo, la comparación de estructura primaria no aporta indicios sobre la vía energética del ancestro.

Finalmente el grupo de las epsilon proteobacteria presentó uno de los *núcleos* estrictos más grandes, con 598 proteínas conservadas, pertenecientes a 18 categorías. La secuencia *query* fue la especie *Nautilia profundicola*, bacterias quimiolitótrofas que habitan las ventilas hidrotermales y que son capaces de utilizar el hidrógeno como donador de electrones y el sulfuro como aceptor. Además pueden fijar carbono, azufre y nitrógeno (Smith *et al.* 2008).

En general el grupo tiene una diversidad fisiológica importante: encontramos bacterias quimiolitótrofas, que pueden reducir azufre o nitrato, otras son quimiorganótrofas, heterótrofas, hay algunas bacterias oxidantes del azufre simbioses de animales invertebrados marinos, en general son microaerófilos o anaerobios. Una especie puede fijar nitrógeno. Asimismo son especies que están cercanamente emparentadas, pues la mayoría pertenecen a un mismo orden.

En resumen, como se puede ver en la tabla 5, los catálogos ancestrales de las alfa proteobacterias subgrupo Rhizobiales, de las beta subgrupo Burkholderiales, de las gamma Enterobacteriales y de las Epsilon, son los de mayor tamaño. Dichas muestras representan grupos cuyas especies están más cercanamente relacionadas, comparadas con el resto de los grupos. Lo anterior indica que a menor distancia filogenética, las especies comparten más cosas y por lo tanto el catálogo ancestral es mayor. De las reconstrucciones mencionadas, la de las beta proteobacterias subgrupo Burkholderiales, fue la que tuvo menos proteínas compartidas, lo que podría explicarse por la pérdida de genes que se ha reportado en la especie que fungió como secuencia *query*.

Por otro lado, en el grupo de las proteobacterias beta no Burkholderiales, se ha reportado la presencia de HGT en la especie de la secuencia *query* (tabla 5), lo que sugiere que la reconstrucción está sobrestimada, es decir, por la distancia filogenética que separa a las especies, se esperaba un núcleo genómico estricto con menos proteínas. En el caso de las gamma no Enterobacteriales sucedió el patrón inverso: se ha reportado pérdida de genes y por lo tanto el núcleo genómico posiblemente era mayor que lo que obtuvimos (tabla 5).

Tabla 5. Comparación de los factores que afectaron las reconstrucciones de los catálogos ancestrales. El tamaño del *núcleo*, se refiere al *núcleo* estricto. La diversidad fisiológica es de la muestra, la cercanía filogenética de las especies es relativa y considerando la filogenia de la figura 15. Finalmente se anota si hubo pérdida o transferencia horizontal en la historia evolutiva, además del estilo de vida de la especie correspondiente a la secuencia *query*.

Grupo	Tamaño del core	Diversidad fisiológica	Cercanía filogenética	Secuencia <i>query</i>		
				Pérdida de genes	HGT	Hábitat
Alfa, subgrupo Rhizobiales	382	Sí	Sí	Sí	Sí	patógeno intracelular
Alfa no Rhizobiales	60	Sí	No	No	No	Vida libre, heterótrofo Oligotrófico acuático
Beta, subgrupo Burkholderiales	261	Sí	Sí	Sí	No	Especializado a nicho
Beta no Burkholderiales	410	Sí	No	No	Sí	Comensal estricto
Gamma, subgrupo Enterobacteriales	585	No	Sí	Sí	No	Endosimbionte
Gamma no Enterobacteriales	173	Sí	No	Sí	No	Parásito obligado
Delta	97	No	No	No	No	Reductor de sulfato
Epsilon	598	Sí	Sí	No	No	Reductor de sulfato

Los núcleos genómicos para los grupos de proteobacterias alfa no Rhizobiales y para las delta, fueron las reconstrucciones con menos sesgos, en el sentido de que las especies seleccionadas como secuencia *query* fueron de vida libre, además de que no hay reportes de pérdida de genes o de HGT. De manera relevante fueron las reconstrucciones más pequeñas, la resolución calculada para dichos análisis fue de 1.8% y 2% respectivamente (ver tabla 3). Lo anterior sugiere de nuevo que la comparación de la secuencia primaria de

proteínas, aún cuando se evitan los sesgos metodológicos, recupera menos del 5% del contenido proteico del ancestro.

Este resultado coincide con todas las reconstrucciones que se han hecho (Mushegian and Koonin 1996; Kyrpides *et al.* 1999; Koonin 2003; Harris *et al.* 2003; Mirkin *et al.* 2003; Yang *et al.* 2005; Delaye *et al.* 2005; Sobolevsky & Trifonov, 2006; Ranea *et al.* 2006, Ouzounis *et al.* 2006; Tuller *et al.* 2010 ; Kim & Caetano-Anollés 2011; Kannan *et al.* 2013), las cuales varían desde 63 genes hasta poco menos de 800 proteínas ancestrales, cantidades que no representan siquiera el tamaño del genoma de un parásito moderno.

No hay que perder de vista que los catálogos ancestrales contienen proteínas que se fueron adquiriendo en momentos distintos de la evolución del linaje (Becerra *et al.* 2007), lo que implica que estamos estudiando un fenómeno con una dimensión temporal. En ese sentido los resultados indican que mientras mayor sea la distancia filogenética entre las especies muestreadas, el ancestro se encontrará más distante en el pasado, y por lo tanto se recuperarán menos proteínas conservadas.

5.4 Alcances y limitaciones del método

Como ya vimos, el método subestima cuantitativamente el contenido del ancestro, lo que implica que éste tendría más proteínas de las que estamos encontrando. Pero además, como se observa en la figura 13, hay un conjunto de funciones que no se recuperan en las reconstrucciones ancestrales, es decir, también hay un sesgo cualitativo. Una posible explicación, es que las proteínas presentes en esas categorías no están conservadas a nivel de estructura primaria y por eso el método no las detecta.

Las categorías en las que aparecieron proteínas conservadas fueron en primer lugar la traducción, la estructura y biogénesis del ribosoma, resultado que coincide con todas las reconstrucciones que se han hecho del LCA (Becerra *et al.* 2007). Esto sugiere que una vez que se estableció la traducción, la evolución no ha dado con vías alternativas para la síntesis de proteínas.

Las otras categorías fueron el metabolismo y transporte de aminoácidos, la replicación, recombinación y reparación, el metabolismo de nucleótidos así como de coenzimas, la biogénesis de la membrana y la pared celular, la producción de energía y

finalmente la modificación postraducciona de proteínas. Sin lugar a dudas, son tareas esenciales para el funcionamiento de cualquier célula, por lo que sugerimos que también las tendría el ancestro de las proteobacterias, e incluso quizá también el LCA ya presentara dichas funciones.

A pesar de las limitaciones cualitativas, los catálogos ancestrales que reconstruimos sugieren que los ancestros de las proteobacterias, tenían una diversidad de funciones semejante a la de las especies modernas. Consideramos que se necesita revisar si las proteínas representan un conjunto funcional, por ejemplo no se sabe si las vías metabólicas están completas.

Cabe decir que se pueden eliminar los sesgos asociados a la selección de la muestra, si se descartan especies endoparásitas, endosimbiontes (Delaye, *et al.* 2005), o aquellas en donde se haya reportado la pérdida o la transferencia horizontal de genes. Asimismo, la reconstrucción de los catálogos ancestrales de proteínas puede mejorar, si se hacen más estrictos los criterios de inclusión al seleccionar la muestra.

5.5 Perspectivas

Ante los resultados obtenidos, se hacen las siguientes sugerencias para refinar el análisis:

- Para evaluar si las rutas metabólicas están completas y poder caracterizar mejor al ancestro, la clasificación funcional se puede hacer con base en las categorías del KEGG.
- Los criterios de inclusión de las especies para la muestra se deben hacer más estrictos.
- La selección de la secuencia *query* es importante. Debe revisarse el estilo de vida de dicha especie, además de que no se haya reportado HGT o reducción en su genoma.

Conclusiones

Al comparar la secuencia primaria de las proteínas, fue posible reconstruir los catálogos ancestrales de las distintas clases de Proteobacteria. Las conclusiones del trabajo se mencionan a continuación:

- Hubo un sesgo inicial en la muestra, debido a que algunos grupos están mejor representados que otros en la base de datos KEGG.
- Los catálogos ancestrales tuvieron entre 69 y 598 proteínas, lo que representa entre el 1.8% y el 25% del número de proteínas promedio de las proteobacterias modernas. Los ancestros de las proteobacterias alfa, beta, gamma, delta y epsilon, tenían más proteínas que las que detectamos.
- Aunque no todas las categorías funcionales están igualmente conservadas, se puede afirmar que los ancestros de las diferentes clases de Proteobacteria, contaban con una diversidad de funciones semejante a las especies modernas.
- La metodología aplicada subestima cuantitativa y cualitativamente el contenido proteico de los ancestros de proteobacterias.
- El tamaño de las reconstrucciones ancestrales depende de factores como la pérdida de genes, la transferencia horizontal, la diversidad fisiológica de las especies dentro de la muestra, y sobretodo de la distancia filogenética entre las especies incluidas.
- Mientras mayor sea la distancia filogenética entre las especies de la muestra, el catálogo ancestral se hace menor, es decir, hay menos proteínas conservadas. Debido a que el LCA representa el nodo más alejado del árbol universal, el método planteado resulta insuficiente para reconstruir su catálogo ancestral.

Referencias

- Abbot, S. L. & J. M. Janda. 2006. The genus *Edwardsiella*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 72-89.
- Alber, D., Oberkotter, M., Suerbaum, S., Claus, H. Frosch, M. & U. Vogel. 2001. Genetic diversity of *Neisseria lactamica* strains from epidemiologically defined carriers. *Journal Of Clinical Microbiology*. 39(5):1710-1715.
- Alsmark, C. M., Frank, A. C., Karlberg, E. O., Legault, B. A., Ardell, D. H., Canback, B., Eriksson, A. S., Naslund, A. K., Handley, S. A., Huvet, M., La Scola, B., Holmberg, M. & S. G. Andersson. 2004. The louse-borne human pathogen *Bartonella quintana* is a genomic derivative of the zoonotic agent *Bartonella henselae*. *Proceedings of the National Academy of Sciences USA*.101(26):9716-9721.
- Audic, S., Robert, C., Campagna, B., Parinello, H., Claverie, J. M., Raoult, D & M. Drancourt. 2007. Genome analysis of *Minibacterium massiliensis* highlights the convergent evolution of water-living bacteria. *PLoS Genetics*. 3(8):e138.
- Bazylnski, D. A., Williams, T. J., Lefèvre, C. T., Berg, R. J., Zhang, C. L., Bowser, S.S., Dean, A. J. & T. J. Beveridge. 2013. *Magnetococcus marinus* gen. nov., sp. nov., a marine, magnetotactic bacterium that represents a novel lineage (Magnetococcaceae fam. Nov., Magnetococcales ord. nov.) at the base of the Alphaproteobacteria. *International Journal of Systematic and Evolutionary Microbiology*. 63(3):801-808.
- Becerra, A., Delaye, L., Islas, S. & Lazcano, A. 2007. The very early stages of biological evolution and the nature of the last common ancestor of the three major cell domains. *Annual Reviews of Ecology, Evolution and Systematics*. 38: 361-379.
- Becerra, A., Islas, S., Leguina, J. I., Silva, E. & A., Lazcano. 1997. Polyphyletic gene losses can bias backtrack characterizations of the cenancestor. *Journal of Molecular Evolution*. 45:115–118.

- Becking, J. H. 2006. The family Azotobacteraceae. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 759-783.
- Bennett, J. S., Griffiths, D. T., McCarthy, N. D., Sleeman, K. L., Jolley, K. A., Crook, D. W. & M. C. Maiden. 2005. Genetic diversity and carriage dynamics of *Neisseria lactamica* in infants. *Infection And Immunity*. 73(4):2424-2432.
- Berglund, E. C. & A. C. Frank. 2012. Horizontal transfer of host-adaptability systems in Bacteria. In ed. Francino, M. P. *Horizontal gene transfer in microorganisms*. Caister Academic Press. pp. 33-46.
- Biebl, H., Allgaier, M., Tindall, B. J., Koblizek, M., Lünsdorf, H., Pukall, R. & I. Wagner-Döbler. 2005. *Dinoroseobacter shibae* gen. nov., sp. nov., a new aerobic phototrophic bacterium isolated from dinoflagellates. *International Journal of Systematic and Evolutionary Microbiology*. 55:1089–1096.
- Borenshtein, D. & D. B. Schauer. 2006. The genus *Citrobacter*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 90-98.
- Bowman, J. 2006. The Methanotrophs —the families Methylococcaceae and Methylocystaceae. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 266-289.
- Bowman, J. P. 2006. The genus *Psychrobacter*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 920-930.
- Brasier, M., McLoughlin, N., Green, O. & D. Wacey. 2006. A fresh look at the fossil evidence for early Archaean cellular life. *Philosophical Transactions of the Royal Society Biological Sciences*. 361:887-902.
- Brisse, S., Grimont, F. & P. A. D. Grimont. 2006. The genus *Klebsiella*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 159-196.

- Butterfield, N. J. 2000. *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology*. 26: 386–404.
- Campbell, B. J., Engel, A. S., Porter, M. L. & K. Takai. 2006. The versatile ϵ -proteobacteria: key players in sulphidic habitats. *Nature Reviews Microbiology*. 4:458-468.
- Canfield, D. E. 2006. Biochemistry: gas with an ancient history. *Nature*. 440:426–27.
- Cao, B., Ma, T., Ren, Y., Ren, Y., Li, G., Li, P., Guo, X., Ding, P. & L. Feng. 2011. Complete genome sequence of *Pusillimonas* sp. T7-7, a cold-tolerant diesel oil-degrading bacterium isolated from the Bohai Sea in China. *Journal of Bacteriology*. 15:4021–4022.
- Christensen, H. & M. Bisgaard. 2006. The genus *Pasteurella*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 1062-1090.
- Cicarelli, F., Doerks, T., von Mering, C., Creevey, C., Snel B. & P. Bork. 2006. Toward automatic reconstruction of a highly resolved tree of life. *Science*. 311:1283-1287.
- Contreras-Moreira, B. & P. Vinuesa. 2013. GET-HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Applied and Environmental Microbiology*. 79(24):7696-7701.
- Darwin, Charles. *El origen de las especies ilustrado. Versión abreviada e introducción de Richard E. Leakey*. Martín Casillas Editores, S. A. 2da edición. México. 1981. pp. 466-467.
- De Boer, W., Leveau, J. H., Kowalchuk, G. A., Klein Gunnewiek, P. J., Abeln, E. C., Figge, M. J., Sjollem, K., Janse, J. D. & J. A. van Veen. 2004. *Collimonas fungivorans* gen. nov., sp. nov., a chitinolytic soil bacterium with the ability to grow on living fungal hyphae. *International Journal Of Systematic And Evolutionary Microbiology*. 54(3):857-64.

- Degnan, P. H., Yu, Y., Sisneros, N., Wing, R. A., Moran, N. A. 2009. *Hamiltonella defensa*, genome evolution of protective bacterial endosymbiont from pathogenic ancestors. *Proceedings of the National Academy of Sciences USA*. 106(22):9063-9068.
- Delaye, L. & A. Becerra. 2012. Cenancestor: the last universal common ancestor. *Evolution: Education and Outreach*. 5(3): 382-388.
- Delaye, L., Becerra, A. & A. Lazcano. 2005. The last common ancestor: what's in a name. *Origins of Life and Evolution of Biospheres*. 35:537-554.
- Delaye, L., Vázquez, H. & A. Lazcano. 2001. The cenancestor and its contemporary biological relics: the case of nucleic acid polymerases. In ed. J Chela-Flores, T Owen, F Raulin. *First Steps in the Origin of Life in the Universe*. Dordrecht: Kluwer Acad. pp. 223–30.
- Di Giulio, M. 2003. The universal ancestor and the ancestor of bacteria were hyperthermophiles. *Journal of Molecular Evolution*. 57:721–30.
- Di Giulio, M. 2007a. The tree of life might be rooted in the branch leading to Nanoarchaeota. *Gene*. 401: 108-113.
- Di Giulio, M. 2007b. The universal ancestor and the ancestors of Archaea and Bacteria were anaerobes whereas the ancestor of the Eukarya domain was an aerobe. *Journal of Evolutionary Biology*. 20(2):543-548.
- Doolittle, W. F. 2000. Uprooting the tree of life. *Scientific American*. 282:90-95.
- Emerson, D., Rentz, J. A., Lilburn, T. G., Davis, R. E, Aldrich, H., Chan, C. & C. L. Moyer. 2007. A novel lineage of proteobacteria involved in formation of marine Fe-oxidizing microbial mat communities. *PLoS One*. 2: e667.
- Ettema, T. J. G. & S. G. E. Andersson. 2009. The α -proteobacteria: the Darwin finches of the bacterial world. *Biology Letters*. 5:429-432.
- Farmer III, J. J. & F.W. Hickman-Brenner. 2006. The Genera *Vibrio* and *Photobacterium*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. &

E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 508-563.

- Ferla, M. P., Thrash, J. C., Giovannoni, S. J. & W. M. Patrick. 2013. New rRNA gene-based phylogenies of the alphaproteobacteria provide perspective on major groups, mitochondrial ancestry and phylogenetic instability. *PLOS ONE*. 8(2):e83383.
- Fitch, W. M. & K. Upper. 1987. The phylogeny of tRNA sequences provides evidence of ambiguity reduction in the origin of the genetic code. *Cold Spring Harbor Symposia on Quantitative Biology*. 52:759–67.
- Forterre P. & H. Philippe. 1999. Where is the root of the universal tree of life?. *Bioessays*. 21(10): 871-879.
- Francino, M. P., Santos, S. R. & H. Ochman. 2006. Phylogenetic relationships of bacteria with special reference to endosymbionts and enteric species. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 41-59.
- Galtier, N., Tourasse, N. & M. Gouy. 1999. A nonhyperthermophilic common ancestor to extant life forms. *Science*. 283:220–21.
- Garrity, G. (Ed.) 2001. *Bergey's Manual of Systematic Bacteriology*. 2nd ed. Springer-Verlag. New York, NY. pp.1:1721.
- Giovannoni, S., Tripp, H. J., Givan, S., Podar, M., Vergin, K. L., Baptista, D., Bibbs, L., Eads, J., Richardson, T. H., Noordewier, M., Rappé, M. S., Short, J. M., Carrington, J. C. & E. J. Mathur. 2005. Genome streamlining in a cosmopolitan oceanic bacterium. *Science*. 309:1242.
- Glansdorff, N., Xu, Y. & B. Labedan. 2008. The Last Universal Common Ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biology Direct*. 3:29.
- Gogarten, J. P., Doolittle, W. F. & J. G. Lawrence. 2002. Prokaryotic evolution in light of gene transfer. *Molecular Biology and Evolution*. 19(12):2226-2238.

- Gogarten, J. P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E. J., Bowman, B. J., Manolson, M. F., Poole, R. J., Date, T. & T. Oshima. 1989. Evolution of the vacuolar H⁺-ATPase, implications for the origin of eukaryotes. *Proceedings of the National Academy of Sciences USA*. 86:6661–65.
- Grimont F. & P. A. D. Grimont. 2006. The genus *Enterobacter*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 197-214.
- Grimont F. & P. A. D. Grimont. 2006b. The genus *Serratia*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 219-244.
- Gupta, R. S. 2000. The natural evolutionary relationships among prokaryotes. *Critical Reviews in Microbiology*. 26:111-131.
- Hahn, M. W., Scheuerl, T., Jezberová, J., Koll, U., Jezbera, J., Šimek, K., Vannini, C. Petroni, G. & Q. L. Wu. 2012. The passive yet successful way of planktonic life: genomic and experimental analysis of the ecology of a free-living Polynucleobacter population. *PLoS ONE*. 7(3):e32772.
- Halbach, M., Koschinsky, A. & P. Halbach. 2001. Report on the discovery of *Gallionella ferruginea* from an active hydrothermal field in the deep sea. *InterRidge News*. 10(1):18-20.
- Harris, J. K., Kelley, S. T., Spiegelman, G. B. & N. R. Pace. 2003. The genetic core of the universal ancestor. *Genome Research*. 13:407-412.
- Hébert, L., Moumen, B., Pons, N., Duquesne, F., Breuil, M., Goux, D., Batto, J., Laugier, C., Renault, P. & S. Petry. 2012. Genomic characterization of the *Taylorella* genus. *PLoS One*. 7(1):e29953.
- Inagaki, F., Takai, K., Kobayashi, H., Nealson, K. H. & K. Horikoshi. 2003. *Sulfurimonas autotrophica* gen. nov., sp. nov., a novel sulfur-oxidizing ϵ -proteobacterium isolated from hydrothermal sediments in the Mid-Okinawa Trough. *International Journal of Systematic Bacteriology*. 53(6):1801-1805.
- Inagaki, F., Takai, K., Nealson, K. & K. Horikoshi. 2004. *Sulfurovum lithotrophicum* gen. nov., sp. nov., a novel sulfur-oxidizing chemolithoautotroph within the ϵ -

Proteobacteria isolated from Okinawa Trough hydrothermal sediments. *International Journal of Systematic Bacteriology*. 54(5):1477-1482.

- Islas, S., Becerra, A., Luigi Luisi, P. & A. Lazcano. 2004. Comparative genomics and the gene complement of a minimal cell. *Origins of Life and Evolution o the Biosphere*. 34:243.256.
- Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S. & T. Miyata. 1989. Evolutionary relationship of archaeobacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proceedings of the National Academy of Sciences USA*. 86:9355–59.
- Janda, J. M. 2006. New Members of the Family Enterobacteriaceae. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 4-50.
- Jeffares & Poole. 2000. Were Bacteria the First Forms of Life on Earth? Disponible en: http://www.actionbioscience.org/evolution/jeffares_poole.html
- Joyce, G. F. 2002. The antiquity of RNA-based evolution. *Nature* 418: 214-221.
- Kado, C. I. 2006. *Erwinia* and related genera. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 443-450.
- Kandler, O. 1994. The early diversification of life. In ed. Bengtson, S. *Early life on earth*. Nobel Symposium No. 84. Columbia University Press, New York pp. 152-160.
- Kaneshia, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M. & M. Hirakawa. 2006. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Research*. 34:D354-D357.
- Kannan, L., Li, H., Rubinstein, B. & A. Mushegian. 2013. Models of gene gain and gene loss for probabilistic reconstruction of gene content in the last universal common ancestor of life. *Biology Direct*. 8:32.

- Kersters, K., Devos, P., Gillis, M., Swings, J., Vandamme, P. & E. Stackebrandt. 2006. Introduction to the Proteobacteria. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. pp. 3-37.
- Kim, K. M. & G. 2011. Caetano-Anollés. The proteomic complexity and rise of the primordial ancestor of diversified life. *BMC Evolutionary Biology*. 11:140.
- Knoll, A. H., Javaux, E. J., Hewitt, D. & P. Cohen. 2006. Eukaryotic organisms in Proterozoic oceans. *Philosophical Transactions of the Royal Society Biological Sciences*. 361: 1023–1038.
- Kodama, Y. & K. Watanabe. 2004. *Sulfuricurvum kujiense* gen. nov., sp. nov., a facultatively anaerobic, chemolithoautotrophic, sulfur-oxidizing bacterium isolated from an underground crude-oil storage cavity. *International Journal of Systematic Bacteriology*. 54(6):2297-2300.
- Kolber, Z. S., Plumley, F. G., Lang, A. S., Beatty, J. T., Blankenship, R. E., VanDover, C. L., Vetriani, C., Koblizek, M., Rathgeber, C. & P. G. Falkowski. 2001. Contribution of aerobic photoheterotrophic bacteria to the carbon cycle in the ocean. *Science*. 292:2492–2495.
- Koonin, E. V. & W. Martin. 2005. On the origin of genomes and cells within inorganic compartments. *Trends in Genetic*. 21:647–54.
- Koonin, E. V. 2003. Comparative genomics, minimal gene set and the last universal common ancestor. *Nature Reviews Microbiology*. 1:127-136.
- Kowalchuk, G. A. & J. R. Stephen. 2001. Ammonia-oxidizing bacteria: A model for molecular microbial ecology. *Annual Review of Microbiology*. 55:485-529.
- Kurland, C. G. 2000. Something for everyone. Horizontal gene transfer in evolution. *EMBO Reports*. 1:92-95.
- Kyripides, N., Overbeek, R. & C. Ouzounis. 1999. Universal protein families and the functional content of the last universal common ancestor. *Journal of Molecular Evolution*. 49:413–23.

- Lazcano, A. 1995. Cellular evolution during the Early Archaean: what happened between the progenote and the ancestor? *Microbiología SEM*. 11:185-198.
- Lazcano, A., Fox, G. E. & J. Oró. 1992. Life before DNA: the origin and early evolution of early Archean cells. In ed. Mortlock, R. P. *The Evolution of Metabolic Function*. Boca Raton, FL: CRC. pp. 237–95.
- Lazcano, A., Guerrero, R., Margulis, L. & J. Oró. 1988b. The evolutionary transition from RNA to DNA in early cells. *Journal of Molecular Evolution*. 27:283–90.
- Leipe, D. D., Aravind, L. & E. V. Koonin. 1999. Did DNA replication evolve twice independently? *Nucleic Acids Research*. 27:3389–401.
- Li, S., Tang, Y., Nie, Y., Cai, M. & X. Wu. 2011. Complete Genome Sequence of *Polymorphum gilvum* SL003B-26A1T, a Crude Oil-Degrading Bacterium from Oil-Polluted Saline Soil. *Journal of Bacteriology*. 193(11):2894-289.
- Line, M. A. 2002. The enigma of the origin of life and its timing. *Microbiology*. 148:21–27.
- Liu, B., Frostegård, A. & J. P. Shapleigh. 2013. Draft genome sequences of five strains in the genus *Thauera*. *Genome Announcements*. 1(1):e00052-12.
- Macinnes, J. I. & E. T. Lally. 2006. The genus *Actinobacillus*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 1094-1118.
- Makarova, K. S., Koonin, E. V. 2003. Comparative genomics of Archaea: How much have we learned in six years, and what's next? *Genome Biology*. 4:115–45.
- Manos, J. & Belas, R. 2006. The Genera *Proteus*, *Providencia*, and *Morganella*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 245-269.
- Mechichi, T., Stackebrandt, E. & G. Fuchs. 2003. *Alicycliphilus denitrificans* gen. nov., sp. nov., a cyclohexanol-degrading, nitrate-reducing β -proteobacterium. *International Journal Of Systematic And Evolutionary Microbiology*. 53:1147-152.

- Mikhailov, V. V., Romanenko, L. A. & E. P. Ivanova. 2006. The genus *Alteromonas* and related Proteobacteria. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 597-645.
- Mirkin, B. G., Fenner, T. I., Galperin, M. Y. & E. V. Koonin. 2003. Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last common ancestor, and dominance of horizontal gene transfer in the evolution of prokaryotes. *BMC Evolutionary Biology*. 3(2).
- Miroshnichenko, M. L., Rainey, F. A. & R. M. Bonch-Osmolovskaya. 1999. *Hippea maritima* gen. nov, sp. nov., a new genus of thermophilic, sulphur-reducing bacterium from submarine hot vents. *International Journal Of Systematic Bacteriology*. 49:1033-1038.
- Moore, E. R. B., Tindall, B. J., Martins Dos Santos, V. A P., Pieper, D. H., Ramos, J. L. N. J. Palleroni. 2006. Nonmedical: *Pseudomonas*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 646-703.
- Morange, M. 2011. Some considerations on the nature of LUCA, and the nature of life. *Research in Microbiology*. 162:5-9.
- Morris, R. M., Rappé, M. S., Connon, S. A., Vergin, K. L., Siebold, W. A., Carlson, C. A. & S. J. Giovannoni. 2002. SAR11 clade dominates ocean surface bacterioplankton communities. *Nature*. 420:806-810.
- Mushegian, A. R. & E. V. Koonin. 1996. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proceedings of the National Academy of Sciences USA*. 93:10268-10273.
- Nakagawa, S., Takai, K., Inagaki, F, Horikoshi, K. & Y. Sako. 2005. *Nitratiruptor tergarcus* gen. nov., sp. nov. and *Nitratifactor salsuginis* gen. nov., sp. nov., nitrate-reducing chemolithoautotrophs of the ϵ -Proteobacteria isolated from a deep-sea hydrothermal system in the Mid-Okinawa Trough. *International Journal of Systematic Bacteriology*. 55(2):925-933.

- Nakatsu, C. H., Hristova, K., Hanada, S. Meng, X., Hanson, J. R., Scow, K. M. & Y. Kamagata. 2006. *Methylibium petroleiphilum* gen. nov., sp. nov., a novel methyl tert-butyl ether-degrading methylotroph of the Betaproteobacteria. *International Journal Of Systematic And Evolutionary Microbiology*. 56(5):983-989.
- Nalin, R., Simonet, P., Vogel, T. M. & P. Normand. *Rhodanobacter lindaniclasticus* gen. nov., sp. nov., a lindane-degrading bacterium. *International Journal of Systematic Bacteriology*. 49:19-23.
- Neelson, K. H. & J. Scott. 2006. Ecophysiology of the Genus *Shewanella*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 1133-1151.
- Norksov-Lauristein, N. & M. Kilian 2006. Reclassification of *Actinobacillus actinomycetemcomitans*, *Haemophilus aphrophilus*, *Haemophilus paraphrophilus* and *Haemophilus segnis* as *Aggregatibacter actinomycetemcomitans* gen. nov., comb. nov., *Aggregatibacter aphrophilus* comb. nov. and *Aggregatibacter segnis* comb. nov., and emended description of *Aggregatibacter aphrophilus* to include V factor-dependent and V factor-independent isolates. *International Journal Of Systematic And Evolutionary Microbiology*. 56(9):2135-2146.
- Ouzounis, A. C., Kunin, V., Darzentas, N. & L. Goldovsky. 2006. A minimal estimate for gene content of the last universal common ancestor- exobiology from a terrestrial perspective. *Research in Microbiology*. 157:57–68.
- Palleroni, N. J. & J. F. Bradbury. 1993. *Stenotrophomonas*, a new bacterial genus for *Xanthoronas maltophilia* (Hugh 1980) Swings *et al.* 1983. *International Journal Of Systematic And Evolutionary Microbiology*. 43(3):606-609.
- Parfrey, L. W., Lahr, D. J., Knoll, A. H. & L. A. Katz. 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proceedings of the National Academy of Sciences USA*. 108: 13624–13629.
- Pays, J. P. 2006. The genus *Moraxella*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 958-987.

- Penny, D., McComish, B. J., Charleston, M. A. & M. D. Hendy. 2001. Mathematical elegance with biochemical realism: the covarion model of molecular evolution. *Journal of Molecular Evolution*. 53:711-723.
- Pinel, N., Davidson, S. K., Stahl, D.A. 2008. *Verminephrobacter eiseniae* gen. nov., sp. nov., a nephridial symbiont of the earthworm *Eisenia foetida* (Savigny). *International Journal Of Systematic And Evolutionary Microbiology*. 58:2147-57.
- Poindexter, J. S. 2006. Dimorphic Prosthecate Bacteria: The Genera *Caulobacter*, *Asticcacaulis*, *Hyphomicrobium*, *Pedomicrobium*, *Hyphomonas* and *Thiodendron*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. pp.72-90.
- Poole, A. M. 2002. My Name is LUCA—The Last Universal Common Ancestor. Disponible en: <http://www.actionbioscience.org/evolution/poolepaper.html>.
- R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna Austria. <http://www.R-project.org>.
- Ranea, A. G., Sillero, A., Thorton, M. J. & A. C. Orengo. 2006. Protein superfamily evolution and the last universal common ancestor (LUCA). *Journal of Molecular Evolution*. 63:513–25.
- Robertson L. A. & J. G. Kuenen. 2006. The genus *Thiobacillus*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 5. pp. 812-827.
- Schopf, J. W. 2006. Fossil evidence of Archaean life. *Philosophical Transactions of the Royal Society Biological Sciences*. 361(1470):869-885.
- Schulze, E.-D. & H. A. Mooney (Eds.) 1994. *Biodiversity and ecosystem function*. Springer-Verlag. Alemania. pp. 67-96.
- Seviour, R. J., Mino, T. & M. Onuki. 2003. The microbiology of biological phosphorus removal in activated sludge systems. *FEMS Microbiology Reviews*. 27(1):99-127.

- Sicheritz-Pontén, T. & S. G. E. Andersson 2001. A phylogenomic approach to microbial evolution. *Nucleic Acids Research*. 29:545-552.
- Smith, J. L., Campbell, B. J., Hanson, T. E., Zhang, C. L., Cary, S. C. 2008. *Nautilia profundicola* sp. nov., a thermophilic, sulfur-reducing epsilonproteobacterium from deep-sea hydrothermal vents. *International Journal Of Systematic And Evolutionary Microbiology*. 58(7):1598-602.
- Sobolevsky, Y. & E. N. Trifonov. 2006. Protein modules conserved since LUCA. *Journal of Molecular Evolution*. 63:622–34.
- Spain, A. M., Krumholz, L. R. & M. S. Elshahed. 2009. Abundance, composition, diversity and novelty of soil Proteobacteria. *The ISME Journal*. 3:992-1000.
- Spring, S. 2006. The genera *Leptothrix* and *Sphaerotilus*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 5. pp. 758-777.
- Stackebrandt, E., Murray, R. G. E. & H. G. Trüper. 1988. Proteobacteria classis nov., a name for the phylogenetic taxon that includes the 'purple bacteria and their relatives'. *International Journal of Systematic Bacteriology*. 38:321–325.
- Steitz, T. A. 1999. DNA polymerases: structural diversity and common mechanisms. *The Journal of Biological Chemistry*. 274:17395–98.
- Stetter, K.O. 2006. Hyperthermophiles in the history of life. *Philosophical Transactions of the Royal Society Biological Sciences*. 361:1837–43.
- Stubbe, J., Ge, J. & C. S. Yee. 2001. The evolution of ribonucleotide reduction revisited. *Trends in Biochemical Sciences*. 26:93–99.
- Su, Q., Xie, W., Wang, S., Wu, Q., Liu, B., Fang, Y., Xu, B. & Y. Zhang. 2014. The endosymbiont *Hamiltonella* increases the growth rate of its host *Bemisia tabaci* during periods of nutritional stress. *PLoS ONE* 9(2):e89002.
- Swings, J. 2006. The genus *Frateuria*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 844-845.

- Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J. & D. A. Natale. 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*. 4:41.
- Tatusov, R. L., Mushegian, A. R., Bork, P., Brown, N. P., Hayes, W. S., Borodovsky, M., Rudd, K. E. & E. V. Koonin. 1996. Metabolism and evolution of *Haemophilus influenzae* deduced from a whole-genome comparison with *Escherichia coli*. *Current Biology*. 6(3):279-291.
- Tekaiia, F., Dujon, B. & A. Lazcano. 1999. Comparative genomics: products of the most conserved protein-encoding genes synthesize degrade or interact with RNA. *Abstracts of the 9th ISSOL meeting, San Diego California, USA* 46:53.
- Thierry, S., Macarie, H., Lizuka, T., Geißdörfer, W., Assih, E. A. A., Spanevello, M., Verhe, F., Thomas, P., Fudou, R., Monroy, O., Labat, M. & A. S. Ouattara. 2004. *Pseudoxanthomonas mexicana* sp. nov. and *Pseudoxanthomonas japonensis* sp. nov., isolated from diverse environments, and emended descriptions of the genus *Pseudoxanthomonas* and of its type species. *International Journal of Systematic Bacteriology*. 54(6):2245-2255.
- Towner, K. 2006. The genus *Acinetobacter*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 746-758.
- Tuller, T., Birin, H., Gophna, U., Kupiec, M. & E. Ruppin. 2010. Reconstructing ancestral gene content by coevolution. *Genome Research*. 20:122-132.
- Valas, R. E. & P. E. Bourne. 2009. Structural analysis of polarizing indels: an emerging consensus on the root of the tree of life. *Biology Direct*. 25: 4-30.
- Vandamme, P. & T. Coenye. 2004. Taxonomy of the genus *Cupriavidus*: a tale of lost and found. *International Journal of Systematic and Evolutionary Microbiology*. 54(6):2285-2289.

- Welch, R. A. 2006. The genus *Escherichia*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 6. pp. 60-71.
- Willems, A. & P. De Vos. 2006. *Comamonas*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. Vol 5. pp. 723-736.
- Willey, J. M. Sherwood, L. M. & C. J. Woolverton. 2008. *Prescott, Harley, and Klein's Microbiology*. 7th edition. McGraw Hill. pp.539-570.
- Williams, T. A., Foster, P. G., Cox, C. J. & M. Embley. 2013. An archaeal origin of eukaryotes supports only two primary domains of life. *Nature*. 504: 231-236.
- Woese, C. R. & G. E. Fox. 1977a. The concept of cellular evolution. *Journal of Molecular Evolution*. 10:1-6.
- Woese, C. R. & G. E. Fox. 1977b. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences USA*. 74(11):5088-5090.
- Woese, C. R. 1987. Bacterial evolution. *Microbiological Reviews*. 51(2): 221-271.
- Woese, C. R. 1998. The universal ancestor. *Proceedings of the National Academy of Sciences USA*. 95:6854-59.
- Woese, C. R., Kandler, O. & Wheelis, M. L. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences USA*. 87(12): 4576-4579.
- Woo, P. C. Y., Lau, S. K. P., Tse, H., Teng, J. L. L., Curreem, S. O. T., Tsang, A. K. L., Fan, R. Y. Y., Wong, G. K. M., Huang, Y., Loman, N. J., Snyder, L. A. S., Cai, J. J., Huang, J., Mak, W., Pallen, M. J., Lok, S. & K. Yuen. 2009. The complete genome and proteome of *Laribacter hongkongensis* reveal potential mechanisms for adaptations to different temperatures and habitats. *PLoS Genetics*. 5(3):e1000416.

- Woods, D. E. & P. A. Sokol. 2006. The genus *Burkholderia*. In eds. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H. & E. Stackebrandt. *The Prokaryotes: A handbook on the biology of bacteria*. New York Springer. pp. 848-859.
- Yang, S., Doolittle, R. F. & P. E. Bourne. 2005. Phylogeny determined by protein domain content. *Proceedings of the National Academy of Sciences USA*. 102:373–78.
- Zhaxybayeva, O. & Gogarten, J. P. 2004. Cladogenesis, coalescence and the evolution of the three domains of life. *Trends in Genetics*. 20:182-187.

Anexos

Anexo A. Listado de proteobacterias alfa, orden Rhizobiales.

Especies de Rhizobiales consideradas para el análisis.			
Especie	Familia	Clave KEGG	Secuencias proteicas
<i>Bartonella grahamii</i>	Bartonellaceae	bgr	1768
<i>Bartonella henselae</i>	Bartonellaceae	bhe	1488
<i>Bartonella tribocorum</i>	Bartonellaceae	btr	2087
<i>Beijerinckia indica</i>	Beijerinckiaceae	bid	3784
<i>Methylocella silvestris</i>	Beijerinckiaceae	msl	3818
<i>Bradyrhizobium japonicum</i>	Bradyrhizobiaceae	bjj	8317
<i>Bradyrhizobium oligotrophica</i>	Bradyrhizobiaceae	aol	7228
<i>Bradyrhizobium sp. BTAi1</i>	Bradyrhizobiaceae	bbt	7621
<i>Bradyrhizobium sp. ORS 278</i>	Bradyrhizobiaceae	bra	6717
<i>Bradyrhizobium sp. S23321</i>	Bradyrhizobiaceae	brs	6892
<i>Nitrobacter hamburgensis</i>	Bradyrhizobiaceae	nha	4326
<i>Nitrobacter winogradskyi</i>	Bradyrhizobiaceae	nwi	3122
<i>Oligotropha carboxidovorans</i>	Bradyrhizobiaceae	oca	3722
<i>Rhodopseudomonas palustris</i>	Bradyrhizobiaceae	rpa	4820
<i>Brucella abortus</i>	Brucellaceae	bmb	3048
<i>Brucella canis</i>	Brucellaceae	bcs	3251
<i>Brucella melitensis</i>	Brucellaceae	bme	3198
<i>Brucella microti</i>	Brucellaceae	bmr	3282
<i>Brucella ovis</i>	Brucellaceae	bov	2890
<i>Brucella pinnipedialis</i>	Brucellaceae	bpp	3324
<i>Brucella suis</i>	Brucellaceae	bms	3271
<i>Ochrobactrum anthropi</i>	Brucellaceae	oan	4799
<i>Hyphomicrobium denitrificans</i>	Hyphomicrobiaceae	hdn	3512
<i>Hyphomicrobium sp. MC1</i>	Hyphomicrobiaceae	hmc	4893
<i>Pelagibacterium halotolerans</i>	Hyphomicrobiaceae	phl	3881
<i>Rhodomicrobium vannielii</i>	Hyphomicrobiaceae	rva	3565
<i>Methylobacterium extorquens</i>	Methylobacteriaceae	mex	4829
<i>Methylobacterium nodulans</i>	Methylobacteriaceae	mno	8308
<i>Methylobacterium populi</i>	Methylobacteriaceae	mpo	5365
<i>Methylobacterium radiotolerans</i>	Methylobacteriaceae	mrd	6431
<i>Methylobacterium sp. 4-46</i>	Methylobacteriaceae	met	6692

<i>Methylocystis sp. SC2</i>	Methylocystaceae	msc	3666
<i>Chelativorans sp. BNC1</i>	Phyllobacteriaceae	mes	4543
<i>Mesorhizobium australicum</i>	Phyllobacteriaceae	mam	5792
<i>Mesorhizobium ciceri</i>	Phyllobacteriaceae	mci	6264
<i>Mesorhizobium loti</i>	Phyllobacteriaceae	mlo	7272
<i>Mesorhizobium opportunistum</i>	Phyllobacteriaceae	mop	6508
<i>Agrobacterium radiobacter</i>	Rhizobiaceae	ara	6684
<i>Agrobacterium sp. H13-3</i>	Rhizobiaceae	agr	5345
<i>Agrobacterium tumefaciens</i>	Rhizobiaceae	atu	5355
<i>Agrobacterium vitis</i>	Rhizobiaceae	avi	5389
<i>Rhizobium etli</i>	Rhizobiaceae	ret	5963
<i>Rhizobium leguminosarum bv. viciae</i>	Rhizobiaceae	rle	7143
<i>Rhizobium tropici</i>	Rhizobiaceae	rtr	6287
<i>Sinorhizobium fredii</i>	Rhizobiaceae	rhi	6362
<i>Sinorhizobium medicae</i>	Rhizobiaceae	smd	6213
<i>Sinorhizobium meliloti</i>	Rhizobiaceae	sme	6218
<i>Parvibaculum lavamentivorans</i>	Rhodobiaceae	pla	3636
<i>Azorhizobium caulinodans</i>	Xanthobacteraceae	azc	4717
<i>Starkeya novella</i>	Xanthobacteraceae	sno	4431
<i>Xanthobacter autotrophicus</i>	Xanthobacteraceae	xau	5035

Anexo B. Listado de proteobacterias alfa no Rhizobiales.

Especies de proteobacterias alfa no Rhizobiales que fueron analizadas.				
Especie	Familia	Orden	Clave KEGG	Secuencias proteicas
<i>Asticcacaulis excentricus</i>	Caulobacteraceae	Caulobacterales	aex	3763
<i>Brevundimonas subvibrioides</i>	Caulobacteraceae	Caulobacterales	bsb	3327
<i>Caulobacter crescentus</i>	Caulobacteraceae	Caulobacterales	ccr	3737
<i>Caulobacter segnis</i>	Caulobacteraceae	Caulobacterales	cse	4139
<i>Caulobacter sp. K31</i>	Caulobacteraceae	Caulobacterales	cak	5438
<i>Phenylobacterium zucineum</i>	Caulobacteraceae	Caulobacterales	pzu	3854
<i>Magnetococcus marinus</i>	Magnetococcaceae	Magnetococcales	mgm	3716
<i>Micavibrio aeruginosavorus</i>	No clasificada	No clasificada	mai	2432
<i>Polymorphum gilvum</i>	No clasificada	No clasificada	pgv	4393
<i>Parvularcula bermudensis</i>	Parvularculaceae	Parvularculales	pbr	2687
<i>Hirschia baltica</i>	Hyphomonadaceae	Rhodobacterales	hba	3187
<i>Hyphomonas neptunium</i>	Hyphomonadaceae	Rhodobacterales	hne	3505

<i>Maricaulis maris</i>	Hyphomonadaceae	Rhodobacterales	mmr	3063
<i>Dinoroseobacter shibae</i>	Rhodobacteraceae	Rhodobacterales	dsh	4192
<i>Jannaschia</i> sp. CCS1	Rhodobacteraceae	Rhodobacterales	jan	4283
<i>Ketogulonicigenium vulgare</i>	Rhodobacteraceae	Rhodobacterales	kvu	3213
<i>Ketogulonigenium vulgare</i>	Rhodobacteraceae	Rhodobacterales	kvl	3054
<i>Octadecabacter antarcticus</i>	Rhodobacteraceae	Rhodobacterales	oat	4492
<i>Octadecabacter arcticus</i>	Rhodobacteraceae	Rhodobacterales	oar	4683
<i>Paracoccus denitrificans</i>	Rhodobacteraceae	Rhodobacterales	pde	5077
<i>Phaeobacter gallaeciensis</i>	Rhodobacteraceae	Rhodobacterales	pga	3875
<i>Pseudovibrio</i> sp. FO-BEG1	Rhodobacteraceae	Rhodobacterales	psf	5468
<i>Rhodobacter capsulatus</i>	Rhodobacteraceae	Rhodobacterales	rcp	3642
<i>Rhodobacter sphaeroides</i>	Rhodobacteraceae	Rhodobacterales	rsp	4242
<i>Roseobacter denitrificans</i>	Rhodobacteraceae	Rhodobacterales	rde	4129
<i>Roseobacter litoralis</i>	Rhodobacteraceae	Rhodobacterales	rli	4537
<i>Ruegeria pomeroyi</i>	Rhodobacteraceae	Rhodobacterales	sil	4252
<i>Ruegeria</i> sp. TM1040	Rhodobacteraceae	Rhodobacterales	sit	3864
<i>Acetobacter pasteurianus</i>	Acetobacteraceae	Rhodospirillales	apt	3049
<i>Acidiphilium cryptum</i>	Acetobacteraceae	Rhodospirillales	acr	3559
<i>Acidiphilium multivorum</i>	Acetobacteraceae	Rhodospirillales	amv	3948
<i>Gluconacetobacter diazotrophicus</i>	Acetobacteraceae	Rhodospirillales	gdi	3851
<i>Gluconacetobacter xylinus</i>	Acetobacteraceae	Rhodospirillales	gxy	3195
<i>Gluconobacter oxydans</i>	Acetobacteraceae	Rhodospirillales	gox	2664
<i>Granulibacter bethesdensis</i>	Acetobacteraceae	Rhodospirillales	gbe	2433
<i>Azospirillum brasilense</i>	Rhodospirillaceae	Rhodospirillales	abs	7557
<i>Azospirillum lipoferum</i>	Rhodospirillaceae	Rhodospirillales	ali	6094
<i>Azospirillum</i> sp. B510	Rhodospirillaceae	Rhodospirillales	azl	6309
<i>Endolissoclinum patella</i>	Rhodospirillaceae	Rhodospirillales	thal	1498
<i>Magnetospirillum magneticum</i>	Rhodospirillaceae	Rhodospirillales	mag	4561
<i>Rhodospirillum centenum</i>	Rhodospirillaceae	Rhodospirillales	rce	4003
<i>Rhodospirillum photometricum</i>	Rhodospirillaceae	Rhodospirillales	rpm	3281
<i>Rhodospirillum rubrum</i>	Rhodospirillaceae	Rhodospirillales	rru	3838
<i>Tistrella mobilis</i>	Rhodospirillaceae	Rhodospirillales	tmo	5783
<i>Orientia tsutsugamushi Ikeda</i>	Rickettsiaceae	Rickettsiales	ott	1967
<i>Rickettsia bellii</i>	Rickettsiaceae	Rickettsiales	rbo	1475
<i>Rickettsia felis</i>	Rickettsiaceae	Rickettsiales	rfe	1512
<i>Alpha proteobacterium HIMB59</i>	SAR11	SAR11	apc	1493
<i>Candidatus Pelagibacter</i> sp. IMCC9063	SAR11	SAR11	pel	1447
<i>Candidatus Pelagibacter ubique</i>	SAR11	SAR11	pub	1354

<i>Candidatus Puniceispirillum marinum</i>	SAR11	SAR11	apb	2543
<i>Erythrobacter litoralis</i>	Erythrobacteraceae	Sphingomonadales	eli	3011
<i>Novosphingobium aromaticivorans</i>	Sphingomonadaceae	Sphingomonadales	nar	3937
<i>Novosphingobium sp. PP1Y</i>	Sphingomonadaceae	Sphingomonadales	npp	4664
<i>Sphingobium chlorophenolicum</i>	Sphingomonadaceae	Sphingomonadales	sch	4072
<i>Sphingobium japonicum</i>	Sphingomonadaceae	Sphingomonadales	sjp	4394
<i>Sphingobium sp. SYK-6</i>	Sphingomonadaceae	Sphingomonadales	ssy	4062
<i>Sphingomonas sp. MM-1</i>	Sphingomonadaceae	Sphingomonadales	sphm	4361
<i>Sphingomonas wittichii</i>	Sphingomonadaceae	Sphingomonadales	swi	5345
<i>Sphingopyxis alaskensis</i>	Sphingomonadaceae	Sphingomonadales	sal	3195
<i>Zymomonas mobilis subsp. mobilis</i>	Sphingomonadaceae	Sphingomonadales	zmo	1736

Anexo C. Listado de proteobacterias beta, orden Burkholderiales.

Especies analizadas de proteobacterias beta del orden Burkholderiales.

Especie	Familia	Clave KEGG	Secuencias proteicas
<i>Achromobacter xylosoxidans</i>	Alcaligenaceae	axy	6815
<i>Advenella kashmirensis</i>	Alcaligenaceae	aka	3933
<i>Bordetella avium</i>	Alcaligenaceae	bav	3381
<i>Bordetella bronchiseptica</i>	Alcaligenaceae	bbh	4781
<i>Bordetella parapertussis</i>	Alcaligenaceae	bpar	4184
<i>Bordetella pertussis</i>	Alcaligenaceae	bpe	3436
<i>Bordetella petrii</i>	Alcaligenaceae	bpt	5027
<i>Pusillimonas sp. T7-7</i>	Alcaligenaceae	put	3773
<i>Taylorella asinigenitalis</i>	Alcaligenaceae	tas	1524
<i>Taylorella equigenitalis</i>	Alcaligenaceae	teq	1556
<i>Burkholderia ambifaria</i>	Burkholderiaceae	bam	6610
<i>Burkholderia cenocepacia</i>	Burkholderiaceae	bch	6919
<i>Burkholderia cepacia</i>	Burkholderiaceae	bct	5825
<i>Burkholderia gladioli</i>	Burkholderiaceae	bgd	7411
<i>Burkholderia glumae</i>	Burkholderiaceae	bgl	5773
<i>Burkholderia mallei</i>	Burkholderiaceae	bma	5023
<i>Burkholderia mallei</i>	Burkholderiaceae	bmv	5184
<i>Burkholderia multivorans</i>	Burkholderiaceae	bmu	6258
<i>Burkholderia phenoliruptrix</i>	Burkholderiaceae	bpx	6496
<i>Burkholderia phymatum</i>	Burkholderiaceae	bph	7496
<i>Burkholderia phytofirmans</i>	Burkholderiaceae	bpy	7241
<i>Burkholderia pseudomallei</i>	Burkholderiaceae	bps	5728
<i>Burkholderia rhizoxinica</i>	Burkholderiaceae	brh	3870

<i>Burkholderia</i> sp. 383	Burkholderiaceae	bur	7716
<i>Burkholderia</i> sp. CCGE1001	Burkholderiaceae	bug	5965
<i>Burkholderia</i> sp. CCGE1002	Burkholderiaceae	bge	6889
<i>Burkholderia</i> sp. CCGE1003	Burkholderiaceae	bgf	5988
<i>Burkholderia</i> sp. KJ006	Burkholderiaceae	buk	6024
<i>Burkholderia</i> sp. YI23	Burkholderiaceae	byi	7804
<i>Burkholderia thailandensis</i>	Burkholderiaceae	bte	5632
<i>Burkholderia vietnamiensis</i>	Burkholderiaceae	bvi	7617
<i>Burkholderia xenovorans</i>	Burkholderiaceae	bxo	8702
<i>Cupriavidus metallidurans</i>	Burkholderiaceae	rme	6477
<i>Cupriavidus necator</i>	Burkholderiaceae	cnc	7832
<i>Cupriavidus taiwanensis</i>	Burkholderiaceae	cti	5896
<i>Polynucleobacter</i> sp. QLW-P1DMWA-1	Burkholderiaceae	pnu	2077
<i>Ralstonia eutropha</i>	Burkholderiaceae	reu	6446
<i>Ralstonia pickettii</i>	Burkholderiaceae	rpi	4952
<i>Ralstonia solanacearum</i>	Burkholderiaceae	rso	5113
<i>Acidovorax avenae</i> subsp. <i>avenae</i>	Comamonadaceae	aaa	4737
<i>Acidovorax citrulli</i>	Comamonadaceae	aav	4709
<i>Acidovorax ebreus</i>	Comamonadaceae	dia	3479
<i>Acidovorax</i> sp. JS42	Comamonadaceae	ajs	4155
<i>Acidovorax</i> sp. KKS102	Comamonadaceae	ack	4742
<i>Alicyclophilus denitrificans</i>	Comamonadaceae	adn	4542
<i>Comamonas testosteroni</i>	Comamonadaceae	ctt	4799
<i>Delftia</i> sp. Cs1-4	Comamonadaceae	del	5861
<i>Polaromonas naphthalenivorans</i>	Comamonadaceae	pna	4929
<i>Polaromonas</i> sp. JS666	Comamonadaceae	pol	5453
<i>Ramlibacter tataouinensis</i>	Comamonadaceae	rta	3880
<i>Rhodoferrax ferrireducens</i>	Comamonadaceae	rfr	4417
<i>Rubrivivax gelatinosus</i>	Comamonadaceae	rge	4693
<i>Thiomonas arsenitoxydans</i>	Comamonadaceae	thi	3596
<i>Variovorax paradoxus</i>	Comamonadaceae	vap	6279
<i>Verminephrobacter eiseniae</i>	Comamonadaceae	vei	4947
<i>Delftia acidovorans</i>	Comamonadaceae	dac	6040
<i>Leptothrix cholodnii</i>	No clasificada	lch	4363
<i>Methylobium petroleiphilum</i>	No clasificada	mpt	4449
<i>Thiomonas intermedia</i>	No clasificada	tin	3172
<i>Collimonas fungivorans</i>	Oxalobacteraceae	cfu	4432
<i>Herbaspirillum seropedicae</i>	Oxalobacteraceae	hse	4735

<i>Herminiimonas arsenicoxydans</i>	Oxalobacteraceae	har	3295
<i>Minibacterium massiliensis</i>	Oxalobacteraceae	mms	3697

Anexo D. Listado de proteobacterias beta no Burkholderiales.

Especies de proteobacterias beta no Burkholderiales analizadas.				
Especie	Familia	Orden	Clave KEGG	Secuencias proteicas
<i>Gallionella capsiferriformans</i>	Gallionellaceae	Gallionellales	gca	2894
<i>Sideroxydans lithotrophicus</i>	Gallionellaceae	Gallionellales	slt	2980
<i>Thiobacillus denitrificans</i>	Hydrogenophilaceae	Hydrogenophilales	tbd	2827
<i>Methylobacillus flagellatus</i>	Methylophilaceae	Methylophilales	mfa	2753
<i>Methylotenera mobilis</i>	Methylophilaceae	Methylophilales	mmb	2338
<i>Methylotenera versatilis</i>	Methylophilaceae	Methylophilales	meh	2764
<i>Methylovorus glucosetrophus</i>	Methylophilaceae	Methylophilales	mei	2909
<i>Methylovorus sp. MP688</i>	Methylophilaceae	Methylophilales	mep	2712
<i>Chromobacterium violaceum</i>	Chromobacteriaceae	Neisseriales	cvi	4405
<i>Laribacter hongkongensis</i>	Chromobacteriaceae	Neisseriales	lhk	3235
<i>Pseudogulbenkiania sp. NH8B</i>	Chromobacteriaceae	Neisseriales	pse	4012
<i>Neisseria gonorrhoeae</i>	Neisseriaceae	Neisseriales	ngo	2002
<i>Neisseria lactamica</i>	Neisseriaceae	Neisseriales	nla	1972
<i>Neisseria meningitidis</i>	Neisseriaceae	Neisseriales	nme	2063
<i>Nitrosomonas europaea</i>	Nitrosomonadaceae	Nitrosomonadales	neu	2461
<i>Nitrosomonas eutropha</i>	Nitrosomonadaceae	Nitrosomonadales	net	2551
<i>Nitrosomonas sp. AL212</i>	Nitrosomonadaceae	Nitrosomonadales	nit	2983
<i>Nitrosomonas sp. Is79A3</i>	Nitrosomonadaceae	Nitrosomonadales	nii	3372
<i>Nitrospira multiformis</i>	Nitrosomonadaceae	Nitrosomonadales	nmu	2805
<i>Accumulibacter phosphatis</i>	No clasificada	No clasificada	app	4562
<i>Beta proteobacterium</i>	No clasificada	No clasificada	bprc	2112
<i>Aromatoleum aromaticum</i>	Rhodocyclaceae	Rhodocyclales	eba	4579
<i>Azoarcus sp. BH72</i>	Rhodocyclaceae	Rhodocyclales	azo	3989
<i>Azoarcus sp. KH32C</i>	Rhodocyclaceae	Rhodocyclales	aza	5188
<i>Dechloromonas aromatica</i>	Rhodocyclaceae	Rhodocyclales	dar	4171
<i>Dechlorosoma suillum</i>	Rhodocyclaceae	Rhodocyclales	dsu	3443
<i>Thauera sp. MZ1T</i>	Rhodocyclaceae	Rhodocyclales	tmz	3978

Anexo E. Listado de proteobacterias gamma, orden Enterobacteriales.

Especies de proteobacterias gamma del orden Enterobacteriales que fueron analizadas.

Especie	Clave KEGG	Secuencias proteicas
<i>Candidatus Hamiltonella defensa</i>	hde	2148
<i>Citrobacter koseri</i>	cko	5006
<i>Citrobacter rodentium</i>	cro	4913
<i>Cronobacter sakazakii</i>	esa	4422
<i>Cronobacter turicensis</i>	ctu	4452
<i>Dickeya dadantii</i>	dda	3970
<i>Dickeya zeae</i>	dze	4163
<i>Edwardsiella ictaluri</i>	eic	3735
<i>Edwardsiella tarda</i>	etr	3591
<i>Enterobacter aerogenes</i>	eae	4912
<i>Enterobacter asburiae</i>	eas	4612
<i>Enterobacter cloacae</i>	enc	5518
<i>Enterobacter lignolyticus</i>	esc	4399
<i>Enterobacter sp. 638</i>	ent	4240
<i>Enterobacter sp. R4-368</i>	enr	4827
<i>Enterobacteriaceae bacterium FGI 57</i>	ebf	4323
<i>Erwinia amylovora</i>	eay	3565
<i>Erwinia billingiae</i>	ebi	4917
<i>Erwinia pyrifoliae</i>	epy	3697
<i>Erwinia sp. Ejp617</i>	erj	3672
<i>Erwinia tasmaniensis</i>	eta	3622
<i>Escherichia coli</i>	eco	4140
<i>Escherichia fergusonii</i>	efe	4318
<i>Klebsiella oxytoca</i>	kox	5488
<i>Klebsiella pneumoniae</i>	kpn	5185
<i>Klebsiella variicola</i>	kva	5057
<i>Morganella morganii</i>	mmk	3526
<i>Pantoea ananatis</i>	pam	4241
<i>Pantoea sp. At-9b</i>	pao	5770
<i>Pantoea vagans</i>	pva	4590
<i>Pectobacterium atrosepticum</i>	eca	4472
<i>Pectobacterium carotovorum</i>	pct	4246
<i>Pectobacterium sp. SCC3193</i>	pec	4705
<i>Pectobacterium wasabiae</i>	pwa	4437

<i>Photorhabdus asymbiotica</i>	pay	4417
<i>Photorhabdus luminescens</i>	plu	4683
<i>Proteus mirabilis</i>	pmr	3662
<i>Providencia stuartii</i>	psi	4099
<i>Rahnella aquatilis</i>	raq	4866
<i>Rahnella sp. Y9602</i>	rah	5111
<i>Raoultella ornithinolytica</i>	ror	4909
<i>Salmonella bongori</i>	sbg	3863
<i>Salmonella enterica</i>	stt	4352
<i>Salmonella enterica</i>	stm	4558
<i>Serratia liquefaciens</i>	slq	4894
<i>Serratia marcescens</i>	smaf	4361
<i>Serratia plymuthica</i>	srr	4952
<i>Serratia proteamaculans</i>	spe	4942
<i>Serratia sp. AS12</i>	srs	4952
<i>Serratia sp. AS13</i>	sra	4951
<i>Shigella boydii</i>	sbo	4281
<i>Shigella dysenteriae</i>	sdv	4501
<i>Shigella flexneri</i>	sfl	4439
<i>Shigella sonnei</i>	ssn	4470
<i>Shimwellia blattae</i>	ebt	3904
<i>Sodalis glossinidius</i>	sgl	2516
<i>Xenorhabdus bovienii</i>	xbo	4258
<i>Xenorhabdus nematophila</i>	xne	4473
<i>Yersinia enterocolitica</i>	yen	4050
<i>Yersinia pestis</i>	ype	4066
<i>Yersinia pseudotuberculosis</i>	yps	4038

Anexo F. Listado de proteobacterias gamma no Enterobacteriales.

Especies de proteobacterias gamma no Enterobacteriales analizadas.				
Especie	Familia	Orden	Clave KEGG	Secuencias proteicas
<i>Colwellia psychrerythraea</i>	Colwelliaceae	Alteromonadales	cps	4910
<i>Idiomarina loihiensis</i>	Idiomarinaceae	Alteromonadales	ilo	2628
<i>Pseudoalteromonas atlantica</i>	Pseudoalteromonadaceae	Alteromonadales	pat	4281
<i>Pseudoalteromonas haloplanktis</i>	Pseudoalteromonadaceae	Alteromonadales	pha	3485
<i>Pseudoalteromonas sp. SM9913</i>	Pseudoalteromonadaceae	Alteromonadales	psm	3712

<i>Shewanella amazonensis</i>	Shewanellaceae	Alteromonadales	saz	3645
<i>Shewanella baltica</i>	Shewanellaceae	Alteromonadales	sbl	4489
<i>Shewanella denitrificans</i>	Shewanellaceae	Alteromonadales	sdn	3754
<i>Shewanella frigidimarina</i>	Shewanellaceae	Alteromonadales	sfr	4029
<i>Shewanella halifaxensis</i>	Shewanellaceae	Alteromonadales	shl	4278
<i>Shewanella loihica</i>	Shewanellaceae	Alteromonadales	slo	3859
<i>Shewanella oneidensis</i>	Shewanellaceae	Alteromonadales	son	4217
<i>Shewanella piezotolerans</i>	Shewanellaceae	Alteromonadales	swp	4911
<i>Shewanella putrefaciens</i>	Shewanellaceae	Alteromonadales	spc	3972
<i>Shewanella sediminis</i>	Shewanellaceae	Alteromonadales	sse	4497
<i>Shewanella violacea</i>	Shewanellaceae	Alteromonadales	svo	4346
<i>Shewanella woodyi</i>	Shewanellaceae	Alteromonadales	swd	4880
<i>Actinobacillus pleuropneumoniae</i>	Pasteurellaceae	Pasteurellales	apl	2013
<i>Actinobacillus succinogenes</i>	Pasteurellaceae	Pasteurellales	asu	2079
<i>Actinobacillus suis</i>	Pasteurellaceae	Pasteurellales	asi	2249
<i>Aggregatibacter actinomycetemcomitans</i>	Pasteurellaceae	Pasteurellales	aat	2214
<i>Aggregatibacter aphrophilus</i>	Pasteurellaceae	Pasteurellales	aap	2219
<i>Bibersteinia trehalosi</i>	Pasteurellaceae	Pasteurellales	bto	2250
<i>Gallibacterium anatis</i>	Pasteurellaceae	Pasteurellales	gan	2500
<i>Haemophilus influenzae</i>	Pasteurellaceae	Pasteurellales	hin	1657
<i>Haemophilus parainfluenzae</i>	Pasteurellaceae	Pasteurellales	hpr	1975
<i>Haemophilus parasuis</i>	Pasteurellaceae	Pasteurellales	hap	2021
<i>Haemophilus somnus</i>	Pasteurellaceae	Pasteurellales	hso	1798
<i>Mannheimia haemolytica</i>	Pasteurellaceae	Pasteurellales	mht	2537
<i>Mannheimia succiniciproducens</i>	Pasteurellaceae	Pasteurellales	msu	2370
<i>Pasteurella multocida subsp. multocida</i>	Pasteurellaceae	Pasteurellales	pmu	2012
<i>Acinetobacter baumannii</i>	Moraxellaceae	Pseudomonadales	acb	3367
<i>Acinetobacter calcoaceticus</i>	Moraxellaceae	Pseudomonadales	acc	3599
<i>Acinetobacter oleivorans</i>	Moraxellaceae	Pseudomonadales	acd	3874
<i>Acinetobacter sp. ADP1</i>	Moraxellaceae	Pseudomonadales	aci	3307
<i>Moraxella catarrhalis</i>	Moraxellaceae	Pseudomonadales	mct	1886
<i>Psychrobacter arcticus</i>	Moraxellaceae	Pseudomonadales	par	2120
<i>Psychrobacter cryohalolentis</i>	Moraxellaceae	Pseudomonadales	pcr	2511
<i>Psychrobacter sp. PRwf-1</i>	Moraxellaceae	Pseudomonadales	prw	2385
<i>Azotobacter vinelandii</i>	Pseudomonadaceae	Pseudomonadales	avn	5050
<i>Cellvibrio japonicus</i>	Pseudomonadaceae	Pseudomonadales	cja	3750
<i>Pseudomonas aeruginosa</i>	Pseudomonadaceae	Pseudomonadales	pae	5571
<i>Pseudomonas brassicacearum</i>	Pseudomonadaceae	Pseudomonadales	pba	6095

<i>Pseudomonas denitrificans</i>	Pseudomonadaceae	Pseudomonadales	pdr	5056
<i>Pseudomonas fluorescens</i>	Pseudomonadaceae	Pseudomonadales	pfo	5722
<i>Pseudomonas fulva</i>	Pseudomonadaceae	Pseudomonadales	pfv	4461
<i>Pseudomonas mendocina</i>	Pseudomonadaceae	Pseudomonadales	pmy	4594
<i>Pseudomonas poae</i>	Pseudomonadaceae	Pseudomonadales	ppz	4796
<i>Pseudomonas protegens</i>	Pseudomonadaceae	Pseudomonadales	pfl	6108
<i>Pseudomonas putida</i>	Pseudomonadaceae	Pseudomonadales	ppu	5350
<i>Pseudomonas resinovorans</i>	Pseudomonadaceae	Pseudomonadales	pre	5875
<i>Pseudomonas stutzeri</i>	Pseudomonadaceae	Pseudomonadales	psa	4128
<i>Aliivibrio fischeri</i>	Vibrionaceae	Vibrionales	vfi	3815
<i>Aliivibrio salmonicida</i>	Vibrionaceae	Vibrionales	vsa	3911
<i>Photobacterium profundum</i>	Vibrionaceae	Vibrionales	ppr	5489
<i>Vibrio anguillarum</i>	Vibrionaceae	Vibrionales	van	3732
<i>Vibrio campbellii</i>	Vibrionaceae	Vibrionales	vha	6041
<i>Vibrio cholerae</i>	Vibrionaceae	Vibrionales	vch	3504
<i>Vibrio furnissii</i>	Vibrionaceae	Vibrionales	vfu	4455
<i>Vibrio parahaemolyticus</i>	Vibrionaceae	Vibrionales	vpa	4832
<i>Vibrio splendidus</i>	Vibrionaceae	Vibrionales	vsp	4432
<i>Frateuria aurantia</i>	Xanthomonadaceae	Xanthomonadales	fau	3101
<i>Pseudoxanthomonas spadix</i>	Xanthomonadaceae	Xanthomonadales	psd	3149
<i>Pseudoxanthomonas suwonensis</i>	Xanthomonadaceae	Xanthomonadales	psu	3070
<i>Rhodanobacter denitrificans</i>	Xanthomonadaceae	Xanthomonadales	rhd	3824
<i>Stenotrophomonas maltophilia</i>	Xanthomonadaceae	Xanthomonadales	sml	4386
<i>Xanthomonas albilineans</i>	Xanthomonadaceae	Xanthomonadales	xal	3208
<i>Xanthomonas campestris</i>	Xanthomonadaceae	Xanthomonadales	xcc	4179
<i>Xanthomonas citri</i>	Xanthomonadaceae	Xanthomonadales	xcu	4760
<i>Xanthomonas oryzae</i>	Xanthomonadaceae	Xanthomonadales	xoo	4065
<i>Xylella fastidiosa</i>	Xanthomonadaceae	Xanthomonadales	xfa	2832

Anexo G. Listado de proteobacterias delta.

Especies de proteobacterias delta consideradas en el análisis.				
Especie	Familia	Orden	Clave KEGG	Secuencias proteicas
<i>Bacteriovorax marinus</i>	Bdellovibrionaceae	Bdellovibrionales	bmx	3231
<i>Bdellovibrio bacteriovorus</i>	Bdellovibrionaceae	Bdellovibrionales	bba	3586
<i>Bdellovibrio bacteriovorus Tiberius</i>	Bdellovibrionaceae	Bdellovibrionales	bbat	3738
<i>Bdellovibrio exovorus</i>	Bdellovibrionaceae	Bdellovibrionales	bex	2618

<i>Desulfarculus baarsii</i>	Desulfarculaceae	Desulfarculales	dbr	3277
<i>Candidatus Desulfococcus oleovorans</i>	Desulfobacteraceae	Desulfobacterales	dol	3265
<i>Desulfatibacillum alkenivorans</i>	Desulfobacteraceae	Desulfobacterales	dal	5252
<i>Desulfobacterium autotrophicum</i>	Desulfobacteraceae	Desulfobacterales	dat	4943
<i>Desulfobacula toluolica</i>	Desulfobacteraceae	Desulfobacterales	dto	4375
<i>Desulfobulbus propionicus</i>	Desulfobulbaceae	Desulfobacterales	dpr	3283
<i>Desulfocapsa sulfexigens</i>	Desulfobulbaceae	Desulfobacterales	dsf	3532
<i>Desulfotalea psychrophila</i>	Desulfobulbaceae	Desulfobacterales	dps	3234
<i>Desulfurivibrio alkaliphilus</i>	Desulfobulbaceae	Desulfobacterales	dak	2620
<i>Desulfohalobium retbaense</i>	Desulfohalobiaceae	Desulfovibrionales	drt	2526
<i>Desulfomicrobium baculatum</i>	Desulfomicrobiaceae	Desulfovibrionales	dba	3436
<i>Desulfovibrio aespoensis</i>	Desulfovibrionaceae	Desulfovibrionales	das	3304
<i>Desulfovibrio africanus</i>	Desulfovibrionaceae	Desulfovibrionales	daf	3725
<i>Desulfovibrio alaskensis</i>	Desulfovibrionaceae	Desulfovibrionales	dde	3258
<i>Desulfovibrio desulfuricans</i>	Desulfovibrionaceae	Desulfovibrionales	dds	2356
<i>Desulfovibrio desulfuricans</i>	Desulfovibrionaceae	Desulfovibrionales	ddn	3454
<i>Desulfovibrio hydrothermalis</i>	Desulfovibrionaceae	Desulfovibrionales	dhy	3548
<i>Desulfovibrio magneticus</i>	Desulfovibrionaceae	Desulfovibrionales	dma	4700
<i>Desulfovibrio piezophilus</i>	Desulfovibrionaceae	Desulfovibrionales	dpi	3354
<i>Desulfovibrio salexigens</i>	Desulfovibrionaceae	Desulfovibrionales	dsa	3807
<i>Desulfovibrio vulgaris</i>	Desulfovibrionaceae	Desulfovibrionales	dvl	3091
<i>Desulfovibrio vulgaris</i>	Desulfovibrionaceae	Desulfovibrionales	dvg	3221
<i>Desulfovibrio vulgaris Hildenborough</i>	Desulfovibrionaceae	Desulfovibrionales	dvu	3536
<i>Hippea maritima</i>	Desulfurellaceae	Desulfurellales	hmr	1677
<i>Geobacter bemidjensis</i>	Geobacteraceae	Desulfuromonadales	gbm	4057
<i>Geobacter lovleyi</i>	Geobacteraceae	Desulfuromonadales	glo	3685
<i>Geobacter metallireducens</i>	Geobacteraceae	Desulfuromonadales	gme	3568
<i>Geobacter sp. FRC-32</i>	Geobacteraceae	Desulfuromonadales	geo	3798
<i>Geobacter sp. M18</i>	Geobacteraceae	Desulfuromonadales	geb	4434
<i>Geobacter sp. M21</i>	Geobacteraceae	Desulfuromonadales	gem	4080
<i>Geobacter sulfurreducens</i>	Geobacteraceae	Desulfuromonadales	gsu	3432
<i>Geobacter sulfurreducens</i>	Geobacteraceae	Desulfuromonadales	gsk	3328
<i>Geobacter uraniireducens</i>	Geobacteraceae	Desulfuromonadales	gur	4358
<i>Pelobacter carbinolicus</i>	Pelobacteraceae	Desulfuromonadales	pca	3304
<i>Pelobacter propionicus</i>	Pelobacteraceae	Desulfuromonadales	ppd	3804
<i>Anaeromyxobacter dehalogenans</i>	Anaeromyxobacteraceae	Myxococcales	ade	4346
<i>Anaeromyxobacter dehalogenans</i>	Anaeromyxobacteraceae	Myxococcales	acp	4473
<i>Anaeromyxobacter sp. Fw109-5</i>	Anaeromyxobacteraceae	Myxococcales	afw	4466

<i>Anaeromyxobacter sp. K</i>	Anaeromyxobacteraceae	Myxococcales	ank	4457
<i>Stigmatella aurantiaca</i>	Cystobacteraceae	Myxococcales	sur	8352
<i>Haliangium ochraceum</i>	Kofleriaceae	Myxococcales	hoh	6719
<i>Corallococcus coralloides</i>	Myxococcaceae	Myxococcales	ccx	8033
<i>Myxococcus fulvus</i>	Myxococcaceae	Myxococcales	mfu	7284
<i>Myxococcus stipitatus</i>	Myxococcaceae	Myxococcales	msd	8043
<i>Myxococcus xanthus</i>	Myxococcaceae	Myxococcales	mxa	7316
<i>Sorangium cellulorum</i>	Polyangiaceae	Myxococcales	scu	10400
<i>Sorangium cellulorum</i>	Polyangiaceae	Myxococcales	scl	9380
<i>Desulfobacca acetoxidans</i>	Syntrophaceae	Syntrophobacterales	dao	2866
<i>Desulfomonile tiedjei</i>	Syntrophaceae	Syntrophobacterales	dti	5494
<i>Syntrophus aciditrophicus</i>	Syntrophaceae	Syntrophobacterales	sat	3166
<i>Syntrophobacter fumaroxidans</i>	Syntrophobacteraceae	Syntrophobacterales	sfu	4064

Anexo H. Listado de proteobacterias epsilon.

Especies de proteobacterias epsilon presentes en la muestra analizada.				
Especie	Familia	Orden	Clave KEGG	Secuencias proteicas
<i>Sulfurimonas denitrificans</i>	Helicobacteraceae	Campylobacterales	tdn	2096
<i>Sulfurimonas autotrophica</i>	Helicobacteraceae	Campylobacterales	sua	2158
<i>Sulfuricurvum kujense</i>	Helicobacteraceae	Campylobacterales	sku	2798
<i>Arcobacter butzleri</i>	Campylobacteraceae	Campylobacterales	abu	2259
<i>Arcobacter butzleri</i>	Campylobacteraceae	Campylobacterales	abt	2158
<i>Arcobacter nitrofigilis</i>	Campylobacteraceae	Campylobacterales	ant	3126
<i>Arcobacter sp. L</i>	Campylobacteraceae	Campylobacterales	arc	2848
<i>Sulfurospirillum deleyianum</i>	Campylobacteraceae	Campylobacterales	sdl	2265
<i>Sulfurospirillum barnesii</i>	Campylobacteraceae	Campylobacterales	sba	2491
<i>Nitratifractor salsuginis</i>	No clasificada	Campylobacterales	nsa	2088
<i>Nitratiruptor sp. SB155-2</i>	No clasificada	No clasificada	nis	1843
<i>Sulfurovum sp. NBC37-1</i>	No clasificada	No clasificada	sun	2438
<i>Nautilia profundicola</i>	Nautiliaceae	Nautiliales	nam	1730