

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
FACULTAD DE CIENCIAS



**ANÁLISIS BIOINFORMÁTICO DE LAS ENZIMAS QUE SINTETIZAN LOS ÁCIDOS
GRASOS EN *Archaea***

TESIS
QUE PARA OBTENER EL TÍTULO DE
BIÓLOGO
PRESENTA
DANIEL DARIO ORTIZ ALDANA

ENSENADA, B. C.

OCTUBRE DE 2009

Agradecimientos

A mi padre, Javier, y a mi madre, Doris, por predicar con el ejemplo y su inagotable generosidad y amor. A mis hermanas, Ana y Montse, por su cariño y tolerancia. A mis tíos, Carlos y Lina y a mis primos, Carlos Francisco y María Andrea, por acogerme como a uno propio. A toda mi familia.

A Ariadna, por su amor y paciencia y por otras tantas razones que no cabrían aquí o no son apropiadas para este tipo de literatura.

Al Dr. Antonio Lazcano por permitirme formar parte del laboratorio de origen de la vida. Al Dr. Arturo Becerra por aceptarme como su pupilo y obsequiarme su amistad. A mis compañeros de laboratorio: Ricardo, Fernando, Mario, Chuchín, Yetzi, Ervin, Ana y Sara.

A todos mis maestros, especialmente al Dr. Carlos Márquez por su confianza e invaluable consejo. También al profesor Pedro Ruiz, por ser mi amigo fuera del aula. Y al profesor Ramón Meza Soria, por enseñarme a estudiar. A mi comité de titulación, por sus comentarios y sugerencias.

A mis amigos: Sergio, Martín, Abraham y Pablo, por su inquebrantable lealtad. A mis compañeros: Adri, Jovani, Luis y Lorei, por los buenos tiempos. A Alán, por sus imperecederas lecciones de vida. A Paloma, la otra amiga que conservo del DF. A todos los que de una manera u otra me han echado la mano.

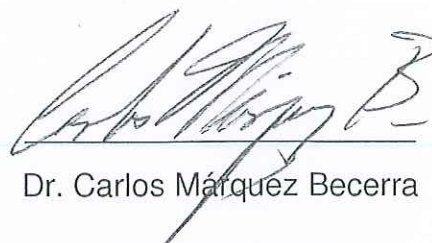
Resumen

El propósito de este trabajo fue detectar y establecer la ubicación de las diferencias y semejanzas de las enzimas que participan en la biosíntesis de ácidos grasos de las arquea utilizando herramientas bioinformáticas como Clustal X, MEGA y BLAST y consultando bases de datos como KEGG, Pfam y SCOP. Sólo se detectaron 6 enzimas involucradas en la síntesis de ácidos grasos en las arquea: 6.3.4.14, 6.4.1.2, 2.3.1.86, 2.3.1.180, 1.1.1.100 y 1.3.1.39. El análisis sustenta la hipótesis de que la ruta metabólica FAS II se ensambló de acuerdo al modelo tipo "patchwork". El análisis filogenético de las enzimas no coincidió con el árbol genealógico de las especies. Los sesgos constitutivos, la tasa evolutiva variable, la heterotaquia y el transporte horizontal de genes pudieron haber alterado el análisis filogenético. La prueba de Kruskal-Wallis indica que la mayoría de las enzimas presentan más homologías que lo esperado. Los análisis estadísticos de regresión lineal y multifactorial descartaron la posibilidad de que el pH, la temperatura y la salinidad hayan propiciado la elevada tasa de duplicación génica.



Dr. Arturo Carlos Il Becerra Bracho

Director de Tesis



Dr. Carlos Márquez Becerra

Codirector de Tesis

Abstract

The main purpose of this study was to detect and place the differences and similarities among the enzymes that synthesize fatty acids in *Archaea* using bioinformatic tools such as Clustal X, MEGA and BLAST and consulting specialized databases such as KEGG, Pfam and SCOP. Only 6 enzymes were detected: 6.3.4.14, 6.4.1.2, 2.3.1.86, 2.3.1.180, 1.1.1.100 and 1.3.1.39. The structural comparison supports the “patchwork” evolution model of the FAS II metabolic pathway. The phylogenetic analysis of the enzymes didn't always match with previous species phylogeny. Compositional biases, long branch attraction, heterotachy and horizontal gene transfer were all factors that modified the internal branching pattern. Kruskal-Wallis test indicated that most enzymes present more homologies than expected. Linear regression and multi-factorial regression analysis ruled out the possibility of pH, temperature and ion concentration as being factors favoring the elevated duplication rate.



Dr. Arturo Carlos Becerra Bracho

Director de Tesis



Dr. Carlos Márquez Becerra

Codirector de tesis

Índice

Votos Aprobatorios

Agradecimientos

Resumen

Abstract

Índice

Índice de Figuras

Índice de Tablas

I Introducción	1
1.1 Biología de las arquea	1
1.2 Características de la membrana citoplasmática de las arquea	3
1.3 Biosíntesis de ácidos grasos en bacterias (FAS II)	6
1.4 Origen y evolución de rutas metabólicas	11
1.5 Las moléculas como fuente de información	16
II Antecedentes	19
III Objetivos	20
IV Hipótesis	21
V Material y Métodos	22
5.1 Análisis funcional y estructural	22
5.2 Análisis filogenético y estadístico	23
VI Resultados	24
6.1 Análisis funcional y estructural	24
6.2 Análisis filogenético y estadístico	33
VII Discusiones	41
7.1 Análisis funcional y estructural	41
7.2 Análisis filogenético y estadístico	43
VI Conclusiones	48
VII Literatura Citada	49
VII Anexos	59

Índice de Figuras

Figura 1. Características de los lípidos de las arquea. Los fosfolípidos de las arquea se caracterizan por (i) su estructura G-1-P, (ii) tener enlaces de éter, (iii) tener cadenas de isoprenoides y (iv) lípidos tetraetéricos bipolares. La mayoría de las cabezas polares son compartidas por arqueas y bacterias. Imagen modificada de Koga y Morii, 2007. 4

Figura 2. Ciclo iterativo de reacciones para la adición de dos carbonos. (1) Condensación, (2) reducción, (3) deshidratación y (4) reducción. Modificado de Gurr *et al.* 2002. 9

Figura 3. Modelo retrógrado de evolución de rutas metabólicas. Un organismo heterotrófico agota el abastecimiento natural del metabolito A. El reclutamiento fortuito de una enzima (Enz 1) capaz de sintetizar A de B y C confiere una ventaja al organismo. Cuando las concentraciones de B y E decaen. Son compensadas por el reclutamiento de las enzimas 'Enz' 2 y 'Enz 3', respectivamente. Modificado de Rison y Thornton, 2002. 12

Figura 4. Modelo "patchwork" de evolución de rutas metabólicas. Este modelo pudo favorecer un mecanismo de sustrato específico (a), sin embargo las enzimas presentaban baja especificidad (b). La duplicación de cualquier gen sería muy ventajosa pues más metabolito sería sintetizado (c). Esta duplicación, seguida por la especialización de la enzima daría lugar a

las rutas que existen actualmente (d). Modificado de Rison y Thornton, 2002.

14

Figura 5. Árbol filogenético de la subunidad carboxil transferasa de la enzima Acetil-CoA carboxilasa (6.4.1.2). Los números en las ramas son las posibilidades bayesianas del clado.

34

Figura 6. Árbol filogenético de la subunidad BCCP de la enzima Acetil-CoA carboxilasa (6.4.1.2). Los números en las ramas son las posibilidades bayesianas del clado.

34

Figura 7. Árbol filogenético de la Biotina Carboxilasa (6.3.4.14). Los números en las ramas son las posibilidades bayesianas del clado.

36

Figura 8. Árbol filogenético de la β -cetoacil-[ACP] sintetasa III (2.3.1.180). Los números en las ramas son las posibilidades bayesianas del clado.

37

Figura 9. Árbol filogenético de las enzimas 3-cetoacil-[ACP] reductasa (1.1.1.100) y enoil-[ACP] reductasa (NADPH) (1.3.1.39). Los números en las ramas son las posibilidades bayesianas del clado.

38

Figura 10. El rango de cada muesca muestra la incertidumbre asociada con el estimado de la mediana de ese grupo. Las muescas son

construidas de tal modo que cualesquiera dos muestras que no se traslapen pueden declararse significativamente distintas ($p < 0.05$). 39

Figura 11. Comparación de la estructura terciaria y de los sitios activos de cinco enzimas de la superfamilia de las tiolosas. (A) Estructura monomérica con las α -hélices en rojo y hojas β en verde. (B) Sitios activos de las enzimas. La cisteína del sitio activo nucleofílico se posiciona en el centro, se puede ver que los residuos de histidina (H) de la tiolasa y la KAS I son equivalentes a los residuos de asparagina (N) de la KAS III, CHS y HMGS. Imagen modificada de Jiang *et al.* 2008. 42

Figura 12. Alineamiento de las secuencias de aminoácidos de la enzima 2.3.1.180 y la enzima HMGS reportada por Jiang *et al.* 2008. 44

Figura 13. Árbol filogenético de 226 secuencias de rRNA 16S de arqueas y bacterias, generado por máxima similitud. Modificado de Brochiere-Armanet *et al.* 2008. 63

Índice de Tablas

Tabla I. Enzimas que catalizan la biosíntesis de ácidos grasos tipo II (FAS II) en *Escherichia coli* HS (ecx), *Mycobacterium tuberculosis* CDC 1551 (mtc) y *Salmonella enterica* subsp. enterica serovar choleraesius (sec). El número EC es un sistema de clasificación numérica de las enzimas basado en la reacción que estas catalizan. 22

Tabla II. Resultados del BLAST. Acetil CoA-Carboxilasa (6.4.1.2) y biotina carboxilasa (6.3.4.14)), Malonil CoA : ACP transacilasa (2.3.1.39), 3-cetoacil-ACP sintetasa III (2.3.1.180), 3-cetoacil-ACP sintetasa II (2.3.1.179), 3-cetoacil-ACP sintetasa I (2.3.1.41), lipoacil-CoA sintetasa (2.3.1.86), 3-cetoacil-ACP reductasa (1.1.1.100) y enoil-ACP reductasa (NADPH) (1.3.1.39). 25

Tabla III. Tabla comparativa de dominios entre las secuencias problema *Escherichia coli* HS y los homólogos detectados por el BLAST en *Archaea*. 26-27

Tabla IV. Tabla comparativa de dominios entre las secuencias problema de *Mycobacterium tuberculosis* CDC1551 y los homólogos detectados por el BLAST en *Archaea*. 28-30

Tabla V. Tabla comparativa de dominios entre las secuencias problema de *Salmonella enterica* subsp. *enterica* serovar Choleraesuis y los homólogos detectados por el BLAST en *Archaea*. 31-32

Tabla VI. Análisis de regresión lineal simple. r es el coeficiente de correlación y p el valor de P 40

Tabla VII. Análisis de regresión multifactorial. r^2 variabilidad observada y p valor de P. 40

Tabla VIII. Genes que codifican las enzimas involucradas en la biosíntesis de ácidos grasos tipo II (FAS II) en *Escherichia coli* HS (ecx), *Mycobacterium tuberculosis* CDC 1551 (mtc), *Salmonella enterica* subsp. *enterica* serovar choleraesuis (sec), según la Enciclopedia de Genes y Genomas (KEGG) y dominios funcionales que integran cada proteína, según Pfam. 59

Tabla IX. Clasificación funcional (Pfam) y estructural (SCOP) de los dominios que integran las enzimas que catalizan la FAS II en *Escherichia coli* HS, *Mycobacterium tuberculosis* CDC 1551 y *Salmonella enterica* subsp. *enterica* serovar choleraesuis. 60-61

Tabla X. Clasificación funcional (Pfam) y estructural (SCOP) de los dominios que integran las enzimas que catalizan la FAS II en *Archaea*. 62

Tabla XI. Biología de las arquea cuyo proteoma fue utilizado en este trabajo.

64-65

I Introducción

1.1 Biología de las arquea

De acuerdo con el análisis filogenético de secuencias de RNA ribosomal (rRNA) 16S, las arquea son un grupo monofilético de procariontes diferente a las bacterias (Woese y Fox, 1977). Además de esta característica molecular, existe otra singularidad estructural que sustenta la clasificación actual de estos organismos, las arquea se distinguen por poseer cadenas de isoprenoides de longitud variable, unidas exclusivamente a través de enlaces de éter a la porción glicerol de los lípidos de la membrana citoplasmática (Tornabene y Langworthy, 1979).

Sin embargo, diversas comparaciones genómicas revelan una estrecha relación entre los rasgos informativos de las arqueas y del nucleocitoplasma eucarionte, pero mayor semejanza con los aspectos metabólicos de las bacterias (Olsen y Woese, 1997; Forterre *et al.* 2002)

Inicialmente, las arquea fueron divididas en dos grupos: Euryarchaeota y Crenarchaeota (Woese *et al.* 1990). El manual de sistemática bacteriológica de Bergey (Boone y Castenholz, 2001) reconoce ocho órdenes para el phylum Euryarchaeota: Termococci, Metanopyri, Metanococci, Metanobacteria, Termoplasmata, Archaeoglobi, Halobacteria y Metanomicrobia; y cuatro órdenes para el phylum Crenarchaeota: Termoproteales, Caldisphaerales, Desulfurococcales y Sulfolobales. Forterre *et al.* (2002) propone una clasificación distinta, reconoce nueve órdenes para el phylum Euryarchaeota: Metanobacteriales, Metanomicrobiales, Metanococcales, Metanosarcinales, Metanopyrales, Halobacteriales,

Termoplasmatales, Termococcales y Archaeoglobales; mientras que reconoce tres órdenes para el phylum Crenarchaeota: Termoproteales, Desulfurococcales y Sulfolabales. Recientemente se han propuesto tres nuevos phyla: el phylum Korarchaeota (Barns *et al.* 1996), con un solo integrante, *Candidatus Korarchaeum Cryptofilum* (Elkins *et al.* 2008); el phylum Nanoarchaeota, también con un único integrante, *Nanoarchaeum equitans* (Huber *et al.* 2002); y el phylum Thaumarchaeota, con *Cenarchaeum symbiosum* y sus parientes (Brochier-Armanet *et al.* 2008).

Aunque en la actualidad se tienen pruebas de que las arquea ocupan distintos nichos ecológicos (Keough *et al.* 2003; Lepp *et al.* 2004; Galand *et al.* 2005; Kim *et al.* 2005; Kleikemper *et al.* 2005; Knittel *et al.* 2005; Mills *et al.* 2005), todavía se les asocia con frecuencia a los ambientes más extremos de nuestro planeta. Las arquea se han establecido con éxito en aquellos sitios que casi ningún otro organismo ha podido colonizar, se les encuentra en el subsuelo del fondo marino (Jørgensen y D'Hondt, 2006), en chimeneas hidrotermales (Jones *et al.* 1983; Takai y Horikoshi 1999, Stefan *et al.* 2000; Huber *et al.* 2002), en manantiales sulfurosos (Whildaber *et al.* 1987; Skirnisdottir *et al.* 2000), lagunas de agua hirviendo (Barns *et al.* 1994; Barns *et al.* 1996, Hobel *et al.* 2005) o estanques y mares hipersalinos (Oren *et al.* 1990; Oren *et al.* 1995, Wainø *et al.* 2000), hasta en el hielo (Battin *et al.* 2001; Ravensschlag *et al.* 2001; Cavicchioli, 2006). Las arquea están perfectamente adaptadas para soportar las condiciones ambientales más adversas: anoxia (Collins *et al.* 2005), estrés energético crónico (Valentine, 2007), numerosas atmósferas de presión, temperaturas muy bajas (Preston

et al. 1996, Murray *et al.* 1998) o exageradamente elevadas, pH igual a 0 (Fütterer *et al.* 2004) o pH superior a 10 e hipersalinidad. Sin embargo, es común hallarlas lidiando con varias de ellas al mismo tiempo (Whildaber *et al.* 1987; Elferink *et al.* 1992; Kaneshiro y Clark, 1995; DeLong, 2000; Fan *et al.* 2004; Fütterer *et al.* 2004; Mesbah y Wiegel, 2008.).

1.2 Características de la membrana citoplasmática de las arquea

La membrana celular es una bicapa de fosfolípidos donde se encuentra embebida una gran cantidad de proteínas (Singer y Nicolson, 1972). Los fosfolípidos son moléculas anfipáticas (moléculas con una región hidrofóbica y una hidrofílica) compuestas por una porción glicerol fosfato unida a dos cadenas laterales de hidrocarburos a través de enlaces de éster (Nelson y Cox, 2005). Sin embargo, existen notables diferencias entre los lípidos de la membrana citoplasmática de las arquea y los organismos pertenecientes a los demás dominios:

(I) Las cadenas de hidrocarburos se unen a las posiciones *sn-2* y *sn-3* de la porción glicerol de los lípidos arqueanos, mientras que en bacterias y eucariontes las cadenas de hidrocarburos se unen a las posiciones *sn-1* y *sn-2*. Esto significa que el glicerol fosfato de los lípidos en las arquea es el *sn-glicerol-1-fosfato*, enantiómero del *sn-glicerol-3-fosfato* de los fosfolípidos bacterianos y eucarióticos (Kates, 1993; Peretó *et al.* 2004; Koga y Morii, 2007).

(II) En los lípidos de las arquea, las cadenas de hidrocarburos se unen a la porción glicerol a través de enlaces de éter exclusivamente. En contraste, los lípidos de las bacterias poseen enlaces de éster entre los

ácidos grasos y la porción glicerol (Kates, 1993; Peretó *et al.* 2004; Koga y Morii, 2007).

(III) Las cadenas de hidrocarburos de los lípidos de las arquea, son isoprenoides altamente metilados, mientras que en su contraparte bacteriana, la gran mayoría son ácidos grasos saturados de cadena larga (Kates, 1993; Peretó *et al.* 2004; Koga y Morii, 2007).

(IV) En un gran número de especies de arquea se encuentran presentes lípidos bipolares con un núcleo tetraetérico. Estos lípidos se expanden para formar una monocapa lipídica (Kates, 1993; Peretó *et al.* 2004; Koga y Morii, 2007).

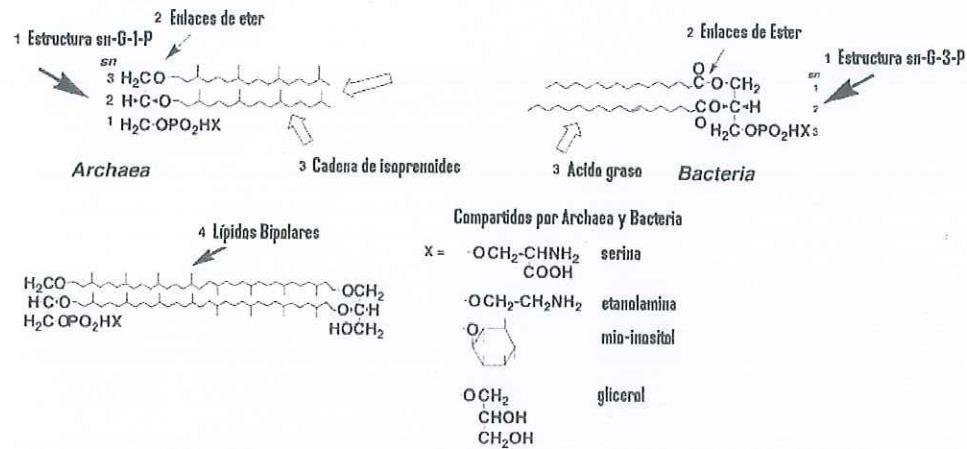


Figura 1. Características de los lípidos de las arquea. Los fosfolípidos de las arquea se caracterizan por (i) su estructura G-1-P, (ii) tener enlaces de éter, (iii) tener cadenas de isoprenoides y (iv) lípidos tetraetéricos bipolares. La mayoría de las cabezas polares son compartidas por arqueas y bacterias. Imagen modificada de Koga y Morii, 2007.

De estas características, la estereoestructura del glicerofosfato es el rasgo filogenético y evolutivo de mayor significado de los lípidos de las arquea. Sin embargo, la diferencia enantiomérica parece ser una propiedad fisicoquímicamente insignificante de la membrana de las arquea, porque los

enantiómeros son fisicoquímicamente equivalentes, exceptuando su quiralidad. Los enlaces de éter, las cadenas de isoprenoides y los lípidos bipolares tetraetéricos son más relevantes para las propiedades fisicoquímicas de los lípidos de la membrana de las arquea (Koga y Morii, 2007).

La estructura de las membranas biológicas se mantiene principalmente unida a través de enlaces no covalentes como los de Van der Waals e interacciones eléctricas (Morrison y Boyd, 1998).

La membrana celular funciona como una barrera entre el citoplasma y el medio, controlando el movimiento de solutos (iones y nutrientes) hacia el interior y el exterior de la célula (Alberts *et al.* 2002). La membrana citoplasmática es crucial para la generación de energía metabólica a través de la transducción de energía iónica. En este proceso, la energía de un gradiente electroquímico generado por el paso de un ión a través de la membrana es transformada en otras formas de energía o viceversa.

Los liposomas compuestos por lípidos tetraetéricos arqueanos son más estables que los liposomas de lípidos bacterianos y su permeabilidad a los protones es menor a una temperatura particular (Elfrink *et al.* 1994; van de Vossenberg *et al.* 1995; van de Vossenberg *et al.* 1998). El núcleo de los isoprenoides es el responsable de esta baja permeabilidad (Yamauchi *et al.* 1993). Los enlaces de éter son más resistentes a la oxidación y a las altas temperaturas que los enlaces de éster. Consecuentemente, los liposomas preparados con lípidos de Arquea son más termoestables (Chang, 1994). Además, a diferencia de los enlaces de éster, los lípidos tetraetéricos no son

susceptibles a la degradación de pH alcalino y degradación enzimática por las fosfolipasas (Choquet *et al.* 1994).

Las bacterias y las arquea pueden responder a los cambios de la temperatura en el ambiente modificando la composición lipídica de su membrana celular (Gaughran, 1947). Estas adecuaciones son necesarias para mantener a la membrana en un estado líquido cristalino (Russell y Fukunaga, 1990) y para limitar la tasa de difusión de protones. A altas temperaturas, las bacterias logran esto incrementando la longitud de la cadena de los lípidos, las ramificaciones y/o el grado de saturación las cadenas acilo (Reizer *et al.* 1985; Svobodová y Svoboda 1988). En algunas arquea el grado de ciclación del isoprenoide C40 en los lípidos tetraetéricos se incrementa a altas temperaturas (De Rosa y Gambacorta, 1988). Al aumentar el grado de ciclación de las cadenas isoprenoides C40 los lípidos pueden compactarse más, lo que ocasiona un movimiento más restringido de los lípidos y provoca que la membrana se torne muy fluida. En otras arquea, además de la ciclación, el incremento de temperatura induce el cambio de lípidos dietéricos por lípidos tetraetéricos (Sprott *et al.* 1991).

1.3 Biosíntesis de Ácidos Grasos en Bacteria (FAS II)

Dependiendo del tipo de enzimas y proteínas que participen en la biosíntesis de ácidos grasos (FAS), existen dos variantes moleculares de este proceso. En la variante del primer tipo (FAS I), los sitios activos que catalizan el ciclo sintético se localizan en los distintos dominios de grandes proteínas multifuncionales. En contraste, en la otra variante (FAS II) cada acción enzimática es catalizada por una proteína discreta. A pesar de esta

importante diferencia, los mecanismos químicos implicados en la biosíntesis de ácidos grasos de ambas vías son virtualmente idénticos (Campbell y Cronan, 2001).

Las enzimas y proteínas involucradas en la biosíntesis de ácidos grasos tipo II, se dividen en dos categorías distintas: aquellas con secuencias proteínicas altamente conservadas y una categoría menos conservada. Las proteínas muy conservadas son frecuentemente codificadas dentro de agrupaciones de genes con un orden bastante conservado. Los genes *fab* están involucrados en la conversión de acetil-CoA y malonil-CoA en ácidos grasos de cadenas largas. Los genes requeridos para la síntesis de malonil-CoA codifican la enzima de cuatro subunidades, acetil-CoA carboxilasa, y se conocen como los genes *acc* (Campbell y Cronan, 2001).

La acetil-CoA carboxilasa cataliza el primer paso de la biosíntesis de ácidos grasos, la conversión de acetil-CoA a malonil-CoA. La reacción se lleva a cabo en dos pasos: la carboxilación ATP dependiente de la biotina con bicarbonato para formar carboxibiotina; y la transferencia del grupo carboxilo de la carboxibiotina a la acetil-CoA para formar malonil-CoA (Waite y Wakil, 1962).

Cada reacción es catalizada por un complejo proteínico diferente. La biotina es acoplada a un residuo de lisina de la proteína acarreadora de biotina-carboxilo (BCCP, codificada por *accB*), mediante un enlace covalente amídico generado por una enzima específica, la apo-biotina ligasa (codificada por *birA*), y es esencial para su actividad. La carboxilación de la

biotina es catalizada por la biotina carboxilasa (codificada por *accC*). Los genes *accB* y *accC* forman un operón. La carboxilación de la biotina implica la reacción del ATP y CO₂ para formar un carboxifosfato de vida corta, que luego interactúa con la biotina en la BCCP para transferir el CO₂ (Li y Cronan, 1992)

La enzima carboxiltransferasa que transfiere el grupo carboxilo de la región biotinilada de la BCCP a la acetil-CoA es un heterotetrámero compuesto por dos copias de dos subunidades diferentes, las subunidades α y β (codificadas por *accA* y *accD*, respectivamente). El análisis de secuencias sugiere que el sitio de unión a la acetil-CoA yace dentro de la subunidad *AccA*. *accA* y *accD* se encuentran alejados el uno del otro y de *accCB* (Li y Cronan, 1992)

Para que el grupo malonato pueda ser usado para la biosíntesis de ácidos grasos, primero debe ser transferido de la malonil-CoA a la malonil-ACP por la malonil-CoA : ACP transacilasa, codificada por el gen *fabD*. Durante el curso de la reacción una se forma la malonil-serina y el subsecuente ataque nucleofílico por el sulfhidrilo de la ACP sobre este éster genera malonil-ACP (Williamson y Wakil, 1968; Ruch y Vagelos, 1973).

Los últimos dos carbonos de la cadena del ácido graso son los primeros en introducirse a la cadena naciente y se puede considerar a la acetil-CoA la molécula templete de la síntesis de ácidos grasos. La condensación inicial, catalizada por la β -cetoacil-ACP sintetasa III (FabH), utiliza acetil-CoA y malonil-ACP para producir la acetoacetil-ACP de cuatro carbonos con la pérdida concomitante de CO₂ (Tsay *et al.* 1992)

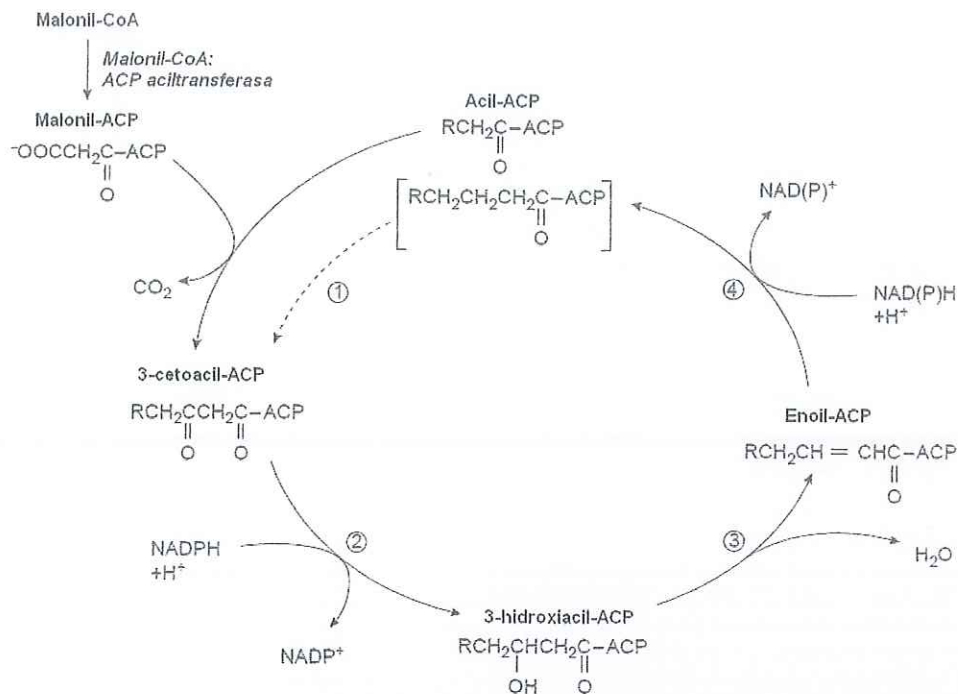


Figura 2. Ciclo iterativo de reacciones para la adición de dos carbonos. (1) Condensación, (2) reducción, (3) deshidratación y (4) reducción. Modificado de Gurr *et al.* 2002.

Cuatro enzimas participan en cada uno de los ciclos iterativos de la elongación (Figura 2). Primero, la β -cetoacil-ACP sintetasa I o II (*fabB* o *fabF*) añade dos carbonos de la malonil-ACP a la creciente acilo-ACP (Garwin *et al.* 1980). El cetoéster resultante es reducido por una β -cetoacil-ACP reductasa (*fabG*) NADHP-dependiente, y una molécula de agua es removida por una β -hidroxiacilo-ACP deshidratasa que pueden ser *fabA* o *fabZ* (Heath y Rock, 1996). El último paso es catalizado por la enoil-ACP reductasa (*fabI*) para formar acilo-ACP saturada, la cual sirve como sustrato para otra reacción de condensación (Bergler, *et al.* 1996).

La sobreproducción de FabB tiene dos efectos: un elevado incremento de ácido *cis*-vaccénico en fosfolípidos y resistencia a la tiolactomicina y cerulenina (Heath *et al.* 2002). La β -cetoacil-ACP sintetasa II

(FabF) es muy similar a FabB. No es esencial para el crecimiento de *Escherichia coli*, pero es indispensable para la regulación de la composición de ácidos grasos como respuesta a las fluctuaciones de la temperatura (Garwin *et al.* 1980). La β -cetoacil-ACP sintetasa III se caracteriza por su preferencia por el templete ligado a CoA, en lugar de acilo-ACP (Tsay *et al.* 1992).

El gen de la β -cetoacil-ACP reductasa (*FabG*) se localiza dentro del grupo de genes *fab*, entre los genes *fabD* y *acpP* y es cotranscrito con *acpP*. La proteína funciona con cadenas de cualquier longitud y exhibe una unión cooperativa a NADHP (Campbell y Cronan, 2001).

E. coli posee dos β -hidroxiacilo-ACP deshidratasas. Una es codificada por *fabZ*, y funciona con cadenas de cualquier longitud, saturadas o insaturadas. Esta enzima es diferente a la β -hidroxidecanoil-ACP deshidratasa/isomerasa (codificada por *fabA*). La enzima FabA deshidrata ácidos grasos saturados, pero no insaturados y cataliza la isomerización clave, en el punto donde la biosíntesis de ácidos grasos insaturados diverge de la biosíntesis de ácidos grasos saturados. Ambas enzimas comparten una débil homología. La histidina del sitio activo se localiza en un largo túnel que actúa como regla molecular para asegurarse que sólo intermediarios de 10 carbonos sean isomerizados (Heath *et al.* 2002).

El último paso en cada ronda de elongación en *E. coli* es la reducción NADH dependiente del doble enlace trans, catalizado por la enoil-ACP reductasa I NADH dependiente, codificada por *fabI* (Heath y Rock, 1995).

La ruta descrita anteriormente basta para producir los ácidos grasos saturados encontrados en los fosfolípidos de la membrana, principalmente ácido palmítico, en *E. coli*. El ratio de ácidos grasos saturados: insaturados es un determinante clave de la fluidez de la membrana, y cambia de acuerdo a la temperatura. La deshidratasa/isomerasa codificada por *fabA* cataliza una reacción vital en el punto donde ambas vías se bifurcan: la isomerización de la *trans*-2-decanoil-ACP a *cis*- β -decanoil-ACP. El compuesto *cis*-3 no es utilizado como sustrato de por la enoil-ACP-reductasa, sino que es rápidamente en una reacción que requiere de FabB (Heath *et al.* 2002). Tanto FabB como FabF son capaces de participar en la síntesis de ácidos grasos saturados como insaturados, y funcionan de manera similar con todas las acilo-ACP de cadena largo, excepto palmitoleil-ACP.

1.4 Origen y evolución de rutas metabólicas

El metabolismo es una compleja red de procesos físico-químicos, la mayoría de ellos catalizados por enzimas, que permite la supervivencia y la perpetuación de los seres vivos (Alves *et al.* 2002). Aunque existen varias hipótesis que intentan explicar el origen y la evolución de las rutas metabólicas (Lazcano y Miller 1999; Rsiion y Thornton 2002; Schmidt *et al.* 2003), el modelo retrógrado y el modelo de "patchwork" son las más citadas para explicar el posible origen y la evolución de esta serie de reacciones en los seres vivos.

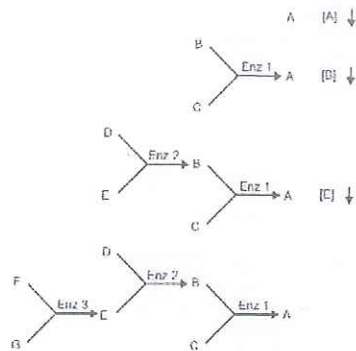


Figura 3. Modelo retrógrado de evolución de rutas metabólicas. Un organismo heterotrófico agota el abastecimiento natural del metabolito A. El reclutamiento fortuito de una enzima (Enz 1) capaz de sintetizar A de B y C confiere una ventaja al organismo. Cuando las concentraciones de B y E decaen. Son compensadas por el reclutamiento de las enzimas 'Enz 2' y 'Enz 3', respectivamente. Modificado de Rison y Thornton, 2002.

El modelo retrógrado establece que las rutas metabólicas evolucionaron "hacia atrás" a partir de un metabolito clave (Horowitz, 1945). Esta teoría se sustenta en la hipótesis heterotrófica de Oparin (1968) y entre la correspondencia uno a uno entre genes y enzimas sugerida por Beadle y Tatum (1941) y da por sentada la existencia de un ambiente químico en el cual el metabolito clave y los posibles intermediarios estaban disponibles. Un organismo heterótrofo dependiente de la molécula A, consumirá el metabolito hasta el punto en el que la disponibilidad decadente limite su crecimiento; en dicho ambiente, un organismo capaz de sintetizar la molécula A a partir de los precursores ambientales B y C, tendrá una ventaja selectiva. Cualquier mutante con una enzima que catalice esta reacción se propagará rápidamente; además en la continua ausencia del metabolito A, cualquier mutación de la enzima será letal, favoreciendo su conservación. A su vez, mientras la concentración de B o C se agota, el proceso será repetido con el reclutamiento similar de otras enzimas. Horowitz también

sugirió que la escasez simultánea de los dos intermediarios (e.g. B o C) favorecería la asociación simbiótica entre dos mutantes, uno capaz de sintetizar B y otro capaz de sintetizar C de otros precursores ambientales. Posteriormente, el descubrimiento de los operones le permitió a Horowitz (1965) replantear su idea original y proponer que el agrupamiento de genes que codifican enzimas de una ruta dada, puede ser explicado como el resultado de duplicaciones tempranas en tándem. Sin embargo, el modelo retrógrado no puede explicar la evolución de rutas metabólicas que incluyen metabolitos inestables, cuya síntesis y acumulación parece improbable en el ambiente prebiótico (Lazcano y Miller, 1999; Rison y Thornton, 2002). Además, muchos de estos intermediarios son compuestos aniónicos fosforilados, los cuales no podrían permear fácilmente las membranas primitivas, carentes de transportadores especializados (Jensen, 1976). Esta teoría sólo puede explicar la evolución de rutas metabólicas en un medio rico en intermediarios metabólicos; la destrucción final del ambiente orgánico evitaría la evolución de rutas metabólicas de manera retrógrada (Rison y Thornton, 2002).

El modelo tipo "patchwork" establece que las rutas metabólicas se ensamblaron a través del reclutamiento de enzimas primitivas que reaccionaban con una gran cantidad de sustratos químicamente relacionados, que después evolucionaron hacia enzimas más eficientes y específicas que catalizan un solo paso de una sola ruta, mediante la duplicación génica y la diversificación (Jensen, 1976). La duplicación génica pudo ser promovida por la falta de mecanismos de reparación del DNA y los

altos niveles de radiación ionizante en el ambiente precámbrico (Koch, 1972). Este modelo explica la diversificación de enzimas ancestrales de poca especificidad hacia familias de catalizadores especializados, se debe mutaciones puntuales, pequeñas supresiones y otros cambios minúsculos. Como los procariontes son organismos haploides, la mayoría de estos pequeños cambios genéticos se expresarían en cuanto hubiesen surgido, resultando una rápida evolución metabólica (Peretó *et al.* 1997). Sin embargo, este modelo sólo puede funcionar luego de la aparición de la biosíntesis de proteínas y el desarrollo de enzimas; después de la aparición del mundo DNA-proteínas. Este mecanismo no aplica a las rutas metabólicas más tempranas porque había pocas, si acaso algunas enzimas disponibles. Pero, puede aplicar para rutas posteriores (Lazcano y Miller, 1999).

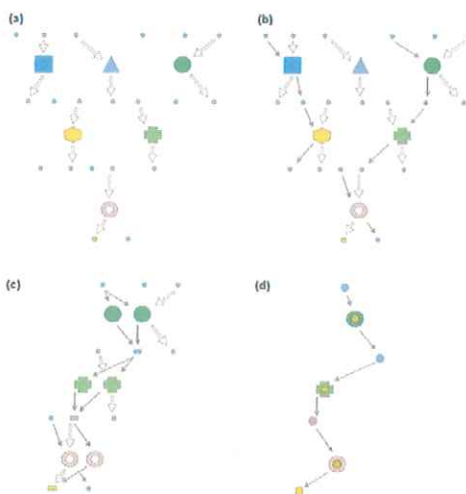


Figura 4. Modelo "patchwork" de evolución de rutas metabólicas. Este modelo pudo favorecer un mecanismo de sustrato específico (a), sin embargo las enzimas presentaban baja especificidad (b). La duplicación de cualquier gen sería muy ventajosa pues más metabolito sería sintetizado (c). Esta duplicación, seguida por la especialización de la enzima daría lugar a las rutas que existen actualmente (d). Modificado de Rison y Thornton, 2002.

A pesar de sus diferencias, estas hipótesis se basan en la idea de que la duplicación génica, seguida por la divergencia, puede originar nuevas rutas metabólicas (Ohno, 1970). La duplicación génica, es el fenómeno en el cual ocurre la duplicación de una o varias secuencias de DNA, que codifican un gen en particular y su principal función es proporcionar material genético nuevo para la adaptación de los seres vivos (Hughes, 2005) y resulta del entrecruzamiento desigual, de la retroposición, o la duplicación cromosómica (Page y Holmes, 1998). El entrecruzamiento desigual usualmente genera duplicaciones génicas en tándem. Dependiendo de la posición del entrecruzamiento, la región duplicada puede contener parte de un gen, un gen entero, o varios genes. Si en los dos casos anteriores la región duplicada presentaba intrones, estos también estarán presentes en los genes duplicados (Zhang, 2003). La retroposición ocurre cuando un RNA mensajero (mRNA) es retrotranscrito a DNA complementario (cDNA) y luego insertado en el genoma. En este proceso hay pérdida de intrones y secuencias reguladoras, presencia de secuencias poli A y presencia de repeticiones cortas flanqueantes (Marques *et al.* 2005). Una importante diferencia con el entrecruzamiento desigual es que los genes generados por retroposición, es que casi nunca se encuentran ligados al gen original, porque la inserción del cDNA al genoma es más o menos azarosa. También es imposible tener bloques de genes duplicados, a menos de que los genes involucrados formen parte de un Operón (Zhang, 2003). La duplicación

cromosómica o genómica probablemente ocurre por una falta de disyunción entre los cromosomas después de la replicación del DNA (Zhang, 2003).

El destino de una duplicación será determinado por la función de los genes duplicados. (i) La redundancia funcional generada por la duplicación de un gen puede resultar no ventajosa. La pseudogenización de un gen ocurre cuando una de las copias pierde su funcionalidad por la acumulación gradual de mutaciones (Haldane, 1933). (ii) Un gen puede ser conservado si las cantidades extras de proteína o RNA producidas son beneficiosas. Los genes parálogos pueden mantenerse funcionales si se encuentran sujetos a evolución concertada, esto quiere decir, que existe una transferencia de información genética no recíproca entre secuencias similares, en el cual no hay intercambio del DNA flanqueante (Liao, 1999); o por selección purificadora, la cual actúa en contra de mutaciones supresoras (Kimura y Ohta, 1974). (iii) La subfuncionalización tiene lugar cuando las copias nuevas adquiere una parte de las funciones del gen ancestral (Jensen, 1976). (iv) La adquisición de una nueva función debida a una gran cantidad de sustituciones de aminoácidos (Ohno, 1970).

1.5 Las moléculas como fuentes de información

Los genes y las proteínas derivadas de los genes, son las moléculas que más historia de los seres vivos conservan (Zuckerlandl y Pauling, 1965). Numerosos métodos computacionales han sido desarrollados para poder analizar la información que contienen.

Existen diversas bases de datos que almacenan las secuencias de nucleótidos de un gen o aminoácidos de un polipéptido, o que clasifican y

ordenan a las proteínas dependiendo de su función o conformación estructural. El GenBank (Benson *et al.* 2007) es probablemente el más conocido de estos bancos de datos, y contiene una enorme cantidad de secuencias nucleotídicas; la enciclopedia de genes y genomas de Kyoto (KEGG) es otra base de datos que ofrece al público recursos similares (Kanehisa y Goto, 2000). Pfam se utiliza para identificar y clasificar los dominios funcionales que integran las proteínas y SCOP las clasifica dependiendo de su acomodo espacial y es muy útil para inferir si comparten un mismo ancestro basándose en esta propiedad (Murzin *et al.* 1995; Finn *et al.* 2008).

En la actualidad existen millones de genes cuya función se desconoce. Una manera ágil de averiguar el papel que desempeñan estas proteínas es comparando sus secuencias y encontrar homologías con las moléculas ya descritas. El método bioinformático más popular para encontrar homologías entre las secuencias es realizando una comparación con el programa BLAST, (Basic Local Alignment Search Tool), el cual detecta similitudes débiles pero biológicamente significativas y es un orden de magnitud más rápido que los algoritmos heurísticos existentes (Altschul *et al.* 1990).

El análisis filogenético de secuencias de DNA o proteínas se ha vuelto una importante herramienta para estudiar la historia evolutiva de las rutas metabólicas. Varios algoritmos matemáticos han sido propuestos para obtener la filogenia más aproximada a la verdadera: de distancia (*neighbor-joining*), máxima parsimonia o de máxima similitud. También se han

implementado diversos análisis estadísticos como el *bootstrap*, para fortalecer los resultados obtenidos de los análisis filogenéticos. Hay muchos programas gratuitos y disponibles en Internet, Mega 4.0 (Tamura *et al.* 2007) y Mr. Bayes (Huelsenbeck y Ronquist, 2001) son un ejemplo.

II Antecedentes: Ésteres de glicerol en Archaea

Los precursores esenciales para la síntesis de isoprenoides que integran la membrana citoplasmática de las arquea son sintetizados por la vía del mevalonato (Smith y Muehleberg, 2000). La importancia adaptativa y taxonómica de la membrana celular de las arquea ha promovido su estudio y existen varias publicaciones al respecto (Tornabene y Langworthy, 1979; Koga y Morii, 2007). Sin embargo, muy poco se sabe sobre los ésteres de glicerol que poseen estos organismos y la función que desempeñan, incluso algunos autores dan por hecho que no tienen (Campbell y Cronan, 2001). En 2004, Peretó *et al.* compararon las secuencias de las enzimas que catalizan la biosíntesis de algunas bacteria con el genoma de algunas arquea y encontró varias homologías. Sin embargo, no se ha reportado un análisis más profundo de esta ruta, que revele las diferencias y semejanzas entre las enzimas que sintetizan estas moléculas y las especies de arquea que las poseen. Por lo tanto, el problema central a resolver consiste en establecer dichas semejanzas y diferencias.

III Objetivos

- 1) Obtener los proteomas de Archaea disponibles en la base de datos KEGG.
- 2) Determinar el mecanismo de ensamblaje de la FAS II en las arquea.
- 3) Establecer las posibles rutas evolutivas de las enzimas que catalizan la biosíntesis de ácidos grasos tipo II en el dominio Archaea.

IV Hipótesis

El análisis bioinformático de las enzimas que sintetizan los ácidos grasos en el grupo Archaea revelará las diferencias y semejanzas que existen entre las especies estudiadas, revelando las regiones conservadas y las más variables.

V Material y Métodos

Las secuencias de aminoácidos de las enzimas que catalizan la biosíntesis de ácidos grasos tipo II en *Escherichia coli*, *Mycobacterium tuberculosis* y *Salmonella enterica* se obtuvieron de la Enciclopedia de Genes y Genomas de Kyoto (<http://www.genome.jp/kegg/>) y se compararon con una recopilación de 49 proteomas de arqueas, para detectar posibles homologías, utilizando BLAST con un valor de $e \leq 0.0001$ (probabilidad de que dos secuencias se parezcan por mero azar;)

Tabla I. Enzimas que catalizan la biosíntesis de ácidos grasos tipo II (FAS II) en *Escherichia coli* HS (ecx), *Mycobacterium tuberculosis* CDC 1551 (mtc) y *Salmonella enterica* subsp. *enterica* serovar choleraesius (sec). El número EC es un sistema de clasificación numérica de las enzimas basado en la reacción que estas catalizan.

Enzima	Número EC	Reacción
Acetil CoA-carboxilasa	6.4.1.2	ATP + acetyl-CoA + HCO(3)(-) = ADP + phosphate + malonyl-CoA
Biotina carboxilasa	6.3.4.14	ATP + biotin-carboxyl-carrier protein + CO(2) = ADP + phosphate + carboxybiotin-carboxyl-carrier protein
3-cetoacil-ACP sintetasa III	2.3.1.180	Acetyl-CoA + malonyl-[acyl-carrier-protein] = acetoacetyl-(acyl-carrier-protein) + CoA + CO(2)
3-cetoacil-ACP sintetasa I	2.3.1.41	Acyl-[acyl-carrier-protein] + malonyl-[acyl-carrier-protein] = 3-oxoacetyl-[acyl-carrier-protein] + CO(2) + [acyl-carrier-protein]
Lipoacil-CoA sintetasa	2.3.1.86	Acetyl-CoA + n malonyl-CoA + 2n NADH + 2n NADPH = long-chain-acyl-CoA + n CoA + n CO2 + 2n NAD(+) + 2n NADP(+)
3-cetoacil-ACP reductasa	1.1.1.100	(3R)-3-hydroxyacyl-[acyl-carrier protein] + NADP(+) = 3-oxoacetyl-[acyl-carrier protein] + NADPH
Enoil-ACP reductasa (NADPH)	1.3.1.39	Acyl-[acyl-carrier protein] + NADP(+) = trans-2,3-dehydroacyl-[acyl-carrier protein] + NADPH
3-cetoacil-ACP sintetasa II	2.3.1.179	(Z)-hexadec-11-enoyl-[acyl-carrier protein] + malonyl-[acyl-carrier protein] = (Z)-3-oxooctadec-13-enoyl-[acyl-carrier protein] + CO(2) + [acyl-carrier protein]
Malonil CoA:ACP transacylase	2.3.1.39	Malonyl-CoA + [acyl-carrier protein] = CoA + malonyl-[acyl-carrier protein]

5.1 Análisis funcional y estructural de la FAS II

Las secuencias generadas por el BLAST se compararon con la base de datos Pfam. Una vez identificadas las regiones funcionales de estas moléculas, las secuencias correspondientes se compararon con la base de datos SCOP.

5.2 Análisis filogenético y estadístico de la FAS II

Las secuencias de aminoácidos se alinearon con el programa Clustal X (Thompson *et al.* 1997). Cuando fue necesario, las secuencias se editaron con el programa BioEdit (Hall, 1999). Los análisis filogenéticos correspondientes se llevaron a cabo con el programa MEGA versión 4.0. Los árboles filogenéticos de las distintas enzimas involucradas, se generaron empleando el método de distancia Neighbor Joining y se evaluó la confiabilidad de su topología con una prueba de bootstrap, utilizando 100 réplicas.

Con el programa StatGraphics Centurion XV se realizó un análisis de variancia para saber si existen diferencias significativas entre el número de duplicaciones de cada enzima. Para determinar si el número de duplicaciones tiene alguna relación con el pH, la temperatura o la salinidad se realizó una regresión lineal simple. Un análisis de regresión multifactorial se utilizó para observar como afectan estas variables en conjunto el número de duplicaciones presente en cada enzima.

VI Resultados

El BLAST detectó 763 homologías en total (Tabla II): 95 para la enzima Acetil CoA-Carboxilasa (6.4.1.2), 86 para la Biotina Carboxilasa (6.3.4.14), 42 para la 3-cetoacil-ACP sintetasa III (2.3.1.180), 41 para la 3-cetoacil-ACP sintetasa I (2.3.1.41), 224 para la lipo-Acil-CoA sintetasa (2.3.1.86), 178 para la 3-cetoacil-ACP reductasa (1.1.1.100) y 97 Pero, no encontró ninguna secuencia homóloga para las enzimas Malonil CoA:ACP transacilasa (2.3.1.39) y 3-cetoacil-ACP sintetasa II (2.3.1.179), en ninguna Arquea.

La mayoría de las secuencias de la enzima 1.3.1.39 generadas por el BLAST son las mismas que el análisis generó para la enzima 1.1.1.100. El BLAST de las enzimas 2.3.1.180 y 2.3.1.41 también generó numerosas redundancias, por ejemplo: el BLAST de los genes MT0557 y SC1141 de la enzima 2.3.1.180 y A1213 de la enzima 2.3.1.41 generó la misma secuencia en *Aeropyrum pernix* (ape:APE_1873.1).

Las enzimas Fatty-Acyl-CoA sintetasa (2.3.1.86) y 3-cetoacil-ACP reductasa (1.1.1.100) son las enzimas que presentaron el mayor número de homologías: 224 y 178, respectivamente.

Haloarcula marismortui fue la especie que presentó más homologías, 54; mientras que *Nanoarchaeum equitans* fue la especie que presentó menos, 1.

5.1 Análisis Funcional y Estructural

El análisis funcional de las enzimas 2.3.1.41 y 2.3.1.180 demostró que las secuencias homólogas obtenidas por el BLAST pertenecen a la enzima

Tabla II. Resultados del BLAST. cetil CoA-Carboxilasa (6.4.1.2) y biotina carboxilasa (6.3.4.14)), Malonil CoA:ACP transacilasa (2.3.1.39), 3-cetoacil-ACP sintetasa III (2.3.1.180), 3-cetoacil-ACP sintetasa II (2.3.1.179), 3-cetoacil-ACP sintetasa I (2.3.1.41), lipoacil-CoA sintetasa (2.3.1.86), 3-cetoacil-ACP reductasa (1.1.1.100) y enoil-ACP reductasa (NADPH) (1.3.1.39).

Especie	6.4.1.2	6.3.4.14	2.3.1.179	2.3.1.180	2.3.1.86	2.3.1.41	2.3.1.39	1.1.1.100	1.3.1.39	Total
<i>Methanococcus jannaschii</i>	mja	4	4	0	1	0	1	0	0	10
<i>Methanococcus maripaludis</i> S2	mmp	3	3	0	2	5	1	0	0	14
<i>Methanococcus maripaludis</i> C5	mmq	3	3	0	1	4	1	0	0	12
<i>Methanosarcina acetivorans</i>	mac	2	3	0	1	5	1	0	7	22
<i>Methanosarcina barkeri</i>	mba	3	3	0	1	3	1	0	2	14
<i>Methanococcoides burtonii</i>	mbu	3	4	0	1	4	1	0	0	13
<i>Methanosaeta thermophila</i>	mtp	3	3	0	1	7	1	0	0	15
<i>Methanocorpusculum labreanum</i>	mia	3	3	0	1	8	1	0	1	18
<i>Methanoculleus marisnigri</i>	mem	3	3	0	1	8	1	0	5	24
<i>Methanobacterium thermoautotrophicum</i>	mth	3	3	0	2	7	1	0	0	16
<i>Methanosphaera stadtmanae</i>	mst	3	3	0	1	2	1	0	2	13
<i>Methanopyrus kandleri</i>	mka	2	3	0	1	0	1	0	0	7
<i>Archaeoglobus fulgidus</i>	afu	5	4	0	1	24	2	0	1	38
<i>Halobacterium</i> sp. NRC-1	hal	3	2	0	1	6	1	0	10	26
<i>Haloarcula marismortui</i>	hma	6	3	0	2	16	1	0	18	54
<i>Haloquadratum walsbyi</i>	hwa	2	3	0	1	3	1	0	11	24
<i>Natronomonas pharaonis</i>	nph	5	3	0	2	12	2	0	17	50
<i>Thermoplasma acidophilum</i>	tac	0	0	0	1	7	2	0	9	24
<i>Thermoplasma volcanium</i>	tvo	0	0	0	1	6	1	0	6	18
<i>Picrophilus torridus</i>	pto	0	1	0	1	5	1	0	4	14
<i>Pyrococcus horikoshii</i>	pho	3	2	0	1	0	1	0	2	10
<i>Pyrococcus abyssi</i>	pab	2	1	0	1	0	1	0	4	11
<i>Pyrococcus furiosus</i>	pfu	3	2	0	1	0	1	0	5	14
<i>Thermococcus kodakaraensis</i>	tko	3	2	0	1	0	1	0	4	13
<i>Aeropyrum pernix</i>	ape	0	0	0	2	6	2	0	2	13
<i>Staphylothermus marinus</i>	smr	4	3	0	1	0	1	0	2	13
<i>Hyperthermus butylicus</i>	hbu	4	3	0	1	0	1	0	4	15
<i>Sulfolobus solfataricus</i>	sso	3	3	0	1	20	1	0	11	45
<i>Sulfolobus tokodaii</i>	sto	3	3	0	1	14	1	0	9	37
<i>Sulfolobus acidocaldarius</i>	sai	4	4	0	2	18	2	0	12	51
<i>Metallosphaera sedula</i>	mse	3	3	0	1	10	1	0	8	31
<i>Pyrobaculum aerophilum</i>	pai	1	1	0	1	8	1	0	5	21
<i>Pyrobaculum islandicum</i>	pis	1	1	0	1	4	1	0	3	12
<i>Pyrobaculum caldifontis</i>	pcl	1	1	0	1	6	1	0	5	19
<i>Pyrobaculum arsenaticum</i>	pas	1	1	0	1	4	1	0	6	19
<i>Thermofilum pendens</i>	tpe	3	2	0	1	2	1	0	2	12
<i>Nanoarchaeum equitans</i>	neq	0	0	0	0	0	0	1	0	1
Total		95	86	0	42	224	41	0	178	763

Tabla III. Cuadro comparativo de dominios entre las secuencias problema *Escherichia coli* HS y los homólogos detectados por el BLAST en Archaea

Especie	Enzima					
	6.4.1.2			2.3.1.41	1.1.1.100	1.3.1.9
	EcHS_A0187	EcHS_A2467	EcHS_A3448	EcHS_A1213	EcHS_A1215	EcHS_A1401
afu	PF01039	PF01039	PF02786 PF00289 PF02785	PF08541		PF00106
ape				PF08541	PF00106	PF00106
cma						
hal	PF01039	PF01039		PF01154 PF08540	PF00106	PF00106
hbu	PF01039	PF01039		PF001154 PF08541	PF00106	PF00106
hma	PF01039	PF01039	PF02786 PF00289 PF02785 PF00364	PF01796	PF00106	PF00106
hwa	PF01039	PF01039	PF02786 PF00289 PF02785 PF00364		PF00106	PF00106
iho						
mac			PF02786 PF00289 PF02785	PF001154 PF08541	PF00106	PF00106
mae						
mba			PF02786 PF00289 PF02785	PF001154 PF08541	PF00106	PF00106
mbn						
mbu			PF02786 PF00289 PF02785 PF00364	PF001154 PF08541		
mem			PF02786 PF00289 PF02785	PF001154 PF08541	PF00106	P00106
mhu						
mja			PF02786 PF00289 PF02785	PF001154 PF08541		
mka				PF001154 PF08541		
mka			PF02786 PF00289 PF02785	PF001154 PF08541	P00106	PF00106
mmp			PF02786 PF00289 PF02785	PF001154 PF08541		
mmq			PF02786 PF00289 PF02785	PF001154 PF08541		
mmx						
mmz						
mse		PF01039		PF08541	PF00106	PF00106
msi						
mst			PF02786 PF00289 PF02785	PF001154 PF08541	PF00106	PF00106
mth			PF02786 PF00289 PF02785	PF001154 PF08541		
mtp			PF02786 PF00289 PF02785	PF08541		
mvn						
neq					PF00106	
nmr						
nph	PF01039	PF01039		PF01796	PF00106	PF00106
pab	PF01039	PF01039		PF001154 PF08541	PF00106	PF00106
pai				PF001154 PF08541	PF00106	PF00106
pas				PF08541	PF00106	PF00106
pcl				PF001154 PF08541	PF00106	PF00106

Tabla III. Cuadro comparativo de dominios entre las secuencias problema *Escherichia coli* HS y los homólogos detectados por el BLAST en Archaea

Especie	Enzima					
	6.4.1.2		EcHS_A3448	2.3.1.41	1.1.1.100	1.3.1.9
	EcHS_A0187	EcHS_A2467		EcHS_A1213	EcHS_A1215	EcHS_A1401
pfu	PF01039	PF01039		PF001154 PF08541	PF00106	PF00106
pho	PF01039	PF01039		PF001154 PF08541	PF00106	PF00106
pis				PF001154 PF08541	PF00106	PF00106
pto				PF001154 PF08541	PF00106	PF00106
rci						
sai	PF01039	PF01039		PF001154 PF08541	PF00106	PF00106
smr	PF01039	PF01039		PF08541	PF00106	PF00106
sso		PF01039		PF08541	PF00106	PF00106
sto	PF01039	PF01039		PF08541	PF00106	PF00106
tac				PF08541	PF00106	PF00106
tko		PF01039		PF001154 PF08541	PF00106	PF00106
tpe	PF01039	PF01039		PF001154 PF08541	PF00106	PF00106
tvo				PF08541	PF00106	PF00106

Tabla IV. Tabla comparativa de dominios entre las secuencias problema de *Mycobacterium tuberculosis* CDC1551 y los homólogos detectados por el BLAST.

Especie	Enzima													
	6.4.1.2		1.1.1.100					1.3.1.9		6.3.4.14		2.3.1.180		
	MT0927	MT0256	MT0793	MT1393	MT1530	MT2836	MT3606	MT3664	MT1531	MT2576	MT0557			
pcl		PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00289	PF02786	PF02787	PF01154	PF08541
pfu	PF01039	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106					PF01154	PF08541
pho	PF01039	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106		PF00289	PF02786	PF02787	PF01154	PF08541
pis		PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106				PF01154	PF08541
pto		PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106					PF01154	PF08541
rci										PF00289	PF02786	PF02787		
sai	PF01039	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00289	PF02786	PF02785	PF01154	PF08541
smr	PF01039	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00289	PF02786	PF02785	PF08541	
ss0	PF01039	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00289	PF02786	PF02785	PF08541	
sto	PF01039	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00289	PF02786	PF02787		
tac		PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106				PF08541	
tko	PF01039	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106		PF00464		PF01154	PF08541
tpe	PF01039	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106				PF01154	PF08541
tvo		PF00106	PF00106	PF00106	PF00106	PF00106	PF00106	PF00106					PF08541	

Tabla V. Tabla comparativa de dominios entre las secuencias problema de *Salmonella enterica* subsp. *enterica* serovar Choleraesuis y los homólogos detectados por el BLAST en Archaea.

Especie	Enzima										
	6.4.1.2					1.1.1.100	1.3.1.9	6.3.4.14		2.3.1.180	
	SC0232	SC2368	SC3317		SC3318	SC1143	SC1694	SC3318		SC1141	
pcl					PF00289 PF02786 PF02787	PF00106	PF00106	PF00289 PF02786 PF02787	PF00289 PF02786 PF02787	PF01154 PF08541	
pfu	PF01039	PF01039		PF00364	PF00289 PF02786 PF02787	PF00106	PF00106	PF00289 PF02786 PF02787	PF00289 PF02786 PF02787	PF01154 PF08541	
pho	PF01039	PF01039	PF00682	PF02436 PF00364		PF00106	PF00106			PF01154 PF08541	
pis					PF00289 PF02786 PF02787	PF00106	PF00106	PF00289 PF02786 PF02787	PF00289 PF02786 PF02787	PF01154 PF08541	
pto						PF00106	PF00106			PF01154 PF08541	
rci											
sai	PF01039	PF01039		PF00364	PF00289 PF02786 PF02785	PF00106	PF00106	PF00289 PF02786 PF02785	PF00289 PF02786 PF02785	PF01154 PF08541	
smr	PF01039	PF01039		PF00364	PF00289 PF02786 PF02785	PF00106	PF00106	PF00289 PF02786 PF02785	PF00289 PF02786 PF02785	PF08541	
ssr	PF01039	PF01039			PF00289 PF02786 PF02785	PF00106	PF00106	PF00289 PF02786 PF02785	PF00289 PF02786 PF02785	PF08541	
sto	PF01039	PF01039			PF00289 PF02786 PF02785	PF00106	PF00106	PF00289 PF02786 PF02785	PF00289 PF02786 PF02785	PF08541	
tac					PF00289 PF02786 PF02787	PF00106	PF00106	PF00289 PF02786 PF02787	PF00289 PF02786 PF02787	PF08541	
tko		PF01039	PF00682	PF02436 PF00364		PF00106	PF00106			PF01154 PF08541	
tpe	PF01039	PF01039			PF00289 PF02786 PF02785	PF00106	PF00106	PF00289 PF02786 PF02785	PF00289 PF02786 PF02785	PF01154 PF08541	
tvo						PF00106	PF00106			PF08541	

2.3.1.180. Sin embargo, no fue suficiente para confirmar la reacción que catalizan las enzimas 1.1.1.100 y 1.3.1.39. También indica que no se encuentran presentes los dominios PF03255 y PF01155 en la enzima 6.4.1.2 y PF08545 en la enzima 2.3.1.180 (Tablas III, IV y V). Además, revela la existencia de varios dominios que no se detectaron en las secuencias de *Escherichia coli* HS, *Mycobacterium tuberculosis* CDC 1551 y *Salmonella enterica* subsp. *enterica* serovar *choleraesuis*: PF00682 y PF02436 en la enzima 6.4.1.2, PF01154 y PF08540 en la enzima 2.3.1.180.

El análisis estructural de los dominios funcionales PF01154, PF08540 y PF08541 indica que estas regiones comparten un arreglo espacial similar y sugiere que proceden de un ancestro común.

5.2 Análisis Filogenético y Estadístico

6.4.1.2 Carboxil Transferasa

La topología que presenta el cladograma obtenido difiere significativamente del árbol filogenético generado con secuencias de rRNA del dominio Archaea (Briochier-Armanet *et al.* 2008) *Thermophilus pendens* (tpe:Tpen0156) y *Staphylothermus marinus* (smr:Smar1426), organismos que pertenecen al orden de las *Crenarchaeota*, fueron emparentados con varias especies del orden *Euryarchaeota*, aunque la prueba del *bootstrap* indica que esta asociación es estadísticamente débil. Además, ninguna de las secuencias de las especies halobacteriales fue agrupada con algún otro organismo *Euryarchaeota*. Lo mismo sucedió con la única secuencia de *Hyperthermus butylicus* (hbu: Hbut1155), miembro de las *Crenarchaeota* (Figura 6).

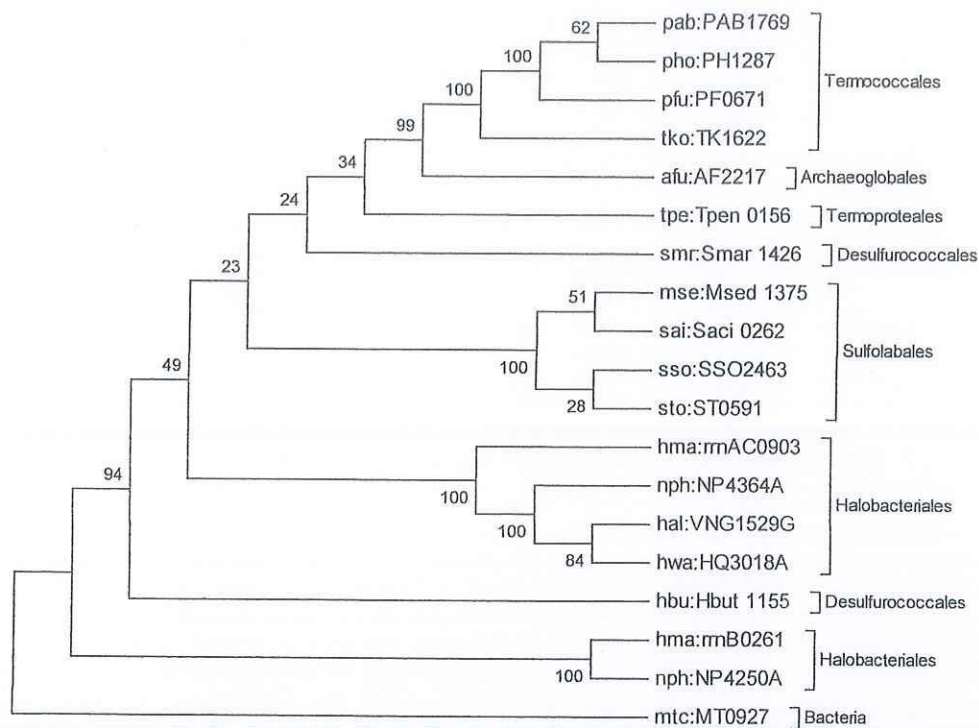


Figura 5. Árbol filogenético de la subunidad carboxil transferasa de la enzima Acetil-CoA carboxilasa (6.4.1.2). Los números en las ramas son las posibilidades bayesianas del clado.

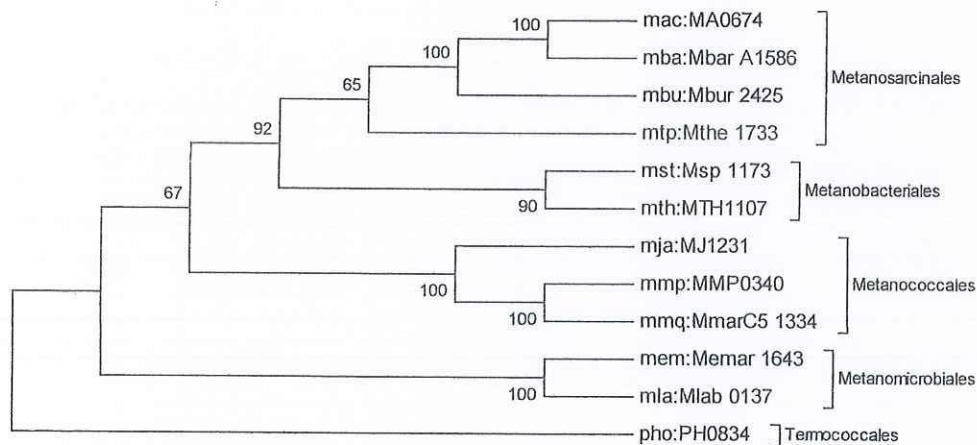


Figura 6. Árbol filogenético de la subunidad BCCP de la enzima Acetil-CoA carboxilasa (6.4.1.2). Los números en las ramas son las posibilidades bayesianas del clado.

6.4.1.2 BCCP

Aunque la topología del árbol filogenético no coincide con análisis filogenéticos de rRNA previos (Briochier-Armanet *et al.* 2008), el soporte estadístico, según la prueba de *bootstrap*, de este cladograma es robusto. El grupo de las *Metanosarcinales* (emparentado con los microbiales) se encuentra, en este caso, más próximo a los *Metanobacterales* y a los *Metanococcales* (Figura 7).

6.3.4.14 Biotina Carboxilasa

Pese a que el análisis filogenético reconoció dos grandes clados, uno que corresponde al orden *Euryarchaeota* y otro al orden *Crenarchaeota*, no todas las secuencias fueron agrupadas dentro del clado al que pertenecen. Tampoco la topología interna coincide con la topología que presenta el árbol de las especies (Briochier-Armanet *et al.* 2008). En este diagrama podemos observar que las secuencias del grupo externo se agruparon con las de las especies *Halobacterales*, lo que sugiere un posible transporte horizontal de genes por parte de estas arquea a este grupo de bacterias. Sin embargo, el *bootstrap* indica que esta relación no es estadísticamente sólida (Figura 8).

2.3.1.180 3-cetoacil-ACP sintetasa III

La mayoría de las secuencias fueron agrupadas en dos clados: *Euryarchaeota* y *Crenarchaeota*. Sin embargo, las secuencias de las enzimas de las especies *Halobacterales* no fueron incluidas dentro de las *Euryarchaeota*, orden al que pertenecen. Por el contrario, *T. pendens*, orden *Crenarchaeota*, fue agrupado dentro del orden *Euryarchaeota*, y su relación es estadísticamente fuerte, lo que sugiere la ocurrencia de transporte

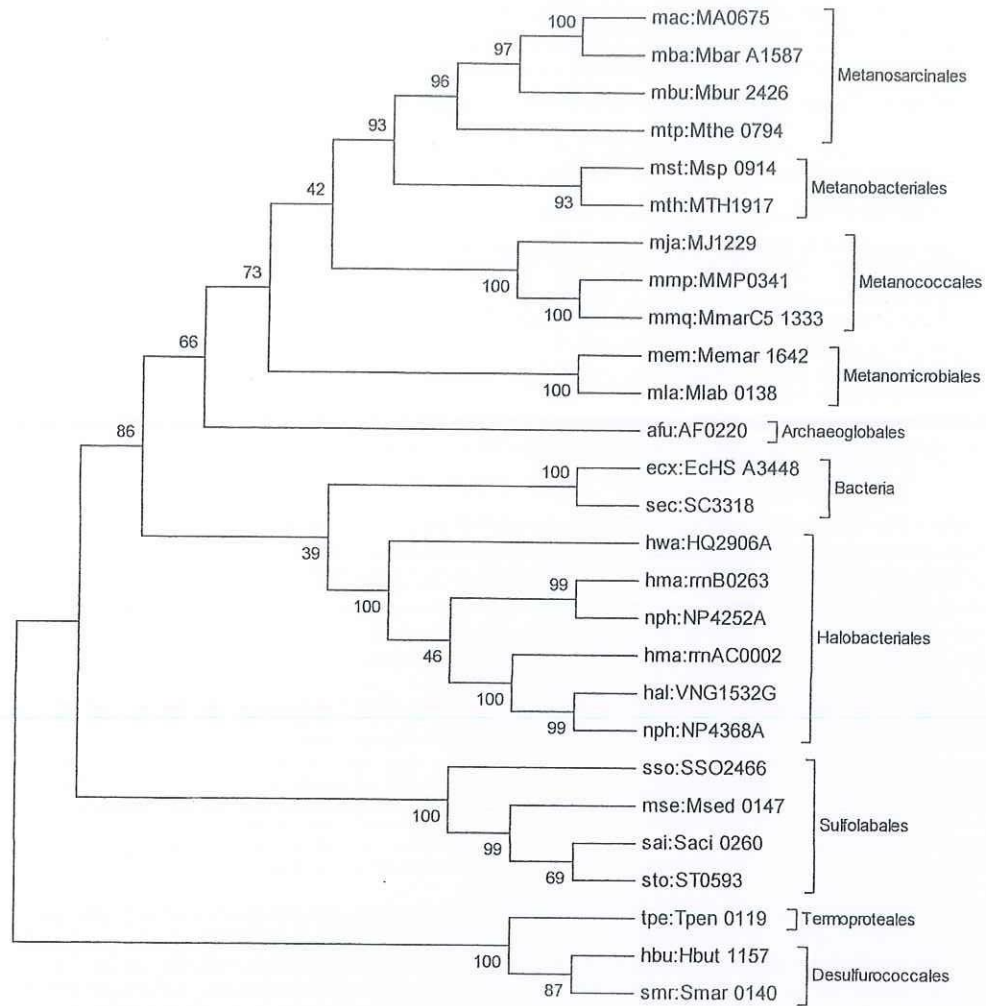


Figura 7. Árbol filogenético de la Biotina Carboxilasa (6.3.4.14). Los números en las ramas son las posibilidades bayesianas del clado.

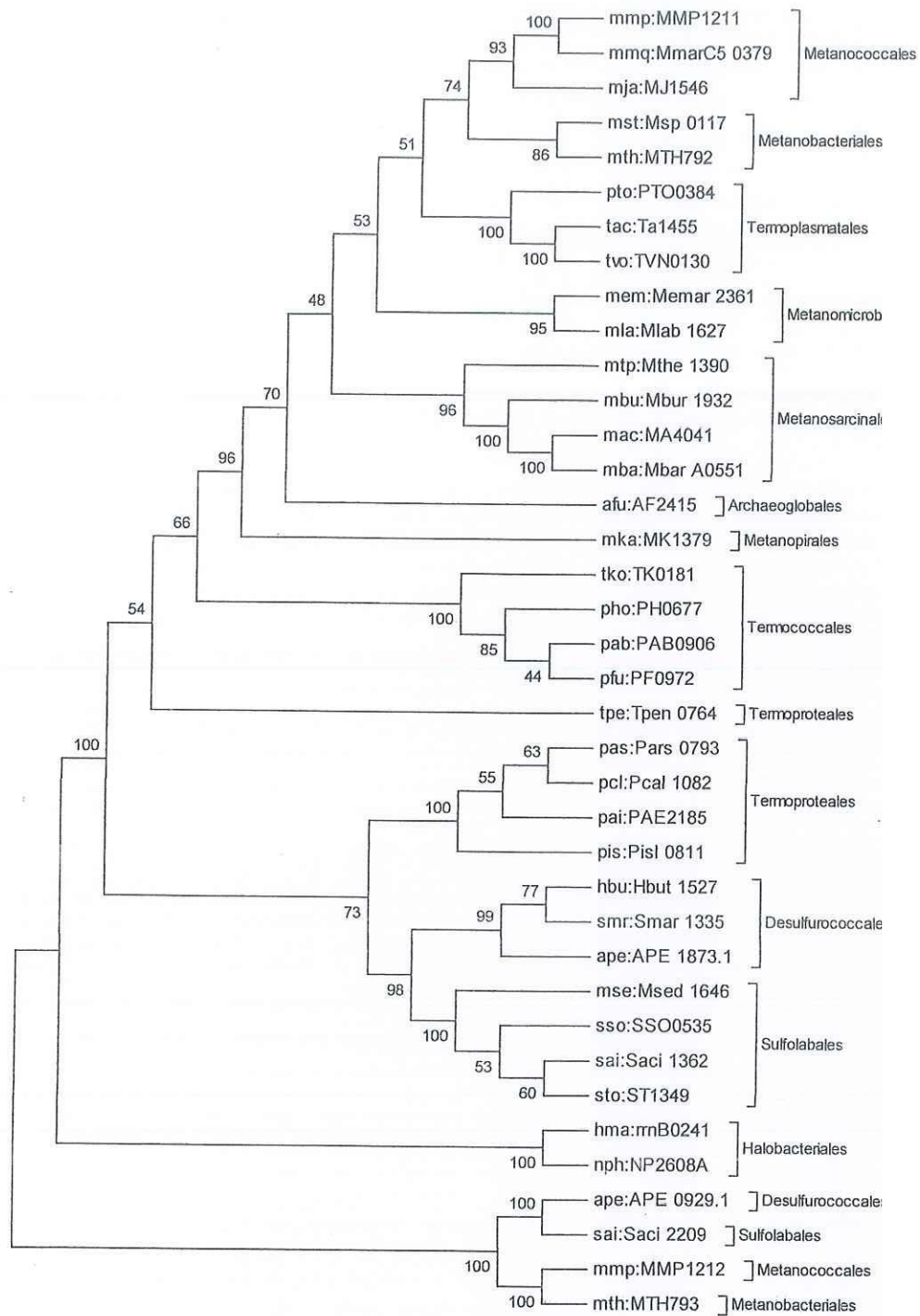


Figura 8. Árbol filogenético de la β -cetoacil-[ACP] sintetasa III (2.3.1.180). Los números en las ramas son las posibilidades bayesianas del clado.

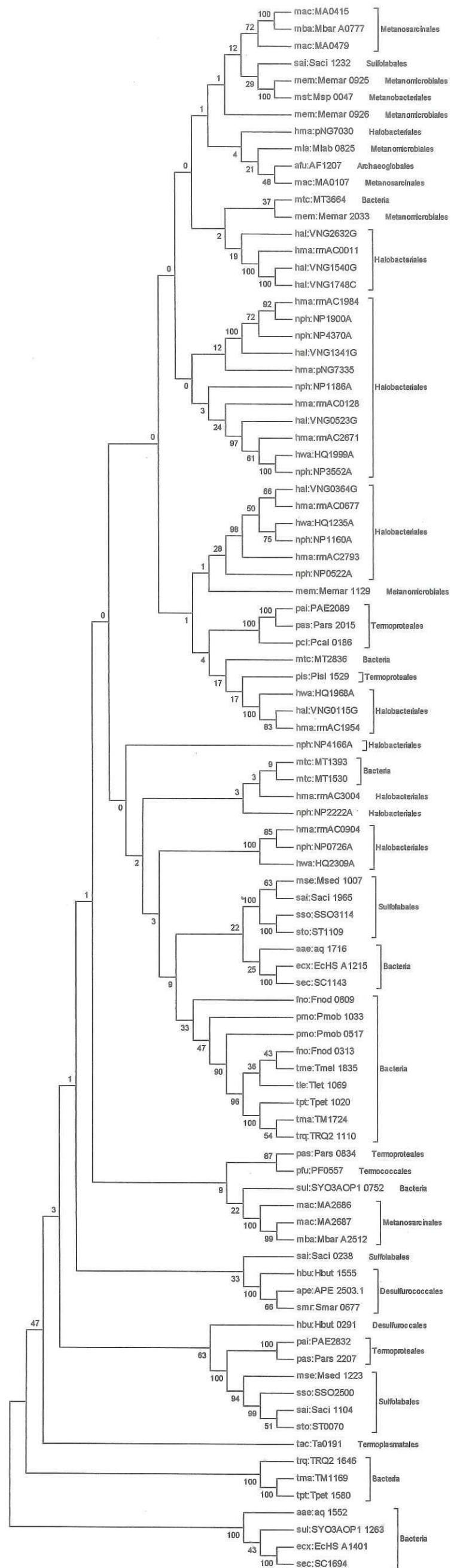


Figura 9. Árbol filogenético de las enzimas 3-cetoacil-[ACP] reductasa (1.1.1.100) y enoil-[ACP] reductasa (NADPH). Los números en las ramas son las posibilidades bayesianas del clado.

horizontal de genes. La ramificación interna del orden de las *Crenarchaeota* no coincide con la filogenia ribosomal, pero se encuentra bien sostenida por la prueba de *bootsrap*. Las posibles duplicaciones no fueron ubicadas dentro de la genealogía principal, debido, probablemente, a variaciones significativas (Figura 9).

El árbol filogenético de las enzimas 1.1.1.100 y 1.3.1.39 (Figura 11) no generó un patrón similar al árbol filogenético obtenido por Brochier-Armanet *et al.* (2008). Fue imposible generar árbol filogenético de la enzima 2.3.1.180.

Debido a que el número de duplicaciones de cada enzima no posee una aparente distribución normal, se realizó una prueba Kruskal-Wallis, la cual indica que existe una diferencia significativa ($p < 0.05$) entre el número de duplicaciones las enzimas. La figura 10 muestra cuales enzimas son significativamente distintas. Según lo observado en la figura 3, la enzima 2.3.1.180 es la más diferente ($p < 0.05$).

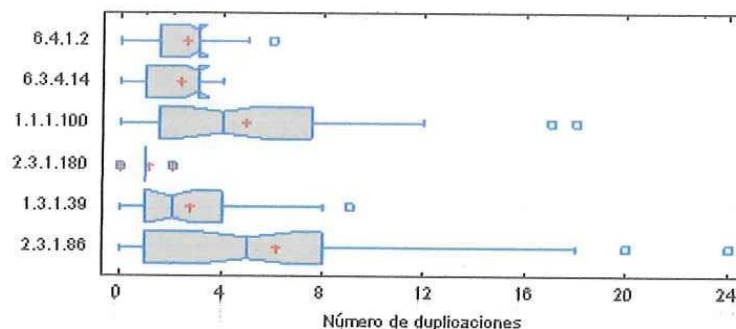


Figura 10. El rango de cada muestra muestra la incertidumbre asociada con el estimado de la mediana de ese grupo. Las muescas son construidas de tal modo que cualesquiera dos muestras que no se traslapen pueden declararse significativamente distintas ($p < 0.05$).

El análisis de regresión lineal simple sugiere que no existe una relación significativa entre el pH, la salinidad y la temperatura y el número de

duplicaciones observado en la mayoría de las enzimas ($p < 0.05$; Tabla VI). Aunque detectó una posible relación entre el número de duplicaciones observadas en la enzima 6.4.1.2 y el pH, el coeficiente de correlación indica que esta dependencia es muy débil. Del mismo modo sucede con la enzima 6.3.4.14 y la temperatura. El análisis también encontró una relación moderada entre el número de duplicaciones de la enzima 1.1.1.100 y la salinidad.

Tabla VI. Análisis de regresión lineal simple. r es el coeficiente de correlación y p el valor de P

Enzima	pH		Salinidad		Temperatura	
	r	p	r	p	R	p
6.4.1.2	0.33	0.04	0.26	0.12	-0.20	0.22
6.3.4.14	0.23	0.16	0.10	0.55	-0.35	0.03
2.3.1.180	0.08	0.60	-0.14	0.40	-0.20	0.23
2.3.1.86	-0.15	0.37	0.12	0.45	-0.17	0.31
1.1.1.100	-0.16	0.32	0.52	0.00	-0.18	0.27
1.3.1.39	-0.28	0.08	0.26	0.14	-0.05	0.76

El análisis de regresión multifactorial señala que no existe un efecto sinérgico entre el pH, la salinidad y la temperatura sobre el número de duplicaciones observado en la mayoría de las enzimas (Tabla VII). A pesar de que detectó una posible relación entre la enzima 1.1.1.100 y las variables estudiadas, sólo explica un 28.84% de este fenómeno. Además, el análisis descarta la temperatura como un factor trascendental en esta relación.

Tabla VII. Análisis de regresión multifactorial. r^2 variabilidad observada y p valor de P .

Enzima	r^2	P
6.4.1.2	9.40 %	0.10
6.3.4.14	10.60 %	0.08
2.3.1.180	0.00 %	0.63
2.3.1.86	0.00 %	0.53
1.1.1.100	28.84 %	0.00
1.3.1.39	9.90 %	0.09

VII Discusiones

6.1 Análisis Funcional y Estructural

De acuerdo con la clasificación funcional y estructural las enzimas 6.4.1.2, 6.3.4.14, 2.3.1.180, 1.1.1.100 y 1.3.1.39 no comparten un origen común, o han divergido tanto que es muy difícil o imposible rastrear su genealogía. La información generada por la clasificación funcional y estructural sustenta la evolución de la ruta por el modelo tipo “patchwork” propuesto por Jensen (1976). Aunque las enzimas 2.3.1.41, 2.3.1.179 y 2.3.1.180 comparten estructuras similares lo que sugiere un mismo origen, no puede aseverarse que la ruta haya evolucionado “hacia atrás” como propone Horowitz (1945), pues sólo catalizan una parte del proceso.

Las enzimas 2.3.1.41, 2.3.1.179 y 2.3.1.180 pertenecen a un grupo de la superfamilia de las tiolasas que cataliza la reacción de condensación de Claisen y comparten una estructura y función similar (Jiang *et al.*, 2008). Todas las enzimas de la superfamilia poseen una cisteína en el sitio activo que actúa como un nucleófilo y ancla la cadena naciente durante las reacciones de condensación. Las enzimas 2.3.1.41 y 2.3.1.179 (Huang *et al.*, 1998; Olsen *et al.*, 1999) conservan, además de la cisteína, dos residuos de histidina invariables (CHH), mientras que en la enzima 2.3.1.180 (Qiu *et al.*, 1999) un residuo de asparagina sustituye la última histidina: CHN (Figura 12).

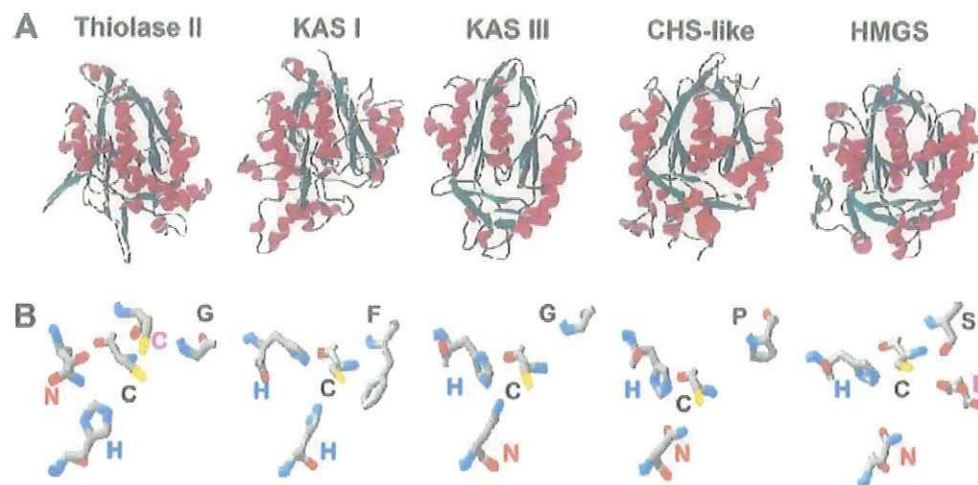


Figura 11. Comparación de la estructura terciaria y de los sitios activos de cinco enzimas de la superfamilia de las tiolasas. (A) Estructura monomérica con las α -hélices en rojo y hojas β en verde. (B) Sitios activos de las enzimas. La cisteína del sitio activo nucleofílico se posiciona en el centro, se puede ver que los residuos de histidina (H) de la tiolasa y la KAS I son equivalentes a los residuos de asparagina (N) de la KAS III, CHS y HMGS. Imagen modificada de Jiang *et al.*, 2008.

Existen varios sitios conservados entre las secuencias de la enzima 2.3.1.180 generadas por el BLAST y las secuencias de la 3-hidroxi-3-metilglutaril-CoA sintetasa (2.3.3.10) de *M. jannaschii*, *M. acetivorans* y *A. pernix*. Además de conservar los residuos CHN, la HMGS contiene un residuo de glutamato que actúa como base en la reacción no descarboxilante (ECHN; Campobasso *et al.*, 2004), el cual también se encuentra presente en las secuencias obtenidas (Figura 13).

En las arquea, la HMGS cataliza la condensación de Claisen de acetil-CoA con acetoacetil-CoA en la ruta del mevalonato. La HMGS forma un homodímero, cada monómero consiste de dos regiones estructurales referidas como región superior e inferior. La región superior de la HMGS y la región superior de la enzima 2.3.1.180 pueden superponerse, implicando un origen común. Las respectivas regiones inferiores muestran el mayor grado de

variación dentro de las enzimas. La HMGS posee más aminoácidos en la región C-terminal, mientras que en la enzima 2.3.1.180 es de menor tamaño y no presenta grandes inserciones (la chalcón sintetasa tiene una inserción de 36 aminoácidos cerca del extremo N-terminal, mientras que la tiolasa contiene una inserción de 100 aminoácidos cerca de la mitad de la molécula).

El análisis funcional de las secuencias generadas por el BLAST, indica que el dominio PF01154 cataliza la función del dominio HMGS N-terminal, lo que robustece la idea de que la HMGS y la enzima 2.3.1.180 comparten un mismo origen; también señala que el dominio PF08541 cataliza la función dominio C-terminal de la 3-oxoacil-[acp] sintetasa III y el dominio PF08540 cataliza la actividad del dominio HMGS C-terminal, lo cual sugiere que la mayoría de estas secuencias pertenece a enzimas que catalizan la reacción 2.3.1.180 y no la 2.3.1.41 o 2.3.3.10. Sin embargo, se recomienda realizar ensayos bioquímicos y genéticos para corroborar esta información.

6.2 Análisis Filogenético y Estadístico

A pesar de que el análisis filogenético de las enzimas 6.3.4.14 y 2.3.1.180 reconoció la presencia de dos grandes clados (Crenarchaeota y Euryarchaeota), la ramificación interna no coincidió con el patrón observado en la filogenia de varias moléculas de rRNA (Brochier-Armanet *et al.*, 2008). Gribaldo y Philippe (2002) identificaron 3 factores que pueden conducir a la interpretación inadecuada de un árbol filogenético:

1. Sesgos constitutivos: especies distantes comparten gran cantidad de su contenido nucleotídico o de aminoácidos y son afiliadas incorrectamente.
2. Tasa evolutiva variable: agrupamiento equivocado de los taxa con mayor tasa de evolución.
3. Heterotaquia: la tasa de sustitución varía a lo largo de la molécula a través del tiempo y los organismos que comparten estos mismos sitios son agrupados sin poseer un parentesco verdadero o cercano.

Aunque la filogenia que presentan estas enzimas no es la misma que observada con marcadores moleculares más precisos, sería aventurado negar que es correcta.

Philippe y Douady (2003) consideran que el transporte horizontal de genes (HGT) entre procariontes es otro factor preponderante al momento de construir una filogenia. Este fenómeno puede ocasionar una inferencia completamente incongruente con la realidad. En la figura 8 se puede observar que los Termococcales y los Archaeoglobales son emparentados directamente con los Desulfurococcales y los Sulfolabales, miembros de las Crenarchaeota. Sin embargo, antes debe descartarse los factores previamente descritos no hayan alterado la filogenia y realizar otros análisis que apoyen la hipótesis del HGT.

La superfamilia de las deshidrogenasas de cadena corta posee una alta variabilidad y consecuentemente ha dado origen a una gran variedad de moléculas de este tipo con distinta especificidad. Esto explica por que las

secuencias de las enzimas 1.1.1.100 y 1.3.1.39 no puedan afiliarse correctamente.

Debido a que la redundancia funcional de un gen duplicado puede no ser ventajosa, casi siempre la copia resultante es perdida por la acumulación gradual de mutaciones (Zhang, 2003). Y aunque este modelo ha sido sometido a numerosos análisis matemáticos, no se ajusta a los datos existentes (Lynch y Force, 1999). Es sostenido que la preponderancia del metabolismo de elementos primarios puede limitar su variabilidad. A diferencia del metabolismo primario, los genes involucrados en el metabolismo de elementos secundarios se distinguen por su alta tasa de variación. Se ha predicho que esta diversidad es promovida por la duplicación génica y la posterior neofuncionalización de las enzimas (Kliebenstein, 2008). A pesar de las diferencias que existen actualmente entre los elementos y la síntesis de los elementos que componen la membrana citoplasmática de las arquea y los demás seres vivos (Smith y Mushegian, 2000; Koga y Morii, 2007), las evidencias sugieren que el último ancestro común poseía la información genética necesaria y la maquinaria enzimática suficiente para sintetizar los precursores estructurales de ambas membranas (Peretó *et al.*, 2004). Aunque las arquea todavía son capaces de sintetizar ácidos grasos, probablemente esta ruta metabólica fue sustituida por otra como consecuencia de una adaptación secundaria a ambientes extremos, propiciando un alto grado de variación dentro de esta vía (Cavallier-Smith, 2006).

Los análisis estadísticos de regresión lineal y multifactorial eliminan la posibilidad de que el gran número de duplicaciones génicas detectadas se deba a factores ambientales como el pH, la temperatura o la salinidad. Sin embargo, no se puede descartar la idea de que los ésteres de glicerol sintetizados por estas enzimas formen parte de la membrana citoplasmática de las arquea, además de mantener la estabilidad física de la célula, los ácidos grasos también tienen otras funciones importantes, pueden actuar como mensajeros a respuestas extracelulares o cofactores de enzimas (Nelsen y Cox, 2005). El hecho de que los mutantes incapaces de sintetizar fosfolípidos relativamente sencillos no sean viables apoya esta hipótesis (Raetz y Dowhan, 1990).

VIII Conclusiones

Sólo se detectaron 6 enzimas que presentan homólogas: Acetil CoA-Carboxilasa (6.4.1.2), Biotina Carboxilasa (6.3.4.14), 3-cetoacil-ACP sintetasa III (2.3.1.180), lipoacil-CoA sintetasa (2.3.1.86), 3-cetoacil-ACP reductasa (1.1.1.100) y enoil-ACP reductasa (NADPH) (1.3.1.39).

La ruta metabólica FAS II se ensambló de acuerdo al modelo tipo "patchwork" propuesto por Jensen.

El análisis filogenético detectó los dos órdenes de *Archaea*. Sin embargo la topología interna no siempre coincidió con el árbol genealógico de las especies.

Los sesgos constitutivos, la tasa evolutiva variable, la heterotaquia y el transporte horizontal de genes afectaron el análisis filogenético modificando ocasionalmente la topología interna.

La mayoría de las enzimas encontradas, excepto la enzima 2.3.1.180, presenta el doble o más homologías que lo esperado.

Los análisis estadísticos de regresión lineal y multifactorial descartan la posibilidad de que el pH, la temperatura y la salinidad hayan propiciado la elevada tasa de duplicación génica.

IX Bibliografía

- Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts y P. Walter. 2002. *Molecular Biology of the Cell*. 4th. ed. 11-12 pp.
- Altschul, S.F., W. Gish, W. Miller, E. W. Myers y D. J. Lipman. 1990. Basic local alignment search tool. *J Mol Biol* 215(3): 4-410.
- Alves, R., R. A. Chaleil y M. J. Sternberg. 2002. Evolution of enzymes in metabolism: a network perspective. *J Mol Biol* 320(4): 751-770.
- Barns, S. M., R. E. Fundyga, M. W. Jeffries y N. R. Pace. 1994. Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc Natl Acad Sci USA* 91(5): 1609-1613.
- Barns, S.M., C.F. Delwiche, Palmer J.D. y Pace N.R. 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc Natl Acad Sci USA* 93(17): 9188-9193.
- Battin, T. J., A. Wille, B. Sattler y R. Psenner. 2001. Phylogenetic and functional heterogeneity of sediment biofilms along environmental gradients in a glacial stream. *Appl Environ Microbiol* 67(2): 799-807.
- Beadle, G. E. y E. L. Tatum. 1941. Genetic control of biochemical reactions in *Neurospora*. *Proc Natl Acad Sci USA* 27(11): 499-506.
- Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell y D. L. Wheeler. 2007. GenBank. *Nucleic Acids Res.* 36: 25-30.
- Bergler, H., S. Fuchsbichler, G. Högenauer y F. Turnowsky. 1996. The enoyl-[acyl-carrier-protein] reductase (FabI) of *Escherichia coli*, which catalyzes a key regulatory step in lipo acid biosynthesis, accepts NADH and NADPH as cofactors and is inhibited by palmitoyl-CoA. *Eur J Biochem* 242(3): 689-694.
- Boone, D.R. y R.W. Castenholz. 2001. *Bergey's manual of systematic bacteriology*. Springer. New York.
- Brochiere-Armanet, C., B. Bassau, S. Gribaldo y P. Forterre. 2008. Mesophilic crenarchaeota Proposal for a third archaeal phylum, the Thaumarchaeota. *Nature*. 6(2): 45-252.
- Campbell, J. W. y J. E. Jr Cronan. 2001. Bacterial lipo acid biosynthesis: targets for antibacterial drug discovery. *Annu Rev Microbiol* 55: 305-332.
- Campobasso, N., M. Patel, I. E. Widing, H. Kallendar, M. Rosenberg y M. N. Gwynn. 2004. *Staphylococcus aureus* 3-hydroxy-3-methylglutaryl-CoA

synthase: crystal structure and mechanism. *J Biol Chem* 279(43): 44883-44888.

Cavallier-Smith, T. 2006. Cell evolution and Earth history: stasis and revolution. *Philos Trans R Soc Lond B Biol Sci* 361(1470): 969-1006.

Cavicchioli, R. 2006. Cold-adapted archaea. *Nat Rev Microbiol* 4(5): 331-343.

Chang, E. L. 1994. Unusual thermal stability of liposomes made from bipolar tetraether lipids. *Biochem Biophys Res Commun.* 202(2): 673-679.

Choquet, C. G., G. B. Patel, T. J. Beveridge y G. D. Sprott. 1994. Stability of pressure-extruded liposomes made from archaeobacterial ether lipids. *Appl Microbiol Biotechnol* 42(2-3): 375-384.

Collins, G., L. O'Connor, T. Mahony, A. Gieseke, D. de Beer y V. O'Flaherty. 2005. Distribution, localization, and phylogeny of abundant populations of Crenarchaeota in anaerobic granular sludge. *Appl Environ Microbiol* 71(11): 7523-7527.

De Rosa, M y A. Gambacorta. 1988. The lipids of archaebacteria. *Prog Lipid Res* 27: 153-175.

DeLong, F. 2000. Extreme Genomes. *Genome Biol* 1(6): 1-3.

Elferink, G. L., G. de Wit, R. Demels, J. M. Driessen y W. N. Konings. 1992. Functional reconstitution of membrane proteins in monolayer liposomes from bipolar lipids of *Sulfolobus acidocaldarius*. *J Biol Chem* 267(2): 1375-1381.

Elferink, M. G., J. G. de Wit, A. J. Driessen, W. N. Konings. 1994. Stability and proton-permeability of liposomes composed of archaeal tetraether lipids. *Biochim Biophys Acta.* 1193(2): 247-54.

Elkins, J.G., M. Podar, D.E. Graham, K.S. Makarova, Y. Wolf, L. Randau, B.P.

Hedlund, C. Brochiere-Armanet, V. Kunin, I. Anderson, A. Lapidus, E. Goltzman, K. Barry, E.V. Koonin, P. Hugenholtz, N. Kyrpides, G. Wanner, P. Richardson, M. Keller y K.O. Stetter. 2008. A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc Natl Acad Sci USA.* 105(23): 8102-8107.

Fan, H., Y. Xue. Y. Ma, A. Ventosa y W. D. Grant. 2004. *Halorubrum tibetense* sp. nov., a novel haloalkaliphilic archaeon from Lake Zabuye in Tibet, China. *Int J Syst Evol Microbiol* 54(4): 1213-1216.

Finn. R. D., J. Tate, J. Mistry, P. C. Coghill, S. J. Sammut, H. R. Hotz, G. Ceric, K. Forslund, S. R. Eddy, E. L. Sonnhammer y A. Bateman. 2008. The Pfam protein families database. *Nucleic Acids Res* 36: 281-8.

Forterre, P., C. Brochiere y H. Philippe. 2002. Evolution of the Archaea. *Theor Popul Biol.* 61(4): 409-422.

Fütterer, O., A. Angelov, H. Liesegang, G. Gottschalk, C. Schepler, B. Schepers, C. Dock, G. Antranikian y W. Liebl. 2004. *Picrophilus torridus* and its implications for life around pH 0. *Proc Natl Acad Sci USA* 101(24): 9091-9096.

Galand, P. E., H. Fritze, R. Conrad y K. Yrjälä. 2005. Pathways for methanogenesis and diversity of methanogenic archaea in three boreal peatland ecosystems. *Appl Environ Microbiol.* 71(4): 2195-2198.

Garwin, J. L., A. L. Klagers y J. E. Cronan Jr. 1980. Structural, enzymatic, and genetic studies of beta-ketoacyl-acyl carrier protein synthases I and II of *Escherichia coli*. *J Biol Chem* 255(24): 11949-11956.

Gaughran, E. R. 1947. The saturation of lipids as a function of temperature. *J Bacteriol* 53: 506-509.

Gribaldo, S. y H. Philippe. 2002. Ancient phylogenetic relationships. *Theor Popul Biol.* 61: 391-408.

Gurr, M. I., J. L. Harwood y K. N. Frayn. 2002. *Lipid Biochemistry*. 5th ed. Blackwell Science. Reino Unido. Pag. 29.

Haldane, J. B. S. 1933. The part played by recurrent evolution. *Am Nat* 67: 5-9.

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

Heath, R. J. y C. O. Rock. 1995. Enoyl-Acyl Carrier Protein Reductase (fabI) Plays a Determinant Role in Completing Cycles of Lipo Acid Elongation in *Escherichia coli*. *J Biol Chem* 270(44): 26538-26542.

Heath, R. J. y C. O. Rock. 1996. Roles of the FabA and FabZ β -Hydroxyacyl-Acyl Carrier Protein Dehydratases in *Escherichia coli* Lipo Acid Biosynthesis. *J Biol Chem* 271(44): 27795-27801.

Heath, R. J., S. Jackowski y C. O. Rock. 2002. Lipo acid and phospholipid metabolism in prokaryotes en *Biochemistry of lipids, lipoproteins and Membranes*. 4th edition. Elsevier Science. 55-91 pp.

Hobel, C. F., V. T. Marteinson, G. Ó. Hreggvidsson y J. K. Kristjánsson. 2005. Investigation of the microbial ecology of intertidal hot springs by using diversity analysis of 16S rRNA and chitinase genes. *Appl Environ Microbiol* 71(5): 2771-2776.

Horowitz, N. H. 1945. On the evolution of biochemical synthesis. *Proc Natl Acad Sci USA* 31(6): 153-157.

Horowitz, N. H. 1965. The evolution of biochemical synthesis -Retrospect and prospect. Academic Press. New York. 15-23 pp.

Huang, W., J. Jia, P. Edwards, K. Dehesh, G. Schneider y Y. Lindquist. 1998. Crystal structure of beta-ketoacyl-acyl carrier protein synthase II from *E.coli* reveals the molecular architecture of condensing enzymes. *EMBO J* 17(5): 1183-1191.

Huber, H., M.J. Hohn, R. Rachel, T. Fuchs, V. C. Wimmer y K. O. Stetter. 2002. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature*. 417(6884): 63-67.

Huelsenbeck, J. P. y F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754-755.

Hughes, A. L. 2005. Gene duplication and the origin of novel proteins. *Proc Natl Acad Sci USA* 102(25): 8791-8792.

Jensen, R. A. 1976. Enzyme recruitment in evolution of new function. *Ann Rev Microbiol* 30: 409-425.

Jiang, C., S. Y. Kim y D. Y. Suh. 2008. Divergent evolution of the thiolase superfamily and chalcone synthase family. *Mol Phylogenet Evol* 49(3): 691-701.

Jones, W. J., J. A. Leigh, F. Mayer, C. R. Woese, y R. S. Wolfe. 1983. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. 136(4): 254-261.

Jørgensen, B. B. y S. D'Hondt. 2006. A starving majority deep beneath the seafloor. *Science*. 314(5801): 932-934.

Kanehisa, M y S. Goto. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28(1): 27-30.

Kaneshiro, S. M. y D. S. Clark. 1995. Pressure effects on the composition and thermal behavior of lipids from the deep-sea thermophile *Methanococcus jannaschii*. *J Bacteriol*. 177(13): 3668-72.

Kates, M. 1993. The Biochemistry of Archaea (Archaeobacteria). Elsevier Science. 261-295 pp.

Keough, B. P., T. M. Schmidt y R. E. Hicks. 2003. Archaeal nucleic acids in picoplankton from great lakes on three continents. *Microbiol Ecol* 46(2): 238-248.

Kim, B.S., H.M Oh, H. Kang, J. Chun 2005. Archaeal diversity in tidal flat sediment as revealed by 16S rDNA analysis. *J Microbiol.* 43: 144-151.

Kimura, M. y T. Ohta. 1974. On some principles governing molecular evolution. *Proc Nat Acad Sci USA* 71(7): 2848-2852.

Kleikemper, J., S. A. Pombo, M. H. Schroth, W. V. Sigler, M. Pesaro y J. Zeyer. 2005. Activity and diversity of methanogens in a petroleum hydrocarbon-contaminated aquifer. *Appl Environ Microbiol.* 71(1): 149-158.

Kliebenstein, D. J. 2008. A role for gene duplication and natural variation of gene expression in the evolution of metabolism. *PLoS ONE* 3(3): e1838.

Knittel, K., T. Lösekann, A. Boetius, R. Kort y R. Amann. 2005. Diversity and distribution of methanotrophic archaea at cold seeps. *Appl Environ Microbiol.* 71(1): 467-479.

Koch, A. L. 1972. Enzyme evolution. I. The importance of untranslatable intermediates. *Genetics* 72(2): 297-316.

Koga, Y. y H. Morii. 2007. Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations. *Microbiol Mol Biol Rev* 71(1): 97-120.

Lazcano, A. y S. L. Miller. 1999. On the origin of metabolic pathways. *J Mol Evol* 49(4): 424-431.

Lepp, P. W., M. M. Brinig, C. C. Ouverney, K. Palm, G. C. Armitage y D. A. Relman. 2004. Methanogenic Archaea and human periodontal disease. *Proc Natl Acad Sci USA.* 101(16): 6176-6181.

Li, S. J. y J. E. Cronan Jr. 1992. The genes encoding the biotina carboxilase subunit of *Escherichia Coli* acetyl-CoA carboxilase. *J Biol Chem.* 267(2): 855-863.

Liao, D. 1999. Concerted evolution: molecular mechanism and biological implications. *Am J Hum Genet* 64(1): 24-30.

Lynch, M. y A. Force. 1999. The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154(1): 459-473.

Marques, A. C., I. Dupanloup, N. Vickenbosch, A. Reymond y H. Kaessmann. 2005. Emergence of Young Human Genes after a Burst of Retroposition in Primates. *PLoS Biol* 3(11): e357.

Mesbah, N. M. y J. Wiegel. 2008. Life at extreme limits the anaerobic halophilic alkalithermophiles. *Ann N Y Acad Sci* 1125: 44-57.

Mills, H. J., R. J. Martinez, S. Story y P. A. Sobecky. 2005. Characterization of microbial community structure in Gulf of Mexico gas hydrates: comparative analysis of DNA- and RNA-derived clone libraries. *Appl Environ Microbiol.* 1: 3235-3247.

Morrison, R. T. y R. N. Boyd. 1998. Química orgánica. 5a. ed. Pearson Educación. México. 1253-1254 pp.

Murray, A. E., C. M. Preston, R. Massana, L. T. Taylor, A. Blakis, K. Wu, E. F. DeLong. 1998. Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters near Anvers Island, Antarctica. *Appl Environ Microbiol* 64(7): 2585-2595.

Murzin, A. G., S. E. Brenner, T. Hubbard y C. Cothia. 1995. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* 247(4): 536-540.

Nelson, D. L. y M. M. Cox. 2005. Lehninger: Principles of Biochemistry. 4th. ed. W. H. Freeman and Company. New York. 343-349 pp.

Ohno, S. 1970. Evolution by gene duplication. Springer-Verlag. 160 pp.

Olsen, G. J. y C.R. Woese. 1997. Archaeal genomics: an overview. *Cell* 89(7): 991-994.

Olsen, J. G., A. Kadziola, P. von Wettstein-Knowles, M. Siggaard-Andersen, Y. Lindquist y S. Larsen. 1999. The X-ray crystal structure of beta-ketoacyl [acyl carrier protein] synthase I. *FEBS Lett* 460(1): 46-52.

Oparin, A. I. 1968. El origen de la vida. Editorial Grijalbo. 154 pp.

Oren, A., M. Ginzburg, B. Z. Ginzburg, Hochstein L. I. y B. E. Volcani. 1990 *Haloarcula marismortui* (Volcani) sp. nov., nom. rev., an extremely halophilic bacterium from the Dead Sea. *Int J Syst Bacteriol* 40(2): 209-210.

Oren, A., P. Gurevich, R. T. Gimmell y A. Teske. 1995. *Halobaculum gomorense* gen. nov., sp. nov., a novel extremely halophilic archaeon from the Dead Sea. *Int J Syst Bacteriol.* 45(4): 747-754.

Page, R. D. y E. C. Holmes. 1998. Molecular evolution: A phylogenetic approach. Blackwell Science Ltd. Massachusetts. 73 p.

Peretó, J., R. Fani, J. I. Leguina y A. Lazcano. 1997. Enzyme evolution and the development of metabolic pathways. Universitat de València. España. 185 p.

Peretó, J., P. López-García y D. Moreira. 2004. Ancestral lipid biosynthesis and early membrane evolution. Trends Biochem Sci. 29(9): 469-77.

Phillipe, H. y C. J. Douady. 2003. Horizontal gene transfer and phylogenetics. Curr Opin Microbiol. 6: 498-505.

Preston, C.M., K. Y. Wu, T. F. Molinskit y E. F. DeLong. 1996. A Psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. Proc Natl Acad Sci USA 93(13): 6241-6246.

Qiu, X. C. A. Janson, A. K. Konstantinidis, S. Nweagwu, C. Silverman, W. W. Smith, S. Khandekar, J. Lonsdale, S. S. Abdel-Mequid. 1999. Crystal structure of beta-ketoacyl-acyl carrier protein synthase III. A key condensing enzyme in bacterial lipo acid biosynthesis. J Biol Chem 274(51): 36465-36471.

Raetz, C. R. y W. Dowhan. 1990. Biosynthesis and function of phospholipids in *Escherichia coli*. J Biol Chem 265(3): 1235-1238.

Ravenschlag, K., K. Sahm y R. Amann. 2001. Quantitative Molecular Analysis of the Microbial Community in Marine Arctic Sediments (Svalbard). Appl Environ Microbiol. 67(1):387-395.

Ray, A., M. S. Da Costa y V. M. Madeira. 1988. Effect of growth temperature on the lipid composition of *Thermus aquaticus*. J Gen Microbiol 134(6): 1553-1560.

Reizer, J., N. Grossowicz y Y. Barenholz. 1985. The effect of growth temperature on the thermotropic behavior of the membranes of a thermophilic *Bacillus*. Composition-structure-function relationships. Biochim Biophys Acta 815: 268-280.

Rison, S.C y J. M. Thornton. 2002. Pathway evolution, structurally speaking. Curr Opin Struct Biol 2(3): 374-382.

Ruch, F. E y P. Vagelos. 1973. Characterization of a malonyl-enzyme intermediate and identification of the malonyl binding site in malonyl coenzyme A-acyl carrier protein transacylase of *Escherichia coli*. J Biol Chem 248(23): 8095-8106.

Rusell, N. J. y N. Fukunaga. 1990. A comparison of thermal adaptation of membrane lipids in psychrophilic and thermophilic bacteria. *FEMS Microbiol Rev* 75: 171-182.

Schmidt, S., S. Sunyaev, P. Bork y T. Dandekar. 2003. Metabolites: a helping hand for pathway evolution? *Trends Biochem Sci* 28(6): 336-341.

Singer, S. J. y G. L. Nicolson. 1972. The fluid mosaic model of the structure of cell membranes. *Science* 175(4023): 720-731.

Skirnisdottir, S., G. O. Hreggvidsson, S. Hjörleifsdottir, V. T. Marteinson, S. K. Petusdottir, O. Holst y J. K. Kristjansson. 2000. Influence of sulfide and temperature on species composition and Community Structure of Hot Spring Microbial Mats. *Appl Environ Microbiol* 66(7): 2835-2841.

Smith, A. y A. Mushegian. 2000. Biosynthesis of isoprenoids via mevalonate in Archaea: the lost pathway. *Genome Res* 10(10): 1468-1484.

Sprott, G. D., M. Maloche y J. C. Richards. 1991. Proportions of diether, macrocyclic diether, and tetraether lipids in *Methanococcus jannaschii* grown at different temperatures. *J Bacteriol* 173(12): 3907-3910.

Stefan, M. S., W. Zebis, J. Kuever y K. Sahm. 2000. Relative abundance of Archaea and Bacteria along a thermal gradient of a shallow-water hydrothermal vent quantified by rRNA slot-blot hybridization. *Microbiology* 146(6): 1287-1293.

Svobodová, J. y P. Svoboda. 1988. Membrane fluidity in *Bacillus subtilis*. Physical change and biological adaptation. *Folia Microbiol (Praha)* 33(3): 161-169.

Takai, K. y K. Horikoshi. 1999. Genetic diversity of archaea in deep-sea hydrothermal vent environments. *Genetics* 152(4): 1285-1297.

Tamura, K., J. Dudley, M. Ney y S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24(8): 1596-1599.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin y D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25(24): 4876-4882.

Tornabene, T. G. y T. A. Langworthy. 1979. Diphytanyl and dibiphytanyl glycerol ether lipids of methanogenic archaeobacteria. *Science* 203(4375): 51-53.

Tsay, J. T., W. Oh, T. J. Larson, S. Jackowski y C. O. Rock. 1992. Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia coli* K-12. *J Biol Chem* 267(10): 6807-6814.

Valentine, D. L. 2007. Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat Rev Microbiol.* 5(4): 316-323.

van de Vossenberg, J. L., T. Ubbink-Kok, M. G. Elferink, A. J. Driessen y W. N. Konings. 1995. Ion permeability of the cytoplasmic membrane limits the maximum growth temperature of bacteria and archaea. *Mol Microbiol* 18(5): 925-932.

van de Vossenberg, J. L., A. J. Driessen y W. N. Konings. 1998. The essence of being extremophilic: the role of the unique archaeal membrane lipids. *Extremophiles.* 2(3): 163-170.

Wainø, M., B. J. Tindall y K. Ingvorsen. 2000. *Halorhabdus utahensis* gen. nov., sp. nov., an aerobic, extremely halophilic member of the Archaea from Great Salt Lake, Utah. *Int J Syst Evol Microbiol* 50(1): 183-190.

Waite, M y S. J. Wakil. 1962. Studies on the mechanism of lipo acid synthesis. XII. Acetyl coenzyme A carboxilase. *J Biol Chem* 237: 2750-2757.

Whildaber, I., U. Santurius y W. Baumeister. 1987. Three-dimensional structure of the surface protein of *Desulfurococcus mobilis*. *J Bacteriol* 169(12): 5563-5568.

Williamson, I. P y S. J. Wakil. 1966. Studies on the mechanism of lipo acid synthesis. XVII. Preparation and general properties of acetyl coenzyme A and malonyl coenzyme A-acyl carrier protein transacylases. *J Biol Chem* 241(10): 2326-2332.

Woese, C.R. y G.E. Fox. 1977. Phylogenetic structure of the prokaryotic domain The primary kingdoms. *Proc Natl Acad Sci USA* 74(11): 5088-5090.

Woese, C.R., O. Kandler y M.L. Wheelis. 1990. Towards a natural system of organisms Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA.* 87(12): 4576-4579.

Yamauchi, K., K. Doi, Y. Yoshida y M. Kinoshita. 1993. Archaeobacterial lipids: highly proton-impermeable membranes from 1,2-diphytanyl-sn-glycero-3-phosphocholine. *Biochim Biophys Acta* 1146(2): 178-182.

Zhang, J. 2003. Evolution by gene duplication: an update. *Trends Ecol Evol* 18(6): 292-298.

Zuckerlandl, E. y L. Pauling. 1965. Molecules as documents of evolutionary history. *J Theor Biol.* 8(2): 357-366.

X Anexos

Tabla VIII. Genes que codifican las enzimas involucradas en la biosíntesis de ácidos grasos tipo II (FAS II) en *Escherichia coli* HS (ecx), *Mycobacterium tuberculosis* CDC 1551 (mtc), *Salmonella enterica* subsp. *enterica* serovar choleraesius (sec), según la Enciclopedia de Genes y Genomas (KEGG) y dominios funcionales que integran cada proteína, según Pfam.

Enzima	Especie		
	<i>E. coli</i> (ecx)	<i>S. enterica</i> (sec)	<i>M. tuberculosis</i> (mtc)
6.4.1.2	EcHS_A018 (PF03255) EcHS_A2467 (PF01155, PF01039) EcHS_A3448 (PF00289, PF00286, PF02785)	SC0232 (PF03255) SC2368 (PF01155, PF01039) SC3317 (PF00364) SC3318 (PF00289, PF02786, PF02785) SC3318 (PF00289, PF02786, PF02785)	MT0927 (PF01039) MT1763 (-)
6.3.4.14			MT2576 (PF00289, PF02786, PF02785, PF00364) MT3384 (PF00289, PF00289, PF02785, PF00364) MT2305 (PF00109, PF02801) MT2306 (PF00109, PF02801) MT0557 (PF08545, PF08541)
2.3.1.179		SC1145 (PF00109, PF02801)	MT1041 (PF00501) MT1244 (PF00501) MT1387 (PF00501) MT1793 (PF00501) MT1976 (PF00501) MT2580 (PF00501) MT3011 (PF00501) MT3023 (PF00501) MT3610 (PF00501) MT3666 (PF00501) MT2306 (PF00109, PF02801) MT2306 (PF00109, PF02801)
2.3.1.180		SC1141 (PF08545, PF08541)	MT3021.1 (PF00109, PF02801, PF08990) MT2303 (PF00698) MT0256 (PF00106) MT0793 (PF00106) MT1393 (PF00106) MT1530 (PF00106) MT2836 (PF00106) MT3606 (PF00106) MT3664 (PF00106)
2.3.1.86			MT1531 (-)
2.3.1.41	EcHS_A1213 (PF08545, PF08541) EcHS_A1217 (PF00109, PF02801) EcHS_A2474 (PF00109, PF02801)	SC2380 (PF00109, PF02801)	
2.3.1.39		SC1142 (PF00698)	
1.1.1.100	EcHS_A1215 (PF00106)	SC1143 (PF00106)	
1.3.1.39	EcHS_A1401 (PF00106)	SC1694 (PF00106)	

Tabla IX. Clasificación funcional (Pfam) y estructural (SCOP) de los dominios que integran las enzimas que catalizan la FAS II en *Escherichia coli* HS, *Mycobacterium tuberculosis* CDC 1551 y *Salmonella enterica* subsp. *enterica* serovar *choleraesuis*.

	Pfam	SCOP
PF03255	Acetil CoA carboxilasa, subunidad alfa de la carboxiltransferasa	Clase: ND Plegamiento: ND Superfamilia: ND Familia: ND
PF01155	Expresión/síntesis de hidrogenasa, familia hypA	Clase: ND Plegamiento: ND Superfamilia: ND Familia: ND
PF01039	Dominio Carboxil Transferasa	Clase: alfa y beta proteína (a/b) Plegamiento: ClpP/Crotonasa Superfamilia: ClpP/Crotonasa Familia: carboxilasa-biotina dependiente, dominio carboxil transferasa
PF00364	Enzima Biotina-requirente	Clase: toda beta proteína Plegamiento: híbrido barril-sandwich Superfamilia: híbrido de un motivo Familia: proteínas y dominios acarreadoras de biotilil/lipoil
PF00289	Carbamoil-fosfato Sintetasa cadena L, dominio N-terminal	Clase: alfa y beta proteína (a/b) Plegamiento: dominio de pre-anclaje de ATP Superfamilia: dominio de pre-anclaje a ATP Familia: dominio similar a BC N-terminal
PF02786	Carbamoil-fosfato Sintetasa cadena L, dominio de unión a ATP	Clase: alfa y beta proteína (a/b) Plegamiento: anclaje a ATP Superfamilia: dominio similar a glutatión sintetasa ATP-binding Familia: dominio similar a BC-ATP binding
PF02785	Biotina carboxilasa, dominio C-terminal	Clase: Toda beta proteína Plegamiento: híbrido barril-sandwich Superfamilia: híbrido de un solo motivo rudimentario Familia: dominio similar a BC C-terminal
PF00109	Beta-cetoacil sintetasa, dominio N-terminal	Clase: alfa y beta proteína (a/b) Plegamiento: Tiolasa similar Superfamilia: Tiolasa similar Familia: Tiolasa similar
PF02801	Beta-cetoacil sintetasa, dominio C-terminal	Clase: alfa y beta proteína (a/b) Plegamiento: Tiolasa similar Superfamilia: Tiolasa similar Familia: Tiolasa similar
PF08890	Proteína similar a fago XkdN	Clase: ND Plegamiento: ND Superfamilia: ND Familia: ND
PF08545	3-oxoacil-[proteína acarreadora de acilo-(acp)] sintetasa III	Clase: alfa y beta proteína (a/b) Plegamiento: tipo tiolasa Superfamilia: tipo tiolasa Familia: tipo chalcon sintetasa
PF08541	3-oxoacil-[proteína acarreadora de acilo-(acp)] sintetasa III C-terminal	Clase: alfa y beta proteína (a/b) Plegamiento: tipo tiolasa Superfamilia: tipo tiolasa Familia: tipo chalcon sintetasa
PF00106	Deshidrogenasa de cadena corta	Clase: alfa y beta proteína (a/b) Plegamiento: dominio plegamiento-Rossmann NAD(P) binding Superfamilia: dominio plegamiento-Rossmann NAD(P) binding Familia: oxidoreductasa tirosina-dependiente
PF00501	Enzima AMP-binding	Clase: proteína de dominios múltiples (alfa y beta) Plegamiento: similar a Acetil-CoA sintetasa

Tabla IX. Clasificación funcional (Pfam) y estructural (SCOP) de los dominios que integran las enzimas que catalizan la FAS II en *Escherichia coli* HS, *Mycobacterium tuberculosis* CDC 1551 y *Salmonella enterica* subsp. *enterica* serovar choleraesius.

Pfam		SCOP
PF00698	Dominio acil transferasa	<p>Superfamilia: similar a Acetil-CoA sintetasa</p> <p>Familia: similar a Acetil-CoA sintetasa</p> <p>Clase: alfa y beta proteína (alfa/beta)</p> <p>Plegamiento: similar a ferredoxina</p> <p>Superfamilia: probable dominio de anclaje a ACP de la malonil-CoA:ACP transacilasa</p> <p>Familia: probable dominio de anclaje a ACP de la malonil-CoA:ACP transacilasa</p>

Tabla X. Clasificación funcional (Pfam) y estructural (SCOP) de los dominios que integran las enzimas que catalizan la FAS II en Arquea.

	Pfam	SCOP
PF01039	Dominio Carboxil Transferasa	Clase: alfa y beta proteína (a/b) Plegamiento: ClpP/Crotonasa Superfamilia: ClpP/Crotonasa Familia: dominio carboxil transferasa de la carboxilasa-biotina dependiente
PF00682	Tipo HMGL	Clase: alfa y beta proteína (a/b) Plegamiento: barril TIM alfa/beta Superfamilia: aldolasa Familia: tipo HMGL
PF02436	Dominio conservado de la carboxilasa	Clase: toda alfa proteína Plegamiento: dominio RuvA C-terminal Superfamilia: dominio tipo post-HMGL Familia: dominio conservado de la carboxilasa
PF00364	Enzima Biotina-requiere	Clase: toda beta proteína Plegamiento: híbrido barril-sandwich Superfamilia: híbrido de un motivo Familia: proteínas y dominios acarreadoras de biotililipoil
PF00289	Carbamoil-fosfato Sintetasa cadena L, dominio N-terminal	Clase: alfa y beta proteína (a/b) Plegamiento: dominio de pre-anclaje de ATP Superfamilia: dominio de pre-anclaje a ATP Familia: dominio tipo BC N-terminal
PF02786	Carbamoil-fosfato Sintetasa cadena L, dominio de unión a ATP	Clase: alfa y beta proteína (a/b) Plegamiento: anclaje a ATP Superfamilia: dominio similar a glutatión sintetasa ATP-binding Familia: dominio tipo BC-ATP binding
PF02785	Biotina carboxilasa, dominio C-terminal	Clase: Toda beta proteína Plegamiento: híbrido barril-sandwich Superfamilia: híbrido de un solo motivo rudimentario Familia: dominio similar a BC C-terminal
PF02787	Carbamoil-fosfato Sintetasa cadena L, dominio de oligomerización	Clase: alfa y beta proteína (a/b) Plegamiento: anclaje a ATP Superfamilia: dominio similar a glutatión sintetasa ATP-binding Familia: dominio similar a BC-ATP binding
PF01154	Hidroximetilglutaril-CoA sintetasa, N-terminal	Clase: alfa y beta proteína (a/b) Plegamiento: tipo tiolasa Superfamilia: tipo tiolasa Familia: tipo chalcon sintetasa
PF08540	Hidroximetilglutaril-CoA sintetasa, C-terminal	Clase: alfa y beta proteína (a/b) Plegamiento: tipo tiolasa Superfamilia: tipo tiolasa Familia: tipo chalcon sintetasa
PF08541	3-oxoacil-[proteína acarreadora de acilo-(acp)] sintetasa III C-terminal	Clase: alfa y beta proteína (a/b) Plegamiento: tipo tiolasa Superfamilia: tipo tiolasa Familia: tipo chalcon sintetasa
PF00106	Deshidrogenasa de cadena corta	Clase: alfa y beta proteína (a/b) Plegamiento: dominio de plegamiento Rossman de anclaje a NAD(P) binding Superfamilia: dominio de plegamiento Rossman de anclaje a NAD(P) binding Familia: oxidoreductasa tirosina-dependiente
PF00501	Enzima AMP-binding	Clase: proteína de dominios múltiples (alfa y beta) Plegamiento: tipo Acetil-CoA sintetasa Superfamilia: tipo Acetil-CoA sintetasa Familia: tipo Acetil-CoA sintetasa

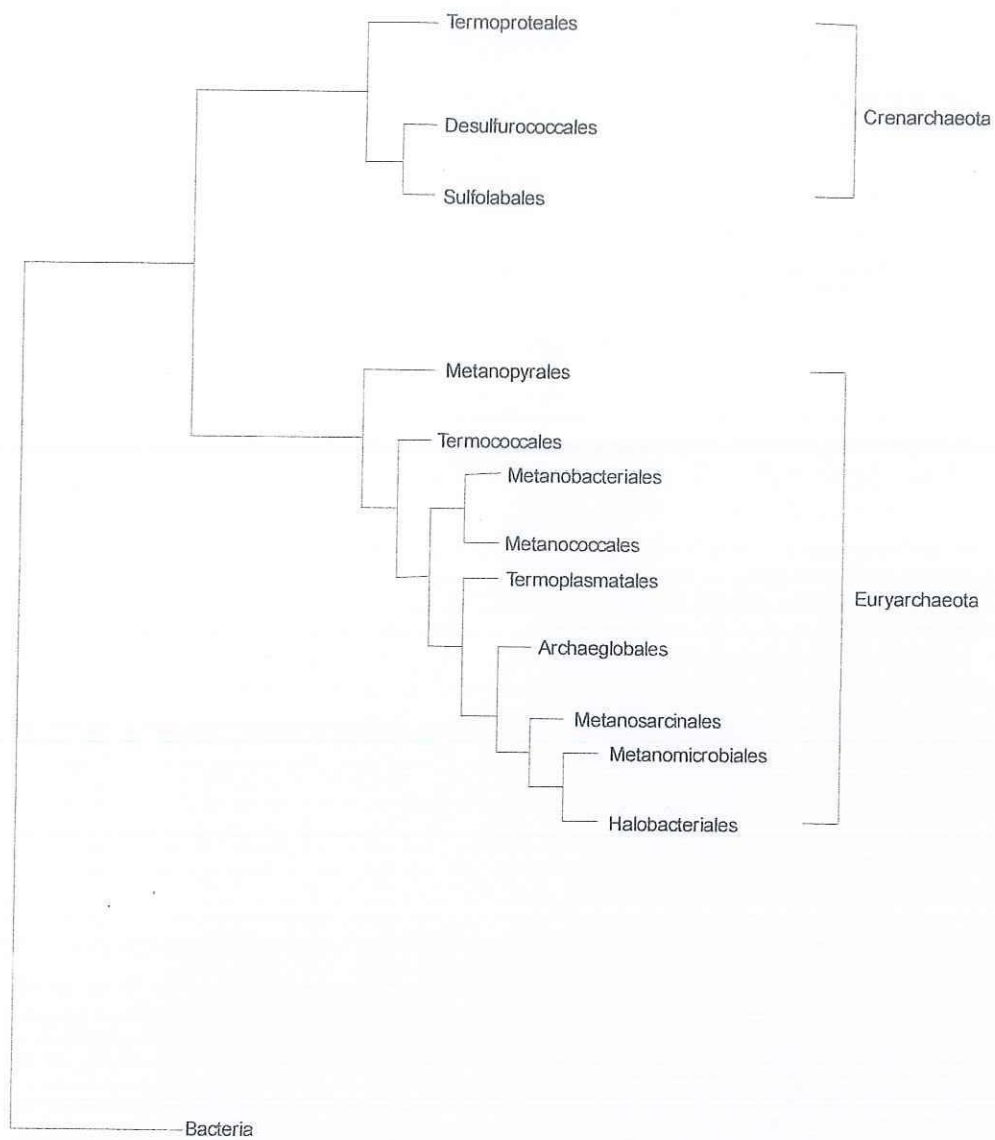


Figura 13. Árbol filogenético de 226 secuencias de rRNA 16S de arqueas y bacterias, generado por máxima similitud. Modificado de Brochiere-Armanet et al. 2008.

Tabla XI. Biología de las arquea cuyo proteoma fue utilizado en este trabajo.

Especie	Estilo de Vida	Temperatura (°C)	Salinidad (NaCl/mM)	pH	Metabolismo	Respiración
<i>Methanococcus jannaschii</i>	vida libre	85	40	6	autótrofo	anaerobia
<i>Methanococcus maripaludis</i>	vida libre	38	15	7	autótrofo	anaerobia
<i>Methanosarcina acetivorans</i>	vida libre	40	20	7	autótrofo	anaerobia
<i>Methanosarcina barkeri</i>	vida libre	37	100	7	autótrofo	anaerobia
<i>Methanococcoides burtonii</i>	vida libre	23	40	7	autótrofo	anaerobia
<i>Methanosaeta thermophila</i>	vida libre	60	0	7	autótrofo	anaerobia
<i>Methanocorpusculum labreanum</i>	vida libre	37	256	7	autótrofo	anaerobia
<i>Methanoculleus marisnigri</i>	vida libre	37	1000	7	autótrofo	anaerobia
<i>Methanobacterium thermoautotrophicum</i>	vida libre	65	15	7	autótrofo	anaerobia
<i>Methanosphaera stadtmanae</i>	vida libre	36	0,01	7	autótrofo	anaerobia
<i>Methanopyrus kandleri</i>	vida libre	98	40	6,5	autótrofo	anaerobia
<i>Archaeoglobus fulgidus</i>	vida libre	83	24	7	autótrofo	anaerobia
<i>Halobacterium sp. NRC-1</i>	vida libre	42	4500	7	heterótrofo	aerobia facultativa
<i>Haloarcula marismortui</i>	vida libre	45	3900	7	heterótrofo	aerobia
<i>Haloquadratum walsbyi</i>	vida libre	45	3300	6,5	heterótrofo	aerobia
<i>Natronomonas pharaonis</i>	vida libre	45	350	8,5	heterótrofo	aerobia
<i>Thermoplasma acidophilum</i>	vida libre	6	0,2	2	heterótrofo	aerobia
<i>Thermoplasma volcanium</i>	vida libre	60	0,01	2	heterótrofo	aerobia facultativa
<i>Picrophilus torridus</i>	vida libre	60	0,01	0,7	heterótrofo	aerobia
<i>Pyrococcus horikoshii</i>	vida libre	98	44	7	heterótrofo	anaerobia
<i>Pyrococcus abyssi</i>	vida libre	96	51	5,8	heterótrofo	anaerobia
<i>Pyrococcus furiosus</i>	vida libre	100	40	7	heterótrofo	anaerobia
<i>Thermococcus kodakaraensis</i>	vida libre	85	51	5,8	heterótrofo	anaerobia
<i>Aeropyrum pernix</i>	vida libre	95	59	7	heterótrofo	aerobia
<i>Staphylothermus marinus</i>	vida libre	92	59	7	heterótrofo	anaerobia
<i>Hyperthermus butylicus</i>	vida libre	100	29	7	heterótrofo	anaerobia
<i>Sulfolobus solfataricus</i>	vida libre	80	0,01	7	heterótrofo	aerobia
<i>Sulfolobus tokodaii</i>	vida libre	80	0,01	2,5	heterótrofo	aerobia
<i>Sulfolobus acidocaldarius</i>	vida libre	80	0,01	2	autótrofo	aerobia
<i>Metallosphaera sedula</i>	vida libre	75	0,01	2	autótrofo facultativo	aerobia

Tabla XI. Biología de las arquea cuyo proteoma fue utilizado en este trabajo.

Especie	Estilo de Vida	Temperatura (°C)	Salinidad (NaCl/mM)	pH	Metabolismo	Respiración
<i>Pyrobaculum aerophilum</i>	vida libre	100	24	7	heterótrofo facultativo	anaerobia
<i>Pyrobaculum islandicum</i>	vida libre	100	0,01	7	heterótrofo facultativo	anaerobia
<i>Pyrobaculum caldifontis</i>	vida libre	95	0	7	heterótrofo	aerobia
<i>Pyrobaculum arsenaticum</i>	vida libre	95	0	7	heterótrofo facultativo	anaerobia
<i>Thermofilum pendens</i>	vida libre	90	0,01	5	heterótrofo	anaerobia
<i>Nanoarchaeum equitans</i>	simbionte obligado	90	40	6	heterótrofo	anaerobia