

Estudios bioquímicos de la ATP citrato liasa de bacterias simbiotes de *Alvinella pompejana* (Anellida: Poliqueta)

Ignacio Pavía

Cátedra de Bioinformática, Área de Biotecnología y Biología Molecular, Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, La Plata, Argentina.

RESUMEN

En el presente trabajo se estudió la ATP citrato liasa de bacterias simbiotes de *Alvinella pompejana*, un poliqueto que habita en chimeneas hidrotermales. Se llevaron a cabo estudios bioinformáticos para determinar sus homólogos, estructura y relaciones filogenéticas. Se determinó que se trata de una proteína formada por dos dominios, uno de unión a citrato y otro con actividad carboxilato-amina ligasa. Posee alto contenido de estructura secundaria, no posee regiones transmembrana ni desordenadas. Se hallaron abundantes homólogos en los diversos taxa de los tres reinos del árbol de la vida. Se demostró que a lo largo de la evolución se conservó el dominio de unión a citrato. El alineamiento múltiple de los homólogos posee información para sustentar varios clados.

PALABRAS CLAVE: ATP citrato liasa, procariontes, bioinformática

INTRODUCCIÓN

La ATP citrato liasa es una enzima clave en el ciclo reductor del ácido tricarbóxico (rTCA) de la fijación de dióxido de carbono durante el crecimiento autótrofo de organismos procariontes. Esta enzima cataliza la escisión del citrato en acetil coenzima A y oxalacetato, en una reacción inversa a la del ciclo de TCA. Una reversión de todo el ciclo genera una molécula de oxalacetato de cuatro moléculas de CO₂.

La mayoría de los quimioautótrofos caracterizados asociados con invertebrados de entornos marinos, incluidos los respiraderos hidrotermales, utilizan la vía Calvin-Benson para la fijación de dióxido de carbono. El ciclo de rTCA se considera una vía alternativa de fijación de carbono y se ha demostrado sólo en unos pocos procariontes.

En el presente trabajo se analizó la ATP citrato liasa de bacterias episimbiotes del gusano poliqueto *Alvinella pompejana*, el cual habita a altas temperaturas en chimeneas hidrotermales de la Dorsal del Pacífico Oriental.

Las caracterizaciones moleculares de la población de simbiotes de *A. pompejana* indican que la mayoría de los simbiotes se agrupan dentro de un solo clado monofilético dentro de la subdivisión épsilon de las Proteobacterias (Bacteria › Proteobacteria › delta/epsilon subdivisions › Epsilonproteobacteria › Campylobacterales › Campylobacteraceae › unclassified Campylobacteraceae)

Los objetivos del presente trabajo fueron analizar la ATP citrato liasa secuencialmente, buscando la presencia de estructuras secundarias, regiones desordenadas, regiones transmembrana y realizar un modelado por homología. Por otro lado, se propuso realizar una búsqueda de secuencias homólogas tanto en taxas cercanos como lejanos filogenéticamente para establecer relaciones y visualizar qué regiones se conservaron a lo largo de la evolución.

MÉTODOS Y RESULTADOS

Predicción de elementos estructurales

Se comenzó con la búsqueda de la proteína en Uniprot (Código: Q6W3M3), se trata de una proteína de 447 aminoácidos.

Para conocer su estructura se realizó una búsqueda de dominios conservados en Pfam (<http://pfam.xfam.org/>), detectándose 2 dominios. El dominio ATP-grasp_2, que va del aminoácido 5 al 228, y se encuentra presente en enzimas que poseen actividad carboxilato-amina ligasa dependiente de ATP, para las cuales es probable que sus mecanismos catalíticos incluyan intermedios de acilfosfato. El otro dominio es el dominio Citrate_bind que comienza en el aminoácido 266 y termina en el 445, dicho dominio posee un sitio de unión a citrato (Fig 1).



Fig 1. Dominios ATP-grasp_2 y Citrate_bind precedidos por Pfam (<http://pfam.xfam.org/>)

A través del sitio web Jpred (<http://www.compbio.dundee.ac.uk/jpred/>) se determinó la presencia de estructuras secundarias. Los resultados obtenidos demuestran la presencia de α -helices y láminas β -plegada en proporciones similares, siendo las primeras más abundantes (Fig 2).

No se detectaron dominios transmembrana ni regiones desordenadas. Para dichos análisis se utilizó Phobius (<https://phobius.sbc.su.se/>) y MobiDM (<https://mobidb.bio.unipd.it/>).

Búsqueda de homólogos

Se buscaron homólogos mediante BLAST en la página del NCBI. Las búsquedas se realizaron diferencialmente para distintos grupos de organismos (tabla 1) y utilizando parámetros por defecto. Se seleccionó al menos una secuencia de cada grupo, seleccionando un total de 30 secuencias. Además, se realizó un PSIBLAST buscando homólogos remotos encontrándose secuencias con hasta un 34% de identidad siendo el valor más bajo registrado.

Alineamiento múltiple

Se realizó un alineamiento múltiple por Clustal Omega para visualizar el grado de conservación de los 2 dominios encontrados en la proteína estudiada. Se utilizaron los parámetros por defecto. Los resultados del alineamiento mostraron una sola región bien conservada que va desde el aminoácido número 279 al 443. Dicha región coincide con el dominio Citrate_bind (Fig 3).

Modelado por homología

Se realizó un modelado por homología utilizando Modeller. Se comenzó con la búsqueda de un template, para lo cual realizamos un PSI-BLAST en el NCBI contra la PDB, siendo seleccionada la proteína Chain A, Citrate lyase, subunit 1 de *Methanotherx soehngeni* (Archaea; Código de acceso 6HXI_A) la cual posee un 49% de identidad con nuestra proteína. Una vez obtenidos 10 modelos, se evaluaron en ProsaII

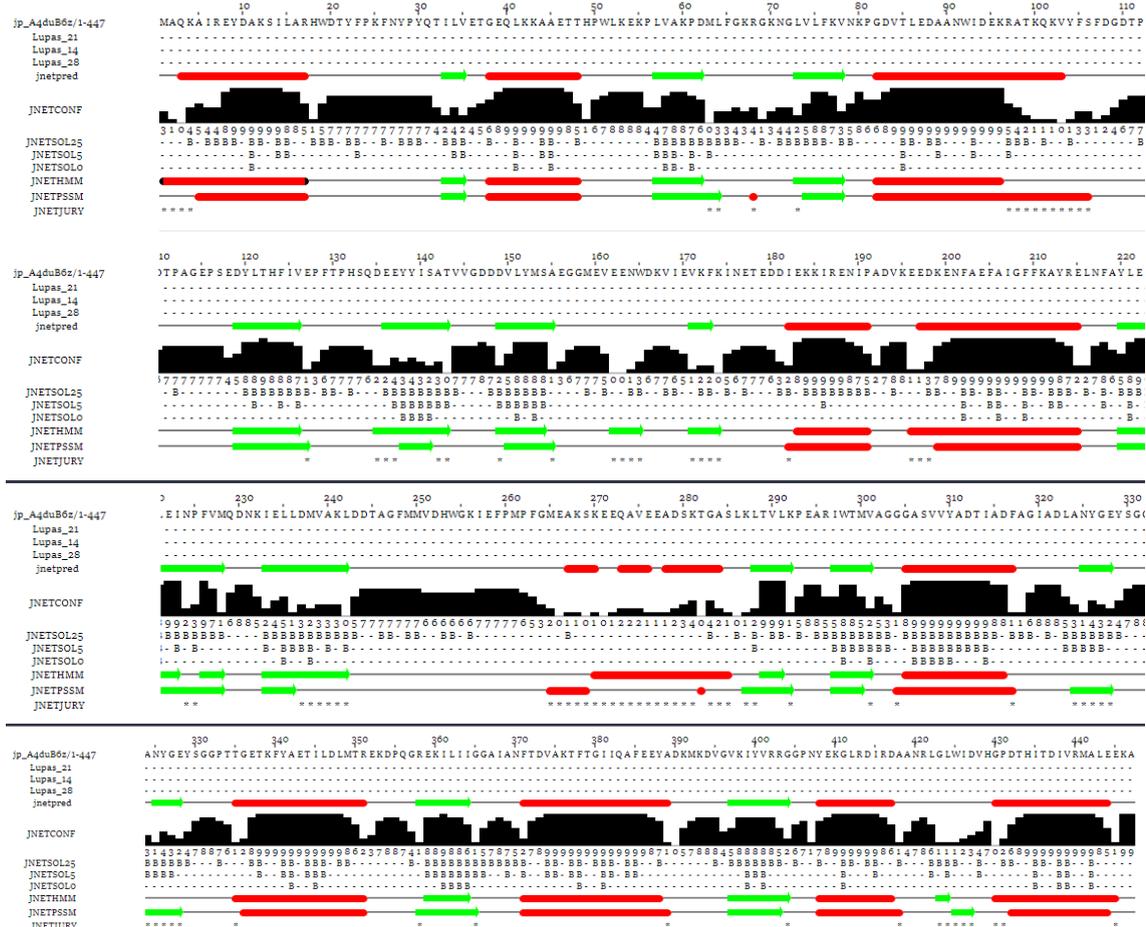


Fig 2. Estimación de estructura secundaria a través de Jpred (<http://www.compbio.dundee.ac.uk/jpred/>). Las líneas rojas y verdes indican estructuras secundarias, α -hélice y lamina β - plegada respectivamente.

Tabla 1. Proteínas de diferentes grupos de organismos seleccionados a partir del BLAST.

Grupo taxonómico	Especie	Código NCBI
Chlorophyta (alga verde)	<i>Micractinium conductrix</i>	PSC76528.1
Annelida	<i>Dimorphilus gyrocoliatius</i>	CAD5123422.1
Annelida	<i>Capitella teleta</i>	ELU08772.1
Amphibia	<i>Xenopus tropicalis</i>	NP_001008028.1
Embryophyta (Plantas terrestres)	<i>Citrus clementina</i>	XP_006438324.1
Embryophyta	<i>Populus alba</i>	XP_034899784.1
Archaea	<i>Candidatus Methanoperedenaceae</i>	CAD7776443.1
Archaea	<i>Thermoplasmatales archaeon</i>	HDM67402.1
Aves	<i>Lepidothrix coronata</i>	XP_017689586.1
Aves	<i>Anhinga rufa</i>	NXT90945.1
Bryozoa	<i>Bugula neritina</i>	KAF6018373.1
Cephalochordata	<i>Branchiostoma belcheri</i>	XP_019635091.1
Arachnida	<i>Tropilaelaps mercedesae</i>	OQR79781.1
Chondrichthyes	<i>Chiloscyllium punctatum</i>	GCC36100.1
Crustacea	<i>Tigriopus kingsejongensis</i>	AUZ82868.1
Bacillariophyta (Diatomea)	<i>Fistulifera solaris</i>	GAX17489.1
Echinodermata	<i>Acanthaster planci</i>	XP_022096812.1
Porifera	<i>Amphimedon queenslandica</i>	XP_003386132.1
Fungi	<i>Podila clonocystis</i>	KAG0018024.1
Fungi	<i>Taphrina deformans</i> PYCC 5710	CCG84442.1
Hemichordata	<i>Saccoglossus kowalevskii</i>	XP_006821438.1
Insecta	<i>Oryctes borbonicus</i>	KRT85288.1
Mammalia	<i>Homo Sapiens</i>	3MWE_A
Mammalia	<i>Mus musculus</i>	BAE36874.1
Mollusca	<i>Biomphalaria glabrata</i>	XP_013071117.1
Nematoda	<i>Caenorhabditis brenneri</i>	EGT55543.1
Testudinata	<i>Pelodiscus sinensis</i>	XP_006131005.1
Teleostei	<i>Perca flavescens</i>	TDG97951.1
Rotifera	<i>Brachionus plicatilis</i>	RMZ99671.1
Tardigrada	<i>Ramazzottius varieornatus</i>	GAV06129.1

(<https://prosa.services.came.sbg.ac.at/prosa.php>) seleccionándose el de menor energía global (Fig 4. A). Al encontrarse nuestro modelo dentro del área sombreada se puede establecer que es un buen modelo, ya que para generar dicha distribución se utiliza una colección curada de proteínas de estructura conocida y con buena resolución. Además, se analizó el gráfico de energías locales mostrando mayormente zonas de baja energía (Fig 4. B). El modelo seleccionado se visualizó mediante PyMol (Fig 5).

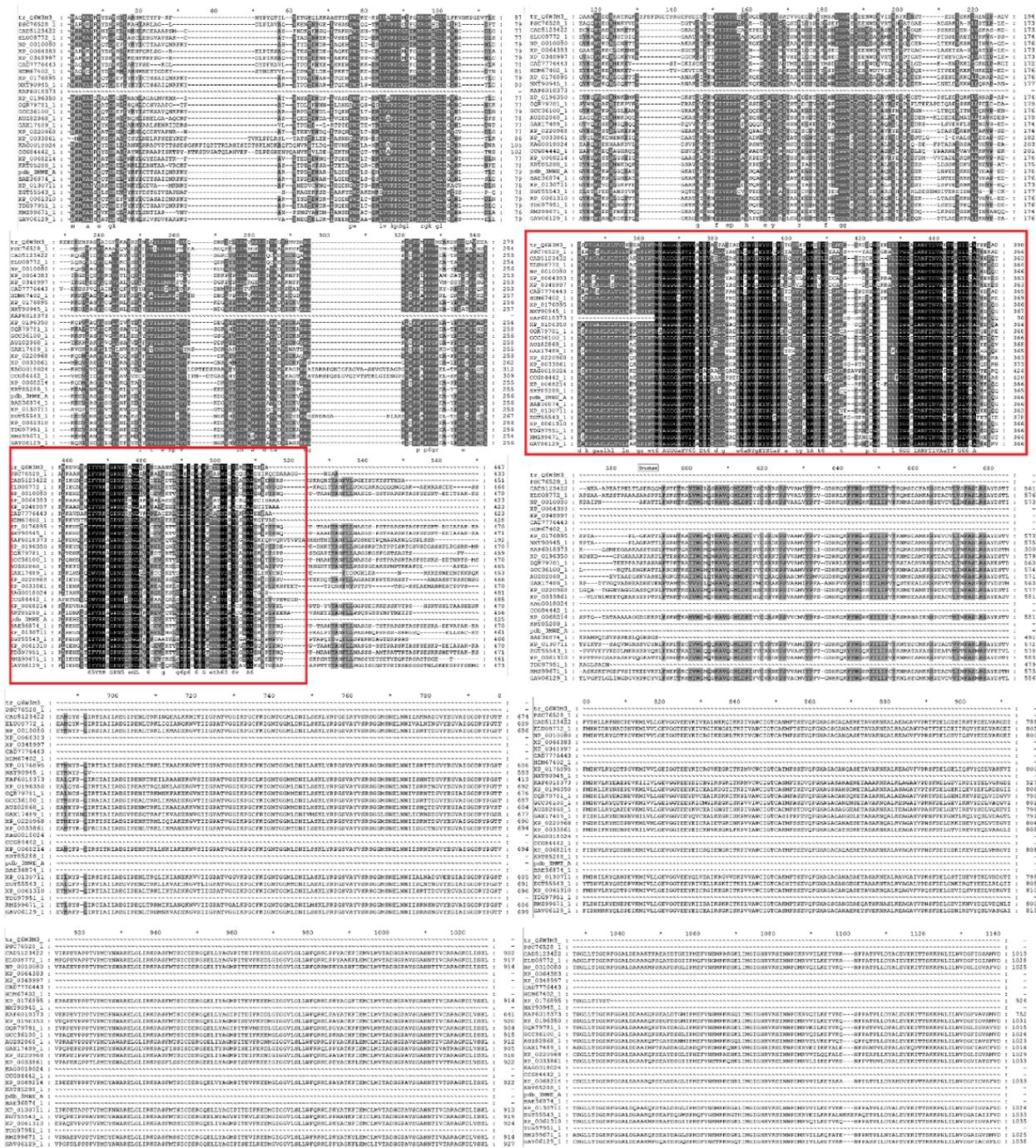


Fig 3. Fragmento del alineamiento múltiple De la ATP citrato liasa de bacterias episimbiontes del gusano poliqueto *Alvinella pompejana* y secuencias de diversos organismos que se muestran en la tabla 1. Los residuos idénticos conservados en todas las secuencias se encuentran sombreados en negro; los conservados en la mayoría de las secuencias están sombreados en gris oscuro; y los cambios conservadores están en gris claro. Las zonas encuadradas en rojo son las más conservadas.

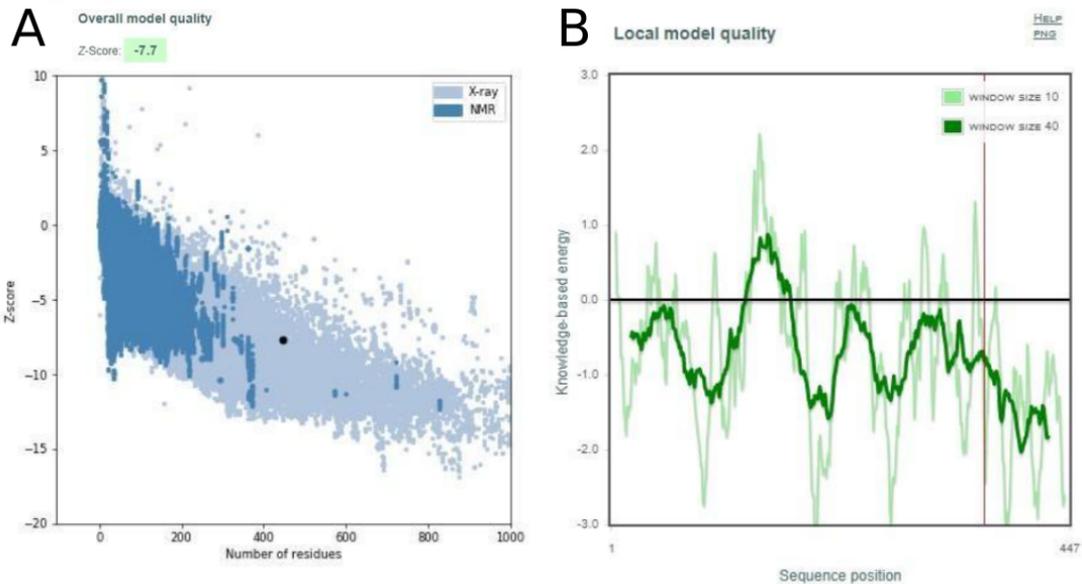


Fig 4. A. Gráfico de distribución de energías en función de la longitud de la estructura, generado con proteínas curadas y de buena resolución. El punto negro representa a nuestro modelo. B. Gráfico de energía local. Muestra la energía de nuestro modelo en función de la posición de los aminoácidos en nuestra secuencia. Se puede observar mayormente zonas de baja energía a excepción de un pico en la posición 150 aproximadamente.

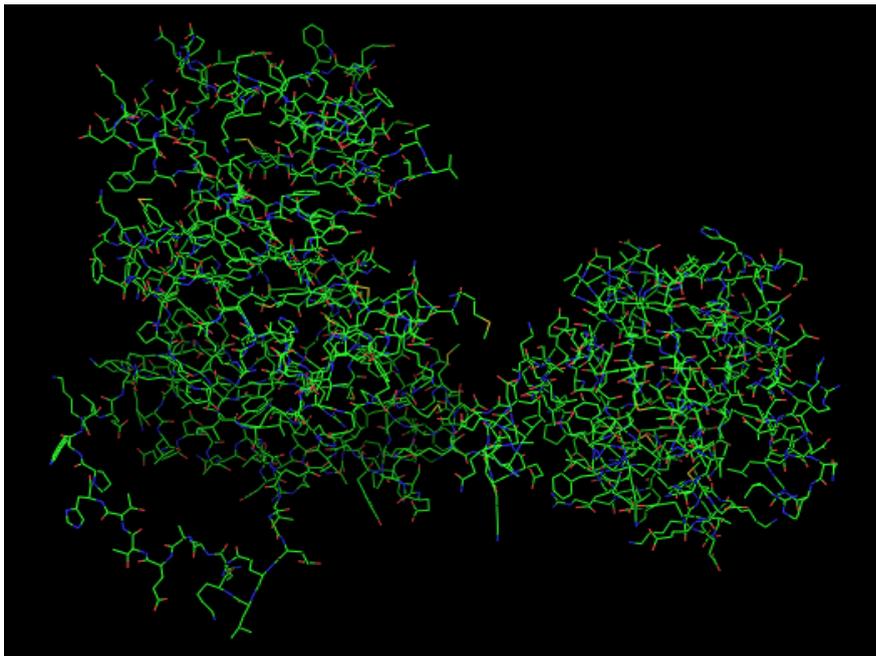


Fig 5. Modelo generado por modelado por homología a través del Modeller.

Inferencia filogenética

El análisis filogenético se llevó a cabo por el método de Maximun Likelihood en el programa PhyML. Para ello se utilizó el alineamiento múltiple realizado anteriormente, un árbol input realizado por Neibor Joining y el modelo evolutivo WAG (Seleccionado como mejor modelo utilizando modeltest en el programa Hyphy). La confianza de las topologías se estimó mediante un bootstraping no paramétrico con 1000 réplicas. Una vez obtenido el árbol se visualizó y editó en el programa FigTree (Fig 6). Se pudo observar que nuestro alineamiento posee mucha información para soportar dos grandes grupos que divergen al principio del árbol, Un primer grupo formado por Bacterias, Arqueas, algas y Plantas terrestres, en el cual están bien soportados todos los nodos entre los diferentes organismos, y un segundo grupo formado por Diatomeas,

Hongos y Animales. En este último grupo hay un subgrupo muy bien soportado el de los Vertebrados (Vertebrata) comenzando con los Chondrictios.

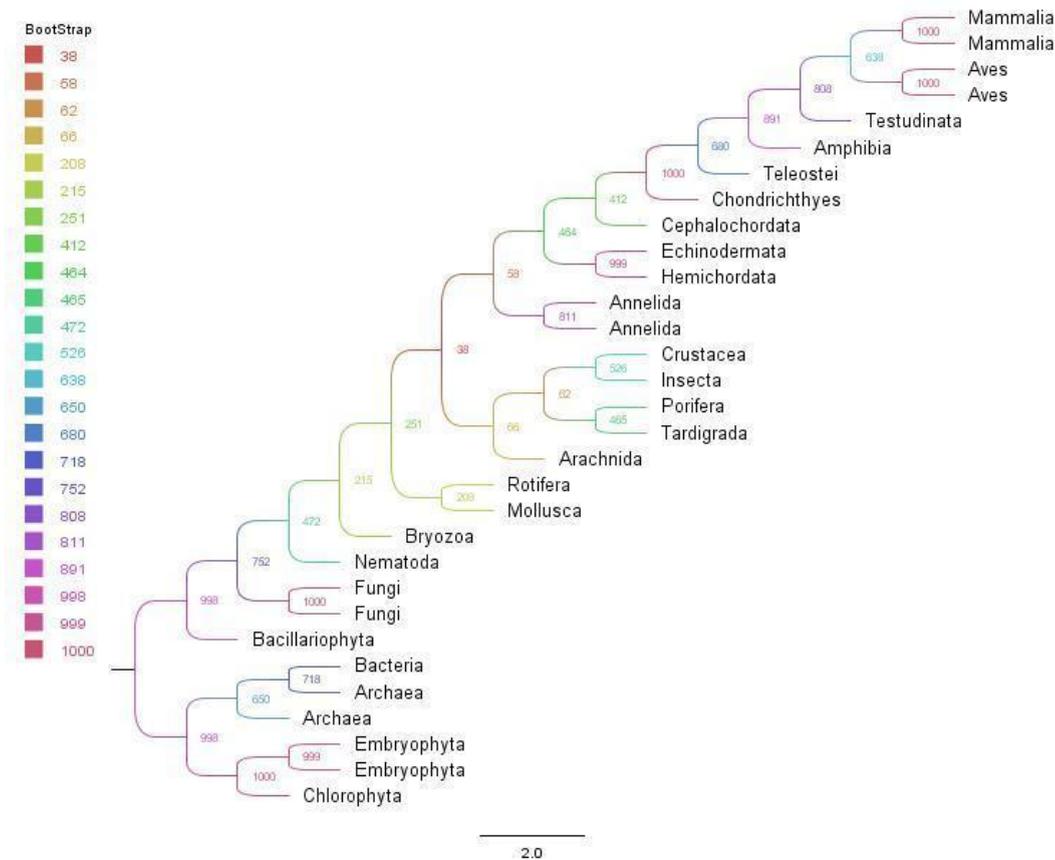


Fig 6. Árbol filogenético de diferentes taxos de organismos (mencionados en la tabla 1).

CONCLUSIONES Y DISCUSIÓN

La proteína Q6W3M3 se trata de una proteína sin regiones transmembrana, que cuenta con alto contenido de estructura secundaria, tanto alfa-hélices como lámina beta-plegada y no posee regiones desordenadas.

Dicha proteína cuenta con gran cantidad de homólogos en diferentes taxones de los 3 dominios (Archaea, Bacteria y Eucharia). Y a lo largo de la evolución pareciese haberse conservado el dominio Citrate_bind.

El análisis filogenético posee información para soportar la relación de varios taxones, incluso en grupos alejados filogenéticamente. En grupos cercanos evolutivamente se puede observar que no ha variado mucho la proteína, ya que, en los grupos duplicados de un mismo taxon siempre se dieron juntos y con un alto valor de bootstrapping.

BIBLIOGRAFÍA

Campbell, B. J., Stein, J. L., Cary, C. S. Evidence of Chemolithoautotrophy in the Bacterial Community Associated with *Alvinella pompejana*, a Hydrothermal Vent Polychaete. 2003. *Applied and environmental microbiology*, vol 69, N° 9, p. 5070–5078.

Di Meo-Savoie, C. A., Luther, G. W., Cary, C. S. Physicochemical characterization of the microhabitat of the epibionts associated with *Alvinella pompejana*, a hydrothermal vent annelid. 2004. *Geochimica et Cosmochimica Acta*, Vol. 68, No. 9, pp. 2055–2066.

Wahlund, T. M., Tabita, F. R., The Reductive Tricarboxylic Acid Cycle of Carbon Dioxide Assimilation: Initial Studies and Purification of ATP-Citrate Lyase from the Green Sulfur Bacterium *Chlorobium tepidum*. 1997. *Journal of Bacteriology*, Vol. 179, No. 15, p. 4859–4867.