RESEARCH ARTICLE

INTERNATIONAL MICROBIOLOGY (2014) 17:91-97
doi:10.2436/20.1501.01.211 ISSN (print): 1139-6709. e-ISSN: 1618-1095
www.im.microbios.org

INTERNATIONAL MICROBIOLOGY
OPEN ACCESS

A phylogenetic approach to the early evolution of autotrophy: the case of the reverse TCA and the reductive acetyl-CoA pathways

Arturo Becerra,[1]* Mario Rivas,[1] Carlos García-Ferris,[2] Antonio Lazcano,[1] Juli Peretó[2]

[1]School of Sciences, National Autonomous University of Mexico, Mexico DF, Mexico.
[2]Department of Biochemistry and Molecular Biology and Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, Valencia, Spain

Received 20 March 2014 · Accepted 4 June 2014

Summary. In recent decades, a number of hypotheses on the autotrophic origin of life have been presented. These proposals invoke the emergence of reaction networks leading from CO or CO_2 to the organic molecules required for life. It has also been suggested that the last (universal) common ancestor (LCA or LUCA) of all extant cell lineages was a chemolitho-autotrophic thermophilic anaerobe. The antiquity of some carbon fixation pathways, the phylogenetic basal distribution of some autotrophic organisms, and the catalytic properties of iron-sulfur minerals have been advanced in support of these ideas. Here we critically examine the phylogenetic distribution and evolution of enzymes that are essential for two of the most ancient autotrophic means of metabolism: the reductive tricarboxylic acid (rTCA) cycle and the reductive acetyl-CoA pathway. Phylogenetic analysis of citryl-CoA synthetase and of citryl-CoA lyase, key enzymatic components of the rTCA cycle, and of CO dehydrogenase/acetyl-CoA synthase, a key enzyme in the reductive acetyl-CoA pathway, revealed that all three enzymes have undergone major lateral transfer events and therefore cannot be used as proof of the LCA's metabolic abilities nor as evidence of an autotrophic origin of life. [**Int Microbiol** 2014; 17(2):91-97]

Keywords: autotrophic pathways · reverse Krebs cycle · Wood–Ljungdahl pathway · origin of life · last common ancestor (LCA, LUCA)

Introduction

Although a heterotrophic origin of life based on the prebiotic synthesis and accumulation of organic compounds is supported by several major lines of evidence [2], other, competing alternatives that advocate an autotrophic emergence of living systems have been suggested as well [20,28,30,34]. An autotrophic origin of life, i.e., the hypothesis that the first organisms fed on CO_2 as sole carbon source, was proposed in the 19th century. However, in recent decades this proposal has been reassessed, based on biochemical analyses and new geochemical data. Both sources invoke the emergence of reaction networks leading from CO or CO_2 to the organic molecules required for life [23]. One of the most well-articulated proposals was that of Wächtershäuser [34], who argued that life began without genetic information and with the appearance of

*Corresponding author: A. Becerra
Facultad de Ciencias
Universidad Nacional Autónoma de México
Ciudad Universitaria
04510 México DF, México
email: abb@ciencias.unam.mx

a prebiotic autocatalytic reductive tricarboxylic acid (rTCA) cycle (also called the reductive citric acid cycle, reverse Krebs cycle, or Arnon cycle), which is assumed to be originally based on the formation of the highly insoluble mineral pyrite in sulfur-rich hydrothermal environments. The fact that this mode of carbon fixation is found in the most deeply divergent bacteria, i.e., Aquificales, has been used as evidence of its primitive character [35]. The geochemical emergence of a primitive version of the reductive acetyl-CoA pathway (or Wood–Ljungdahl pathway) associated with the FeS-rich mineral boundaries of alkaline hydrothermal vents was proposed by Russell and Hall [20,28]. Accordingly, it has also been suggested that the last common ancestor (LCA) of all extant cell lineages was a chemolitho-autotrophic thermophilic anaerobe [19,31,36].

Nonetheless, the hypothesis that the first organisms on Earth could fix carbon is far from proven, since the ultimate nature of the first life form is unknown. Phylogenetic analysis using comparative genomics offers clues to the nature of the LCA and could provide evidence of its ability to fix carbon. Even so, the attributes of the first living entities are unknown and a cladistic approach to the origin of life is not feasible, given that all possible intermediate organisms that may have once existed have long since vanished [2,17]. Moreover, it is not possible to extend an investigation beyond the threshold that corresponds to a period of cellular evolution in which protein biosynthesis was already in operation, i.e., the RNA/protein world [3].

With these caveats, a few clues about the early evolution of autotrophy have been acquired, by searching in modern genomes. Extant beings are able to fix carbon in at least six different ways, namely: the reductive pentose phosphate cycle (Calvin–Benson cycle), the rTCA cycle, the reductive acetyl-CoA pathway, the dicarboxylate/4-hydroxybutyrate cycle, the 3-hydroxypropionate/4-hydroxybutyrate cycle, and the 3-hydroxybutyrate bicycle [13,14]. Here we critically examine the distribution and phylogenetic evolution of enzymes that are essential for the purportedly most ancient autotrophic pathways, i.e., the rTCA cycle and the reductive acetyl-CoA pathway [6, 22]. We analyzed sequences from citryl-CoA synthetase and citryl-CoA lyase, two enzymes participating in the rTCA cycle, and CO dehydrogenase/acetyl-CoA synthase, from the reductive acetyl-CoA pathway. Our results show that the early evolution of autotrophy was not free from horizontal gene transfer. Thus, these pathways cannot be invoked as proof of the LCA's metabolic abilities, nor as evidence of an autotrophic origin of life in volcanic environments rich in transition-metal sulfides.

Materials and methods

Sequences and genomes. The query sequences citryl-CoA synthetase (hth:HTH_1737), citryl-CoA lyase (hth:HTH_0311), acetyl-CoA synthase (mta:Moth_1202), and CO dehydrogenase subunits (mta:Moth_1203) were retrieved, respectively, from the completely sequenced *Hydrogenobacter thermophilus* and *Moorella thermoacetica* genomes available in the Kyoto Encyclopedia of Genes and Genomes (KEGG) [16]. Reports on biochemical information were collected from KEGG, BRENDA [29], MetaCyc [7], and PDB [5] databases.

Search for homologous genes. Searches for homologous proteins were carried out by comparing the query sequences to the genomes database in the KEGG, using BLAST searches of its platform [http://www.genome.jp/tools/blast/]. The BLASTP cutoff value for homologous identification was set at $e \leq 1 \times 10^{-7}$ and an identity ≥34 %.

Phylogenetic analysis. The amino acids sequences of each enzyme were aligned using MUSCLE 3 [12] software with default parameters. Neighbor-joining (500 bootstrap replications, maximum composite likelihood distance estimation, and uniform rates among sites) and maximum-likelihood (model Kimura-2P plus gamma distribution with invariant sites, selected according to the Bayesian information criterion, 500 bootstrap replications) phylogenetic trees were constructed using MEGA5 software [33]. The root of the trees was placed using the midpoint method.

Results and Discussion

The rTCA cycle can be defined as the (oxidative) Krebs cycle running in reverse (Fig. 1A). While the Krebs cycle is a central pathway in many organisms and is used to oxidize acetyl-CoA to CO_2 and to generate intermediates for biosynthesis, the rTCA cycle allows the inverse process, i.e., the biosynthesis of acetyl-CoA from two molecules of CO_2 [14]. Following its discovery in the anaerobic green sulphur photosynthetic bacterium *Chlorobium limicola*, this autotrophic pathway has been detected in strict anaerobic (and microaerobic) bacteria, such as some members of Aquificales [13]. Many of the enzymes involved in the rTCA and Krebs pathways are the same, with the exception of the key enzymes that allow the cycle to run in reverse, namely: 2-oxoglutarate synthase (2-oxoglutarate:ferredoxin oxidoreductase), fumarate reductase, and the citrate-cleaving enzymes [13,14].

The key step of the rTCA cycle is the ATP-dependent cleavage of citrate into acetyl-CoA and oxaloacetate [14]. In different species, this essential reaction is catalyzed by ATP citrate lyase (ACL; EC 2.3.3.8) in one step or by citryl-CoA synthetase (CCS; EC 6.2.1.18) and citryl-CoA lyase (CCL; EC 4.1.3.34) in two steps. Sequence analysis supports the origin of ACL through the fusion of the gene encoding CCS and CCL. Hence, ACL is a derived sequence and our study focused on

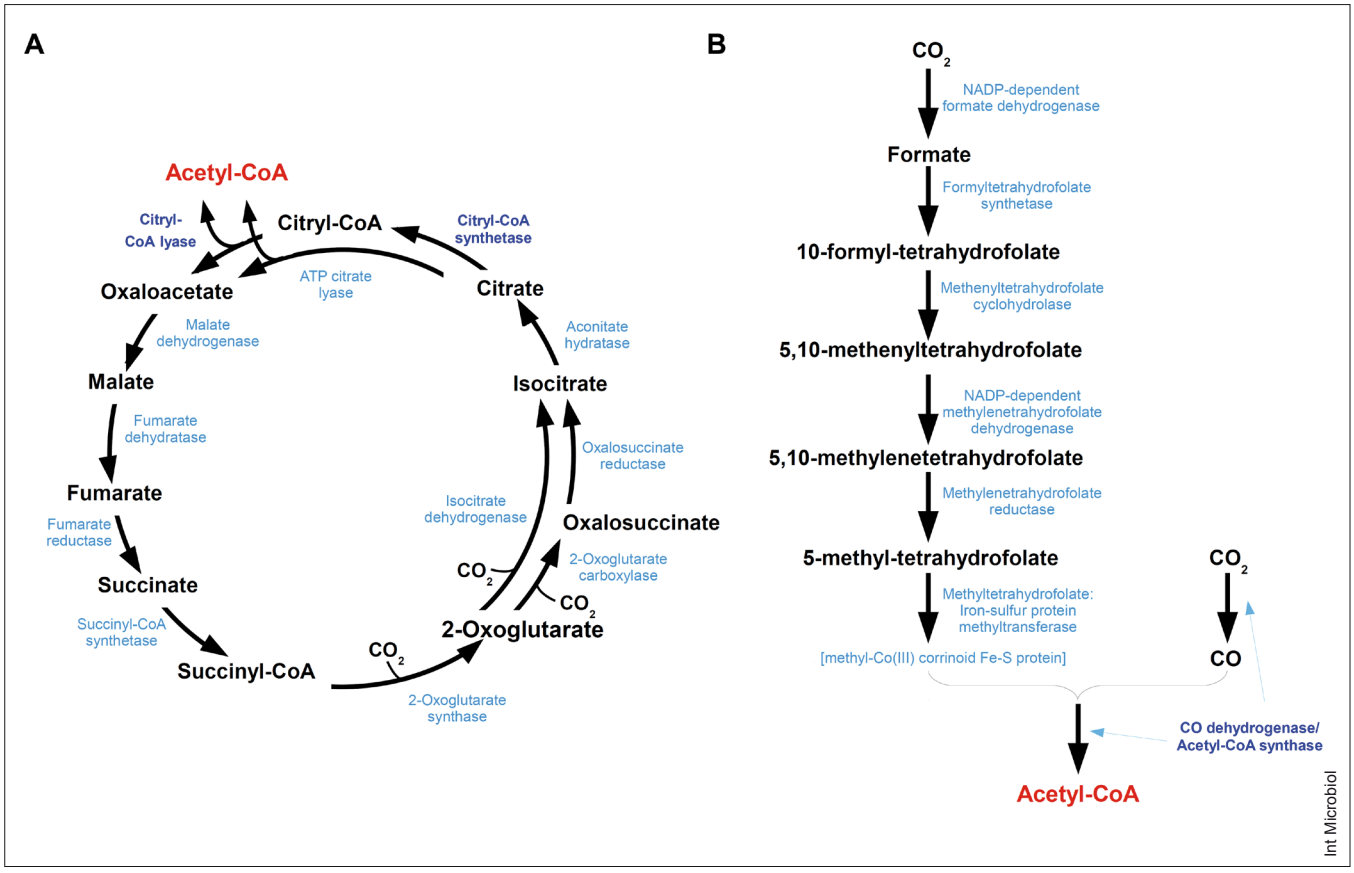

Fig. 1. Reaction schemes of (**A**) the rTCA cycle and (**B**) the Wood–Ljungdahl pathway.

its parental enzymes (CCS/CCL). Furthermore, a common ancestry has been demonstrated between CCS and succinyl-CoA synthetase (SCS) and between CCL and citrate synthase (CS), both of which participate in the Krebs cycle [1].

A feasible evolutionary history of the reaction was proposed by Aoshima [1], who suggested that CCS and CCL activities would be the plesiomorphic state of SCS and CS, respectively. This idea is congruous with the hypothesis of an earliest appearance of a primitive version of the rTCA cycle, as proposed by Wächtershäuser [34]. However, our phylogenetic analysis based on the primary structure of the enzymes does not support Aoshima's hypothesis [1]; instead, the results indicate a horizontal gene transfer between archaea and hyperthermophilic bacteria. As shown in Fig. 2A, the CCS clade from hyperthermophilic bacteria does not appear as a sister group of the SCS bacterial clade. Similar results were obtained with the phylogenetic tree of CCL and CS (Fig. 2B).

The reductive acetyl-CoA pathway was discovered and described in acetogenic bacteria by the laboratories of Wood, Ljungdahl, Thauer, and others [4,13]. This pathway is a linear metabolic route in which two molecules of CO_2 (or CO_2 and CO) are combined directly to form acetyl-CoA (Fig. 1B). The pathway can be divided into two branches: the methyl or "eastern" branch, in which CO_2 is sequentially reduced to a cofactor-bound methyl residue, and the carbonyl or "western" branch, in which another molecule of CO_2 is reduced to an enzyme-bound carbonyl residue [14,25]. The key enzyme in this pathway is CO dehydrogenase/acetyl-CoA synthase (CODH/ACS; EC 1.2.7.4/2.3.1.169), which has a metallic cluster integrated by iron, sulfur, nickel, zinc, and/or copper [10], although the most active form seems to be Ni-Ni-[4Fe-4S] [32]. CODH/ACS is a bifunctional catalyst that reduces CO_2 to carbon monoxide, forming the carbonyl group of acetyl-CoA, and catalyzes the synthesis of acetyl-CoA [10,24]. This enzyme is found both in anaerobic archaea and in chemoautotrophic bacteria [4,13,14].

Based on its catalytic activity and metabolic function, the acetyl-CoA synthase of CODH/ACS is a class I enzyme. Since this class is thought to be older than class II [18], CODH/ACS was the focus of our study. Moreover, it has been

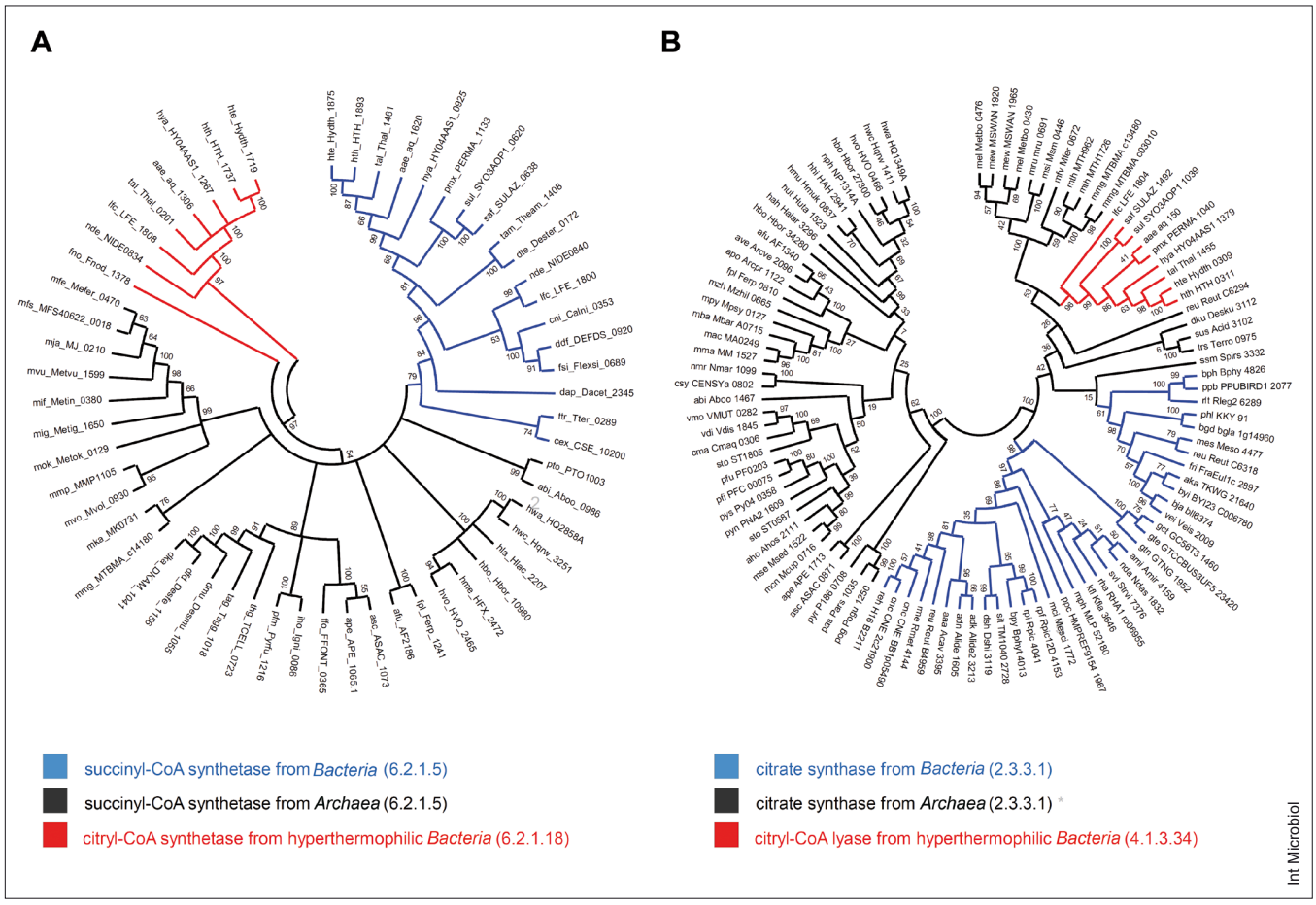

Fig. 2. Maximum-likelihood tree (bootstrap 500 replicates) from the amino acid sequences of (**A**) citryl-CoA synthetase and succinyl-CoA synthetase and (**B**) citryl-CoA lyase and citrate synthase. The colors of the clades describe the function and taxonomy of the OTUs. The root of the trees was placed using the midpoint method.

proposed as one of the oldest enzymes of life, responsible for the ability of early organisms to live in CO_2-rich atmospheres [11].

CODH/ACS is a $\alpha_2\beta_2$ heterotetramer that catalyzes two different reactions using seven metalloclusters. The β-subunit is involved in the CODH activity that generates CO from CO_2, and the α subunit in the ACS activity that synthesizes acetyl-CoA [9]. The subunits were found to differ in their phylogenetic distribution (Fig. 3). Also, there were more β-subunits and the topology of the archaeal subunit was more intricate. These findings can be explained in part by considering CODH as occurring within a large group of bacteria that includes carboxydotrophic bacteria, species of anaerobic acetogenic bacteria, sulfate-reducing bacteria and archaea, phototrophic bacteria, hydrogenogenic bacteria, and methanogenic archaea [8]. Moreover, CODH is present in a larger number of organisms as a monofunctional enzyme than as a bifunctional activity together with the ACS subunit [15]. The topology of the α-subunit tree is defined by two distinct major clades, one containing the majority of archaeal subunits and another that includes all bacteria. However, a well-supported branch of four methanoarchaea appears as the root of the bacterial clade, suggesting horizontal gene transfer from *Archaea* to *Bacteria* (Fig. 3A).

Furthermore, CODH/ACS is an enzymatic complex that requires Ni-S-Fe clusters to transfer electrons, i.e., the 2Ni–[4Fe–4S] cluster (A-cluster) of the α-subunit and the [Ni–4Fe–5S] cluster (C-cluster) of the β-subunit. Both CODH and ACS are highly specialized enzymes that require specific chaperone systems for their assembly [21,27]. The role of the Fe–S clusters in catalysis is to funnel electrons onto an assembly line that opens and closes in order to catalyze the condensation of CO with a methyl moiety and then "pump out" acetyl-CoA [9]. This complex molecular machinery implies that the activity of CODH/ACS is an evolutionary innovation,

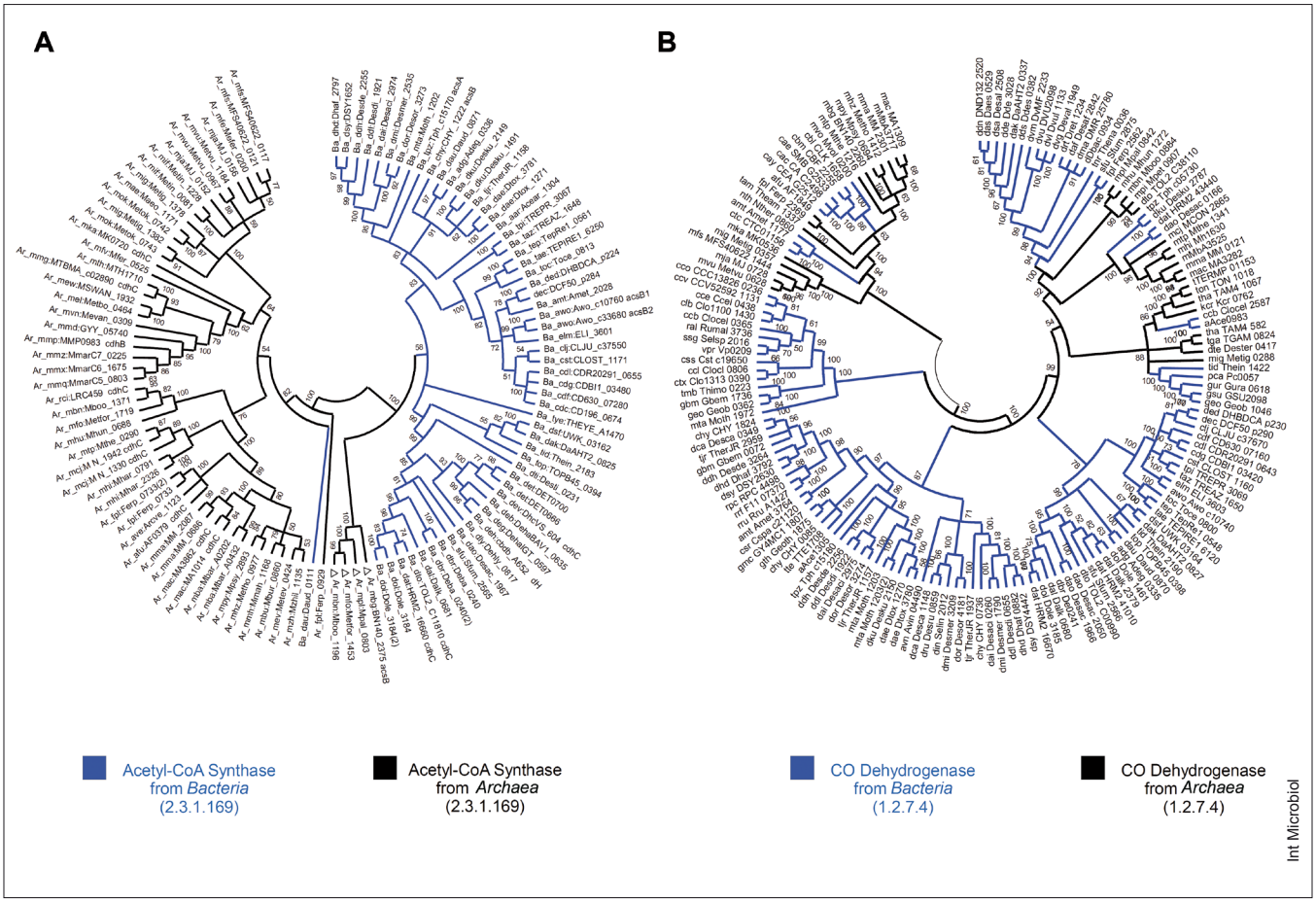

Fig. 3. Maximum-likelihood tree (bootstrap 500 replicates) from the amino acid sequences of CO dehydrogenase/acetyl-CoA synthase: (**A**) the α-subunit ACS (EC 2.3.1.169) and (**B**) the β-subunit of CODH (EC 1.2.7.4; 1.2.99.2). The colors of the clades describe the taxonomy of the OTUs. The root of the trees was placed using the midpoint method.

involving the repurposing of a large number of pre-existing components. It also implies that CODH/ACS was not present at the very early stages of the evolution of life, but was preceded by a long evolutionary history. This possibility is supported by the phylogenetic distribution of the subunits, which suggests that the enzyme appeared after the divergence of the ancestral LCA/ LUCA population.

Our phylogenetic analysis of the citrate-cleaving enzymes participating in the rTCA cycle argue against the hypothesis of Aoshima [1], because these enzymes are not the plesiomorphic version of their homologous counterparts in the oxidative Krebs cycle. Instead, if the key enzymes that allow the Krebs cycle to run in reverse derive from older versions, then it is reasonable to propose that the enzymes required for the oxidative citric acid cycle were present on Earth earlier than those of the rTCA cycle. Phylogenetic analysis of the fundamental citrate-cleaving enzymes, CCS and CCL, do not support the presence of a complete rTCA cycle in the LCA, since both were acquired by horizontal gene transfer, from archaeal homologs participating in Krebs-cycle-like activities to bacteria. Moreover, the autotrophic growth of the strictly anaerobic Thermoproteales can be explained in full by the dicarboxylate/4-hydroxybutyrate cycle rather than by the rTCA cycle, as was initially proposed [4,13,26]. In summary, current evidence strongly suggests that the rTCA cycle is an idiosyncratic pathway of the bacterial domain that evolved after the LCA.

Similar conclusions can be derived from the phylogenetic analysis of the subunits of the bifunctional enzyme CODH/ACS, an essential component of the Wood–Ljungdahl pathway. Despite its wide distribution and diverse activities, there is no evidence that the β-subunit of CODH was present in the LCA. The same holds true for the α-subunit (ACS), which was clearly acquired from methanoarchaea. The bifunctional

CO dehydrogenase/acetyl-CoA synthase must have emerged after its components and could have not been present during the LCA's epoch. Since both subunits are part of the Cdh-A-E complex of methanogens, then the CODH/ACS enzymes of the extant reductive acetyl-CoA pathway must have derived from those utilized by methanogens.

It is of course possible that the LCA was an autotrophic organism. However, this hypothesis is not supported by our results on the key enzymes from the oldest CO_2 fixation pathways. More studies, including a better phylogenetic representation of the most basal lineages in both prokaryotic domains, are needed to reveal the metabolic abilities of the LCA, and especially to determine whether it was a heterotrophic or an autotrophic organism.

As noted herein and discussed elsewhere [3,17], the origin of life is not amenable to phylogenetic analysis, since molecular cladistics and comparative genomics cannot be extended beyond a threshold that corresponds to a period of cellular evolution in which protein biosynthesis was already in operation. Accordingly, the genome distribution of enzymes from the rTCA cycle and the reductive acetyl-CoA pathway or even the presence of these metabolic routes is not evidence for an autotrophic origin of life.

Acknowledgements. Financial support of CONACYT (100199) to AB and (50520-Q) to AL is gratefully acknowledged. This work was completed during a sabbatical leave of absence of AB, with support from the DGAPA-UNAM and the University of Valencia, where he enjoyed the hospitality of Amparo Latorre at the Institut Cavanilles (Valencia, Spain). CGF and JP acknowledge the financial support of MINECO (Grant ref. BFU2012-39816-C02-01). We acknowledge the technical support of Sara Islas-Graciano and Ricardo Hernández-Morales.

Competing interests. None declared.

References

1. Aoshima M (2007) Novel enzyme reactions related to the tricarboxylic acid cycle: Phylogenetic/functional implications and biotechnological applications. Appl Microbiol Biot 75:249-255
2. Bada J, Lazcano A (2009) The origin of life. In: Ruse M, Travis J (eds) Evolution: The first four billion years. Belknap/Harvard Univ. Press, Cambridge, MA, USA, pp 49-79
3. Becerra A, Delaye L, Islas S, Lazcano A (2007) The very early stages of biological evolution and the nature of the last common ancestor of the three major cell domains. Annu Rev Ecol Evol Sci 38:361-379
4. Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hügler M, Alber BE, Fuchs G (2010) Autotrophic carbon fixation in archaea. Nat Rev Microbiol 8:447-460
5. Berman HM, Bhat TN, Bourne PE, Feng Z, Gilliland G, Weissig H, Westbrook J (2000) The protein data bank and the challenge of structural genomics. Nat Struct Biol 7:957-959
6. Braakman R, Smith E (2012) The emergence and early evolution of biological carbon-fixation. PLoS Comput Biol 8:e1002455
7. Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 40:D742-D753
8. Dobbek H, Svetlitchnyi V, Gremer L, Huber R, Meyer O (2001) Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni–4Fe–5S] cluster. Science 293:1281-1285
9. Doukov TI, Blasiak LC, Seravalli J, Ragsdale SW, Drennan CL (2008) Xenon in and at the end of the tunnel of bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. Biochemistry 47:3474-3483
10. Doukov TI, Iverson TM, Seravalli J, Ragsdale SW, Drennan CL (2002) A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. Science 298:567-572
11. Drennan CL, Doukov TI, Ragsdale SW (2004) The metalloclusters of carbon monoxide dehydrogenase/acetyl-CoA synthase: a story in pictures. J Biol Inorg Chem 9:511-515
12. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797
13. Fuchs G (2011) Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? Annu Rev Microbiol 65:631-658
14. Hügler M, Sievert SM (2011) Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. Annu Rev Mar Sci 3:261-289
15. Jeoung JH, Dobbek H (2012) *n*-Butyl isocyanide oxidation at the [$NiFe_4S_4OH_x$] cluster of CO dehydrogenase. J Biol Inorg Chem 17: 167-173
16. Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27-30
17. Lazcano A (2010) Which way to life? Origins Life Evol Biol 40:161-167
18. Lindahl PA, Chang B (2001) The evolution of acetyl-CoA synthase. Origins Life Evol Biol 31:403-434
19. Martin W, Baross J, Kelley D, Russell MJ (2008) Hydrothermal vents and the origin of life. Nat Rev Microbiol 6:805-814
20. Martin W, Russell MJ (2003) On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. Philos Trans R Soc Lond B Biol Sci 358:59-83
21. Meyer J (2008) Iron-sulfur protein folds, iron-sulfur chemistry, and evolution. J Biol Inorg Chem 13:157-170
22. Peretó JG, Velasco AM, Becerra A, Lazcano A (1999) Comparative biochemistry of CO_2 fixation and the evolution of autotrophy. Int Microbiol 2:3-10
23. Peretó J (2012) Out of fuzzy chemistry: from prebiotic chemistry to metabolic networks. Chem Soc Rev 41:5394-5403
24. Ragsdale SW, Clark JE, Ljungdahl LG, Lundie LL, Drake HL (1983) Properties of purified carbon monoxide dehydrogenase from *Clostridium thermoaceticum*, a nickel, iron-sulfur protein. J Biol Chem 258:2364-2369
25. Ragsdale SW (2008) Enzymology of the Wood–Ljungdahl pathway of acetogenesis. Ann NY Acad Sci 1125:129-136
26. Ramos-Vera WH, Berg IA, Fuchs G (2009) Autotrophic carbon dioxide assimilation in *Thermoproteales* revisited. J Bacteriol 191:4286-4297
27. Rees DC (2002) Great metalloclusters in enzymology. Annu Rev Biochem 71:221-246
28. Russell MJ, Hall AJ (1997) The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. J Geol Soc 154:377-402

29. Schomburg I, Chang A, Hofmann O, Ebeling C, Ehrentreich F, Schomburg D (2002) BRENDA: a resource for enzyme data and metabolic information. Trends Biochem Sci 27:54-56
30. Smith E, Morowitz HJ (2004) Universality in intermediary metabolism. Proc Natl Acad Sci USA 101:13168-13173
31. Stüeken EE, Anderson RE, Bowman JS, Brazelton WJ, Colangelo-Lillis J, Goldman AD, Som SM, Baross JA (2013) Did life originate from a global chemical reactor? Geobiology 11:101-126
32. Svetlitchnyi V, Dobbek H, Meyer-Klaucke W, Meins T, Thiele B, Romer P, Huber R, Meyer O (2004) A functional Ni–Ni–[4Fe–4S] cluster in the monomeric acetyl-CoA synthase from *Carboxydothermus hydrogenoformans*. Proc Natl Acad Sci USA 101:446-451
33. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739
34. Wächtershäuser G (1988) Before enzymes and templates: theory of surface metabolism. Microbiol Rev 52:452-484
35. Wächtershäuser G (1990) Evolution of the first metabolic cycles. Proc Natl Acad Sci USA 87:200-204
36. Wächtershäuser G (2007) On the chemistry and evolution of the pioneer organism. Chem Biodiver 4:584-602