Molecular Phylogenetics and Evolution 91 (2015) 12–16

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Short Communication

Gene fusion, fission, lateral transfer, and loss: Not-so-rare events in the evolution of eukaryotic ATP citrate lyase ☆

Ryan M.R. Gawryluk [1,2], Laura Eme [1], Andrew J. Roger *

Centre for Comparative Genomics and Evolutionary Bioinformatics, Department of Biochemistry and Molecular Biology, Dalhousie University, 5850 College Street, Halifax, Nova Scotia B3H 4R2, Canada

ARTICLE INFO

Article history:
Received 17 December 2014
Revised 7 May 2015
Accepted 11 May 2015
Available online 27 May 2015

Keywords:
ATP citrate lyase
Gene fusion
Gene fission
Lateral gene transfer
Eukaryotic evolution
Protist

ABSTRACT

ATP citrate lyase (ACL) is an enzyme critical to the generation of cytosolic acetyl-CoA in eukaryotes. In most studied organisms, ACL activity is conferred in combination by two proteins, ACLA and ACLB (dsACL); however, animals encode a single-subunit ACL (ssACL) – the result of a gene fusion event. Through phylogenetic analyses, we investigated the evolution of ACL in a broad range of eukaryotes, including numerous microbes (protists). We show that the fused form is not restricted to animals, and is instead widely distributed among eukaryotes. Furthermore, *ssACL* and *dsACL* are patchily distributed and appear to be mutually exclusive; both types arose early in eukaryotic evolution. Finally, we present several compelling hypotheses of lateral gene transfer and gene loss, along with the secondary gene fission of *ssACL* in Ascomycota. Collectively, our in-depth analyses suggest that a complex suite of evolutionary events, usually considered rare, has shaped the evolution of ACL in eukaryotes.

1. Introduction

Acetyl coenzyme A (acetyl-CoA) is a high-energy metabolite that is a product of carbohydrate, amino acid, and lipid catabolism, and the precursor of numerous anabolic pathways (Oliver et al., 2009). Given its role at the 'hub' of cellular metabolism, a thorough understanding of acetyl-CoA biosynthesis is critical; this is especially true of eukaryotes, as acetyl-CoA is membrane-impermeable, and distinct biosynthetic mechanisms are therefore required in the various subcellular compartments, including mitochondria, chloroplasts, peroxisomes, and the cytosol (Oliver et al., 2009).

In animals (Elshourbagy et al., 1990, 1992), land plants (Fatland et al., 2002), a glaucophyte alga (Ma et al., 2001), and filamentous fungi (Hynes and Murray, 2010; Son et al., 2011), the major cytosolic source of acetyl-CoA is ATP-citrate lyase (ACL; EC 2.3.3.8), an enzyme that catalyzes the ATP-dependent cleavage of citrate into oxaloacetate and acetyl-CoA. Typically, ACL's substrate is mitochondrion-derived citrate; this enzyme therefore plays a role in the 'citrate shuttle' that effects the net transfer of acetyl-CoA equivalents to the cytosol for fatty acid biosynthesis. The acetyl-CoA generated by ACL is a key substrate of myriad other downstream anabolic processes in eukaryotes, including the biosynthesis of sterols, waxes, isoprenoids, and flavonoids (Oliver et al., 2009), and nuclear histone acetylation (Wellen et al., 2009).

ACL is also encoded in the genomes of prokaryotes; however, it is sparsely distributed, and is found only in a few species belonging to ε-proteobacteria, Aquificae, Chlorobi, and Euryarchaeota, many of which are thermophiles living near deep-sea vents (Campbell and Cary, 2004). Although the chemical reaction catalyzed by ACL is the same as in eukaryotes, the physiological context is different: in prokaryotes, ACL is a component of the reverse TCA cycle, a reductive, carbon-fixing pathway that serves as an alternative to the Calvin–Benson–Bassham reductive pentose phosphate cycle (Buchanan and Arnon, 1990). Thus, there has been a functional modification in eukaryotes, with ACL shifting from a role in permitting autotrophic growth, to supplying a key intermediate in various eukaryotic anabolic processes (Fatland et al., 2002).

In bacteria (Hügler et al., 2007; Kanao et al., 2001), a glaucophyte alga (Ma et al., 2001), green algae/land plants (Fatland et al., 2002), and filamentous fungi (Nowrousian et al., 2000), ACL enzyme activity requires ACLA, and ACLB (referred to here as dual-subunit ACL, or dsACL) (Kanao et al., 2001). ACLA is

Abbreviations: ACL, ATP citrate lyase; ssACL, single-subunit ACL; dsACL, dual-subunit ACL; ACLA, dsACL protein homologous to the N-terminal portion of ssACL; ACLB, dsACL protein homologous to the C-terminal portion of ssACL; CS, citrate synthase; SCS, succinyl-CoA synthetase; ML, Maximum Likelihood; LGT, lateral gene transfer; EF 1-alpha, elongation factor 1-alpha; EF-L, elongation factor-like.

☆ This paper was edited by the Associate Editor Austin Hughes.

* Corresponding author.
E-mail address: andrew.roger@dal.ca (A.J. Roger).
[1] These authors contributed equally to this work.
[2] Current address: Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada.

http://dx.doi.org/10.1016/j.ympev.2015.05.010
1055-7903/

homologous to the β subunit of succinyl-CoA synthetase (SCS), while ACLB is homologous to the α subunit of SCS, fused to a small portion homologous to citrate synthase (CS). An evolutionary model for the origin of ACL from the aforementioned TCA cycle enzymes *via* gene duplication, fusion, and divergence has been suggested (Fatland et al., 2002). In contrast, animal ACL is a fusion protein (referred to here as single-subunit ACL, or ssACL) (Elshourbagy et al., 1990, 1992); the N-terminal portion is homologous to ACLA, and the C-terminal portion to ACLB. It was suggested that ssACL represents a molecular synapomorphy of animals (Fatland et al., 2002), although ssACL homologs were recently identified in some non-ascomycete fungi (Hynes and Murray, 2010).

Here, we have undertaken a comprehensive phylogenetic analysis of ACL across eukaryotes. We demonstrate that ssACL and dsACL constitute ancient, distinct monophyletic lineages, that *dsACL* and *ssACL* have been laterally transferred and lost numerous times, and, contrary to previous analyses, that *ssACL* is likely the product of a gene fusion event that occurred very early in eukaryotic evolution.

2. Methods

2.1. Taxon sampling and multiple alignment

ssACL, ACLA and ACLB homologs were identified *via* BLAST queries of public databases (NCBI, Sanger Institute and MMETSP (Keeling et al., 2014)). In lineages with large numbers of ACL homologs (e.g., fungi, and animals) a subset of phylogenetically representative sequences was manually selected. Alignments were generated with MAFFT L-INS-i v7 (Katoh and Standley, 2013). Individual ACLA/ssACL and ACLB/ssACL alignments were concatenated, and trimmed automatically with BMGE 1.0, using the BLOSUM50 similarity matrix (Criscuolo and Gribaldo, 2010).

2.2. Phylogenetics

Maximum Likelihood (ML) phylogenies were estimated with RAxML version 8.0.19 (Stamatakis, 2014), under the PROTGAMMALGF model. Bootstrap support values estimated from 1000 replicates were mapped onto the estimated ML tree (obtained by 100 heuristic searches).

Bayesian analyses were carried out with PhyloBayes version 3.3f (Lartillot et al., 2009) by running two chains under the catfix C20 + Poisson model, until convergence (maxdiff $\ll$ 0.1) after discarding 3000 burn-in trees. Posterior probabilities were mapped onto the ML tree.

3. Results and discussion

3.1. ACL gene structure and taxonomic distribution

We confirm here the presence of two separate genes for ACLA and ACLB proteins in all prokaryotes encoding ACL, as well as all dsACL-encoding members of the eukaryotic supergroup Archaeplastida, and ascomycete fungi (Fig. 1). Additionally, we identified *ACLA* and *ACLB* in the jakobid *Andalucia godoyi* (Excavata), several members of Amoebozoa (e.g., Mycetozoa and certain Discosea), and the apusomonad *Thecamonas trahens* – a putative sister of opisthokonts (Brown et al., 2013).

Surprisingly, in addition to the extensively characterized ssACL of animals, and the ssACL of some fungi, we identified ssACL in all ACL-encoding members of the vast SAR (Stramenopiles + Alveolata + Rhizaria) supergroup, as well as other opisthokonts, haptophytes, the putatively basal cryptophyte *Palpitomonas bilix* (Yabuki et al., 2014), and numerous amoebozoans (e.g., Tubulinea and a few Discosea; Fig. 1). It is therefore apparent that *ssACL* is not an animal-specific trait (Fatland et al., 2002).

Hereafter, we reconstruct the evolutionary history of *ACL*, updating and extending upon earlier analyses. We clarify the relationships between *dsACL* and *ssACL* by carrying out in-depth phylogenetic analyses using a broad taxonomic sample of eukaryotes.

3.2. Eukaryotic ssACL and dsACL resolve into two monophyletic groups

We confirmed that all the subunits/domains constituting ACL apparently share the same evolutionary histories by performing separate phylogenetic analyses of homologs of: (1) ACLA and SCS β; (2) ACLB and SCS α; and (3) ACLB and CS (not shown). In agreement with previous analyses (Fatland et al., 2002; Hügler et al., 2007), our Maximum Likelihood (ML) and Bayesian phylogenetic reconstructions suggest that ACL homologs from green sulfur bacteria (i.e., Chlorobi) are basal to all others, and that *bona fide* ACL likely originated in this clade. Although the long-branch leading to this group does not preclude a long-branch attraction artifact, it is likely that ACL is ancestrally bipartite, and that ssACL, as observed in numerous eukaryotes, is a derived feature.

To improve phylogenetic resolution, we concatenated ACLA and ACLB datasets, using Chlorobi as outgroup (Fig. 2). We found that each of ssACL and eukaryotic dsACL constitute highly supported monophyletic groups, with the exception of ACL from ascomycetes (discussed in Section 3.4). It is thus reasonable to conclude that eukaryotes acquired *ACLA* and *ACLB* from a prokaryotic source – although the backbone is not sufficiently resolved to infer the exact nature of the donor. Similarly, the low backbone resolution prevents us from favoring an origin of *ssACL* through gene duplication within the eukaryotic branch over a second LGT from prokaryotes to eukaryotes before LECA. In either case, a single ancient gene fusion event took place early in the eukaryotic line and led to the establishment of *ssACL*.

3.3. ssACL and dsACL likely co-existed in LECA

Although ACL can be found in virtually all major eukaryotic groups, the patchy distribution of both *ssACL* and *dsACL* makes it difficult to ascertain precisely the timing of their emergence in eukaryotes (Fig. 1). We suggest that both isoforms of *ACL* likely arose early in eukaryotes, possibly prior to the last eukaryotic common ancestor (LECA).

The internal relationships within main eukaryotic groups in the ssACL clade are congruent with organismal phylogeny, and are usually strongly supported (Fig. 2). In combination with its broad distribution, this suggests that ssACL was likely present in the ancestor of SAR, Opisthokonta, and Haptophyta (i.e., LECA), and was subsequently inherited vertically (with a few exceptions discussed in Section 3.4). An alternative explanation is that *ssACL* was dispersed by horizontal transfer between eukaryotic lineages. However, this would have had to happen shortly after LECA, but before the diversification within eukaryotic supergroups (i.e., within ~300 million years (Eme et al., 2014)). Regardless, the fusion of *ACLA* and *ACLB* likely occurred early in the eukaryotic lineage.

We have identified *dsACL* in Archaeplastida, in some distantly related amoebozoans (e.g., Mycetozoa and Discosea), in the excavates *Andalucia godoyi* and *Bodo saltans*, and in the apusomonad, *Thecamonas trahens* (Fig. 1). Thus, *dsACL* likely appeared in eukaryotes prior to the ancestors of Amoebozoa and Archaeplastida, and therefore, before LECA. Alternatively, *ssACL* is the ancestral isoform of ACL in eukaryotes, and its distribution is the result of *ssACL* fission, for instance in the ancestor of Archaeplastida, with subsequent lateral gene transfers (LGTs) to other eukaryotes.

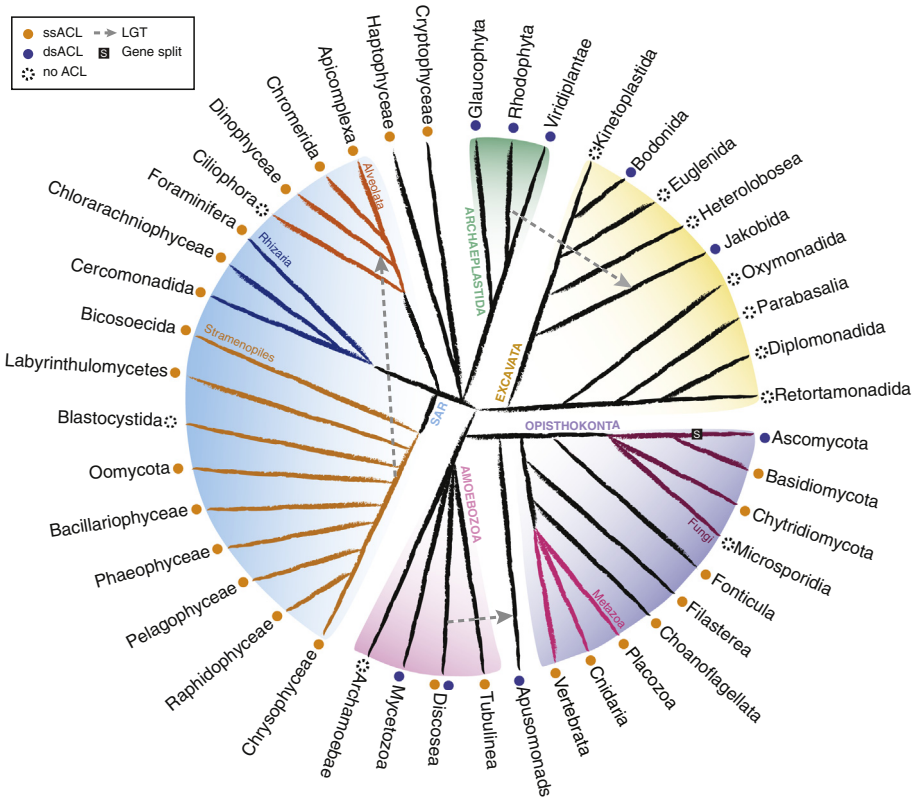

Fig. 1. Distribution of ACL isoforms among eukaryotes. Yellow circles show the presence of ssACL, blue circles show the presence of dsACL, and empty dotted circles show the absence of ACL homologs. Arrows represent lateral gene transfer (LGT) events. The black boxed "S" indicates a gene splitting event. The tree topology was adapted from (Burki, 2014). The phylogenetic depth displayed for each eukaryotic group was chosen to represent ACL evolution optimally.

Amoebozoans encode either, but not both, of the ACL variants, and the distribution of these forms conflicts with amoebozoan phylogeny (Cavalier-Smith et al., 2014; Lahr et al., 2011); fairly closely related organisms do not share the same character state (e.g., *Pessonella* sp. and *Vannella robusta*; family Vannellidae), whereas more distant groups do. It is difficult to envision how both versions of *ACL* would have persisted until so recently, with one subsequently disappearing altogether in multiple lineages independently. This indicates that vertical inheritance with differential loss is unlikely to provide a comprehensive explanation for the distribution of *ACL*; LGT events are presumably also responsible for the observed pattern.

3.4. ACL evolution is littered with loss, gene transfer, and secondary fission events

Although ancient LGT events are difficult to assess, we have identified several more recent LGT candidates involving *ACL*: (1) from stramenopiles to dinoflagellates; (2) from Discosea (Amoebozoa) to the apusozoan *Thecamonas*; and (3) from red algae to the jakobid *Andalucia godoyi* (Figs. 1 and 2). The apparent concomitant transfer of two genes is surprising, as gene order is commonly shuffled rapidly in eukaryotic genomes. Interestingly, however, *ACLA* and *ACLB* are adjacent in the genome of the rhodophyte *Galdieria sulphuraria*, as well as in *Dictyostelium discoideum*. Unfortunately, it was not possible to investigate gene organization in Discosea, as only transcriptome data are generally available (Keeling et al., 2014). Nevertheless, that *ACLA* and *ACLB* are neighboring genes in some eukaryotes indicates that the transfer of both genes through a single LGT event is feasible.

In contrast, the ascomycete dsACL branches with its ssACL fungal counterparts (Fig. 2), clearly demonstrating reversion to the ancestral state *via* gene fission, rather than LGT from another *dsACL*-carrying lineage. We did not identify *ssACL* in any other ascomycete, indicating that *ssACL* split early in the evolution of Ascomycota and that opisthokonts ancestrally possessed *ssACL*, in contradiction to prior hypotheses that proposed a fusion of *ACLA* and *ACLB* in animals (Fatland et al., 2002).

Along with LGT, fusion, and fission events, *ACL* is absent from numerous individual lineages, suggesting secondary losses (Fig. 1). For instance, while most SAR members encode ACL (typically ssACL), ciliates possess no ACL homologs. In addition, many lineages that have lost *ACL* are anaerobic (e.g., *Blastocystis*, *Entamoeba* and *Mastigamoeba*) and/or parasitic (e.g., *Trypanosoma*, *Giardia* and *Trichomonas*); unsurprisingly, this absence correlates fairly well with the absence of *CS*, and, potentially, the absence of ACL's substrate, citrate.

A reasonable scenario accounting for the distribution of *ACL* in eukaryotes is that LECA possessed both variants, with subsequent differential losses. Nonetheless, we have not found any instances of *ssACL* and *dsACL* co-occurring in a single species. This is somewhat reminiscent of the case of *EF1-α* (elongation factor 1-alpha) and *EFL* (elongation factor-like), functionally overlapping core components of the eukaryotic translation machinery that are patchily distributed across eukaryotes. To date, only a dozen species that encode both *EF-1α* and *EFL* have been identified; in these organisms, *EF-1α* is transcriptionally suppressed relative to *EFL* (Kamikawa et al., 2013). A similar scenario may account for the distribution of *ssACL* and *dsACL*: given that both isoforms overlap functionally, one of the two – depending on the lineage – might have become progressively less expressed, leading to its eventual loss. In this vein, as much of our analysis is based on transcriptome data, it is possible that alternative isoforms might not be identified if poorly expressed. Nuclear genomic sequences from diverse

Fig. 2. Rooted Maximum Likelihood phylogenetic tree of ACL homologs (110 sequences, 820 sites). The alignment contains ssACL and concatenated dsACL sequences. The root of the tree was positioned based on preliminary analyses, which included the α and β subunits of succinyl-CoA synthetase, and citrate synthase as outgroups. The light and dark gray shadings indicate dsACL and ssACL isoforms, respectively. Taxon names are colored according to their taxonomic affiliation: black represents Bacteria; gray, Archaea; and other colors represent various eukaryotic groups. LGT recipient lineages discussed in the main text are indicated with a red star. Black "F" and "S" boxes indicate gene fusion and splitting events, respectively. Numbers at branches are bootstrap values (BV) and posterior probabilities (PP). For the sake of clarity, only BV > 70% and PP > 0.7 are indicated. Black dots indicate maximum support (BV = 100% and PP = 1.0). Scale bar represents the average number of substitutions per site. The branch leading to Chlorobi was reduced to a third of its length. Accession numbers are provided for all sequences to the right of the taxon names. The *Imantonia sp.* sequence is marked with an asterisk to indicate that the authors suspect it is the result of data contamination from an amoebozoan organism. For this reason, we did not further consider it as a LGT candidate.

eukaryotic microbes will be required to definitively resolve this issue.

4. Conclusions

As sequence data accumulate, more prokaryote-derived gene fusions are being identified in eukaryotic nuclear genomes (Gawryluk et al., 2014; Maguire et al., 2014; Stairs et al., 2014; Stechmann and Cavalier-Smith, 2002); in many cases, these genes derive from relatively recent transfer/fusion events. In contrast, we have demonstrated here that the fusion of *ACLA* and *ACLB* into *ssACL* is not a curious feature of animals; rather *ssACL*, along with *dsACL*, likely represents an ancestral feature of eukaryotic genomes.

Our phylogenetic analyses demonstrate that the evolution of *ACL* in eukaryotes has involved vertical inheritance, LGT, extensive gene loss, and gene fusion/fission. Moreover, these results emphasize the insidiousness of homoplasy in supposedly rare genomic changes (Maguire et al., 2014), and how the propensity to overstate the importance of such characters (Stechmann and Cavalier-Smith, 2002) may ultimately impair the interpretation of evolutionary relationships.

Acknowledgments

LE was supported by a CGEB postdoctoral fellowship from the Tula Foundation. RMRG's postdoctoral fellowship and this work were supported by a CIHR-NSHRF RPP Grant (FRN# 62809) awarded to AJR.

References

Brown, M.W., Sharpe, S.C., Silberman, J.D., Heiss, A.A., Lang, B.F., Simpson, A.G.B., Roger, A.J., 2013. Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. Proc. Biol. Sci. 280, 20131755.

Buchanan, B.B., Arnon, D.I., 1990. A reverse KREBS cycle in photosynthesis: consensus at last. Photosynth. Res. 24, 47–53.

Burki, F., 2014. The eukaryotic tree of life from a global phylogenomic perspective. Cold Spring Harb. Perspect. Biol. 6, a016147.

Campbell, B.J., Cary, S.C., 2004. Abundance of reverse tricarboxylic acid cycle genes in free-living microorganisms at deep-sea hydrothermal vents. Appl. Environ. Microbiol. 70, 6282–6289.

Cavalier-Smith, T., Fiore-Donno, A.M., Chao, E., Kudryavtsev, A., Berney, C., Snell, E.A., Lewis, R., 2014. Multigene phylogeny resolves deep branching of Amoebozoa. Mol. Phylogenet. Evol.

Criscuolo, A., Gribaldo, S., 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol. Biol. 10, 210.

Elshourbagy, N.A., Near, J.C., Kmetz, P.J., Sathe, G.M., Southan, C., Strickler, J.E., Gross, M., Young, J.F., Wells, T.N., Groot, P.H., 1990. Rat ATP citrate-lyase. Molecular cloning and sequence analysis of a full-length cDNA and mRNA abundance as a function of diet, organ, and age. J. Biol. Chem. 265, 1430–1435.

Elshourbagy, N.A., Near, J.C., Kmetz, P.J., Wells, T.N.C., Groot, P.H.E., Saxty, B.A., Hughes, S.A., Franklin, M., Gloger, I.S., 1992. Cloning and expression of a human ATP-citrate lyase cDNA. Eur. J. Biochem. 204, 491–499.

Eme, L., Sharpe, S.C., Brown, M.W., Roger, A.J., 2014. On the age of eukaryotes: evaluating evidence from fossils and molecular clocks. Cold Spring Harb. Perspect. Biol. 6, a016139.

Fatland, B.L., Ke, J., Anderson, M.D., Mentzen, W.I., Cui, L.W., Allred, C.C., Johnston, J.L., Nikolau, B.J., Wurtele, E.S., 2002. Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in Arabidopsis. Plant Physiol. 130, 740–756.

Gawryluk, R.M.R., Chisholm, K.A., Pinto, D.M., Gray, M.W., 2014. Compositional complexity of the mitochondrial proteome of a unicellular eukaryote (*Acanthamoeba castellanii*, supergroup Amoebozoa) rivals that of animals, fungi, and plants. J. Proteomics 109, 400–416.

Hügler, M., Huber, H., Molyneaux, S.J., Vetriani, C., Sievert, S.M., 2007. Autotrophic CO_2 fixation via the reductive tricarboxylic acid cycle in different lineages within the phylum Aquificae: evidence for two ways of citrate cleavage. Environ. Microbiol. 9, 81–92.

Hynes, M.J., Murray, S.L., 2010. ATP-citrate lyase is required for production of cytosolic acetyl coenzyme A and development in *Aspergillus nidulans*. Eukaryot. Cell 9, 1039–1048.

Kamikawa, R., Brown, M.W., Nishimura, Y., Sako, Y., Heiss, A.A., Yubuki, N., Gawryluk, R., Simpson, A.G.B., Roger, A.J., Hashimoto, T., 2013. Parallel re-modeling of EF-1alpha function: divergent EF-1alpha genes co-occur with EFL genes in diverse distantly related eukaryotes. BMC Evol. Biol. 13, 131.

Kanao, T., Fukui, T., Atomi, H., Imanaka, T., 2001. ATP-citrate lyase from the green sulfur bacterium Chlorobium limicola is a heteromeric enzyme composed of two distinct gene products. Eur. J. Biochem. 268, 1670–1678.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Keeling, P.J., Burki, F., Wilcox, H.M., Allam, B., Allen, E.E., Amaral-Zettler, L.A., Armbrust, E.V., Archibald, J.M., Bharti, A.K., Bell, C.J., Beszteri, B., Bidle, K.D., Cameron, C.T., Campbell, L., Caron, D.A., Cattolico, R.A., Collier, J.L., Coyne, K., Davy, S.K., Deschamps, P., Dyhrman, S.T., Edvardsen, B., Gates, R.D., Gobler, C.J., Greenwood, S.J., Guida, S.M., Jacobi, J.L., Jakobsen, K.S., James, E.R., Jenkins, B., John, U., Johnson, M.D., Juhl, A.R., Lovejoy, C., Lynn, D.H., Marchetti, A., McManus, G., Nedelcu, A.M., Menden-Deuer, S., Miceli, C., Mock, T., Montresor, M., Moran, M.A., Murray, S., Nadathur, G., Nagai, S., Ngam, P.B., Palenik, B., Pawlowski, J., Petroni, G., Piganeau, G., Posewitz, M.C., Rengefors, K., Romano, G., Rumpho, M.E., Rynearson, T., Schilling, K.B., Schroeder, D.C., Simpson, A.G.B., Slamovits, C.H., Smith, D.R., Smith, G.J., Smith, S.R., Sosik, H.M., Stief, P., Theriot, E., Twary, S.N., Umale, P.E., Vaulot, D., Wawrik, B., Wheeler, G.L., Wilson, W.H., Xu, Y., Zingone, A., Worden, A.Z., 2014. The marine microbial eukaryote transcriptome sequencing project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. PLoS Biol. 12, e1001889.

Lahr, D.J.G., Grant, J., Nguyen, T., Lin, J.H., Katz, L.A., 2011. Comprehensive phylogenetic reconstruction of amoebozoa based on concatenated analyses of SSU-rDNA and actin genes. PLoS One 6, e22780.

Lartillot, N., Lepage, T., Blanquart, S., 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25, 2286–2288.

Ma, Y., Jakowitsch, J., Maier, T., Bayer, M., Müller, N., Schenk, H., Löffelhardt, W., 2001. ATP citrate lyase in the glaucocystophyte alga *Cyanophora paradoxa* is a cytosolic enzyme: characterisation of the gene for the large subunit at the cDNA and genomic levels. Mol. Genet. Genomics 266, 231–238.

Maguire, F., Henriquez, F.L., Leonard, G., Dacks, J.B., Brown, M.W., Richards, T.A., 2014. Complex patterns of gene fission in the eukaryotic folate biosynthesis pathway. Genome Biol. Evol. 6, 2709–2720.

Nowrousian, M., Kück, U., Loser, K., Weltring, K.-M., 2000. The fungal *acl*1 and *acl*2 genes encode two polypeptides with homology to the N-and C-terminal parts of the animal ATP citrate lyase polypeptide. Curr. Genet. 37, 189–193.

Oliver, D.J., Nikolau, B.J., Wurtele, E.S., 2009. Acetyl-CoA—life at the metabolic nexus. Plant Sci. 176, 597–601.

Son, H., Lee, J., Park, A.R., Lee, Y.-W., 2011. ATP citrate lyase is required for normal sexual and asexual development in *Gibberella zeae*. Fungal Genet. Biol. 48, 408–417.

Stairs, C.W., Eme, L., Brown, M.W., Mutsaers, C., Susko, E., Dellaire, G., Soanes, D.M., van der Giezen, M., Roger, A.J., 2014. A SUF Fe-S cluster biogenesis system in the mitochondrion-related organelles of the anaerobic protist *Pygsuia*. Curr. Biol. 24, 1176–1186.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Stechmann, A., Cavalier-Smith, T., 2002. Rooting the eukaryote tree by using a derived gene fusion. Science 297, 89–91.

Wellen, K.E., Hatzivassiliou, G., Sachdeva, U.M., Bui, T.V., Cross, J.R., Thompson, C.B., 2009. ATP-citrate lyase links cellular metabolism to histone acetylation. Science 324, 1076–1080.

Yabuki, A., Kamikawa, R., Ishikawa, S.A., Kolisko, M., Kim, E., Tanabe, A.S., Kume, K., Ishida, K.-I., Inagki, Y., 2014. *Palpitomonas bilix* represents a basal cryptist lineage: insight into the character evolution in Cryptista. Sci. Rep. 4, 4641.