

Published in final edited form as:

Eur J Protistol. 2012 January ; 48(1): 89–93. doi:10.1016/j.ejop.2011.10.003.

A genomic survey shows that the haloarchaeal type tyrosyl tRNA synthetase is not a synapomorphy of opisthokonts

John D.L. Shadwick^a and Iñaki Ruiz-Trillo^{a,b,*}

^aDepartament de Genètica & Institut de Recerca en Biodiversitat (Irbio), Universitat de Barcelona, Barcelona, Spain

^bInstitució Catalana per a la Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys, 23, 08010 Barcelona, Spain

Abstract

The haloarchaeal-type tyrosyl tRNA synthetase (tyrRS) have previously been proposed to be a molecular synapomorphy of the opisthokonts. To re-evaluate this we have performed a taxon-wide genomic survey of tyrRS in eukaryotes and prokaryotes. Our phylogenetic trees group eukaryotes with archaea, with all opisthokonts sharing the haloarchaeal-type tyrRS. However, this type of tyrRS is not exclusive to opisthokonts, since it also encoded by two amoebozoans. Whether this is a consequence of lateral gene transfer or lineage sorting remains unsolved, but in any case haloarchaeal-type tyrRS is not a synapomorphy of opisthokonts. This demonstrates that molecular markers should be re-evaluated once a better taxon sampling becomes available.

Keywords

Molecular markers; Opisthokonts; Lateral gene transfer; Amoebozoa; Lineage sorting

Introduction

The Opisthokonta is an evolutionary eukaryotic super-clade composed of metazoans, fungi, and their unicellular and colonial relatives, i.e., the choanoflagellates, the filastereans, the ichthyosporeans, the nucleariids, the aggregating amoeba *Fonticula alba*, and the difficult to place *Corallochytrium* (Adl et al. 2005; Cavalier-Smith and Chao 2003; Steenkamp and Baldauf 2004; see Paps and Ruiz-Trillo 2010 for a recent review). The monophyly of the opisthokonts is supported by single gene, multiple gene and phylogenomic datasets (Brown et al. 2009; Lang et al. 2002; Medina et al. 2003; Ruiz-Trillo et al. 2004; Ruiz-Trillo et al. 2006; Ruiz-Trillo et al. 2008; Shalchian-Tabrizi et al. 2008; Steenkamp et al. 2006; Parfrey et al. 2010; Torruella et al. 2011) and by morphological characters including posteriorly directed flagella in flagellated forms, and flat mitochondrial cristae (Cavalier-Smith and

© 2011 Elsevier GmbH. All rights reserved.

*Corresponding author at: Universitat de Barcelona-Parc Científic de Barcelona Edifici Hèlix, Planta 1, Lab A04 Baldiri Reixac, 15-21 08028 Barcelona, Spain. Tel.: +34 93 402 04 82; fax: +34 93 403 44 20. inaki.ruiz@icrea.es (I. Ruiz-Trillo).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejop.2011.10.003.

Chao 2003). Molecular synapomorphies for the group have also been proposed including a 12 amino acid insertion in the elongation 1 alpha (EF1- α) gene (Baldauf and Palmer 1993), and a tyrosyl tRNA synthetase (tyrRS) gene that is most closely related to halophylic archaea and represents a possible ancient lateral gene transfer to the earliest opisthokonts (Huang et al. 2005).

With the advent of high throughput sequencing methods, dozens of eukaryotes have had all or a large part of their genomes sequenced. As new data become available it is worthwhile to reevaluate previously proposed molecular synapomorphies of the opisthokonts and see if they hold up under the weight of new data. The 12 amino acid insertion in the EF1- α gene has recently been re-evaluated and it continues to be exclusive to opisthokonts, and not present in any of the analyzed amoebozoan, apusozoan or bikont genomes (Sebé-Pedrós et al. 2010). In contrast, the tyrRS has so far only been re-analyzed using a very limited taxon sampling and some rather incomplete sequences (Huang and Gogarten 2009). To fill this gap, we have here reanalyzed the evolution of tyrRS with the broader possible taxon sampling.

Material and Methods

A large dataset of the tyrRS amino acid sequence was collected including 97 eukaryotes, 80 Archaea, and 197 Eubacteria, using a combination of BLASTp and tBLASTn and with different sequences and tyrRS types used as query. Two datasets were used for phylogenetic analyses, one with all 374 taxa from all three domains of life, and another with just the Archaea and eukaryotes. Only sequences that were at least 500 amino acids long were taken into account. Alignments were refined by eye using the program Geneious Pro for alignment editing. Conserved sections of the alignments were selected using Gblocks (Castresana 2000). After using Gblocks in the 374 taxon dataset, 182 conserved amino acid positions were recovered, while 187 conserved positions were recovered from the smaller Archaea and eukaryote dataset. An extra alignment including partial, putative tyrRS sequences of the choanoflagellate *Monosiga ovata* and the amoebozoan *Hartmannella vermiformis* was also done. Final alignments can be downloaded from the webpage www.multicellgenome.com or upon request. Phylogenetic trees were built using RaxML (Stamatakis 2006) for maximum likelihood and, for the smallest dataset, MrBayes (Ronquist and Huelsenbeck 2003) for Bayesian methods. In the RaxML and MrBayes runs we used the WAG substitution matrix with the gamma model of rate substitutions, and proportion of invariable sites estimated (WAG + Γ + I). In the RaxML runs 1000 bootstraps were estimated with the rapid bootstrapping option (Stamatakis et al. 2008).

Results and Discussion

To reevaluate tyrRS as a molecular marker we assembled the tyrRS amino acid sequences of a large number of Eubacteria, Archaea, and all of the eukaryotes of which we had knowledge. This added numerous eukaryotic lineages to the dataset of Huang et al. (2005) and Huang and Gogarten (2009), including ichthyosporeans, filastereans, and choanoflagellates in the opisthokonts, as well as ciliates, jakobids, stramenopiles, apusozoans, several more apicomplexans, kinetoplastids, many more plants and green algae,

and several more *Plasmodium* species. We created trees from a large dataset that included all Eubacteria, Archaea, and eukaryotes (Fig. S1, Supplementary Material), and from a smaller dataset that included just Archaea and eukaryotes (Fig. 1).

The tyrRS trees that included taxa from all three domains (Fig. S1, Supplementary Material) had a monophyletic eubacterial clade with very long branches clustering within the Archaea clade. The eukaryotes were split into two well supported groups (Bootstrap Value (BV) = 100%), both of them branching within archaeal clades. We thus made trees with just the Archaea and eukaryotes to get a better estimate of relationships in this clade (Fig. 1). The resultant tree had again the eukaryotes broken up into two monophyletic clades. One clade includes all opisthokonts and has a tyrRS gene most closely related to the haloarchaeal group with BV = 100% and Bayesian posterior probability (PP) = 1.0. A second clade containing most other eukaryotes has a tyrRS gene more closely related to a clade containing crenarchaeotes, euryarchaeotes, and nanoarchaeotes with BV = 100% and PP = 1.0. Interestingly, as in Huang and Gogarten (2009) the eukaryotic + haloarchaea clade also includes some amoebozoans, specifically *Acanthamoeba castellanii* and *Mastigamoeba balamuthi* invalidating the haloarchaeal form of tyrRS as a synapomorphy of the opisthokonts. The non-opisthokont eukaryotic clade, includes, however, all the remaining Amoebozoa sampled, i.e., *Entamoeba dispar*, *Entamoeba histolytica*, *Dictyostelium discoideum*, and *Dictyostelium purpureum*. It also includes the apusozoan *Thecamonas trahens*, the putative sister-group to the opisthokonts (Torruella et al. 2011). It is worth mentioning that Huang and Gogarten (2009) included expressed sequence tag data from the amoebozoan *Hartmannella vermiformis* and the choanoflagellate *Monosiga ovata* in their analysis, but we did not include them in our final analysis due to their short sequence (see section 'Material and Methods') and the inability to check for contamination artifacts. However, a tree containing both those partial sequences is shown in Fig. S2, in which *H. vermiformis* also groups with the haloarchaeal type (as *Acanthamoeba* and *Mastigamoeba*). *M. ovata* sequence appears, as in Huang and Gogarten (2009), with the non-opisthokonts, within a clade of Viridiplantae (see Fig. S2). With the current data and given that the sequence is partial, we can not discard, however, a contamination artifact. Moreover, there is no tyrRS sequence from other choanoflagellates, even the current genome assembly of the choanoflagellate *Salpingoeca rosetta* does not contain any tyrRS.

The presence of the haloarchaeal type of tyrRS gene in *Acanthamoeba* and *Mastigamoeba* (and maybe *Hartmannella*) but not in *Entamoeba*, *Dictyostelium* and *Thecamonas* may be due to one of several hypotheses: (1) this form of the tyrRS gene is ancestral to *Acanthamoeba*, *Mastigamoeba*, and opisthokonts, (2) lateral gene transfer (LGT) events, or (3) incomplete lineage sorting due to hidden paralogy. The first hypothesis, that the haloarchaeal tyrRS type is ancestral to *Acanthamoeba*, *Mastigamoeba*, and opisthokonts, will imply the opisthokonts are more closely related to *Acanthamoeba* and *Mastigamoeba* than to *Entamoeba* and *Dictyostelium*, and even *Thecamonas*, and it will imply amoebozoans are not monophyletic. This goes against new phylogenetic (Cavalier-Smith and Chao 2010) as well as phylogenomic (Torruella et al. 2011) data that strongly show the apusozoans, rather than the amoebozoans, as the sister group to opisthokonts. It also goes against phylogenomic data that show that *Mastigamoeba* is most closely related to

Entamoeba (and *Dictyostelium*) than to *Acanthamoeba* (see panel B in Fig. 1) (Minge et al. 2009; Parfrey et al. 2010). Having a clear root of the eukaryotes and a better sampling of taxa in Amoebozoa may help clarify this issue. Our data cannot confidently discern between LGT or hidden paralogy. The opisthokonts appear monophyletic in the tyrRS tree, although with low statistical values, while *Acanthamoeba* and *Mastigamoeba* both appear as paraphyletic, branching earlier than the opisthokonts (Fig. 1). Although we cannot rule out LGT, a possible explanation is that the ancestral eukaryote had the two types of tyrRS and that the haloarchaeal-type copy was lost in most eukaryotic lineages, except for opisthokonts, *Acanthamoeba* and *Mastigamoeba* (and maybe, *Hartmannella*). Additional data from Amoebozoa, choanoflagellates and other eukaryotic lineages are needed to get a clear picture of tyrRS evolutionary history. In any case our data represent a cautionary tale about proposed molecular synapomorphies when the taxon sampling is scarce. Suggested rare genomic changes or molecular markers should from time to time be reevaluated when new genomic information is gathered.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The genome sequences of *C. owczarzaki*, *A. macrogynus*, *S. punctatus*, and *T. trahens* are being determined by the Broad Institute of MIT/Harvard University under the auspices of the National Human Genome Research Institute (NHGRI) and within the UNICORN initiative. We thank JGI, BI, and BCM for making data publicly available. We thank Kim C. Worley and the team of the *A. castellanii* genome project for accession to the genome data. This work was supported by an ICREA contract, an European Research Council Starting Grant (ERC-2007-StG-206883), and a grant (BFU2008-02839/BMC) from Ministerio de Ciencia e Innovación (MICINN) to IR-T.

References

- Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MF. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* 2005; 52:399–451. [PubMed: 16248873]
- Baldauf SL, Palmer JD. Animals and fungi are each other's closest relatives: congruent evidence from multiple proteins. *Proc. Natl. Acad. Sci. U.S.A.* 1993; 90:11558–11562. [PubMed: 8265589]
- Brown MW, Spiegel FW, Silberman JD. Phylogeny of the “forgotten” cellular slime mold, *Fonticula alba*, reveals a key evolutionary branch within Opisthokonta. *Mol. Biol. Evol.* 2009; 26:2699–2709. [PubMed: 19692665]
- Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 2000; 17:540–552. [PubMed: 10742046]
- Cavalier-Smith T, Chao EE. Phylogeny of choanozoa, apusozoa, and other protozoa and early eukaryote megaevolution. *J. Mol. Evol.* 2003; 56:540–563. [PubMed: 12698292]
- Cavalier-Smith T, Chao EE. Phylogeny and evolution of *Apusomonadida* (Protozoa: Apusozoa): new genera and species. *Protist.* 2010; 161:549–576. [PubMed: 20537943]
- Huang J, Gogarten JP. Ancient gene transfer as a tool in phylogenetic reconstruction. *Methods Mol. Biol.* 2009; 532:127–139. [PubMed: 19271182]
- Huang J, Xu Y, Gogarten JP. The presence of a haloarchaeal type tyrosyl-tRNA synthetase marks the opisthokonts as monophyletic. *Mol. Biol. Evol.* 2005; 22:2142–2146. [PubMed: 16049196]

- Lang BF, et al. The closest unicellular relatives of animals. *Curr. Biol.* 2002; 12:1773–1778. [PubMed: 12401173]
- Medina M, Collins AG, Taylor JW, Valentine JW, Lipps JH, Amaral-Zettler L, Sogin ML. Phylogeny of Opisthokonta and the evolution of multicellularity and complexity in fungi and metazoa. *Int. J. Astrobiol.* 2003; 2:203–211.
- Minge MA, Silberman JD, Orr RJ, Cavalier-Smith T, Shalchian-Tabrizi K, Burki F, Skjaeveland A, Jakobsen KS. Evolutionary position of breviate amoebae and the primary eukaryote divergence. *Proc. Biol. Sci.* 2009; 276:597–604. [PubMed: 19004754]
- Paps J, Ruiz-Trillo I. *Encyclopedia of Life Sciences (ELS)*. John Wiley & Sons, Ltd.; Chichester: 2010. Animals and Their Unicellular Ancestors.
- Parfrey LW, Grant J, Tekle YI, Lasek-Nesselquist E, Morrison HG, Sogin ML, Patterson DJ, Katz LA. Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. *Syst. Biol.* 2010; 59:518–533. [PubMed: 20656852]
- Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics.* 2003; 19:1572–1574. [PubMed: 12912839]
- Ruiz-Trillo I, Inagaki Y, Davis LA, Sperstad S, Landfald B, Roger AJ. *Capsaspora owczarzaki* is an independent opisthokont lineage. *Curr Biol.* 2004; 14:R946–R947. [PubMed: 15556849]
- Ruiz-Trillo I, Roger AJ, Burger G, Gray MW, Lang BF. A phylogenomic investigation into the origin of metazoa. *Mol. Biol. Evol.* 2008; 25:664–672. [PubMed: 18184723]
- Ruiz-Trillo I, Lane CEM, Roger AJ. Insights into the evolutionary origin and genome architecture of the unicellular opisthokonts *Capsaspora owczarzaki* and *Sphaeroforma arctica*. *J. Eukaryot. Microbiol.* 2006; 53:1–6.
- Sebé-Pedrós A, Roger AJ, Lang FB, King N, Ruiz-Trillo I. Ancient origin of the integrin-mediated adhesion and signaling machinery. *Proc. Natl. Acad. Sci. U.S.A.* 2010; 107:10142–10147. [PubMed: 20479219]
- Shalchian-Tabrizi K, Minge MA, Espelund M, Orr R, Ruden T, Jakobsen KS, Cavalier-Smith T. Multigene phylogeny of choanozoa and the origin of animals. *PLoS ONE.* 2008; 3:e2098. [PubMed: 18461162]
- Stamatakis A. RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 2006; 22:2688–2690. [PubMed: 16928733]
- Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAXML Web servers. *Syst. Biol.* 2008; 57:758–771. [PubMed: 18853362]
- Steenkamp, ET.; Baldauf, SL. Origin and evolution of animals, fungi and their unicellular allies (Opisthokonta). In: Hirt, RP.; Horner, DS., editors. *Organelles, Genomes and Eukaryote Phylogeny: An Evolutionary Synthesis in the Age of Genomics*. CRC Press; Boca Raton: 2004. p. 109-129.
- Steenkamp ET, Wright J, Baldauf SL. The protistan origins of animals and fungi. *Mol. Biol. Evol.* 2006; 23:93–106. [PubMed: 16151185]
- Torruella G, Derelle R, Paps J, Lang BF, Roger AJ, Shalchian-Tabrizi K, Ruiz-Trillo I. Phylogenetic relationships within the Opisthokonta based on phylogenomic analyses of conserved single copy protein domains. *Mol. Biol. Evol.* 2011 doi:10.1093/molbev/msr185 Epub ahead of print.

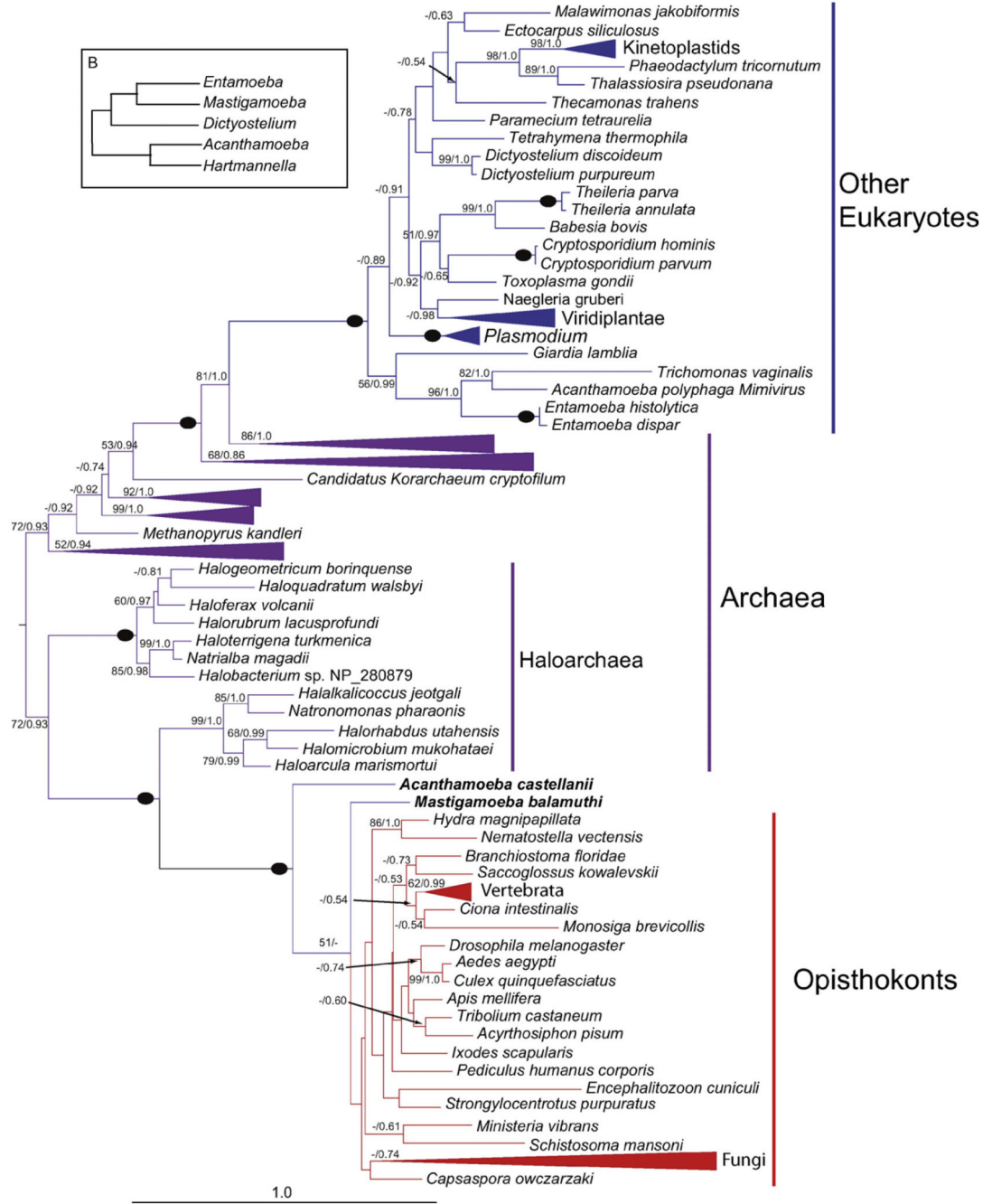


Fig. 1. Maximum likelihood tree of the tyrRS gene of the dataset containing archaea and eukaryotes only. Colors of the branches represent groups to which the taxa belong (purple = archaea, red = opisthokonts, and blue = non-opisthokont eukaryotes). The percentage of maximum likelihood bootstrap values and Bayesian posterior probabilities (BV/PP) are to left of each node above or below the branch or indicated with an arrow if there is not enough space. Support values of 100%/1.0 are represented with a black oval. Values less than 50%/0.5 are represented with a “-” or not reported. Several nodes are represented in cartoon form to save

space. In panel B, a schematic tree of the putative phylogenetic relationships between the different amoebozoan taxa used in this study is shown based on Minge et al. (2009) and Parfrey et al. (2010).