

FORMAL COMMENT

Asgard archaea are the closest prokaryotic relatives of eukaryotes

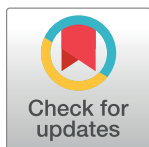
Anja Spang^{1☯}, Laura Eme^{1☯}, Jimmy H. Saw², Eva F. Caceres¹, Katarzyna Zaremba-Niedzwiedzka¹, Jonathan Lombard¹, Lionel Guy³, Thijs J. G. Ettema^{1*}

1 Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden, **2** Department of Microbiology, Oregon State University, Corvallis, Oregon, United States of America, **3** Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

☯ These authors contributed equally to this work.

✉ Current address: NIOZ, Royal Netherlands Institute for Sea Research, Department of Marine Microbiology and Biogeochemistry, and Utrecht University, NL-1790 AB Den Burg, The Netherlands

* thijs.ettema@icm.uu.se



OPEN ACCESS

Citation: Spang A, Eme L, Saw JH, Caceres EF, Zaremba-Niedzwiedzka K, Lombard J, et al. (2018) Asgard archaea are the closest prokaryotic relatives of eukaryotes. *PLoS Genet* 14(3): e1007080. <https://doi.org/10.1371/journal.pgen.1007080>

Editor: Antonis Rokas, Vanderbilt University, UNITED STATES

Received: August 21, 2017

Accepted: October 17, 2017

Published: March 29, 2018

Copyright: © 2018 Spang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work is supported by grants of the European Research Council (ERC Starting Grant 310039-PUZZLE_CELL), the Swedish Foundation for Strategic Research (SSF-FFL5), and the Swedish Research Council (VR grant 2015–04959), awarded to TJGE. LE is funded by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 704263. JL is supported by a postdoctoral fellowship for foreign researchers from the Wenner-Gren Foundations in Stockholm (UPD2016-0072). The funders had no role in study design, data collection

In Spang et al. [1], we reported the discovery of Lokiarchaeum (Loki1) and two related lineages (Loki2 and Loki3, now referred to as Heimdallarchaeote LC_2 and LC_3). These were the first known representatives of the subsequently described Asgard superphylum [2], which branches as a sister clade to the TACK archaea [3]. We provided extensive phylogenomic evidence for the emergence of eukaryotes from within the Asgard archaea (thus supporting the two-domains [2D] tree of life), and the investigation of the Lokiarchaeum genome revealed the presence of more eukaryotic signature proteins (ESPs) than in any other described archaeal lineage [1].

In the recent study by Da Cunha et al. [4], the authors argue against the findings reported in Spang et al. Their main claims are that (i) the published genome data contains extensive contamination from distantly related organisms; (ii) a subset of universal genes supports a three-domains (3D) tree of life; and (iii) removing the elongation factor 2 (EF2) from the phylogenomic analysis breaks the Lokiarchaeota–Eukarya affiliation.

Below, we show that the claims by Da Cunha et al. are unfounded.

There is no evidence of significant contamination in the lokiarchaeal genomes

Da Cunha et al. determined the quality of the Loki1 genome bin with CheckM [5] and observed 78.21% heterogeneity, which they incorrectly interpreted as a sign of contamination. CheckM measures heterogeneity based on the detection of multiple copies of single-copy marker genes, and in fact, 78.21% of the redundant markers share at least 90% amino acid identity. This confirms our original report, which reported that the Lokiarchaeum genome bin comprises sequences from closely related strains [1].

The authors argue that single-gene trees indicate contamination because Asgard homologs are not always monophyletic. However, given that Loki1–3 representatives belong to the distinct phyla Lokiarchaeota and Heimdallarchaeota, this is not surprising [2]. Furthermore, most deep nodes in these single-gene trees are not significantly supported, which hinders reaching strong conclusions regarding the placement of these sequences.

Da Cunha et al. also state that the single gene–encoding RNA polymerase subunit A (*rpoA*) represents a contaminant of the Heimdallarchaeote LC_3 bin, because *rpoA* is encoded by two genes in other Asgard archaea, while some lineages within TACK encode a single gene [2]. BLAST searches for the N- and C-termini of the Heimdallarchaeote LC_3 *rpoA* gene sequence actually recover best hits to other Asgard archaea, thus indicating that it represents a bona fide

and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Heimdallarchaeote gene. In addition, Heimdallarchaeote LC_3 RpoA forms a strongly supported monophyletic group with the homologs of all other Asgard archaea in phylogenetic analyses [2]. Finally, split/fusion events of archaeal RpoA subunits appear to have occurred more frequently than assumed previously [2], rendering them unsuited for the elucidation of distant evolutionary events.

Moreover, the authors claim that lokiarchaeal EF-2 homologs contain indels indicative of chimerism. However, it is unlikely that chimerism has affected the same gene in all Loki- and Heimdallarchaeota independently while leaving all other genes in these genomes unaffected. The evolution of the EF-2 homologs in archaea is actually a much more complex subject than previously acknowledged. An analysis of the archaeal EF-2 paralogs and their characteristics is already available in [2] and will be discussed in greater detail in a devoted article (Narrowe et al., submitted).

Finally, most of the work by Da Cunha et al. limits itself to the first lokiarchaeal genome bin [1] and disregards the eight additional metagenomic bins from a diversity of Asgard archaea [2]. The currently available Asgard genomes stem from five different samples and metagenomes that were assembled and binned by three independent laboratories using different methodologies. All of them corroborate our original phylogenetic placement of Asgard lineages, their relationship with eukaryotes, and the unique presence of ESPs in all of these genomes.

The 3D topology recovered in a subset of universal genes is likely due to long-branch attraction artefacts

After running topology tests, Da Cunha et al. argue that only 6 and 11 universal markers statistically rejected 2D and 3D topologies, respectively. They split these two datasets into respective concatenations and reconstructed the corresponding phylogenies (S7 and S8 Figs in [4]). Unsurprisingly, they found a 3D topology for the 6-protein dataset, and a 2D topology, where eukaryotes and Asgard archaea are monophyletic, for the 11-protein concatenation. As the number of individual trees showing a 2D topology is higher when additional longer-branching taxa are included (S1 Fig [4]), the authors incorrectly argue that the topology obtained by the 11-protein dataset is likely the result of a long-branch attraction (LBA) artefact. By definition, LBA artefactually attracts long-branching lineages together, irrespective of their correct phylogenetic placement. Thus, if LBA was affecting the phylogenetic reconstruction, the eukaryotic clade would be attracted outside the archaea, to the distantly related bacterial outgroup, and not well nested within the archaeal clade. Rather, we argue that the 3D topology obtained by Da Cunha et al. for the 6-protein dataset is itself the result of LBA: the branch length connecting the bacterial outgroup to the archaeal and eukaryotic clade is strikingly longer in the 6-protein phylogeny compared to the 11-protein tree (2.1 versus 1.3 average substitution per site). The existence of artefacts in the 3D topology is further supported by the fact that the internal relationships among eukaryotes are less consistent with the accepted phylogeny of organisms than in the 2D tree. For example, the fast-evolving *Entamoeba histolytica* is attracted toward the base in the 3D tree, whereas it is correctly placed as sister to *Dictyostelium discoideum* in the 2D tree. The 3D topology reported in Da Cunha et al. is thus most likely incorrect.

Finally, it is important to acknowledge that the 2D topology has been obtained by numerous studies worldwide using different datasets and methods [3,6–14]; several of these studies were published before the publication of the Lokiarchaeum genome.

Phylogenetic analyses excluding EF2 support the monophyly of Asgard and eukaryotes

The claim “removal of the EF-2, is sufficient to break the Eukaryotes-Lokiarchaea affiliation” [4] has been disproven by phylogenetic analyses performed on three different concatenated

datasets [2], none of which included elongation factor EF-2 (48 universal single-copy genes, 55 ribosomal proteins, and 16S and 23S rRNA genes, respectively). All analyses clearly retrieved a strongly supported affiliation of eukaryotes with the Asgard superphylum.

Conclusions

The origin of eukaryotes represents a controversial topic, and while we believe it to be refreshing and insightful, our recent work will certainly not be the final word in this debate. However, Da Cunha et al. use inadequate methodology, misinterpret data and ignore previous work [1,2] that already addressed many of their criticisms. In summary, they provide no evidence to falsify the conclusions drawn in Spang et al. and Zaremba-Niedzwiedzka et al.

Author Contributions

Conceptualization: Anja Spang, Laura Eme, Thijs J. G. Ettema.

Formal analysis: Anja Spang, Laura Eme, Jimmy H. Saw, Eva F. Caceres, Katarzyna Zaremba-Niedzwiedzka, Jonathan Lombard, Lionel Guy, Thijs J. G. Ettema.

Funding acquisition: Thijs J. G. Ettema.

Supervision: Thijs J. G. Ettema.

Writing – original draft: Anja Spang, Laura Eme, Jimmy H. Saw, Eva F. Caceres, Katarzyna Zaremba-Niedzwiedzka, Jonathan Lombard, Lionel Guy, Thijs J. G. Ettema.

Writing – review & editing: Anja Spang, Laura Eme, Jimmy H. Saw, Eva F. Caceres, Katarzyna Zaremba-Niedzwiedzka, Jonathan Lombard, Lionel Guy, Thijs J. G. Ettema.

References

1. Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, et al. (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521: 173–179. <https://doi.org/10.1038/nature14447> PMID: 25945739
2. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström D, Juzokaite L, et al. (2017) Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541: 353–358. <https://doi.org/10.1038/nature21031> PMID: 28077874
3. Guy L, Ettema TJG (2011) The archaeal “TACK” superphylum and the origin of eukaryotes. *Trends Microbiol* 19: 580–587. <https://doi.org/10.1016/j.tim.2011.09.002> PMID: 22018741
4. Da Cunha V, Gaia M, Gadelle D, Nasir A, Forterre P (2017) Lokiarchaea are close relatives of Euryarchaeota, not bridging the gap between prokaryotes and eukaryotes. *PLoS Genet* 13: e1006810. <https://doi.org/10.1371/journal.pgen.1006810> PMID: 28604769
5. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25: 1043–1055. <https://doi.org/10.1101/gr.186072.114> PMID: 25977477
6. Pisani D, Cotton JA, McInerney JO (2007) Supertrees disentangle the chimerical origin of eukaryotic genomes. *Mol Biol Evol* 24: 1752–1760. <https://doi.org/10.1093/molbev/msm095> PMID: 17504772
7. Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM (2008) The archaeobacterial origin of eukaryotes. *Proc Natl Acad Sci U S A* 105: 20356–20361. <https://doi.org/10.1073/pnas.0810647105> PMID: 19073919
8. Foster PG, Cox CJ, Embley TM (2009) The primary divisions of life: a phylogenomic approach employing composition-heterogeneous methods. *Philos Trans R Soc Lond, B, Biol Sci* 364: 2197–2207. <https://doi.org/10.1098/rstb.2009.0034> PMID: 19571240
9. Kelly S, Wickstead B, Gull K (2011) Archaeal phylogenomics provides evidence in support of a methanogenic origin of the Archaea and a thaumarchaeal origin for the eukaryotes. *Proc Biol Sci* 278: 1009–1018. <https://doi.org/10.1098/rspb.2010.1427> PMID: 20880885
10. Williams TA, Foster PG, Cox CJ, Embley TM (2013) An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* 504: 231–236. <https://doi.org/10.1038/nature12779> PMID: 24336283

11. Lasek-Nesselquist E, Gogarten JP (2013) The effects of model choice and mitigating bias on the ribosomal tree of life. *Mol Phylogenet Evol* 69: 17–38. <https://doi.org/10.1016/j.ympev.2013.05.006> PMID: [23707703](https://pubmed.ncbi.nlm.nih.gov/23707703/)
12. Williams TA, Embley TM (2014) Archaeal “dark matter” and the origin of eukaryotes. *Genome Biol Evol* 6: 474–481. <https://doi.org/10.1093/gbe/evu031> PMID: [24532674](https://pubmed.ncbi.nlm.nih.gov/24532674/)
13. Guy L, Saw JH, Ettema TJ (2014) The archaeal legacy of eukaryotes: a phylogenomic perspective. *Cold Spring Harbor perspectives in biology* 6:a016022. <https://doi.org/10.1101/cshperspect.a016022> PMID: [24993577](https://pubmed.ncbi.nlm.nih.gov/24993577/)
14. Raymann K, Brochier-Armanet C, Gribaldo S (2015) The two-domain tree of life is linked to a new root for the Archaea. *Proc Natl Acad Sci U S A* 112: 6670–6675. <https://doi.org/10.1073/pnas.1420858112> PMID: [25964353](https://pubmed.ncbi.nlm.nih.gov/25964353/)