nature ecology & evolution

Matters arising

https://doi.org/10.1038/s41559-022-01896-8

Reply to: Phylogenetic affiliation of mitochondria with Alpha-II and Rickettsiales is an artefact

Received: 5 February 2022

Accepted: 1 September 2022

Published online: 24 October 2022

Check for updates

Lu Fan[1,2] ✉, Dingfeng Wu[3], Vadim Goremykin[4], Katharina Trost[5], Michael Knopp[5], Chuanlun Zhang[1,2], William F. Martin[5] ✉ and Ruixin Zhu[3] ✉

REPLYING TO: J. Martijn et al. *Nature Ecology & Evolution* https://doi.org/10.1038/s41559-022-01871-3 (2022)

Martijn et al.[1] argue that phylogenetic data support a sister group relationship between Alphaproteobacteria and mitochondria, in contrast with our findings[2] that mitochondria branch within Alphaproteobacteria. We disagree with their conclusion and below we discuss multiple criticisms of their approach, including data filtration, interpretation of results, datasets used and taxonomic declarations.

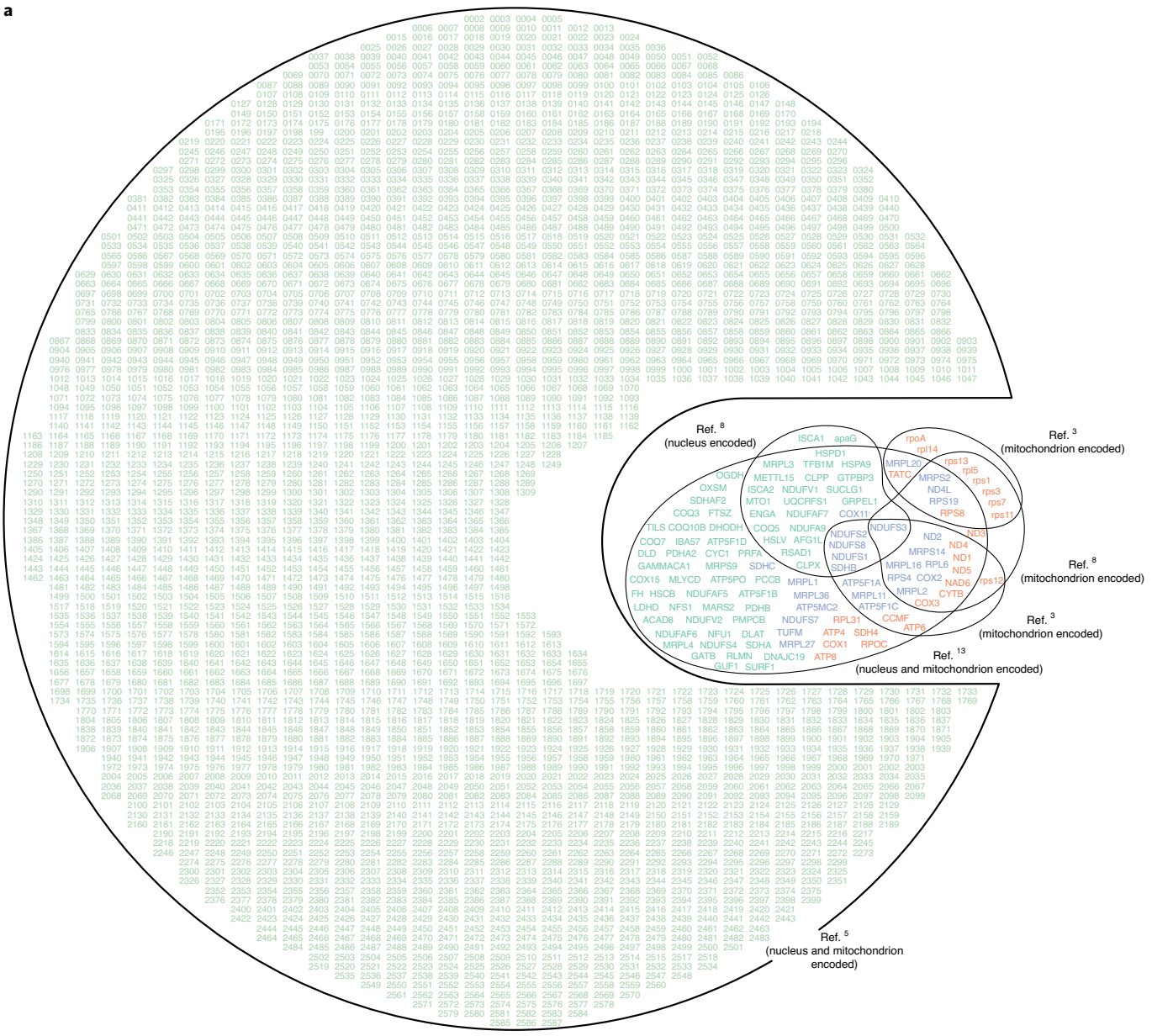
The main technical problems in their present analysis[1] are the same as in their previous analysis[3], namely that their site exclusion procedures are arbitrary. Martijn et al.[1] again used ad hoc criteria and terminated the site exclusion procedure—after 5, 10, 20, 30 or 40% of the sites were removed—when a particular result (mitochondria branching outside the Alphaproteobacteria) was obtained. Selecting a threshold to remove highly variable sites from alignments is challenging[2,4] and we disagree with the authors' justification. In selecting sites for exclusion, Martijn et al. relied on Fig. 1a of Fan et al.[2] to indicate that methods that were better at improving fit yielded Alphaproteobacteria-sister trees whereas those that were worse at improving fit yielded Rickettsiales-sister or unresolved trees. However, we disagree with their interpretation of our figure. The figure clearly shows that Bowker's score performed as well as the χ^2 score and Stuart's score in improving the model fit. At the same time, it also generated a tree topology of *Tistrella mobilis* sisterhood to mitochondria when the model fit was substantially improved. Therefore, the central conclusions of our paper[2] remain unaddressed by their correspondence.

Martijn et al. argue against our suggestion of possible outgroup attraction[2], but we disagree with their arguments for three reasons. (1) They conclude[1] that the outgroup removal test[3] applies only if mitochondria truly branch sister to Rickettsiales or if non-Rickettsiales are outgroups (according to this test) that are able to falsely attract mitochondria. Their conclusion is actually in line with our speculation that if Rickettsiales sisterhood might be true for these data, site removal might have led to signal loss and consequently outgroup attraction[2]. (2) They admit that outgroup attraction may exist but must be weak[1]. Given the nature of this test, by simulating alignments under a Rickettsiales-sister constraint tree and using them to infer rooted trees[3], it is highly likely that the strength of the Rickettsiales-sister constraint (parameter) they set up in the simulated data can essentially determine how much the Rickettsiales-sister topology will be challenged by outgroup attraction. Thus, the result of the test[3] is probably parameter sensitive but Martijn et al.[1] have not investigated this possibility. (3) We commented[2] on their random sequence test[3] that in the ten trees they provided, eight, zero, three, eight, five, zero, zero, six, zero and three of the ten random sequences are attracted by the outgroup, respectively. However, Martijn et al. think these are branch-within instead of attraction events[1]. We assume they mean that the random sequences might have found their close relatives in the outgroup. However, we find it unlikely that eight in ten random sequences would find their close relatives in a small outgroup containing six taxa while only two would do the same in the remaining 71 taxa. If these are really random sequences, only long-branch attraction can explain their presence in the outgroup with high frequency.

Martijn et al.[1] argue that the placement of mitochondria within Alpha-II (Supplementary Figs. 35b and 36 in ref. [2]) was caused by the attraction of Rickettsiales or FEMAG-II by a hypothetical problematic taxon of Alpha-II—MarineAlpha9 Bin5. FEMAG-II refers to a group of fast-evolving metagenome-assembled genomes in Alphaproteobacteria and was proposed by Fan et al.[2]. FEMAG-II includes MarineAlpha6 CompositeBin56 (MarineAlpha6 Bin5), MarineAlpha6 Bin1, MarineAlpha8, MarineAlpha7, MarineAlpha5 Bin12, MarineAlpha5 Bin9, MarineAlpha5 Bin5, MarineAlpha5 CompositeBin678

[1]Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, China. [2]Shenzhen Key Laboratory of Marine Archaea Geo-Omics, Department of Ocean Science and Engineering, Southern University of Science and Technology (SUSTech), Shenzhen, China. [3]The Shanghai Tenth People's Hospital, School of Life Sciences and Technology, Tongji University, Shanghai, China. [4]Research and Innovation Centre, Fondazione E. Mach, San Michele all'Adige, Italy. [5]Institute of Molecular Evolution, Heinrich Heine University, Düsseldorf, Germany. ✉e-mail: fanl@sustech.edu.cn; bill@hhu.de; rxzhu@tongji.edu.cn

https://doi.org/10.1038/s41559-022-01896-8



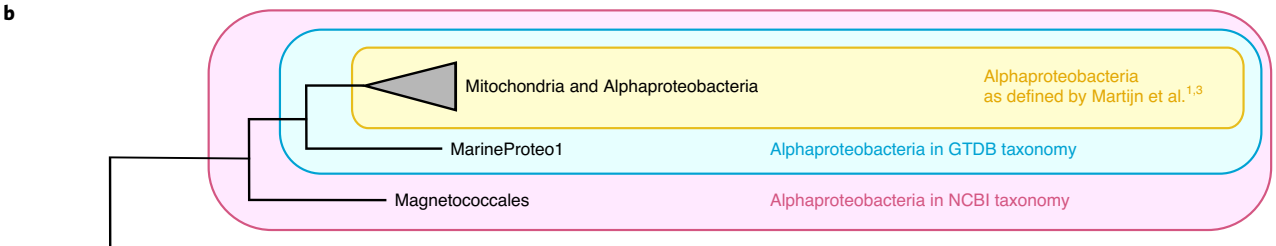


Fig. 1 | Gene sampling and taxonomic declarations by Martijn et al. pose problems when inferring the origin of mitochondria. a, Diagram showing the relative sample sizes of protein studies in phylogenetic investigations of the origin of mitochondria. The largest set of 2,587 protein clusters is that of Nagies et al.[5]. The smaller sets included therein are from Extended Data Fig. 1 of Muñoz-Gómez et al.[13]. The gene assignments for the numbers (used here for comparable scale) are given in Supplementary Table 1. **b**, The Alphaproteobacteria-sister hypothesis forwarded by Martijn et al.[3] in 2018 is based on a definition of the class Alphaproteobacteria that we find problematic. Shown is a schematized phylogenetic tree showing the phylogenetic relationship of Alphaproteobacteria and mitochondria, as well as the definition of the class Alphaproteobacteria by Martijn et al.[1,3] and two representative taxonomic databases. The tree topology shown is consistently recovered by Martijn et al.[1,3], by others[13] and by us[2]. GTDB, Genome Taxonomy Database; NCBI, National Center for Biotechnology Information.

(MarineAlpha5 Bin7), MarineAlpha5 CompositeBin1011 and MarineAlpha5 CompositeBin123(MarineAlpha5 Bin3). Removal of this taxon broke the direct connection between mitochondria and Alpha-II[1].

The exact phylogenetic relationship between mitochondria and major alphaproteobacterial clades is a challenging problem and we don't think that including or excluding potential sister taxa will solve it.

It is often reported in some difficult cases of phylogenetic study that the presence or absence of one taxon changes the entire tree topology. However, such an observation does not necessarily mean either topology is correct but only that possibly (but not certainly) more evidence is required to obtain a robust conclusion.

Martijn et al.[1] argue that the Rickettsiales-sister topology based on our Modified18 dataset[2] was the result of an artefact caused by secondary convergent evolution of selected Rickettsiales and mitochondria from AT-rich ancestors towards the present GC-enriched state. They then tried to develop a new strategy to identify and remove compositional heterogeneous sites that could potentially be masked by this convergent factor[1]. However, their assumption that all extant mitochondria (including those with higher GC content mentioned here) have an AT-enriched common ancestor is an unproven premise. Furthermore, the protocol that Martijn et al.[1] developed to treat their data is problematic. In their alignment matrix, orthologues of more GC-rich Rickettsiales and mitochondria were replaced by more AT-rich representatives of other Rickettsiales and mitochondria (from outside the dataset) to identify heterogeneous sites and for downstream phylogenetic inference[1]. Martijn et al. created a chimeric dataset with concatenated sequences each comprising orthologues of different taxa, which we think makes their finding unlikely to be robust.

Another systematic issue with the approaches used by Martijn et al.[1,3] is the limited sampling of genes for phylogenetic study. One needs to consider all of the data, not just the data that fit a particular model. In Fig. 1a, we show that Martijn et al.'s sample[1,3] comprises less than 2% of the 2,587 protein clusters analysed previously in sequenced eukaryotic genomes—clusters whose alignments and trees address the phylogenetic affinity of mitochondria. Those 2,587 trees speak to the issue[5] and highlight the uncomfortable reality that lateral gene transfer among prokaryotes imposes on studies of the endosymbiotic origin of organelles: the Alphaproteobactera, like Delta-, Gamma- and Betaproteobacteria, are among the least vertically evolving of all major prokaryotic groups known[5] (only 21% of their genes recover monophyly of the group because of the impact of lateral gene transfer[5]).

Finally, the Alphaproteobacteria-sister hypothesis for the origin of mitochondria forwarded by Martijn et al.[1,3] concerns their arbitrary exclusion of Magnetococcales and the so-called MarineProteo1 clade from Alphaproteobacteria, as shown in Fig. 1b. In the original literature[6–10] and NCBI Taxonomy[11], members of the Magnetococcales are included within the Alphaproteobacteria, and the group MarineProteo1 is assigned to the Alphaproteobacteria in the Genome Taxonomy Database[12]. However, Martijn et al.[1,3] excluded both groups from the Alphaproteobacteria because they are not closely related to other Alphaproteobacteria in their phylogenetic tree. Martijn et al. did not explain their criteria for excluding those taxa from the Alphaproteobacteria.

The continuing discovery of novel alphaproteobacterial genomes[3,6,12,13] offers an opportunity to revisit the phylogenetic relationship of mitochondria and extant alphaproteobacterial lineages. One important avenue of research is to search for close relatives of mitochondria and for relationships that are robust and insensitive to parameter selection. In future studies, the inclusion of lineages ancestral to intracellular Rickettsiales, along with additional basal alphaproteobacterial lineages, can help to address the position of mitochondria within the tree of Alphaproteobacteria, and furthermore to identify modern alphaproteobacterial lineages that harbour similar collections of genes[5,14] to the ancestor of mitochondria.

References

1. Martijn, J., Vosseberg, J., Guy, L., Offre, P. & Ettema, T. J. G. Phylogenetic affiliation of mitochondria with Alpha-II and Rickettsiales is an artefact. *Nat. Ecol. Evol.* https://doi.org/10.1038/s41559-022-01871-3 (2022).
2. Fan, L. et al. Phylogenetic analyses with systematic taxon sampling show that mitochondria branch within Alphaproteobacteria. *Nat. Ecol. Evol.* **4**, 1213–1219 (2020).
3. Martijn, J., Vosseberg, J., Guy, L., Offre, P. & Ettema, T. J. G. Deep mitochondrial origin outside the sampled Alphaproteobacteria. *Nature* **557**, 101–105 (2018).
4. Goremykin, V. V., Hansmann, S. & Martin, W. F. Evolutionary analysis of 58 proteins encoded in six completely sequenced chloroplast genomes: revised molecular estimates of two seed plant divergence times. *Plant Syst. Evol.* **206**, 337–351 (1997).
5. Nagies, F. S. P., Brueckner, J., Tria, F. D. K. & Martin, W. F. A spectrum of verticality across genes. *PLoS Genet.* **16**, e1009200 (2020).
6. Bazylinski, D. A. et al. *Magnetococcus marinus* gen. nov., sp. nov., a marine, magnetotactic bacterium that represents a novel lineage (*Magnetococcaceae* fam. nov., *Magnetococcales* ord. nov.) at the base of the Alphaproteobacteria. *Int. J. Syst. Evol. Microbiol.* **63**, 801–808 (2013).
7. Dziuba, M. V., Kolganova, T. V., Gorlenko, V. M. & Kuznetsov, B. B. Species diversity of magnetotactic bacteria from the Ol'khovka River, Russia. *Microbiology* **82**, 335–340 (2013).
8. Wang, Z. & Wu, M. An integrated phylogenomic approach toward pinpointing the origin of mitochondria. *Sci Rep.* **5**, 7949 (2015).
9. Ferla, M. P., Thrash, J. C., Giovannoni, S. J. & Patrick, W. M. New rRNA gene-based phylogenies of the Alphaproteobacteria provide perspective on major groups, mitochondrial ancestry and phylogenetic instability. *PLoS ONE* **8**, e83383 (2013).
10. Morillo, V. et al. Isolation, cultivation and genomic analysis of magnetosome biomineralization genes of a new genus of South-seeking magnetotactic cocci within the Alphaproteobacteria. *Front. Microbiol.* **5**, 72 (2014).
11. Schoch, C. L. et al. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database (Oxford)* **2020**, baaa062 (2020).
12. Parks, D. H. et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).
13. Muñoz-Gómez, S. A. et al. Site-and-branch-heterogeneous analyses of an expanded dataset favour mitochondria as sister to known Alphaproteobacteria. *Nat. Ecol. Evol.* **6**, 253–262 (2022).
14. Esser, C., Martin, W. & Dagan, T. The origin of mitochondria in light of a fluid prokaryotic chromosome model. *Biol. Lett.* **3**, 180–184 (2007).

Author contributions

L.F., K.T., W.F.M. and R.Z. drafted the reply. All co-authors contributed to the text. D.W., V.G., M.K. and C.Z. revised the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-022-01896-8.

Correspondence and requests for materials should be addressed to Lu Fan, William F. Martin or Ruixin Zhu.

https://doi.org/10.1038/s41559-022-01896-8

Peer review information *Nature Ecology & Evolution* thanks the anonymous reviewers for their contribution to the peer review of this work.