nature portfolio

Peer Review File

Phylogenomics of Asgard archaea reveals a unique blend of prokaryotic-like horizontal transfer and eukaryotic-like gene duplication.

Corresponding Author: Dr Toni Gabaldon

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
In this article, the authors performed phylogenetic analyses centered around the gene repertoire of Asgard archaea, and suggested that the Asgard archaeal genomes contain a substantial amount of duplication, which is more similar to eukaryotes than prokaryotes. The authors also examined the rates of HGT from other prokaryotic lineages and found that metabolic genes are enriched with such events. I find the topic of interest as a detailed gene content profiling of Asgard archaea is likely to provide some missing information to further our understanding in eukaryogenesis. The authors are relatively careful about conclusions on the second part of the paper regarding HGT, which I appreciate.

However, I do have major concerns regarding the use of small number of taxonomically biased genomes to infer duplication events and make the major conclusion of enriched duplication in Asgard archaea: The authors used only 61 prokaryotic genome set, where from what I can find in the texts - 31 is from Asgard, 17 from TACK, 10 other Archaea, and 5 Bacteria – which by btw add up to 63, not 61 as stated in the paper. In addition, given that a selected gene may not exist in some of the 63 genomes examined, an actual gene tree in this study very likely included genes from less than 63 genomes. This has a number of consequences:

1. This provides little, if not completely no, chance for the discovery of HGT events from outside Asgard/TACK, and hence a strong over-estimate of duplication events. For example, if a Lokiarchaeal genome encode two genes that belong to the same gene family, each were HGTed from a different bacterial lineage (or from archaea outside TACK), but this gene family does not exist in TACK. These two gene copies would have been reported as a duplication within Asgard using the authors' approach.
2. Such a selection restricted the paper to make comparisons only with TACK for the extent of duplication. The comparison itself has two main issues: 1) The authors selected 17 TACK vs 31 Asgard, and while the gene family selection was based on 12 Asgard without considering TACK-specific genes. 2) The selected TACK lineages mostly have small genomes averaging 1.5Mb, almost half of the average smaller than most Asgard. These factors can lead to strong statistical bias. More importantly, even without considering the statistical bias, the main conclusion of the paper was made based on a comparison with TACK without considering other bacteria and archaea with larger genomes. I feel that it is far from sufficient.

To make the conclusions written in this paper, I think it is important that the authors make a broader and more even taxonomic sampling, as well as providing sufficient evidence to exclude potential HGT that disguise as duplication from outside the selected lineages. Please do pay attention to potential biases in genome sizes and gene counts when making quantitative comparisons.

Other major concerns:
Novelty of the paper: 1) On the topic of gene duplication in Asgard archaea, Vosseberg et al 2020 Nature Eco & Evo, where the last authors of this article T. Gabaldon was a corresponding author, has specifically highlighted such a case. It is strange to me that the authors did not refer to this paper. In this 2020 aper, it was already concluded that Asgard archaea exhibited most duplications, so I suggest that the authors discuss the current work in the context of their previous paper, and state clearly what is new and different. 2) The main conclusion for the HGT side seems to be that the metabolic functions are more prominently involved, which does not seem too surprising. Beyond that, the authors seem to follow with various detailed

descriptions, which I do appreciate but unfortunately show little implications for eukaryogenesis. Can the authors perhaps provide more insights into the functional consequences of those HGTs and potential link to eukaryogenesis?

The taxonomic sampling is also a little strange: Even with as few as 61 or 63 genomes, the authors selected MAGs that belong to the same species or genus: two Baldr from g__DXJG01, two Heimdall from g__UBA460, 3 Hermod from g__Hermodarchaeum, 3 Hod from g__Hodarchaeum, two Jord from g__DTBI01, two Odin from g__LCB-4, 2 Borr from the same species s__Borrarchaeum weybense, two Wokong from species s__Wukongarchaeum yapensis. Same goes for two g__Nitrosopumilus. In addition, is there a reason not to use circularized Freya/Atabey/Heimdall genomes available for analyses? In the SI table file, I can only find the information for the Asgard and TACK genomes but not the remaining 15, so I cannot judge for the rest.

The authors suggest that there is higher rate of intra-Asgard HGT of ESPs than other gene families. Given that most of the ESPs only exist in the Asgard archaea, could this also be statistically biased?

Other minor and general issues:
The two parts of the paper used different approaches. Have the authors cross-compared the duplication events concluded from the method 1 with the HGT events concluded from method 2?
Line 321, PMIDs, replace with real reference.
Line 269-271, please consistently use Desulfoproteobacteriota for SRBs.
Line 340-343, This is interesting. who are the hosts of those viruses? Can you provide some clear examples?
Line 254: (D)->(E)

Reviewer #2

(Remarks to the Author)
In this work, Manzano-Morales and Gabaldon performed a comparative phylogenomic analysis of Asgard archaea and their sister group TACK and found that genome evolution in Asgard archaea is characterized by high rates of gene duplication and horizontal gene transfer (HGT). Such blend of prokaryotic-like HGT rates and eukaryotic-like gene duplication is a remarkable finding that, in my opinion, will be of interest to many potential readers of Nature Communications.

The analysis is comprehensive and thorough, and the authors made a great effort to ensure that their findings are not due to contamination or genome incompleteness by applying rigorous filters and combining multiple lines of evidence. There are some details that require clarification or could be better explained (see below), but I do not think that these compromise the integrity of the manuscript.

Some general comments and questions:

1. I wonder if the authors could go one step further and provide some insight into the evolutionary regime that characterizes genome evolution in Asgard archaea. To contextualize this comment: Most bacterial genomes evolve in a regime of deletion-driven genome reduction, such that useful genes are maintained by selection. Some manifestations of that regime are a positive association between genome size and effective population size, and a scarcity of psedogenes in bacterial genomes. Instead, eukaryotic genomes evolve in a non-adaptive regime driven by gene duplication and the proliferation of transposons and other mobile genetic elements. Such non-adaptive regime results in a negative association between genome size and effective population size, as well as accumulation of noncoding DNA. Estimating the effective population size is complicated, possibly unfeasible if only one high-quality genome per species is available. However, it would be interesting to use the ancestral reconstructions to investigate how genome size has changed along the evolution of Asgard archaea and compare it with the TACK group. Have high duplication rates resulted in a net increase in genome size? Were duplications compensated for by higher rates of gene loss? If there has been a net increase in genome size, did it result from higher numbers of functional genes or from pseudogenes and mobile genetic elements? In that regard, it may be worth assessing the abundance of pseudogenes and the rates of pseudogenization and pseudogene decay.

2. I found it surprising that there were no signs of transposons, proviruses, and other integrated mobile genetic elements among the duplicated and horizontally transferred genes. Is the lack of MGE duplication and transfer a consequence of the conservative filters applied to identify HGT and duplications or is it a genuine finding?

Some points that need to be clarified or better explained:

3. Lines 69-73. After reading this paragraph it is not clear if those MAGs were used or not.

4. Lines 85-86: How is it that only 17K out of 31K ***Asgard archaeal*** OGs contain at least 1 Asgard representative? What about the other 14K? Perhaps I am missing something: what do you exactly mean by Asgard representative?

5. Lines 91-92: Did you only include KEGG pathways or also KEGG ontologies?

6. Lines 108-109: The finding that Chao's lower bound and core genome size are sensitive to genome incompleteness is not surprising. Chao's lower bound depends on the number of singletons and OGs with 2 members, hence it is a not quite robust estimator. As for the core genome, the reasons are trivial (but see for example PMID:37904249 for a quantitative assessment). In the light of these results, it may be worth running mOTUpan on both sets of genomes to get a more robust estimate of the core genome size and pangenome fluidity.

7. Line 110: I am not sure of what you mean by the "full set". Do you mean all available MAGs? It has been shown that orthofinder performs quite poorly when applied to incomplete MAGs. I also wonder if the analysis of the full dataset is really necessary: given the openness of the pangenome and looking at the rarefaction curve in Fig 1 it is sufficiently clear that there are a lot of Asgard archaeal genes yet to discover and that non-HQ MAGs will contain genes that are not present in the representative set.

8. Figure 1A: This plot would be more informative if it used a logarithmic scale for the y-axis.

9. Lines 120-121: Here "entire genome set" means the representative HQ genomes, right? Which is not the same as the "full genome set" mentioned in line 110. That is confusing.

10. Figure 1C: What is the hub of the PPI network (K12158)? What are the proteins that form the 2 triangles?

11. Line 129: In the absence of a null model, to say that the network is highly connected is quite subjective. Less than half of the proteins have interactions and, if we only look at those, more than half of the interactions have a low score (around 200).

12. Line 187: I am not sure of what you mean by "a value of 1 would indicate a whole genome duplication". In the figure there are cases of 10: I do not suppose that those represent 10 whole genome duplications, right? Adding to the confusion, the caption says "mean number of duplications per node" but the label of the x-axis says "mean number of duplications per OG". That requires clarification.

13. Figure 2: What are the rate units for the duplication and transfer rates? Events per unit of time (what unit of time)? Per gene? Per genome?

14. Figure 2E: Do the distributions correspond to nodes or OGs? Equivalently, do the tails of the distribution correspond to OGs that duplicate a lot or genomes with a lot of duplications?

15. Figure 2F: What is the definition of verticality?

16. Figure 2G (the caption says F): Something looks wrong with the boxplot and overlapping scatter plots. The "other" category is clearly not normally distributed, so it would probably be better to use logarithmically transformed data. The same applies to the statistical test: did you use a robust test or transformed data? What are the effect size and p-value? These are important details that must be disclosed (not only for panel G, but also for panels C, D, F and Fig 3E). The biological implications (higher transfer of ESP than other genes), if confirmed by statistical analyses, are certainly relevant.

17. Lines 193-196: When comparing the duplication rate of lineage signature proteins within and outside the lineage, did you control for the lineage-wise duplication rates? Did lineage signature proteins also duplicate more than other (non-signature) OGs within the same lineage?

18. Line 232 and below: Do the methods used to infer inter-domain HGT account for the fact that proteins can evolve at very different rates in bacteria, archaea and eukaryotes? Would it be possible to obtain reliable estimates of the fraction of false positives?

19. Figure 3: Do all panels in Fig 3 refer to Bacteria-to-Asgard HGT?

20. Fig 3A: what does "contribution of different seeds" mean?

21. Figure 3D: Besides the KEGG and KO terms, it would be helpful to see a higher-order classification of the transferred genes, for example in terms of COG categories.

22. Fig 3E, bottom panel: Why are the degrees always <1? Does it mean that most proteins have degree 0 (no interactions)? But then, how is it that in the top panel the Q25 is already well above 0?

23. Line 266: Figure 3C would be more informative if instead of the number of taxa it showed which taxa acted as donors. That information is not currently shown here or the supplement, although some donor taxa are mentioned in the main text. If adding that to Fig 3 is not feasible, a summary table in the supplement would work too.

24. Line 294: The degree is only one of several possible measurements of centrality. Have you checked other centrality indices, such as the eigenvector centrality?

25. Line 303: Based on phylogenomic reconstructions, did the transfer of these enzymes lead to the replacement of a preexisting enzyme carrying out the same function? Otherwise, did HGT lead to qualitatively new or broader metabolic potential?

Online Methods

26. Line 26, regarding the lowest CheckM contamination criterion: Was it only applied to break ties? Did you apply any hard threshold on the contamination score? Contamination would result in apparent duplication and HGT, so, for these purposes,

minimizing contamination seems more important than maximizing completeness.

27. Line 41: How did you harmonize the need to work with homogeneous genome annotations (to control for biases and annotation-related confounding factors) while keeping the original genome annotations?

28. Line 51: Please, provide a brief summary of the annotation workflow (Ref 33).

29. Line 53: FLUH has not been defined. Furthermore, this paragraph is quite cryptic. What does it refer to?

30. Line 59 and below: Were the alignments and trees based on nucleotide or amino acid sequences?

31. Line 176: Did you check that putative instances of tandem duplications are not due to fragments of the same gene interrupted by a stop codon (either real or resulting from a sequencing error)?

32. Extended Data Figure 3: The distribution of the number of transfers (upper left) would be more informative if it was plotted in log-log or log-linear scale.

33. Line 311: This caption refers to folate metabolism, not carbon metabolism.

Typos

- Line 264: Data Figure 4A (instead of 3C).
- Line 345: High (instead of highly).
- Online methods, lines 68-69: the sentence in parentheses is repeated.

Reviewer #3

(Remarks to the Author)
The paper presents the results of a phylogenomic study focused on Asgard archaea. The main findings relate to patterns of gene duplication and HGT and their functional and genomic impacts. The findings are interesting and shed light on some unique aspects of Asgard evolution with implications for understanding eukaryogenesis. My comments primarily have to do with thoroughness of the analysis and appropriateness of the methods used.

Major comments:
- There are several cutoff thresholds used throughout the paper but these seem to have been chosen somewhat arbitrarily. As an example, Asgard signature proteins are defined using thresholds of at least 70% presence in Asgard and at most 15% presence in other archaea or bacteria. Why were these specific thresholds chosen and how do results/findings change if other reasonable thresholds, such as 80% and 10%, are used instead? The same applies to the thresholds used on lines 139-140.

- It is mentioned that two different species phylogenies, constrained and unconstrained, were generated and used for HGT and gene duplication analyses. However, it seems that results are only presented for one of those phylogenies. Which phylogeny was used for the reported results? And how do findings change when the alternative phylogeny is used? Results should be shown for both phylogenies, and the phylogenies themselves should be provided as well.

- It is mentioned that two methods were used for duplication analysis, but results are reported only for the reconciliation-based method. Results should also be shown for the other (species overlap) method. Are all reported results supported by both methods?

- Same as above, there also appear to be multiple methods used for HGT inference. Are all reported results consistent between these different analyses (e.g., reconciliation-based and non-reconciliation-based)?

- The AleRax tool is used for reconciliation-based analysis. It appears that the species tree generated by AleRax was used to perform this analysis. Do we now have a third species tree that is different from the constrained and unconstrained species phylogenies used for the other analyses? If so, this needs to be clearly stated in the paper. And how do results change if the two previous species phylogenies are used for this reconciliation analysis? The use of multiple species trees makes it all the more important to clear report how findings change as different reasonable species trees are used for the analysis.

- Some of the results, such as elevated duplication rates at the base of the Asgards, may be artifacts of errors in the gene trees. Paying further attention to gene tree quality may help resolve some of these artifacts.

- In the time of acquisition analysis of Lines 259-264, using last common ancestors of all genomes carrying the transferred gene is almost guaranteed to provide incorrect time estimates since such genes may have been acquired multiple times or may have been transferred again after initial acquisition. A much more accurate way would be to directly use the reconciliation to estimate when a gene was first acquired.

- Line 87: where are the 8365 OGs found only in a single Asgard genome coming from? A BLAST search against a larger collection of prokaryotic and eukaryotic genomes could help resolve this question.

- Line 83 of online methods supplementary text: why are only the top 150 hits chosen per protein? Artificial cutoffs such as this can result in under-clustering, leading to significant errors in downstream analyses and especially in reconciliation. This analysis should be repeated with a more biologically grounded approach for identifying gene families.

- Lines 95-101 of online methods supplementary text: The description of how and on what data AleRax was run is unclear. What is meant by “we selected one sequence per OG present in the Asgard clade”?

Minor comments:
- Increase font size of node annotations in Figure 1.

- Typo on line 345, highly dynamism.

- Lines 133-134: why were only 39 out of the 83 previously identified proteins found? Is it because only a subset of available Asgard genomes was used in the analysis?

- Lines 146-147: Why not use all 31 Asgard genomes for this analysis? How were the 12 genomes chosen?

- In the analysis of tandem duplications (Lines 198-208), were phylogenetic relationships between the genomes considered? Ancestral tandem duplications would be inherited by multiple genomes and may be counted multiple times, possibly impacting the results shown in Extended Data Figure 2.

- Line 59 of online methods supplementary text. Which markers were used?

- Line 142 of online methods supplementary text. There are better rooting methods than midpoint rooting. The authors could consider using Minimum Ancestor Deviation (MAD) rooting.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have addressed the previous comments and made substantial improvement to the analyses and clarification of the paper. I have only a few minor comments on the revised manuscript:

1. I recommend that the authors refer to the original articles reporting the 32 representative Asgard archaeal genomes.
2. The authors used Em et al 2023 tree for reconciliation, does it indeed include all the species used in this study?
3. Note that Deltaproteria is now designated into various Desulfobacterota clades in GTDB, Desulfoproteobacterota-> Desulfobacterota. Please be consistent throughout the paper.
4. line 431, 'high dynamism in Asgard genome'. It is unclear what it means, please rephrase.
5. Figure 3A. What do all the colors of the pie charts stand for? what is Njordarchaeia.1?
6. Please properly label the species names and accession numbers on the tips of phylogenomic trees in Fig. S1, S3 and S6.

(Remarks on code availability)
The page does not seem to be accessible at the moment.

Reviewer #2

(Remarks to the Author)
Most of my previous comments have been addressed in the revised version. I only have a minor request:

Could you please comment on the (putative) functionality of duplicated genes? Is there any evidence of pseudogenization (e.g. truncating mutations, high variance of ORF lengths)? Are they subject to positive, relaxed, or negative selection? With bacteria in mind, one would expect that, if duplicated genes have persisted for a long time, they are probably functional (simply because pseudogenes are quickly eroded in bacteria). But these are not bacteria and it seems that gene loss is less frequent than gene duplication/gain in Asgard, so it might be that persistence does not imply functionality.

Other than that, I detected some typos and would suggest some edits for clarity:

l. 77: I would suggest to add some connector in the beginning of this paragraph to stress that restricting the pangenome to 32 high-quality Asgard genomes was a necessary compromise to circumvent the potential problems of working with a larger set of incomplete MAGs. One can figure it out from the context, but a more explicit statement would be helpful.

l. 84: "17,429 contained at least one Asgard protein, 12,397 were exclusively present in Asgard archaea, and 10,470 were present in a single Asgard genome and no other prokaryote in our dataset." Are these numbers nested within each other? I interpret that the 10,470 singletons are included within the 12,397 Asgard-only OGs, and those are included within the 17,429 Asgard OGs. Just in case, it may be a good idea to rephrase a little bit (perhaps adding "of those") to make that point totally clear.

l. 86: "figAsgard" -> possible typo?

l. 103: "not in terms of core genome" -> I find that statement confusing because the estimated strict core is 18 in both cases. Perhaps the statement refers to the observed strict and relaxed cores (but those are not reported for the random dataset) of for the mOTUpan estimates (but those are not presented until the following paragraph).

l. 141: "Figure 1D" -> Figure 1C in the updated version.

l. 188: missing space before 23.87%

l. 202: "Figure 2E" -> Figure 2D?

l. 239: "(F)" -> (G)

l. 312: The pie charts appear below (not above) the nodes. The legend linking colors to seeds is missing.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
The authors have made many changes to the manuscript and have addressed most of my concerns. But there is a related issue that has not been adequately addressed: While the authors have performed additional analysis and reported their findings, some of these new findings are not well integrated into the manuscript. For example, the authors now provide the results of the reconciliation analysis using the alternative topology in Supplementary Figure 1. However, there is no discussion of these results elsewhere in the manuscript.

The same applies to results presented in Supplementary Figure 3.

It would be great if the authors could better integrate the results of these alternative analyses into the manuscript, clearly explaining which findings are supported and not supported by the alternative analyses.

Minor comment:
The two topologies shown in Supplementary Figure 3 are missing leaf labels.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.
The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.
To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this article, the authors performed phylogenetic analyses centered around the gene repertoire of Asgard archaea, and suggested that the Asgard archaeal genomes contain a substantial amount of duplication, which is more similar to eukaryotes than prokaryotes. The authors also examined the rates of HGT from other prokaryotic lineages and found that metabolic genes are enriched with such events. I find the topic of interest as a detailed gene content profiling of Asgard archaea is likely to provide some missing information to further our understanding in eukaryogenesis. The authors are relatively careful about conclusions on the second part of the paper regarding HGT, which I appreciate.

Response: We thank the reviewer for appreciating the interest of our contribution.

However, I do have major concerns regarding the use of small number of taxonomically biased genomes to infer duplication events and make the major conclusion of enriched duplication in Asgard archaea: The authors used only 61 prokaryotic genome set**,** where from what I can find in the texts - 31 is from Asgard, 17 from TACK, 10 other Archaea, and 5 Bacteria – which by btw add up to 63, not 61 as stated in the paper. In addition, given that a selected gene may not exist in some of the 63 genomes examined, an actual gene tree in this study very likely included genes from less than 63 genomes. This has a number of consequences:

1. This provides **little, if not completely no, chance for the discovery of HGT events from outside Asgard/TACK**, and hence a **strong over-estimate of duplication events**. For example, if a Lokiarchaeal genome encode two genes that belong to the same gene family, each were HGTed from a different bacterial lineage (or from archaea outside TACK), but this gene family does not exist in TACK. These two gene copies would have been reported as a duplication within Asgard using the authors' approach.
2. Such a **selection restricted the paper to make comparisons only with TACK for the extent of duplication**. The comparison itself has two main issues: 1) The authors selected **17 TACK vs 31 Asgard**, and while the gene family selection was based on 12 Asgard without considering TACK-specific genes. 2) **The selected TACK lineages mostly have small genomes averaging 1.5Mb,** almost half of the average smaller than most Asgard. These factors can lead to strong statistical bias. More importantly, even without considering the statistical bias, the main conclusion of the paper was made based on a comparison with TACK without considering other bacteria and archaea with larger genomes. I feel that it is far from sufficient.

To make the conclusions written in this paper, I think it is important that the authors make a **broader and more even taxonomic sampling**, as well as providing **sufficient evidence to exclude potential HGT that disguise as duplication** from outside the selected lineages. Please do pay attention to **potential biases in genome sizes and gene counts** when making quantitative comparisons.

Response: We thank the reviewer for their comments. We would like to emphasize that the purpose of our study is not to provide an exhaustive catalogue of HGT in Asgards, but rather provide a comparative assessment of the relative impact of HGT with respect to duplication, in this group. The reviewer rightly points that the taxonomic scope of the phylogenetic analysis limits the range of HGT that could be detected. We agree with that. However, we note that the long-range (inter-domain) HGT is covered by a different, multi-step analysis that starts with blast searches and ends with phylogenetic assessment. In this analysis there is no limit other than the entire sequence database.

Nevertheless the reviewer is also right in that the imbalance between the selected Bacteria, Asgard and TACK genomes in the phylogenetic dataset may affect our results. We also agree on this point. Our choice was the result of the focal interest in Asgards and the computational constraints to reconstruct large phylogenies of sufficient quality. We have now re-assessed this choice considering the reviewer's concerns, as explained below.

We have now crafted a new dataset of 195 genomes, comprising 32 Asgard archaea (one per family, except for the Lokiarchaeia MK-D1 family that contains both cultured isolates) and 28 TACK (one per order, as TACK is a superphylum-level clade and selecting one genome per family would become unfeasible with 108 genomes). We complement this with one representative per phylum for the rest of prokaryotes. In this way we significantly increase the diversity of other prokaryotes, decreasing the odds of HGT being misdirected as a duplication This dataset is also more balanced in terms of the Asgard-to-TACK ratio, both for the reconciliation approach (which uses the gene trees computed from the complete dataset) and the phylomes (for which we select the 60 proteomes belonging to Asgard and TACK, and select 10 seeds for Asgard and 10 for TACK to perform 20 phylomes). .

We are aware that Asgard archaea have significantly larger genome sizes than TACK (in fact, they have significantly larger genome sizes than any other archaeal phylum), as this fact triggered our question about their genome evolution. We want to note that when comparing duplications, we normalize by using the duplication rate (that is, number of duplications divided by number of gene families present), not the absolute number of duplication events, as we do for the transfer rate, to account for differences in genome size. We tested normalizing the duplications and transfers by speciations, as performed in (https://doi.org/10.1101/2025.11.11.687807), with no effect in the conclusions.

Additionally, to minimise potential HGT that disguise as duplication, we now remove the basal nodes for each clade (Asgard and TACK) from the nodes we consider for the duplication rate assessment.

Finally, the reviewer mentions that we base our conclusions on the TACK-versus-Asgard comparison, irrespective of other prokaryotes with large genomes. This is correct and this was exactly our purpose. In this work we are not addressing what mechanisms render large genomes across prokaryotes (an interesting question but out of our scope), but rather specifically on how Asgards specifically enlarged their genome. That is why we consider that comparison to the closest sister group is the best approach here. More specifically we address this question in relation to how this change in genome evolution may relate to the origin of Eukaryotes. We hope the reviewer understands our focus and agrees that a global comparison of duplication rates across prokaryotes is not necessary here. We now rewrote part of our

introduction to better stress the purpose of our analyses and the adequacy of the datasets used.

Other major concerns:
Novelty of the paper: 1) On the topic of gene duplication in Asgard archaea, Vosseberg et al 2020 Nature Eco & Evo, where the last authors of this article T. Gabaldon was a corresponding author, has specifically highlighted such a case. It is strange to me that the authors did not refer to this paper. In this 2020 aper, it was already concluded that Asgard archaea exhibited most duplications, so I suggest that the authors **discuss the current work in the context of their previous paper,** and state clearly what is new and different.

Response: The focus of the Vosseberg et. al. paper and this one are totally different. In that paper only duplications within the Last Eukaryotic Common Ancestor (LECA) lineage are considered. So the reference to Asgard in that paper relates to the origin of these eukaryotic lineages, and all quantified duplications occur within eukaryotes, not in Asgard. This work, in contrast, focuses on Asgard evolution and the duplications tracked are within the Asgard lineage and therefore distinct from the ones reported in Vosseberg et. al. This is why that article was not mentioned, and we do not think that study detracts any novelty from this one. Nevertheless, we now comment on that finding, as we think is interesting to discuss that the duplication trend in Asgard genes is maintained in their descendants within eukaryotes.

2) The main conclusion for the HGT side seems to be that the metabolic functions are more prominently involved, which does not seem too surprising. Beyond that, the authors seem to follow with various detailed descriptions, which I do appreciate but unfortunately show **little implications for eukaryogenesis**. Can the authors perhaps provide more insights into the functional consequences of those HGTs and potential link to eukaryogenesis?

Response: We believe that the conclusion that HGT predominantly involves metabolic functions is, on itself, a result with implications for eukaryogenesis. Given that, for the reconstructions of the Last Eukaryotic Common Ancestor, the gene families with metabolic functions seem to be predominantly from bacteria, whereas the archaeal contribution is enriched in informational genes (see https://doi.org/10.1101/2024.10.14.618062), the fact that most HGT in Asgard archaea is metabolic, combined with the notion that we did not detect among these HGT genes any that we could confidently trace to LECA leads us to believe that, for the families in LECA whose origin can be traced to bacteria outside alphaproteobacteria, we have to imagine scenarios outside HGT to Asgard and subsequent vertical inheritance in eukaryotes through the Asgard ancestor, and therefore lend likelihood to direct transfers to the proto-eukaryote. We now better stress this in the revised manuscript.

The taxonomic sampling is also a little strange: Even with as few as 61 or 63 genomes, the authors **selected MAGs that belong to the same species or genus**: two Baldr from g__DXJG01, two Heimdall from g__UBA460, 3 Hermod from g__Hermodarchaeum, 3 Hod from g__Hodarchaeum, two Jord from g__DTBI01, two Odin from g__LCB-4, 2 Borr from the same species s__Borrarchaeum weybense, two Wokong from species s__Wukongarchaeum yapensis. Same goes for two g__Nitrosopumilus. In addition, is there a r**eason not to use circularized Freya/Atabey/Heimdall genomes available for analyses**? In the SI table file, I

can only find the information for the Asgard and TACK genomes but not the remaining 15, so I cannot judge for the rest.

Response: We agree with this criticism. The selection was made purely following genome quality criteria, without assessing the final taxonomic choice within a lineage. To solve this, we have now performed a more taxonomically balanced sampling of Asgard genomes (one per family-level clade *sensu* GTDB), in addition to the increased sampling mentioned above. We assigned each assembly a quality score (CheckM completeness - 5*CheckM contamination) and selected one genome per family, with the exception of the replacement of f__MK-D1 representative with the isolates *Ca.* Prometheoarchaeum syntrophicum and *Ca.* Lokiarchaeum ossiferum. We have additionally taken note of the reviewer's comment and added the highest quality circularized Heimdall genomes from the two available at (PMID: 35027677), the highest quality Atabey and the Freya from (PMID: 39406503).

We now include the associated GTDB metadata for all the selected genomes (bacterial and archaeal), along with their quality score, in Supplementary Table 1.

The authors suggest that there is higher rate of intra-Asgard HGT of ESPs than other gene families. Given that most of the ESPs only exist in the Asgard archaea, could this also be statistically biased?

Response: When comparing the transfers of ESPs vs the general background of transfers, we only take into account the transfer events that have occurred within Asgard nodes. Therefore that they are Asgard-specific should not be a bias as we only compare with proteins present, and sum the number of events across all Asgard nodes per OG. Additionally, given that the number of OGs per category (ESPs vs non-ESPs) is very different, we apply a permutation test rather than a t-test or Wilcoxon rank sum test, which we also hope alleviates the potential statistical biases.

Other minor and general issues:
The two parts of the paper used different approaches. Have the authors cross-compared the duplication events concluded from the method 1 with the HGT events concluded from method 2?

Response: The different HGT-detection approaches have different target scopes and are necessarily different as explained above. One of the methods (reconciliation) jointly predicts both HGT and duplications. When two alternative approaches can be used in the same dataset (as species-overlap and reconciliation based duplication inference) we compare results. The reviewer seems to ask for a comparison of results from a duplication-inference and HGT inference. We have now done this comparison. The HGT events from method 2 affect 980 Orthogroups that, on average, have a significantly higher number of duplications ($p = 1.0413e-67$) as well as transfers ($p = 2.2000e-16$). We now report these results.

Line 321, PMIDs, replace with real reference.

Response: We apologize, this is now corrected.

Line 269-271, please consistently use Desulfoproteobacteriota for SRBs.

Response: We have clarified this in the manuscript.

Line 340-343, This is interesting. who are the hosts of those viruses? Can you provide some clear examples?

Response: They belong to phylum Nucleocytoviricota (Algavirales, Imitervirales, Pimascovirales) which mainly infect microbial eukaryotes of varied phylogenetic affiliations (amoebozoa, and other protists). We now discuss this finding and provide the trees and the taxonomies of the viruses in the data repository.

Line 254: (D)->(E)

Response: This has now been corrected.

Reviewer #2 (Remarks to the Author):

In this work, Manzano-Morales and Gabaldon performed a comparative phylogenomic analysis of Asgard archaea and their sister group TACK and found that genome evolution in Asgard archaea is characterized by high rates of gene duplication and horizontal gene transfer (HGT). Such blend of prokaryotic-like HGT rates and eukaryotic-like gene duplication is a **remarkable finding** that, in my opinion, will be **of interest to many potential readers of Nature Communications**.

The analysis is **comprehensive and thorough**, and the authors made a **great effort to ensure that their findings are not due to contamination or genome incompleteness** by applying **rigorous filters** and combining **multiple lines of evidence**. There are some details that require **clarification or could be better explained** (see below), but **I do not think that these compromise the integrity of the manuscript**.

Response: We thank the reviewer for the appreciation of our work, we now made an effort to clarify the points raised by the reviewer.

Some general comments and questions:

1. I wonder if the authors could **go one step further and provide some insight into the evolutionary regime that characterizes genome evolution** in Asgard archaea. To contextualize this comment: Most **bacterial genomes evolve in a regime of deletion-driven** genome reduction, such that **useful genes are maintained** by selection. Some manifestations of that regime are a **positive association between genome size and effective population size**, and a **scarcity of psedogenes in bacterial genomes**. Instead, **eukaryotic genomes evolve in a non-adaptive regime** driven by gene duplication and the **proliferation of transposons and other mobile genetic elements**. Such non-adaptive regime results in a **negative association between genome size and effective population size**, as well as accumulation of noncoding DNA. **Estimating the effective population size is**

complicated, possibly unfeasible if only one high-quality genome per species is available. However, it would be interesting to use the **ancestral reconstructions to investigate how genome size has changed along the evolution of Asgard** archaea and **compare it with the TACK group**. Have **high duplication rates resulted in a net increase** in genome size? **Were duplications compensated for by higher rates** of gene loss? If there has been a **net increase in genome size**, did it result from higher numbers of functional genes or from **pseudogenes and mobile genetic elements**? In that regard, it may be worth assessing the abundance of pseudogenes and the rates of pseudogenization and pseudogene decay.

Response: The reviewer points out a very interesting biological insight that can be inferred from the reconciliation analysis, and we thank them for the suggestion. As the reviewer mentions, it is unfeasible to estimate population sizes in the Asgard clade, given the scarcity of data. We can, however, assess the genome size in relation to the duplication rate across both the TACK group and Asgard group. We now show these results in Supplementary Figure 2, and discuss them in a new section in the Supplementary Discussion. In brief, we see that high duplication rates tend to result in a net increase in genome size. Number of duplications and number of orthogroups present tend to linearly correlate both for TACK and for Asgards (but less so for TACK), but the number of duplications is not correlated with the number losses for both cases.

As for the comment on Mobile Genetic Elements (MGEs), we analyzed the copy number of orthogroups associated with MGEs (see Supplementary Methods) and found that Asgard archaea tend to have more copies than TACK. We have now added this information to the manuscript.

2. I found it surprising that there were **no signs of transposons, proviruses, and other integrated mobile genetic elements among the duplicated and horizontally transferred genes**. Is the lack of MGE duplication and transfer a consequence of the conservative filters applied to identify HGT and duplications or is it a genuine finding?

Response: For the functional analysis of duplicated genes, we have based our conclusions in the KEGG Pathways. These unfortunately do not have a particular pathway for MGEs and therefore they would go undetected in our enrichments. When analyzing the COG categories of the Orthogroups, however, we see that X (mobilome) is the highest category after U (Intracellular trafficking, secretion and vesicular transport) with a mean of 6.68 duplications across Asgard nodes for these orthogroups. We now comment on this in the article. Also, related to the previous question, these MGEs seem to accumulate more in Asgard than in its sister clade Thermoproteota.

For the HGT genes, we could add GO term enrichment, which does include terms for integrases and trasposons, some of which are enriched in respect to the background of their respective lineagesdoes not pass FDR correction. We also see an increase in the proportion of the COG category X in HGT genes related to the genomic background for some lineages.
Some points that need to be clarified or better explained:

3. Lines 69-73. After reading this paragraph it is not clear if those MAGs were used or not.

Response: We mention that MAGs can suffer from genome incompleteness and this might affect the downstream conclusions obtained from their analysis. However, there is a scarcity of genomes available in the Asgard clade that come from cultured isolates, which forces analyses on Asgard archaea to use MAGs, therefore, we do use them. However, being aware of this possible caveat, we performed the same analyses on alternative datasets (chosen at random, instead of following quality criteria) to ensure robustness of the results.

4. Lines 85-86: How is it that only 17K out of 31K ***Asgard archaeal*** OGs contain at least 1 Asgard representative? What about the other 14K? Perhaps I am missing something: what do you exactly mean by Asgard representative?

Response: The reviewer is right, this was an error on our part in the writing of the manuscript. There are 31K orthogroups total, out of which 17K contain at least 1 Asgard representative, the rest of them being orthogroups formed by representatives of the other bacteria and archaea in the dataset. This has been corrected in the new version of the manuscript.

5. Lines 91-92: Did you only include KEGG pathways or also KEGG ontologies?

Response: We annotated each protein with the KEGG Ortholog (KO) assigned by eggNOG-mapper, calculated the majority KO per orthogroup (see Methods) and performed a KEGG Pathway enrichment analysis of these KOs. So, in this sense, for the functional enrichment we took into account KEGG pathways.

6. Lines 108-109: The finding that Chao's lower bound and core genome size are sensitive to genome incompleteness is not surprising. Chao's lower bound depends on the number of singletons and OGs with 2 members, hence it is a not quite robust estimator. As for the core genome, the reasons are trivial (but see for example PMID:37904249 for a quantitative assessment). In the light of these results, it may be worth running mOTUpan on both sets of genomes to get a more robust estimate of the core genome size and pangenome fluidity.

Response: We thank the reviewer for the suggestion. We have now included the analysis of mOTUpan in the paper.

7. Line 110: I am not sure of what you mean by the "full set". Do you mean all available MAGs? It has been shown that orthofinder performs quite poorly when applied to incomplete MAGs. I also wonder if the analysis of the full dataset is really necessary: given the openness of the pangenome and looking at the rarefaction curve in Fig 1 it is sufficiently clear that there are a lot of Asgard archaeal genes yet to discover and that non-HQ MAGs will contain genes that are not present in the representative set.

Response: The ‘full set’ refers to the full collection of species representatives available in this version of GTDB for Asgard archaea. We agree that MAG incompleteness could cause biases in the estimates, that is why we mainly considered the set of high-quality representatives and only used this ‘full set’ as a complementary analysis. We have removed this part from the manuscript.

8. Figure 1A: This plot would be more informative if it used a logarithmic scale for the y-axis.

Response: Thank you for the suggestion, we have now modified the figure accordingly

9. Lines 120-121: Here "entire genome set" means the representative HQ genomes, right? Which is not the same as the "full genome set" mentioned in line 110. That is confusing.

Response: The reviewer is correct, it means the set of HQ genomes. We apologize for the confusion and have re-written this in the manuscript for clarity.

10. Figure 1C: What is the hub of the PPI network (K12158)? What are the proteins that form the 2 triangles?

Response: K12158 is an ubiquitin-like protein (Nedd8). For clarity, we have now labeled the nodes with their protein name, rather than the KO ID. As we have increased the dataset, the network has changed slightly.

11. Line 129: In the absence of a null model, to say that the network is highly connected is quite subjective. Less than half of the proteins have interactions and, if we only look at those, more than half of the interactions have a low score (around 200).

Response: We agree with the reviewer that the statement can come across as subjective. We have now performed a PPI enrichment score of the protein set against the full set of PPIs in the proteome as implemented in STRING.

12. Line 187: I am not sure of what you mean by "a value of 1 would indicate a whole genome duplication". In the figure there are cases of 10: I do not suppose that those represent 10 whole genome duplications, right? Adding to the confusion, the caption says "mean number of duplications per node" but the label of the x-axis says "mean number of duplications per OG". That requires clarification.

Response: Figure 2B represents the frequency distribution across orthogroups of the number of duplication events per orthogroup. A value of 1 implies that the given orthogroup has duplicated on average once in the Asgard clade, whereas 10 means a mean number of 10 duplication events. The reviewer is right that a value of 1 does not mean a whole genome duplication: this was a mistake on our part from previous versions of the figure when Figure 2B represented the mean number of duplication events per node (not per orthogroup).We apologize for this mistake and have now corrected the text.

13. Figure 2: What are the rate units for the duplication and transfer rates? Events per unit of time (what unit of time)? Per gene? Per genome?

Response: The duplication and transfer rates are calculated as the total number of duplications, or transfers, relative to the number of OGs present at the given node. We also tested normalizing by speciations, as in https://doi.org/10.1101/2025.11.11.687807, with no effect on the comparisons versus its sister clade. This is now clarified in the text.

14. Figure 2E: Do the distributions correspond to nodes or OGs? Equivalently, do the tails of the distribution correspond to OGs that duplicate a lot or genomes with a lot of duplications?

Response: The distributions correspond to duplication rates per node, that is, number of duplication events relative to the total number of orthogroups present at that given node. The tails represent nodes in which a lot of duplication events have occurred. This is now better clarified in the text.

15. Figure 2F: What is the definition of verticality?

Response: We have now removed this panel, as the information it provides was redundant with the duplication rate and transfer rate (Fig 1C,D) and created confusion.

16. Figure 2G (the caption says F): Something looks wrong with the boxplot and overlapping scatter plots. The "other" category is clearly not normally distributed, so it would probably be better to use logarithmically transformed data. The same applies to the statistical test: did you use a robust test or transformed data? What are the effect size and p-value? These are important details that must be disclosed (not only for panel G, but also for panels C, D, F and Fig 3E). The biological implications (higher transfer of ESP than other genes), if confirmed by statistical analyses, are certainly relevant.

Response: We were using Wilcoxon rank sum test as the data was not normally distributed. To ensure that we are performing the correct tests for statistical significance of the data, we have now selected the appropriate test using the R package “groupcompare”. We have now included the information about p-values and effect sizes in the text. We agree that this result is relevant and we highlight this more now in the manuscript.

17. Lines 193-196: When comparing the duplication rate of lineage signature proteins within and outside the lineage, did you control for the lineage-wise duplication rates? Did lineage signature proteins also duplicate more than other (non-signature) OGs within the same lineage?

Response: We did not. We appreciate this point. In the updated version, we normalize the duplications each of these OGs suffers per node by the duplication rate of that given node. We thank the reviewer for noting this.

18. Line 232 and below: Do the methods used to infer inter-domain HGT account for the fact that proteins can evolve at very different rates in bacteria, archaea and eukaryotes? Would it be possible to obtain reliable estimates of the fraction of false positives?

Response: Our method for the inference of inter-domain HGT events is based on the topology of the gene trees. For the inference of the trees, we incorporate into the model selection mixture models that take into account site heterogeneity, with which we expect to alleviate Long Branch Attraction, which could be an effect of the differential evolutionary rates. We cannot think of a way to estimate the fraction of FPs, and the analysis we apply are the state of the art in the field.

19. Figure 3: Do all panels in Fig 3 refer to Bacteria-to-Asgard HGT?

Response: All panels correspond to inter-domain HGT. Figures A, B and C do not differentiate directionality (both Bacteria-to-Asgard and Asgard-to-Bacteria are considered) whereas D and E, as they involve analysis within the Asgard proteome, deal only with Bacteria-to-Asgard transfer events. We use 'partner' throughout the text when directionality is not considered, and 'donor/acceptor' when it is. This is now better clarified in the text.

20. Fig 3A: what does "contribution of different seeds" mean?

Response: We analyze these HGT events starting from a set of 'seed' genomes, from which we pre-select candidate proteins of HGT origin and then build a gene tree based on a BLAST search of that protein. Therefore, the gene trees are seed based. The pie chart represents the proportion of proteins from each given 'seed' proteome whose gene trees point to HGT events occurring at that given node. We wanted to confirm that the HGT signal for the internal nodes does not come from only one lineage, but rather, that it is backed by the signal in several proteins independently.

21. Figure 3D: Besides the KEGG and KO terms, it would be helpful to see a higher-order classification of the transferred genes, for example in terms of COG categories.

Response: Thank you for the suggestion. We have added this in the new version of the manuscript (Figure 3)

22. Fig 3E, bottom panel: Why are the degrees always <1? Does it mean that most proteins have degree 0 (no interactions)? But then, how is it that in the top panel the Q25 is already well above 0?

Response: Given that there are less HGT proteins than proteins of vertical inheritance, we normalized the degrees by the total number of each category. We have clarified this in the legend and apologize for the inconvenience.

23. Line 266: Figure 3C would be more informative if instead of the number of taxa it showed which taxa acted as donors. That information is not currently shown here or the supplement, although some donor taxa are mentioned in the main text. If adding that to Fig 3 is not feasible, a summary table in the supplement would work too.

Response: Thank you again, we have now modified it to show the main contributor phyla

24. Line 294: The degree is only one of several possible measurements of centrality. Have you checked other centrality indices, such as the eigenvector centrality?

Response: We have added eigenvector centrality and betweenness to the analysis. They yield qualitatively the same results. We add this to Supplementary Figure 7.

25. Line 303: Based on phylogenomic reconstructions, did the transfer of these enzymes lead to the replacement of a preexisting enzyme carrying out the same function? Otherwise, did HGT lead to qualitatively new or broader metabolic potential?

Response: Based on manual analysis of the KEGG pathways with KEGG mapping, we saw a mixture of new metabolic potential and functional redundancy.

To perform this analysis in a more systematic way, we have now assessed the KOs brought by the HGT events (Bacteria-to-Asgard only) into the genome and found that, on average, around 40% of the KOs found in HGT genes are also found in a copy with vertical inheritance in the same genome, suggesting at least a degree of functional redundancy. Despite this, the majority of them still suppose novel KOs.

To go deeper into the extent to which these novel KOs constitute novel metabolic capabilities, we took the KOs and estimated the metabolic pathways they might form with anvi'o (anvi-estimate-metabolism). The resulting modules have a mean pathwise completeness of 0.17(of 1), reinforcing the view of an overall addition to parts of existing metabolic pathways rather than the addition of whole new metabolic pathways to the organism. Nevertheless, we found, some modules in *P. syntrophicum* and *L. ossiferum*that could be considered stepwise complete and are comprised entirely of KOs of HGT origin. These new results are now commented on in the text.

Online Methods

26. Line 26, regarding the lowest CheckM contamination criterion: Was it only applied to break ties? Did you apply any hard threshold on the contamination score? Contamination would result in apparent duplication and HGT, so, for these purposes, minimizing contamination seems more important than maximizing completeness.

Response: We agree with the reviewer and we now used an index that better balances completeness and contamination, based on the quality score of the GTDB (PMID: 38512750) which is CheckM completeness - 5*(CheckM contamination).

27. Line 41: How did you harmonize the need to work with homogeneous genome annotations (to control for biases and annotation-related confounding factors) while keeping the original genome annotations?

Response: We apologize for the confusion this may have caused. As for 'genome annotation' we mean the detection of protein-coding genes within a genome, for which Prokka is a widely used tool. As for the functional annotation of these protein-coding genes, we re-annotated them from scratch by mapping them to eggNOG OGs with eggNOG-mapper and to GO terms and Pfam domains with InterProScan, disregarding any annotations performed by Prokka or other tools so that all of them were annotated by the same method. As for the gene calling, we have harmonized this by annotating all the genomes with Prokka. We would have liked to take the protein files from GTDB directly, but unfortunately they do not provide the GFFs we need for the genomic context analysis.

We compared the number of called proteins with the number of proteins in the FASTA files from GTDB, resulting in a median difference of -47.00 proteins, or around -1.94% (relative to the number in the GTDB file). Given that the GTDB annotation is done by Prodigal v2.6.3 and Prokka uses Prodigal under the hood, we expect the differences to stem from the fact that

Prokka applies various filtering to its Prodigal results, and has a mode more tailored to metagenomes, and thus we believe the Prokka annotation to be more accurate.

28. Line 51: Please, provide a brief summary of the annotation workflow (Ref 33).

Response: We now provide a brief explanation in the text, after the citation.

29. Line 53: FLUH has not been defined. Furthermore, this paragraph is quite cryptic. What does it refer to?

Response: We have now clarified that FLUH refers to “Free-Living Unicellular Heterotroph”. We thank the reviewer for pointing this out as it may be confusing.

30. Line 59 and below: Were the alignments and trees based on nucleotide or amino acid sequences?
Response: We only employed amino acid sequences for all phylogenies.

31. Line 176: Did you check that putative instances of tandem duplications are not due to fragments of the same gene interrupted by a stop codon (either real or resulting from a sequencing error)?

Response: That is theoretically possible, although we expect it to be rare. For this to happen, both fragments should have hits in BLAST to the same protein from the seed genome A which are in the top 150 hits that pass e-value < 1e–05 and a continuous overlap over 50% of the query sequence. To check for this, we assessed whether the two proteins in the tandem have their hit with e-value < 1e-5 and continuous overlap >= 50% in the same range of the query protein. For the cases with overlap, we assessed what percentage of the total alignment with the query protein (with respect to the shortest alignment of the two) they represent. We found that there’s only 4.28% where there’s no overlap, and when there’s overlap, the median overlap is 94.93% (sd 10.57). We filter out cases where this overlap is <= 50%, to err on the side of caution.

32. Extended Data Figure 3: The distribution of the number of transfers (upper left) would be more informative if it was plotted in log-log or log-linear scale.

Response: Thank you for this suggestion. We have modified the figure accordingly

33. Line 311: This caption refers to folate metabolism, not carbon metabolism.

Response: Corrected

Typos

- Line 264: Data Figure 4A (instead of 3C).

Response: Corrected

- Line 345: High (instead of highly).

Response: Corrected

- Online methods, lines 68-69: the sentence in parentheses is repeated.

Response: Corrected

Reviewer #3 (Remarks to the Author):

The paper presents the results of a phylogenomic study focused on Asgard archaea. The main findings relate to patterns of gene duplication and HGT and their functional and genomic impacts. The findings are **interesting and shed light on some unique aspects of Asgard evolution** with implications for understanding eukaryogenesis. My comments primarily have to do with **thoroughness of the analysis** and appropriateness of the methods used.

Response: We thank the reviewer for the insightful comments on our work.

Major comments:
- There are **several cutoff thresholds** used throughout the paper but these seem to have been chosen somewhat arbitrarily. As an example, **Asgard signature proteins** are defined using thresholds of at least 70% presence in Asgard and at most 15% presence in other archaea or bacteria. Why were these specific thresholds chosen and how do results/findings change if other reasonable thresholds, such as 80% and 10%, are used instead? The same applies to the thresholds used on lines 139-140.

Response: We agree that the cut-off thresholds may come across as arbitrary, we now justify the choice better and have changed one of the thresholds. Pangenomes tend to have an 'U' shape, where most gene families are either very rare or very common. The threshold for the presence 'outside' (15%) corresponds to that typically defined as the cloud component of a pangenome. For the presence 'inside', since the estimation of the core component of the pangenome is sensitive to genome incompleteness, we assessed the prevalence of orthogroups annotated as the COG category "J" (Translation, ribosomal structure and biogenesis) and OGs annotated as KOs from archaeal ribosomal proteins. We used the prevalence of these categories, which we assume to be housekeeping genes, to select the threshold of 86 instead of 70% for what we consider 'widespread' in the Asgard clade (see **Supplementary Methods**).

- It is mentioned that **two different species phylogenies**, constrained and unconstrained, were generated and used for HGT and gene duplication analyses. However, it seems that **results are only presented for one of those phylogenies**. Which phylogeny was used for the reported results? And how do findings change when the alternative phylogeny is used? Results should be shown for both phylogenies, and the phylogenies themselves should be provided as well.

Response: The results shown are the ones for the topology that follows the species tree for Eme et al. (2023). However, since the Asgard species tree is a highly debated and still unresolved topic, we tested the robustness of the conclusions by employing an alternative

topology for the Asgard clade, for which we used the GTDB archaeal tree topology. In this version of the manuscript, we could not let the topology fully unconstrained as it broke the monophyly of Thermoproteota. The results for the alternative topology are qualitatively similar. We have added them as Supplementary Figure 1.

- It is mentioned that **two methods were used for duplication** analysis, but results are reported only for the reconciliation-based method. Results **should also be shown for the other (species overlap) method.** Are all reported results supported by both methods?

Response: We assessed the correlation between duplication rates for the reconciliation and species-overlap approach and found them to be correlated, and the results to be qualitatively the same. We now add the resulting plot for the species-overlap method as a supplementary image (Supplementary Figure 3) so readers can compare.

- Same as above, there also appear to be multiple methods used for **HGT inference**. Are all reported results consistent between these different analyses (e.g., reconciliation-based and non-reconciliation-based)?

Response: We believe the results for the two approaches for HGT inference are complementary, yet not directly comparable. For example, when we apply the search against a broad-DB and posterior phylogenetic pipeline, we are exclusively looking for inter-domain (Bacteria-Asgard) HGT events, and therefore we consciously overlook intra-Archaea or intra-Asgard HGT. On the contrary, due to the relatively scarce sampling of bacteria for the reconciliation approach, we focus on intra-Asgard events, as Bacteria-Asgard transfers are likely underestimated due to scarcity of bacterial genomes. However, we did see that the OGs to which the proteins flagged as HGT belong have a significantly higher number of transfer events than the rest of OGs.

- The AleRax tool is used for **reconciliation-based** analysis. It appears that the **species tree generated by AleRax was used to perform this analysis**. Do we now have a third species tree that is different from the constrained and unconstrained species phylogenies used for the other analyses? If so, this needs to be clearly stated in the paper. And how do results change if the two previous species phylogenies are used for this reconciliation analysis? The use of multiple species trees makes it all the more important to clear report how findings change as different reasonable species trees are used for the analysis.

Response: We apologize for the misunderstanding. The AleRax analysis was performed with the Eme et al. (2023) topology, and with the 'alternative' topology built from the GTDB archaeal tree. There is no third species tree topology. We now better clarify this in the manuscript.

- Some of the results, such as **elevated duplication rates at the base** of the Asgards, may be **artifacts of errors in the gene trees**. Paying further attention to **gene tree quality** may help resolve some of these artifacts.

Response: We agree that the duplication rates at the base of Asgard may be inflated. To minimize the impact of these artefacts, we now have removed the MRCA of Asgard and Thermoproteota, respectively, from the duplication rate analysis. Moreover, we expect that

with the increased size in the dataset these errors are diminished, as we increase the sampling outside the Asgard group.

As for the tree quality comment. Given that AleRax considers the breadth of bootstrap trees in its analysis, we think that the variability in gene tree quality is inherently taken into account (the more noisy trees the lower bootstrap and variability across bootstrap trees).

- In the **time of acquisition** analysis of Lines 259-264, using **last common ancestors** of all genomes carrying the transferred gene is almost guaranteed to provide **incorrect time estimates** since such genes may have been acquired multiple times or may have been transferred again after initial acquisition. A much more accurate way would be to **directly use the reconciliation** to estimate when a gene was first acquired.

Response: The computation time for AleRax is roughly cubic to the number of species. Given that the database we employ for the homolog search of putative HGTs in Asgard contains all species representatives of GTDB v202, as well as eukaryotic species, this becomes computationally prohibitive. Moreover, our aim when analyzing the HGT events in Asgard is to assess how much of their genome is comprised of genes of HGT origin and the transfer partners. We agree that the last common ancestor of all transfers (i.e. maximum parsimony) might not be totally accurate. However, given the computational barriers of using reconciliation at this scale we think it is a good proxy. Moreover, the only conclusion we infer from this temporal analysis is that transfers have occurred as a continuum rather than in a single wave. We do not think a reconciliation analysis would change that qualitative conclusion in any way. Nevertheless, we now discuss the limitation of the Maximum Parsimony assumption.

- Line 87: where are the 8365 OGs found only in a single Asgard genome coming from? A BLAST search against a larger collection of prokaryotic and eukaryotic genomes could help resolve this question.

Response: This is an interesting question, but as we performed this analysis for a broad characterization of the Asgard gene repertoire, we felt that an in-depth analysis of the singletons was beyond the scope of the paper.

- Line 83 of online methods supplementary text: why are only the **top 150 hits chosen per protein?** Artificial cutoffs such as this can result in under-clustering, leading to significant errors in downstream analyses and especially in reconciliation. This analysis should be repeated with a more biologically grounded approach for identifying gene families.

Response: This cut-off is due to constraints of the phylome approach and the computational load of the metaphylome, this threshold is justified elsewhere (PMID: 34718760). We agree that it is not ideal. In this version of the manuscript, while we maintain this cut-off as part of the phylome pipeline, the trees for the reconciliation step are built from OrthoFinder orthogroups, providing a complementary source of information that lessens these biases.

- Lines 95-101 of online methods supplementary text: The description of how and on what data AleRax was run is unclear. What is meant by “we selected one sequence per OG present in the Asgard clade”?

Response: Initially, the gene trees employed by the reconciliation were chosen among the gene trees calculated by the phylome. We have now changed this and calculated the gene trees directly from the OrthoFinder OGs. We have improved this section of the methods so it is clearer.

Minor comments:
- Increase font size of node annotations in Figure 1.
Response: We have corrected this in the manuscript

- Typo on line 345, highly dynamism.
Response: We appreciate the reviewer for pointing out this typo. It has been corrected in the manuscript

- Lines 133-134: why were only 39 out of the 83 previously identified proteins found? Is it because only a subset of available Asgard genomes was used in the analysis?

Response: Our intuition is that the rest of the proteins were not found due to the search being constrained to a subset of genomes, yes. With the new set of genomes, we have found more (44). We have additionally relaxed our score criteria to encompass those that may have been priorly overlooked due to poor scores.

- Lines 146-147: Why not use all 31 Asgard genomes for this analysis? How were the 12 genomes chosen?

Response: We restricted the set of genomes out of computational constraints. We selected the highest-quality genomes while maintaining a taxonomic balance.

- In the analysis of tandem duplications (Lines 198-208), were phylogenetic relationships between the genomes considered? Ancestral tandem duplications would be inherited by multiple genomes and may be counted multiple times, possibly impacting the results shown in Extended Data Figure 2.

Response: Extended Data Figure 2 (now Supplementary Figure 5) represents the percentage of proteins in an organism's proteome that have been flagged as a tandem duplication, by dividing the proteins in the GFF into two categories. Therefore each protein is counted only once per genome, irrespective of the number of duplication nodes it appears in. Out of this percentage per proteome, a given proportion will be shared with other proteomes and the rest would be specific, but as we are seeing the proportion per proteome, each protein is counted only once

- Line 59 of online methods supplementary text. Which markers were used?

Response: We used the markers employed by GTDB to build their archaeal taxonomy.

- Line 142 of online methods supplementary text. There are better rooting methods than midpoint rooting. The authors could consider using Minimum Ancestor Deviation (MAD) rooting.

Response: We have taken this into account. We tested both rooting methods and found an overlap of 94% in the genes we detect as HGT. In any case, we take note of the reviewer's advice and use MAD rooting in the resubmitted version. We have added a section in the Supplementary Methods to discuss this.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have addressed the previous comments and made substantial improvement to the analyses and clarification of the paper. I have only a few minor comments on the revised manuscript:

1. I recommend that the authors refer to the original articles reporting the 32 representative Asgard archaeal genomes.

Response: We appreciate the reviewer's comment. We note that some of the assemblies do not have associated publications. For these, the BioProject is supplied in ST1. For the genomes that have publications we have added the corresponding citations to the “Data Availability” section.

2. The authors used Em et al 2023 tree for reconciliation, does it indeed include all the species used in this study?

Response: We used the Eme et al. tree for the class-level topology of the lineages, as it indeed does not contain all the species used in this study. We took the GTDB archaeal tree and re-arranged the class-level branches to fit the Eme et al. topology, as the topology of the GTDB archaeal tree was somewhat discordant. We have clarified this in the text.

3. Note that Deltaproteria is now designated into various Desulfobacterota clades in GTDB, Desulfoproteobacterota-> Desulfobacterota. Please be consistent throughout the paper.

Response: We changed “Desulfobacterota” to “Desulfoproteobacterota” upon a previous request for consistent use of “Desulfoproteobacterota” for SRBs. However, we are strictly employing the GTDB taxonomy framework, in which the phylum name is “Desulfobacterota”. We are now using “Desulfobacterota” throughout to comply with the taxonomy *sensu* GTDB.

4. line 431, 'high dynamism in Asgard genome'. It is unclear what it means, please rephrase.

Response: We have modified the text to clarify this respect.

5. Figure 3A. What do all the colors of the pie charts stand for? what is Njordarchaeia.1?

Response: The colors in the pie chart represent the different seed genomes. We have colored the tip labels as a legend. We apologize for the .1 in Njordarchaeia, it is indeed a typo.

6. Please properly label the species names and accession numbers on the tips of phylogenomic trees in Fig. S1, S3 and S6.

Response: SF1 contains the clade names, we feel that adding the species names and accession numbers would make the figure too overcrowded. In any case, we have added the requested labels to SF3, which contains both the Eme et al (2023) and the GTDB topologies. Given that the tree is the same, it should not be difficult for a reader to find the label(s) of their interest in SF3 and extrapolate that to SF1. We have added accession numbers to SF6 to not make the figure too crowded as well.

Reviewer #1 (Remarks on code availability):

The page does not seem to be accessible at the moment.

Response: We apologize for the inconvenience. The GitHub repository is currently set to private. We have left a copy of the code in the Zenodo for review. https://zenodo.org/records/15789115?preview=1&token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6Ijc5OWY3YjZkLWJlNWYtNGY0My1hYmViLTQyM2VlNmExYzRjMCIsImRhdGEiOnt9LCJyYW5kb20iOiI5YjQzYjUyMzU5NTI3NWFjNWZiZGEyYzU0OGQ1ZTFjYyJ9.MVQ7lZimHu4O9lK2xWsxnzDqwd_SbiAzPKp_WKO2E_GGkAFFtlKEdk_M72p_5leISqCqfN6ZuH-TMWs8aOnb2Q

Reviewer #2 (Remarks to the Author):

Most of my previous comments have been addressed in the revised version. I only have a minor request:

Could you please comment on the (putative) functionality of duplicated genes? Is there any evidence of pseudogenization (e.g. truncating mutations, high variance of ORF lengths)? Are they subject to positive, relaxed, or negative selection? With bacteria in mind, one would expect that, if duplicated genes have persisted for a long time, they are probably functional (simply because pseudogenes are quickly eroded in bacteria). But these are not bacteria and it seems that gene loss is less frequent than gene duplication/gain in Asgard, so it might be that persistence does not imply functionality.

Response: To gain insight into the putative functionality of the duplicated genes, we have assessed the variance in ORF lengths (the ratio between the CDS, measured as longest/shortest)

and their dn/ds ratio. For this, we have taken the tandem duplications as a representative set. For the dn/ds ratio, we see that for Asgard duplications for which we could obtain this ratio (n = 8694), the dn/ds ratio is of median 0.367 (sd 0.137)

Other than that, I detected some typos and would suggest some edits for clarity:

l. 77: I would suggest to add some connector in the beginning of this paragraph to stress that restricting the pangenome to 32 high-quality Asgard genomes was a necessary compromise to circumvent the potential problems of working with a larger set of incomplete MAGs. One can figure it out from the context, but a more explicit statement would be helpful.

Response: We have modified line 77 to improve readability as suggested by the reviewer.

l. 84: "17,429 contained at least one Asgard protein, 12,397 were exclusively present in Asgard archaea, and 10,470 were present in a single Asgard genome and no other prokaryote in our dataset." Are these numbers nested within each other? I interpret that the 10,470 singletons are included within the 12,397 Asgard-only OGs, and those are included within the 17,429 Asgard OGs. Just in case, it may be a good idea to rephrase a little bit (perhaps adding "of those") to make that point totally clear.

Response: We have modified this sentence to make it easier to understand.

l. 86: "figAsgard" -> possible typo?

Response: The reviewer is right to point it out as a typo, we have corrected this in the new version of the manuscript.

l. 103: "not in terms of core genome" -> I find that statement confusing because the estimated strict core is 18 in both cases. Perhaps the statement refers to the observed strict and relaxed cores (but those are not reported for the random dataset) of for the mOTUpan estimates (but those are not presented until the following paragraph).

Response: Yes, the statement referred to the observed strict and relaxed cores, which are 57 and 269 for the random set, respectively. The reviewer is right to point out that this wasn't discussed in the text. We apologize for this. We have added this data and clarified this comment in the text.

l. 141: "Figure 1D" -> Figure 1C in the updated version.

Response: We have modified the text accordingly.

l. 188: missing space before 23.87%

Response: We have corrected this.

l. 202: "Figure 2E" -> Figure 2D?

Response: We have corrected this.

l. 239: "(F)" -> (G)

Response: We have corrected this.

l. 312: The pie charts appear below (not above) the nodes. The legend linking colors to seeds is missing.

Response: We have colored the tip labels with the corresponding color to act as a legend, as adding a legend for the color would make the figure too crowded.

Reviewer #3 (Remarks to the Author):

The authors have made many changes to the manuscript and have addressed most of my concerns. But there is a related issue that has not been adequately addressed: While the authors have performed additional analysis and reported their findings, some of these new findings are not well integrated into the manuscript. For example, the authors now provide the results of the reconciliation analysis using the alternative topology in Supplementary Figure 1. However, there is no discussion of these results elsewhere in the manuscript.

The same applies to results presented in Supplementary Figure 3.

It would be great if the authors could better integrate the results of these alternative analyses into the manuscript, clearly explaining which findings are supported and not supported by the alternative analyses.

Response: We have now made more effort to comment on the alternative analyses.

Minor comment:

The two topologies shown in Supplementary Figure 3 are missing leaf labels.

Response: We have now added the corresponding leaf labels