Phanerozoic Radiation of Ammonia Oxidizing Bacteria

LM Ward[1,2], DT Johnston[1], and PM Shih[3,4,5]

1. Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, United States
2. Earth-Life Science Institute, Tokyo Institute of Technology, Tokyo, Japan
3. Department of Plant Biology, University of California, Davis, Davis, CA, United States
4. Department of Energy, Feedstocks Division, Joint BioEnergy Institute, Emeryville, CA, United States
5. Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States

Abstract:
The modern nitrogen cycle consists of a web of microbially mediated redox transformations. Among the most crucial steps in this cycle is the oxidation of ammonia to nitrite, an obligately aerobic process performed by a limited number of lineages of bacteria (AOB) and archaea (AOA). As this process has an absolute requirement for O_2, the timing of its evolution – specially as it relates to the Great Oxygenation Event ~2.3 billion years ago – remains contested. To estimate the antiquity of bacterial ammonia oxidation, we performed phylogenetic and molecular clock analyses of AOB. Surprisingly, bacterial ammonia oxidation appears quite young, with crown group clades having originated during Neoproterozoic time (or later) with major radiations occurring during Paleozoic time. These results place the evolution of AOB broadly coincident with the pervasive oxygenation of the deep ocean. The late evolution AOB challenges earlier interpretations of the ancient nitrogen isotope record, predicts a more substantial role for AOA during Precambrian time, and may have implications for understanding of the size and structure of the biogeochemical nitrogen cycle through geologic time.

Keywords: nitrification, comammox, GOE, NOE, Nitrospira, Betaproteobacteria, Gammaproteobacteria

Introduction:

The biogeochemical nitrogen cycle is second only to carbon in size and, arguably, importance for the biosphere (e.g. Tyrrell 1999). The nitrogen cycle supplies fixed nitrogen for biomass while also fueling diverse microbial metabolisms, with fluxes hundreds of teramoles of nitrogen per year (e.g. Canfield et al. 2010). Nitrogen enters this cycle by way of reduced forms (i.e. ammonia fixed from N_2 by the enzyme nitrogenase), and so biological nitrification (i.e. the oxidation of ammonia to nitrite and nitrate) is an essential step for enabling downstream processes such as anammox and denitrification (Zerkle and Mikhail 2017). No metabolism has yet been discovered that is capable of oxidizing ammonia in the absence of O_2 or O_2-derived compounds like nitrite or NO (Hu et al. 2019)—therefore the modern nitrogen cycle where oxidized forms are regenerated and recycled are necessarily tied to O_2.

Aerobic ammonia oxidation is found in a limited, polyphyletic set of Bacteria (AOB) and Archaea (AOA). In both AOB and AOA, the first step in ammonia oxidation is performed via ammonia monooxygenase (AMO), a member of the copper membrane monooxygenase (CuMMO) family. The CuMMO family includes the related particulate methane monooxygenases (pMMO) and enzymes that oxidize other small hydrocarbons (Khadka et al.

2018). CuMMO enzymes have an absolute requirement for O_2, suggesting that metabolic pathways utilizing these enzymes—including ammonia oxidation—evolved after the evolution of oxygenic photosynthesis provided significant O_2 to the biosphere (e.g. Zerkle and Mikhail 2017). The evolution of oxygenic photosynthesis in Cyanobacteria led to the accumulation of atmospheric O_2 to biologically meaningful concentrations ~2.3 billion years ago (Ga) during the Great Oxygenation Event (GOE), and it has been suggested that the aerobic nitrogen cycle was initiated shortly thereafter (Zerkle et al. 2017). However, others have argued from isotopic evidence that an aerobic nitrogen cycle was in place much deeper in Earth history (e.g. Garvin et al. 2009). Distinguishing between these possibilities from the rock record alone is difficult due to the poor preservation of Archean strata and the lack of a robust framework for interpreting the ancient nitrogen isotope record. Instead, the biological record can provide opportunities for querying the antiquity of organisms and metabolisms responsible for driving the nitrogen cycle.

Here, we estimate the antiquity of AOB via phylogenetic and molecular clock analyses. We show that ammonia oxidation in bacteria has evolved convergently at least twice and that crown group AOB clades originated <1 Ga, with major radiations occurring within the last ~500 million years (Ma). This suggests that bacteria did not drive ammonia oxidation until late in Earth history—more than 1.5 Ga after O_2 first accumulated in the atmosphere. The predicted appearance of AOB at a time when Earth surface environments underwent oxygenation to modern-like levels points to the potential role for niche expansion in fostering evolution and boosting turnover of the marine fixed nitrogen inventory. This suggests a substantial difference in scale or structure of the biogeochemical nitrogen cycle during Precambrian time, potentially with a more dominant role for AOA.

Phylogenetic distribution of proteins involved in ammonia oxidation:

Phylogenetic analysis of the distribution of genes associated with ammonia oxidation shows that this metabolism is restricted to four clades of characterized ammonia oxidizers (Figure 1). These include members of the Nitrosphaeria class of Crenarchaeota/Thaumarchaeota, the Nitrococcaceae family within the Gammaproteobacteria, the Nitrosomonadaceae family within the Betaproteobacteria, and some members of *Nitrospira* (Figure 2). These results are based on ammonia monooxygenase and homologous proteins from the copper membrane monooxygenase family (Figure 2, Supplemental Figures 1 and 2). While CuMMO sequences were recovered from diverse lineages including some that have not previously been characterized to possess the capacity for methanotrophy (e.g. members of the UBP10 and Myxococcota phyla), these proteins are most closely related to enzymes that are characterized as performing carbon oxidation (e.g. pMMO, butane monooxygenase), and no organisms outside of characterized clades of ammonia oxidizers were found to encode AMO. In all cases, ammonia oxidation appears to be a derived trait, with basal members of the clades and closely related outgroups lacking the capacity for ammonia oxidation. Importantly, these clades of ammonia oxidizing microorganisms are not closely related and are phylogenetically separated by many lineages incapable of ammonia oxidation (Figure 1).

Phylogenetic analysis of proteins involved in ammonia oxidation, compared to organismal relationships among AOB, provide evidence for convergent evolution and horizontal gene transfer as major drivers for the extant diversity of organisms with the genetic capacity for ammonia oxidation. These relationships are consistent with major clades of AOB acquiring the capacity for ammonia oxidation through separate evolutionary events, followed largely by vertical inheritance within each AOB clade. These data are not consistent with a much more

ancient acquisition of ammonia oxidation (e.g. in the last common ancestor of Nitrosococcales and Nitrosomonadaceae) followed by extensive loss. As a result, the age of total group Nitrosococcales, Nitrosomonadaceae, and the extant diversity of comammox *Nitrospira* can confidently be used to constrain the timing of acquisition of the capacity for ammonia oxidation in each lineage. Additionally, our data are consistent with hypotheses for ammonia oxidation evolving from earlier aerobic methane oxidation pathways (Supplemental Information).

Molecular clock evidence for the late evolution of ammonia oxidizing bacteria

To connect the evolution history of AOB described above to events in Earth history, we performed molecular clock analyses to determine when AOB clades diverged from non-ammonia oxidizing relatives (i.e. age of total groups) and when AOB clades subsequently radiated (i.e. age of crown groups). Molecular clocks estimate the origin of each AOB clade to Neoproterozoic or Phanerozoic time, with each stem group AOB lineage emerging between 238 (comammox *Nitrospira*) and 894 Ma (Nitrosococcaceae) and radiation of crown groups occurring after 538 Ma (Figure 3, Table 1). The 95% confidence intervals of divergence times introduce uncertainty of +/- 150 Ma to these estimates (Table 1, Supplemental Figure 4), but in all cases firmly place the major radiation of extant AOB to Phanerozoic time even if the origin of crown groups was in late Neoproterozoic time. Between the uncertainty in the ages of total group AOB clades, and the fact that the acquisition of ammonia oxidation could in theory occur at any point along stem lineages prior to the divergence of crown groups, the range of 95% confidence intervals of ages of origin of the first AOB consistent with our data is between 1169 and 414 Ma. This age range includes scenarios involving the first evolution of ammonia oxidation in the earliest stem group Nitrosococcaceae (1169 Ma) or at the base of crown group Nitrosomonadaceae (414 Ma). This also accommodates the possibility that these groups acquired ammonia oxidation roughly simultaneously between 414 and 490 Ma. The analysis therefore does not allow for a unique determination of which proteobacterial lineage first acquired the capacity for ammonia oxidation. However, all scenarios consistent with our data involve a later acquisition of ammonia oxidation within the *Nitrospira,* after the radiation of ammonia oxidizing Nitrosomonadaceae and nitrite oxidizing *Nitrospira.*

The Neoproterozoic to early Phanerozoic origin of crown group AOB suggested by our data is ~1.5 Ga later than previously suggestions which placed ammonium oxidation at or before the GOE ~2.3 Ga (e.g. Garvin et al. 2009, Zerkle et al. 2017). Our estimate for the origin of bacterial ammonia oxidation is during the time in Earth history that saw the biosphere transition from a low-productivity, exclusively microbial state characteristic of Proterozoic time (Dick et al. 2018, Ward and Shih 2019) to a more modern system fueled by eukaryotic algae and supporting complex multicellular organisms including animals (Erwin et al. 2011, Brocks et al. 2017). Net primary productivity of the biosphere is thought to have increased significantly at this time (Ward et al. 2019a, Ward and Shih 2019), along with a rise in atmospheric oxygen concentrations to near-modern levels along with the oxygenation of the deep ocean and the loss of euxinic or ferruginous open ocean conditions (Lyons et al. 2014); the timing of this "Neoproterozoic Oxygenation Event" (NOE) is not well constrained and may have occurred during Paleozoic time ~420 Ma (e.g. Stolper and Keller 2018). The divergence of stem group proteobacterial AOB during Neoproterozoic time and the radiation of crown group proteobacterial AOB clades during Paleozoic time—broadly coincident with the NOE—suggests that these evolutionary innovations may be causally linked. Increased oxygenation of the oceans would have provided additional O_2 for ammonia oxidation, while higher NPP necessitates higher

fluxes of fixed nitrogen through the biosphere leading to higher rates of N_2 fixation to reduced forms (e.g. Ward et al. 2019a). As a result, the necessary substrates for ammonia oxidation would have been more abundant by the NOE than earlier in Proterozoic time, potentially opening additional niche space which enabled the radiation of AOB. This hypothesis is further supported by the adaptation of AOB to significantly higher substrate concentrations than AOA (e.g. Hink et al. 2018).

The delayed arrival of AOB may also be a consequence of limited copper availability in Proterozoic oceans. Aerobic ammonia oxidation has a relatively high requirement for copper for enzyme cofactors (e.g. Amin et al. 2013), and copper availability in Proterozoic oceans may have been limited due to the insolubility of copper sulfide minerals in periodically euxinic oceans (e.g. Saito et al. 2003). The expansion of some metabolic pathways may therefore have been impeded by the availability of trace metals necessary as enzyme cofactors (Saito et al. 2003). Extreme copper limitation would also be expected to inhibit the expansion of AOA, and so may retard the entire oxidative arm of the nitrogen cycle for much of Proterozoic time.

Implications for the Proterozoic nitrogen cycle

The late evolution of crown group AOB clades suggests that bacteria were not playing a dominant role in driving the oxidative arm of the nitrogen cycle during Proterozoic time, though it is always possible that there is a deeper history of bacterial ammonia oxidation by other lineages that remain undiscovered or that are now extinct. The antiquity of AOA is not yet constrained, and so it is a reasonable hypothesis that AOA evolved earlier in Earth history, perhaps in time to drive the aerobic nitrogen cycle shortly after the GOE ~2.3 Ga (e.g. Zerkle et al. 2017). Consistent with this hypothesis is the low oxygen requirements of AOA compared to proteobacterial ammonia oxidizers, leading to the continued dominance of AOA in modern oxygen minimum zones (e.g. Bristow et al. 2016). This is also in keeping with AOA evolving at a time with one to several orders of magnitude lower O_2 than was present during the origin of AOB (e.g. Lyons et al. 2014). The relative contribution of AOA and AOB to ammonia oxidation fluxes through time is not constrained as the nitrogen isotope signatures of these groups overlap (Santoro and Casciotti 2011) and the relative abundance and activity of AOA and AOB in modern environments is only determined roughly on a local scale via sequencing-based approaches that are not applicable to deep time (e.g. Bristow et al. 2016).

It is important to note that AOA and AOB utilize different biochemical pathways downstream of AMO, so their relative contribution to ammonia oxidation through time has significant implications for modeling of the productivity and atmospheric impact of the ancient biosphere. For example, extant AOA fix carbon using a uniquely energy-efficient O_2-tolerant carbon fixation pathway (a variant of the hydroxypropionate/hydroxybutyrate pathway), while AOB typically utilize the Calvin Cycle (or, in comammox *Nitrospira*, the rTCA cycle) (Könneke et al. 2014, Ward and Shih 2019). This allows AOA to fix 1.3g of dry cell mass for every mole of ammonia oxidized, in contrast to only 0.8g/mol in ammonia oxidizing Proteobacteria (Könneke et al. 2014). Further, nitrification currently accounts for ~75% of non-photosynthetic carbon fixation in aquatic environments (Raven 2009), so a nearly twofold difference in efficiency of carbon fixation in AOA versus AOB may lead to significant differences in predictions of net primary productivity of the biosphere through time. This is particularly important in the Proterozoic when photosynthetic carbon fixation rates are thought to have been much lower than today (e.g. Crockford et al. 2018, Ward and Shih 2019). Furthermore, the typical release of N_2O by AOA is significantly lower than from AOB (Hink et al. 2018),

particularly under low-oxygen conditions (Stieglmeyer et al. 2014). As a result, an increased contribution of AOA to nitrification during Proterozoic time would likely be associated with a lower biogenic N_2O flux, potentially at levels sufficiently low to prevent N_2O from accumulating as an important greenhouse gas in the Proterozoic atmosphere as previously proposed (e.g. Buick 2007).

However, further work will be necessary to constrain the origin of AOA. While it is a reasonable working hypothesis that the AOA evolved much earlier than AOB and fueled ammonia oxidation through Proterozoic time, it remains possible that AOA also significantly postdate the GOE. Before the origin of the first ammonia oxidizers, the nitrogen cycle would have consisted primarily of a vector toward reduced forms, with perhaps some oxidized nitrogen produced abiotically via processes like lightning (e.g. Wong et al. 2017). This reduced biogeochemical nitrogen cycle is thought to have persisted through Archean time (e.g. Ward et al. 2019a) and may have continued into Proterozoic time until the evolution of the first ammonia oxidizers. Due to the stability of fixed nitrogen in the oceans as ammonia at this time, the nitrogen demands of phototrophic primary productivity would have been readily met. Following the evolution of the first ammonia oxidizers, however, the Proterozoic biosphere may have undergone nitrogen limitation as the conversion of ammonia to nitrite and nitrate in oxygenated surface oceans would likely be followed by substantial loss of fixed nitrogen via dentrification and anammox in anoxic bottom waters (e.g. Fennel et al. 2005). Determining the antiquity of AOA, and therefore the age of the oxidative arm of the biogeochemical nitrogen cycle overall, will provide an estimate of when nitrogen loss may have first led to a fixed nitrogen crisis for the biosphere.

Conclusions

The molecular clock evidence for broadly coincident radiations of multiple convergently evolved crown group AOB clades (Nitrosomonadaceae, Nitrosococcaceae, and comammox *Nitrospira*) shown here is largely unprecedented in molecular clock studies, which typically address the age of a single clade (e.g. acquisition of phototrophy within a bacterial phylum, Shih et al. 2017b) or show multiple evolutionary events scattered through time (e.g. evolution of C_{30} sterols in sponges and algae, Gold et al. 2016). This adds strength to interpretations that the convergent evolutionary transitions to ammonia oxidation in these groups and/or their subsequent radiation is causally linked to increases in ocean oxygenation at this time and highlights the evolutionary response of biogeochemically relevant microorganisms to major environmental perturbations.

The necessity of reevaluating the structure and size of the Proterozoic nitrogen cycle in light of evidence for late-evolving AOB highlights a recurring problem in assessing the antiquity of microbial lineages—the rock record typically records the indirect effects of a (bio)geochemical process or the metabolism driving it, not directly the organisms that perform it. As a result, care must be taken in applying a strictly uniformitarian interpretation of the biological drivers of geochemical processes in deep time, as this overlooks evolutionary processes such as convergent evolution or horizontal gene transfer of metabolic pathways that lead to incongruent histories and potentially different combinations of traits in ancient drivers of biogeochemical cycles from the organisms responsible for these processes today.

Methods:

Genomes were downloaded from the NCBI Genbank and WGS databases. Completeness and contamination of metagenome-assembled genomes (MAGs) was estimated based on presence and copy number of conserved single-copy proteins by CheckM (Parks et al. 2015). Protein sequences used in analyses (see below) were identified locally with the *tblastn* function of BLAST+(Camacho et al. 2009), aligned with MUSCLE (Edgar 2004), and manually curated in Jalview (Waterhouse 2009). Positive BLAST hits were considered to be full length (e.g. >90% the shortest reference sequence from an isolate genome) with *e*-values greater than $1e^{-20}$. Presence of metabolic pathways of interest in incomplete MAGs was predicted with MetaPOAP (Ward et al. 2018b) to check for False Positives (contamination) or False Negatives (genes present in source genome but not recovered in metagenome-assembled genomes). Phylogenetic trees were calculated using RAxML (Stamatakis 2014) on the Cipres science gateway (Miller et al. 2010). Transfer bootstrap support values were calculated by BOOSTER (Lemoine et al. 2018), and trees were visualized with the Interactive Tree of Life viewer (Letunic and Bork 2016). Taxonomic assignment was confirmed with GTDB-Tk (Parks et al. 2018). Histories of vertical versus horizontal inheritance of metabolic genes was inferred by comparison of organismal and metabolic protein phylogenies (Doolittle 1986, Ward et al. 2018a).

A concatenated protein alignment was generated by extracting protein sequences for marker genes from genomes of interest via the *tblastn* function of BLAST+ (Camacho et al. 2009), aligning protein sequences with MUSCLE (Edgar 2004), and then concatenating aligned sequences. Concatenated alignments were curated with Gblocks (Castresana 2000) and manually in Jalview (Waterhouse 2009). Taxa included in this alignment consist of all available AOB genomes on the NCBI GenBank and WGS databases as well as sister groups and outgroups spanning the full diversity of the Proteobacteria and Nitrospirota as well as closely related phyla (e.g. Methylomirabilota and Nitrospinota) as assessed by GTDB (Parks et al. 2018) and concatenated ribosomal protein phylogenies of the tree of life (Figure 1, Hug et al. 2016), as well as Cyanobacteria, Plastids, and Mitochondria. Phylogenetic markers were chosen as conserved proteins across bacteria, plastids, and mitochondria, as previously reported (Shih et al 2017a, Shih et al. 2017b), and consisted of AtpA, AtpB, EfTu, AtpE, AtpF, AtpH, AtpI, Rpl2, Rpl16, Rps3, and Rps12 protein sequences. Bayesian molecular clock analyses were carried out using BEAST v2.4.5 (Bouckaert et al. 2019) using the Cyberinfrastructure for Phylogenetic Research (CIPRES) Science Gateway v 3.3 server (Miller et al. 2010). As previously reported, the CpREV model was chosen as the best-fitting amino acid substitution model for the concatenated protein dataset based on ProtTest analysis (Shih et al 2017b) Cross-calibration techniques utilizing plastid and mitochondrial endosymbiosis events were used as priors, utilizing time constraints for the most recent common ancestor of Angiosperms (normal distribution with a mean of 217 Ma and sigma of 40 Ma) and of land plants (normal distribution with a mean of 477 Ma and sigma of 70 Ma) as has been previously described by Smith et al (Smith et al. 2010, Shih and Matzke 2013). We also constrained the most recent common ancestor of Rhodophytes with as a more recent study and precise estimate of the fossil constraint Bangiomorpha pubescens utilizing Re-Os isotopic measurements of the Bylot Supergroup of Baffin Island where the fossil was first described (Gibson et al. 2017). Taking into account previously reported ages of Bangiomorpha, we set this constraint as a uniform prior from 1030-1200 Ma, in order to account both Re-Os and Pb-Pb isotopic measurements estimating the age of Bangiomorpha (Butterfield 2000). A conservative uniform prior between 2300-3800 Ma was set on the divergence between Cyanobacteria and Melainabacteria, as oxygenic photosynthesis evolved prior to the Great Oxygenation Event and most likely evolved sometime after the Late Heavy Bombardment.

Finally, a uniform prior for all taxa was again set conservatively between 2400-3800 Ma, assuming that the Last Bacterial Common Ancestor most likely evolved after the Late Heavy Bombardment. Wide uniform priors were used as a means to provide very conservative upper and lower limits. Three Markov chain Monte Carlo chains were run for 100 million generations sampling every 10,000th generation, and the first 50% of generations were discarded as burn-in. TreeAnnotator v1.7.5 (Bouckaert et al. 2019) was used to generate maximum clade credibility trees.

Figures:

1. Tree of life built with concatenated ribosomal proteins following methods from Hug et al. 2016. Clades of ammonia oxidizing organisms highlighted in orange and labeled. The distribution of ammonia oxidation is polyphyletic, spread across one lineage within the Archaea (Nitrososphaeria) and three within the Bacteria (Nitrosococcales and Nitrosomonadaceae in the Proteobacteria phylum, and some members of the genus *Nitrospira* within the Nitrospirota phylum).

2. Phylogeny of concatenated protein sequences of CuMMO A and B subunits (i.e. AmoA and AmoB, PmoA and PmoB), with ammonia monooxygenases highlighted in orange. Major clades are collapsed and labeled by the taxonomy of the organisms in which they are found as determined with GTDB-Tk (Parks et al. 2018). The highly divergent archaeal ammonia monooxygenase is placed as an outgroup, though the placement of the root is indeterminate. Transfer Bootstrap Expectation (TBE) support values shown as calculated by BOOSTER (Lemoine et al. 2018).

3. Molecular clock showing estimated age of clades of ammonia oxidizing bacteria. Phylum-level clades highlighted in gray and brown, with ammonia oxidizing clades highlighted in orange. Approximate timing of Great Oxygenation Event (~2.45-2.3 Ga) and Neoproterozoic/Paleozoic Oxygenation Event (~635-420 Ma) shown with light blue bars. While stem lineages of AOB clades may predate the NOE, the radiation of crown groups all occur broadly coincident or subsequent to the NOE, suggesting that evolutionary radiations of nitrogen-cycling organisms may have been causally linked with expansions in ocean oxygenation and/or productivity during this time period.

Tables:

1. Age ranges for key divergences discussed in the text in millions of years (Ma).

Clade	Estimated age of crown group	95% Confidence Interval	Estimated age of total group	95% confidence interval
Nitrosococcaceae	325	194-490	894	596-1169
Nitrosomonadaceae	538	414-662	657	510-821
Nitrospira	124	59-197	238	149-329

Acknowledgements:

References:

1.   Amin, S.A., Moffett, J.W., Martens-Habbena, W., Jacquot, J.E., Han, Y., Devol, A., Ingalls, A.E., Stahl, D.A. and Armbrust, E.V., 2013. Copper requirements of the ammonia☐oxidizing archaeon Nitrosopumilus maritimus SCM1 and implications for nitrification in the marine environment. *Limnology and Oceanography*, *58*(6), pp.2037-2045.

2.   Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N. and Matschiner, M., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS computational biology*, *15*(4), p.e1006650.

3.   Bristow, L.A., Dalsgaard, T., Tiano, L., Mills, D.B., Bertagnolli, A.D., Wright, J.J., Hallam, S.J., Ulloa, O., Canfield, D.E., Revsbech, N.P. and Thamdrup, B., 2016. Ammonium and nitrite oxidation at nanomolar oxygen concentrations in oxygen minimum zone waters. *Proceedings of the National Academy of Sciences*, *113*(38), pp.10601-10606.

4.   Brocks, J. J., Jarrett, A. J. M., Sirantoine, E., Hallmann, C., Hoshino, Y., & Liyanage, T. (2017). The rise of algae in Cryogenian oceans and the emergence of animals. Nature, 548(7669), 578–581.

5.   Buick, R., 2007. Did the Proterozoic 'Canfield Ocean'cause a laughing gas greenhouse?. *Geobiology*, *5*(2), pp.97-100.

6.   Butterfield, N.J., 2000. Bangiomorpha pubescens n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology*, *26*(3), pp.386-404.

7.   Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer and T. L. Madden (2009). "BLAST+: architecture and applications." Bmc Bioinformatics **10**: 421. DOI: 10.1186/1471-2105-10-421.

8.   Canfield, D.E., Glazer, A.N. and Falkowski, P.G., 2010. The evolution and future of Earth's nitrogen cycle. *science*, *330*(6001), pp.192-196.

9.   Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution*, *17*(4), pp.540-552.

10.   Crockford, P.W., Hayles, J.A., Bao, H., Planavsky, N.J., Bekker, A., Fralick, P.W., Halverson, G.P., Bui, T.H., Peng, Y. and Wing, B.A., 2018. Triple oxygen isotope evidence for limited mid-Proterozoic primary productivity. *Nature*, *559*(7715), p.613.

11.   Dick, G. J., Grim, S. L., & Klatt, J. M. (2018). Controls on O2 Production in Cyanobacterial Mats and Implications for Earth's Oxygenation. Annual Review of Earth and Planetary Sciences. https://doi.org/10.1146/annurev-earth-082517-010035

12.   Doolittle, R. F. (1986). Of urfs and orfs: A primer on how to analyze derived amino acid sequences. Mill Valley, California, University Science Books.

13.   Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Res **32**(5): 1792-1797. DOI: 10.1093/nar/gkh340.

14.   Erwin, D. H., Laflamme, M., Tweedt, S. M., Sperling, E. A., Pisani, D., & Peterson, K. J. (2011). The Cambrian conundrum: early divergence and later ecological success in the early history of animals. Science, 334(6059), 1091–1097.

15.   Fennel, K., Follows, M. and Falkowski, P.G., 2005. The co-evolution of the nitrogen, carbon and oxygen cycles in the Proterozoic ocean. *American Journal of Science*, *305*(6-8), pp.526-545.

16.   Garvin, J., Buick, R., Anbar, A.D., Arnold, G.L. and Kaufman, A.J., 2009. Isotopic evidence for an aerobic nitrogen cycle in the latest Archean. *Science*, *323*(5917), pp.1045-1048.

17.   Gibson, T.M., Shih, P.M., Cumming, V.M., Fischer, W.W., Crockford, P.W., Hodgskiss, M.S., Wörndle, S., Creaser, R.A., Rainbird, R.H., Skulski, T.M. and Halverson, G.P., 2017. Precise age of Bangiomorpha pubescens dates the origin of eukaryotic photosynthesis. *Geology*, *46*(2), pp.135-138.

18.   Gold, D.A., Grabenstatter, J., de Mendoza, A., Riesgo, A., Ruiz-Trillo, I. and Summons, R.E., 2016. Sterol and genomic analyses validate the sponge biomarker hypothesis. *Proceedings of the National Academy of Sciences*, *113*(10), pp.2684-2689.

19.   Hink, L., Gubry-Rangin, C., Nicol, G.W. and Prosser, J.I., 2018. The consequences of niche and physiological differentiation of archaeal and bacterial ammonia oxidisers for nitrous oxide emissions. *The ISME journal*, *12*(4), p.1084.

20.   Hu, Z., Wessels, H.J., van Alen, T., Jetten, M.S. and Kartal, B., 2019. Nitric oxide-dependent anaerobic ammonium oxidation. *Nature Communications*, *10*.

21.   Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., Butterfield, C.N., Hernsdorf, A.W., Amano, Y., Ise, K. and Suzuki, Y., 2016. A new view of the tree of life. *Nature microbiology*, *1*(5), p.16048.

22.   Khadka, R., Clothier, L., Wang, L., Lim, C.K., Klotz, M.G. and Dunfield, P.F., 2018. Evolutionary history of copper membrane monooxygenases. *Frontiers in microbiology*, *9*.

23.   Könneke, M., Schubert, D.M., Brown, P.C., Hügler, M., Standfest, S., Schwander, T., von Borzyskowski, L.S., Erb, T.J., Stahl, D.A. and Berg, I.A., 2014. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *Proceedings of the National Academy of Sciences*, *111*(22), pp.8239-8244.

24.   Lemoine, F., J. B. Domelevo Entfellner, E. Wilkinson, D. Correia, M. Davila Felipe, T. De Oliveira and O. Gascuel (2018). "Renewing Felsenstein's phylogenetic bootstrap in the era of big data." Nature **556**(7702): 452-456. DOI: 10.1038/s41586-018-0043-0.

25.   Letunic, I. and P. Bork (2016). "Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees." Nucleic Acids Res **44**(W1): W242-W245. DOI: 10.1093/nar/gkw290.

26.   Lyons, T. W., Reinhard, C. T., & Planavsky, N. J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. Nature. https://doi.org/10.1038/nature13068

27.   Miller, M. A., W. Pfeiffer and T. Schwartz (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE).

28.   Parks, D. H., M. Imelfort, C. T. Skennerton, P. Hugenholtz and G. W. Tyson (2015). "CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes." Genome Res **25**(7): 1043-1055. DOI: 10.1101/gr.186072.114.

29.   Parks, D. H., M. Chuvochina, D. W. Waite, C. Rinke, A. Skarshewski, P. A. Chaumeil and P. Hugenholtz (2018). "A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life." Nat Biotechnol **36**(10): 996-1004 DOI: 10.1038/nbt.4229.

30.   Raven, J. A. (2009). Contributions of anoxygenic and oxygenic phototrophy and chemolithotrophy to carbon and oxygen fluxes in aquatic environments. *Aquatic Microbial Ecology*, *56*(2-3), 177-192.

31.   Saito, M.A., Sigman, D.M. and Morel, F.M., 2003. The bioinorganic chemistry of the ancient ocean: the co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean–Proterozoic boundary?. *Inorganica Chimica Acta*, *356*, pp.308-318.

32.   Santoro, A.E. and Casciotti, K.L., 2011. Enrichment and characterization of ammonia-oxidizing archaea from the open ocean: phylogeny, physiology and stable isotope fractionation. *The ISME journal*, *5*(11), p.1796.

33.   Shih, P.M. and Matzke, N.J., 2013. Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins. *Proceedings of the National Academy of Sciences*, *110*(30), pp.12355-12360.

34.   Shih, P.M., Hemp, J., Ward, L.M., Matzke, N.J. and Fischer, W.W., 2017a. Crown group Oxyphotobacteria postdate the rise of oxygen. *Geobiology*, *15*(1), pp.19-29.

35.   Shih, P.M., Ward, L.M. and Fischer, W.W., 2017b. Evolution of the 3-hydroxypropionate bicycle and recent transfer of anoxygenic photosynthesis into the Chloroflexi. *Proceedings of the National Academy of Sciences*, *114*(40), pp.10749-10754.

36.   Smith, S.A., Beaulieu, J.M. and Donoghue, M.J., 2010. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proceedings of the National Academy of Sciences*, *107*(13), pp.5897-5902.

37.   Stamatakis, A. (2014). "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies." Bioinformatics **30**(9): 1312-1313. DOI: 10.1093/bioinformatics/btu033.

38.   Stieglmeier, M., Mooshammer, M., Kitzler, B., Wanek, W., Zechmeister-Boltenstern, S., Richter, A. and Schleper, C., 2014. Aerobic nitrous oxide production through N-nitrosating hybrid formation in ammonia-oxidizing archaea. *The ISME journal*, *8*(5), p.1135.

39.   Stolper, D.A. and Keller, C.B., 2018. A record of deep-ocean dissolved O 2 from the oxidation state of iron in submarine basalts. *Nature*, *553*(7688), p.323.

40.   Tyrrell, T., 1999. The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature*, *400*(6744), p.525.

41.   Ward, L.M., Kirschvink, J.L. and Fischer, W.W., 2016. Timescales of oxygenation following the evolution of oxygenic photosynthesis. *Origins of Life and Evolution of Biospheres*, *46*(1), pp.51-65.

42.   Ward, L. M., J. Hemp, P. M. Shih, S. E. McGlynn and W. W. Fischer (2018a). "Evolution of phototrophy in the Chloroflexi phylum driven by horizontal gene transfer." Front Microbiol **9**: 260. DOI: 10.3389/fmicb.2018.00260.

43.   Ward, L. M., P. M. Shih and W. W. Fischer (2018b). "MetaPOAP: presence or absence of metabolic pathways in metagenome-assembled genomes." Bioinformatics **34**(24): 4284-4289. DOI: 10.1093/bioinformatics/bty510.

44.   Ward, L.M. and Shih, P.M., 2019. The evolution and productivity of carbon fixation pathways in response to changes in oxygen concentration over geological time. *Free Radical Biology and Medicine*.

45.   Ward, L.M., Rasmussen, B. and Fischer, W.W., 2019a. Primary productivity was limited by electron donors prior to the advent of oxygenic photosynthesis. *Journal of Geophysical Research: Biogeosciences*, *124*(2), pp.211-226.

46.   Ward, L., Shih, P.M., Hemp, J., Kakegawa, T., Fischer, W.W. and McGlynn, S.E., 2019b. Phototrophic Methane Oxidation in a Member of the Chloroflexi Phylum. *bioRxiv*, p.531582.

47.   Waterhouse, A. M., J. B. Procter, D. M. A. Martin, M. Clamp and G. J. Barton (2009). "Jalview Version 2-a multiple sequence alignment editor and analysis workbench." Bioinformatics **25**(9): 1189-1191. DOI: 10.1093/bioinformatics/btp033.

48.   Wong, M.L., Charnay, B.D., Gao, P., Yung, Y.L. and Russell, M.J., 2017. Nitrogen oxides in early Earth's atmosphere as electron acceptors for life's emergence. *Astrobiology*, *17*(10), pp.975-983.

49.   Zerkle, A.L. and Mikhail, S., 2017. The geobiological nitrogen cycle: from microbes to the mantle. *Geobiology*, *15*(3), pp.343-352.

50.   Zerkle, A.L., Poulton, S.W., Newton, R.J., Mettam, C., Claire, M.W., Bekker, A. and Junium, C.K., 2017. Onset of the aerobic nitrogen cycle during the Great Oxidation Event. *Nature*, *542*(7642), p.465.
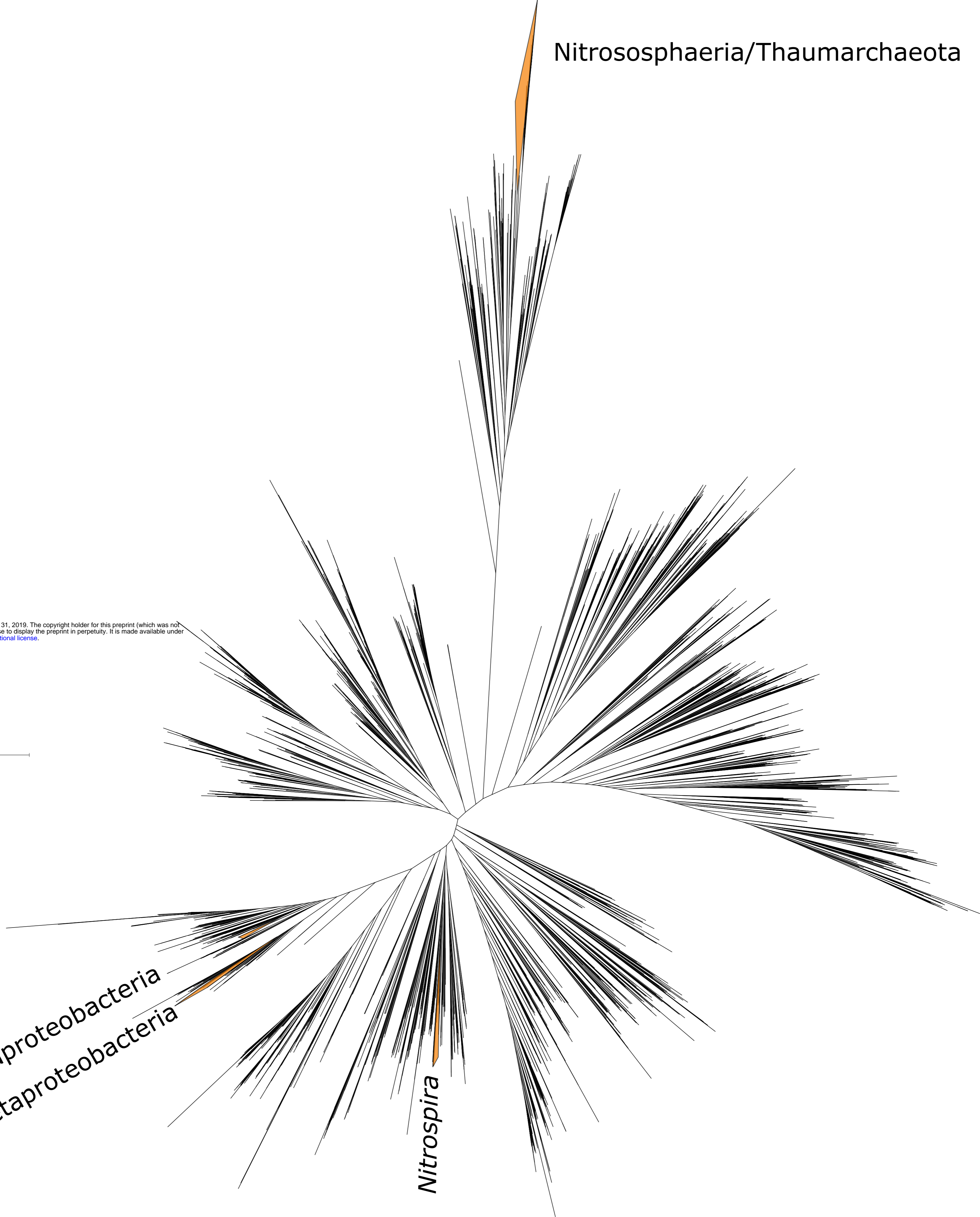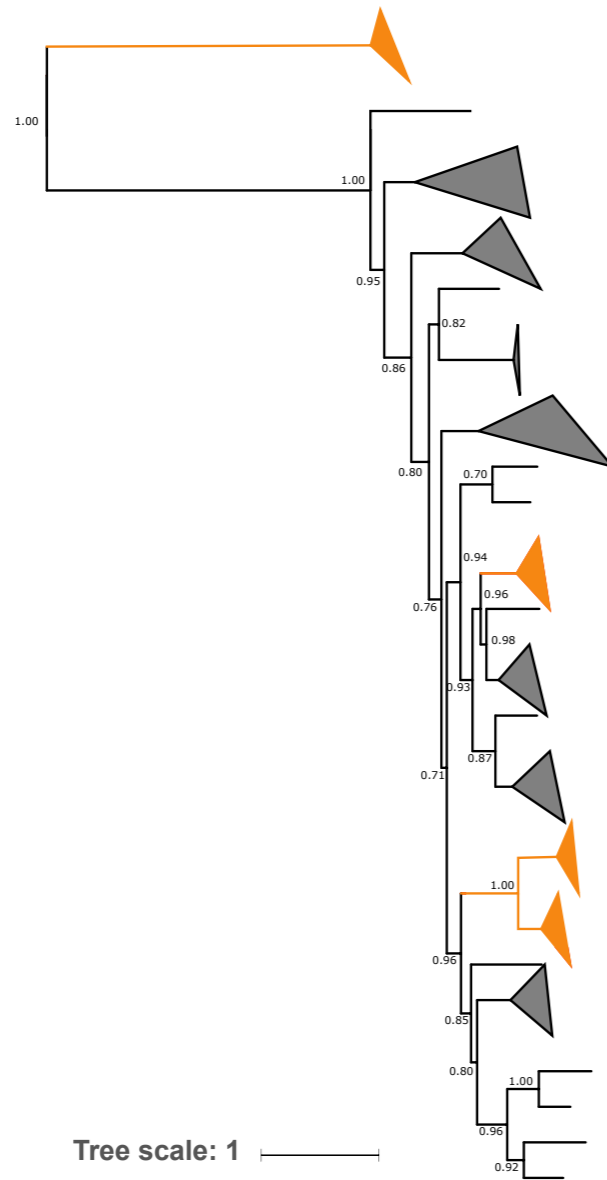
Nitrososphaeria/Thaumarchaeota

Nitrosococcales, Gammaproteobacteria

Nitrosomonadaceae, Betaproteobacteria

Nitrospira

Tree scale: 1

d Archaea p Crenarchaeota c Nitrososphaeria o Nitrososphaerales

d Bacteria p UBP10 c GR-WP33-30 o bin18 f bin18 g bin18

d Bacteria p Actinobacteriota c Actinobacteria

d Bacteria p Verrucomicrobiota c Verrucomicrobiae o Methylacidiphilales f Methylacidiphilaceae

d Bacteria p Myxococcota c UBA9160 o UBA9160 f UBA6930 g UBA6930

d Bacteria p Methylomirabilota c Methylomirabilia o Methylomirabilales f Methylomirabilaceae g Methylomirabilis

d Bacteria p Proteobacteria c Gammaproteobacteria
d Bacteria p Proteobacteria c Alphaproteobacteria o Rhizobiales f Beijerinckiaceae g Methylocystis s Methylocystis rosea
d Bacteria p Proteobacteria c Gammaproteobacteria o Methylococcales f Methylomonadaceae g Methyloglobulus

d Bacteria p Proteobacteria c Gammaproteobacteria o Nitrosococcales f Nitrosococcaceae

d Bacteria p Proteobacteria c Gammaproteobacteria o UBA7966 f UBA7966 g USCg-Taylor

d Bacteria p Proteobacteria c Gammaproteobacteria o Methylococcales

d Bacteria p Proteobacteria c Alphaproteobacteria o Azospirillales f Azospirillaceae g Skermanella s Skermanella aerolata

d Bacteria p Proteobacteria c Alphaproteobacteria o Rhizobiales f Beijerinckiaceae

d Bacteria p Nitrospirota c Nitrospiria o Nitrospirales f Nitrospiraceae g Nitrospira

d Bacteria p Proteobacteria c Gammaproteobacteria o Betaproteobacteriales f Nitrosomonadaceae

d Bacteria p Proteobacteria c Gammaproteobacteria o Methylococcales f Cycloclasticaceae g UBA2780
d Bacteria p Proteobacteria c Gammaproteobacteria o Betaproteobacteriales f Rhodocyclaceae g Rugosibacter

d Bacteria p Proteobacteria c Gammaproteobacteria o Nevskiales f Nevskiaceae g Solimonas s Solimonas aquatica
d Bacteria p Proteobacteria c Gammaproteobacteria o Nevskiales f Nevskiaceae g Polycyclovorans
d Bacteria p Proteobacteria c Alphaproteobacteria o Rhizobiales f Xanthobacteraceae g Bradyrhizobium s Bradyrhizobium manausense
d Bacteria p Proteobacteria c Gammaproteobacteria o Betaproteobacteriales f Ga0077523 g Ga0077526

Tree scale: 1

Nitrosococcaceae

Proteobacteria

Nitrosomonadaceae

Mitochondria

Nitrospiraceae

Nitrospira

Nitrospina/Methylomirabilota

Cyanobacteria

Plastids

Archean	Paleoproterozoic	Mesoproterozoic	Neoproterozoic	Phanerozoic

3000   2500   2000   1500   1000   500   0

300.0