Palaeogeography, Palaeoclimatology, Palaeoecology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Palaeogeography, Palaeoclimatology, Palaeoecology

journal homepage: www.elsevier.com/locate/palaeo

Molecular clock evidence for an Archean diversification of heme-copper oxygen reductase enzymes☆

Fatima Husain [a,*], Haitao Shang [b,c], Stilianos Louca [b], Gregory P. Fournier [a]

[a] Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, 55-101, Cambridge, MA 02139, United States of America
[b] Department of Biology, University of Oregon, 77 Klamath Hall, 1210 University of Oregon, Eugene, OR 97403, United States of America
[c] Department of Earth, Environmental and Resource Sciences, The University of Texas at El Paso, 500 W University, El Paso, Texas 79902, United States of America

ARTICLE INFO

Editor: H Falcon-Lang

Keywords:
Archean
Great Oxidation Event
Atmospheric evolution
Aerobic respiration
Phylogenomics, phylogenetic trees
Molecular clocks

ABSTRACT

Age estimates for the emergence of oxygenic photosynthesis derived from molecular clock analyses reach back to the Archean, substantially predating geochemical evidence of persistent, irreversible atmospheric oxygenation following the Great Oxidation Event (GOE). Additional evidence for the Archean biological cycling of oxygen would provide valuable insight into environmental oxygen concentrations and dynamics during this time. Here, we present phylogenetic reconstructions and molecular clock analyses of heme-copper oxygen reductases, enzymes that use oxygen as a terminal electron acceptor in many aerobic electron transport chains. These analyses reveal that these enzymes likely emerged and diversified in the Mesoarchean, following the earlier emergence of oxygenic photosynthesis. These results are consistent with oxygen concentrations being sufficient for aerobic physiologies to persist in at least some environments prior to the GOE, adding a key component to our understanding of the history of planetary oxygenation independent of the geochemical record.

1. Introduction

During the Archean eon, biological innovations initiated a cascade of geochemical and ecological changes that set the stage for the emergence of complex life. Recent molecular clock analyses estimate that the Cyanobacteria, microorganisms capable of oxygenic photosynthesis, emerged before Earth's Great Oxidation Event (GOE) (Boden et al., 2021; Fournier et al., 2021; Sánchez-Baracaldo, 2015), a transition which records the first significant oxygenation of the Earth's atmosphere. The temporal disconnect between the evolution of oxygenic photosynthesis and its geochemical impact is often principally attributed to the local, abiotic sequestration of molecular oxygen by reduced chemical species in the Archean environment. Phylogenetic reconstructions and analyses of heme-copper oxygen reductases (HCO), enzymes that catalyze the reduction of oxygen to water and contribute to the generation of a proton gradient in respiratory electron transport chains, can aid in the evaluation of the hypothesis that biological sinks, such as aerobic respiration, additionally consumed oxygen during the Archean, forestalling its accumulation.

The HCO enzymes constitute a family of multi-subunit enzymes found in cytoplasmic membranes, cyanobacterial thylakoid membranes, and mitochondrial membranes (Calhoun et al., 1994; García-Horsman et al., 1994; Saraste, 1990; Saraste et al., 1991). As the terminal enzymes in aerobic electron transport chains, they are often referred to as Complex IV or cytochrome oxidases, though some members of the enzyme family use electron donors other than cytochromes (Pereira et al., 2001; Sousa et al., 2012). During aerobic respiration in mitochondria, electrons are carried from an electron donor, the protein cytochrome *c*, to molecular oxygen through a series of electron acceptors located within the enzyme's subunits. In this process, cytochrome *c* delivers four electrons to a copper center, Cu_A, which are transferred to a heme *a*, then transferred to a heme a_3-Cu_B binuclear center, where they are ultimately transferred to an oxygen molecule bound to heme a_3 (Nelson and Cox, 2013; Wikström, 1989). During this process, four protons from the mitochondrial matrix are used to convert oxygen to water, while another four protons are pumped across the membrane to the intermembrane space, contributing to the generation of a proton-motive force which eventually drives the synthesis of the energy storage molecule adenosine triphosphate, also known as ATP (Nelson and Cox, 2013; Wikström, 1989).

☆ This article is part of a Special issue entitled: 'Genomic palaeoecology' published in Palaeogeography, Palaeoclimatology, Palaeoecology.

* Corresponding author.
E-mail addresses: fhusain@mit.edu (F. Husain), hshang@utep.edu (H. Shang), g4nier@mit.edu (G.P. Fournier).

https://doi.org/10.1016/j.palaeo.2025.113531
Received 26 July 2025; Received in revised form 18 December 2025; Accepted 28 December 2025
Available online 5 January 2026

A general equation for this process is summarized below, where *D* refers to the electron donor:

$$4D_{reduced} + 4H^{+}_{in(chemistry)} + nH^{+}_{in(pumped)} + O_2 \rightarrow 2H_2O + nH^{+}_{out} + 4D_{oxidized}$$

Within the enzyme family, constituents may vary in their specific electron donors, hemes, proton translocation channels, and taxonomic distribution (Table 1). The enzymes may also have different numbers of subunits; prokaryotic oxygen reductases may have as few as three subunits, while many eukaryotic oxygen reductases may have up to thirteen (García-Horsman et al., 1994). Despite these differences, membership in the heme-copper oxygen reductase family is predicated on the presence of the catalytic subunit, subunit I, which contains a low-spin heme as well as a binuclear center comprised of a high-spin heme, denoted with a subscripted 3, and a Cu_B copper ion. Six strictly conserved histidine residues in subunit I act as ligands for both hemes and Cu_B (Sousa et al., 2012). In 2001, Pereira and colleagues proposed classifying the oxygen reductases into A, B, or C families based on the amino acids which comprise their core sequences and are involved in the enzymes' proton transfer mechanisms (Pereira et al., 2001). Within this framework, the A-type oxygen reductases are further comprised by A1-type, or mitochondrial-like oxygen reductases, and A2-type oxygen reductases (Pereira et al., 2001). Bioenergetic studies of the oxygen reductases have revealed that the subtypes differ in terms of their oxygen binding affinities: A-type oxygen reductases have low affinities, while B-type and C-type oxygen reductases have comparatively higher affinities (Degli Esposti et al., 2019; Morris and Schmidt, 2013; Trojan et al., 2021). These affinities are sometimes associated with organisms' oxygen requirements in environmental settings, though Sousa and colleagues in 2012 noted that the different oxygen reductases are not necessarily confined to specific environmental niches (Sousa et al., 2012). Many prokaryotic microbes have also been shown to use multiple terminal oxidases, including multiple heme-copper oxygen reductase subtypes (Sousa et al., 2012; Trojan et al., 2021). This flexibility is a function of branched respiratory chains, which confer on organisms the ability to survive varying oxygen concentrations (Sousa et al., 2012).

The phylogenetic boundaries defining heme-copper oxygen reductase families remain somewhat uncertain: the catalytic subunits of nitric oxide reductases (NORs), membrane-bound enzymes involved in the reduction of nitric oxide to nitrous oxide, are structurally similar to those in oxygen reductases but do not contain copper in their binuclear centers (Hendriks et al., 1998). The similarities between these subunits in HCO enzymes and NORs have led to many studies on the potential evolutionary relationships between the enzymes (Chen and Strous, 2013; de Vries and Schröder, 2002; Ducluzeau et al., 2008; Murali et al., 2024; Saraste and Castresana, 1994; Sharma and Wikström, 2014; van der Oost et al., 1994). Indeed, broad phylogenetic reconstructions of the heme-copper superfamily and NORs consistently show that the enzymes form distinct groups, with C-type oxygen reductases and NORs grouping closely (Ducluzeau et al., 2014; Sousa et al., 2012; Sousa et al., 2011) (Fig. 1).

There have been many hypotheses posed for the evolution of HCO enzymes (Degli Esposti, 2020), which are thought to be monophyletic

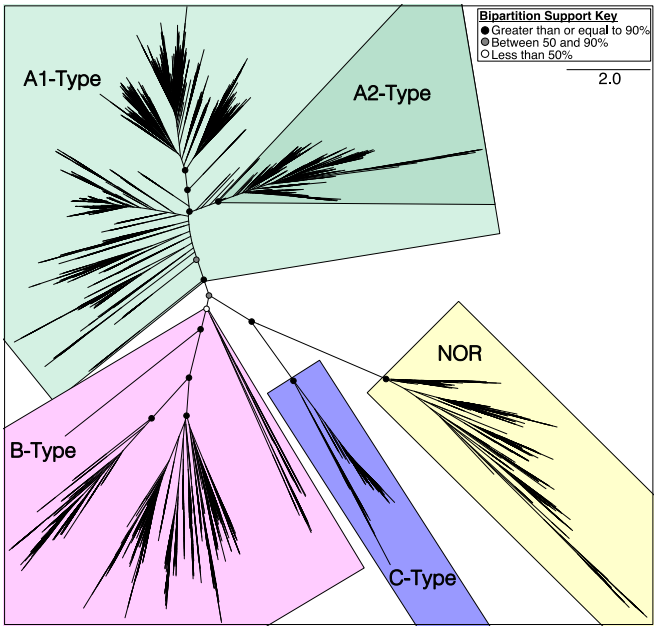


Fig. 1. Maximum-likelihood phylogenetic tree of the heme-copper oxygen reductase family. The tree includes 5360 sequences from all domains of life and shows the recovery of subtype classification grouping, as well as the identification of a group of A2-type within the A-type oxygen reductases. Bipartition support values are coloured according to bootstrap support values, with nodes coloured black having ≥90 % support, nodes coloured gray having between 50 % and 90 % support, and nodes coloured white having <50 % support.

Table 1
General characteristics of the heme-copper oxygen reductase subtypes and nitric oxide reductases. The A1-, A2-, B-, and C-type members comprise the oxygen reductases. The catalytic subunits of nitric oxide reductases (NOR) are structurally similar to those of the oxygen reductases, and their primary function is to reduce nitric oxide to nitrous oxide. *One electrogenic qNOR was characterized by Al-Attar and de Vries in 2015 (Al-Attar and De Vries, 2015).

Classification (Pereira et al., 2001)	**A1-Type**	**A2-Type**	**B-Type**	**C-Type**	**NOR**
Catalytic Subunit	Subunit I	Subunit I	Subunit I	Subunit I	NorB
Metals in Binuclear Center (Hendriks et al., 1998; Saraste, 1990)	Heme-Cu_B	Heme-Cu_B	Heme-Cu_B	Heme-Cu_B	Heme-Fe_B
Generation of Electrochemical Gradient (Al-Attar and De Vries, 2015; García-Horsman et al., 1994; Hendriks et al., 2002; Wikström, 1989)	Yes	Yes	Yes	Yes	No*
Translocation Ratio(s) (Han et al., 2011; Pereira and Teixeira, 2004; Sousa et al., 2012)	$1H^+/e^-$	$1H^+/e^-$	0.5 - $1H^+/e^-$	0.5 - $1H^+/e^-$	*n.a.*
Proton Translocating K-Channel (Pereira and Teixeira, 2004)	Yes [KTSY]	Yes [KTSY]	Alternative [TSYY]	Alternative [-SY-]	No
Proton Translocating D-Channel (Pereira and Teixeira, 2004)	Yes [E]	Yes [Y]	No	No	No
Primary Reduction Activity	O_2 to H_2O	O_2 to H_2O	O_2 to H_2O	O_2 to H_2O	NO to N_2O
Domains (Sousa et al., 2012)					
Eukaryotes	Yes	No	No	No	No
Bacteria	Yes	Yes	Yes	Yes	Yes
Archaea	Yes	No	Yes	No	Yes
Electron Donor(s) (Pereira et al., 2001; Pitcher and Watmough, 2004)					
Cytochrome c	Yes	Yes	Yes	Yes	Yes
Quinol	Yes	No	Yes	No	Yes
Oxygen Binding Affinity (Degli Esposti et al., 2019; Morris and Schmidt, 2013; Saraste and Castresana, 1994; van der Oost et al., 1994)	Low	Low	Intermediate-to-High	High	*n.a.*

(Castresana et al., 1994; Pereira et al., 2001) and descendants of an ancestral proto-oxidase or uroxidase (Castresana et al., 1994), with different subtypes emerging via gene duplication events (Castresana et al., 1994). Inferring the deep evolutionary history of HCO enzyme subtypes relies on where the root, representing the common ancestor, is placed within phylogenetic reconstructions. In previous studies, this root was obtained via either manual rooting based on substrate utilization (i.e., between NOR and oxygen-utilizing types), or midpoint rooting (Gribaldo et al., 2009; Saraste and Castresana, 1994), recovering dissimilar evolutionary histories. In other studies, proposed scenarios for the emergence of the oxygen reductases have been informed by enzyme structure and functionality and relating those features to substrate availability during the Archean eon (Ducluzeau et al., 2014; Sharma and Wikström, 2014). Molecular clock analyses based on a phylogeny of diverse oxygen reductase representatives from all domains of life, including prokaryotic representatives who may have originally acquired HCO enzymes by horizontal gene transfer (Brochier-Armanet et al., 2009; Castresana et al., 1994; Degli Esposti, 2020; Gribaldo et al., 2009; Pereira et al., 2001), and improved methodological approaches may better resolve the evolutionary history of oxygen reductase subtypes. To construct a general history of the HCO enzymes, the conserved catalytic subunit was used as a phylogenetic marker. This resulting survey tree was subsampled to be amenable to molecular clock analysis, maintaining both the representative diversity of these enzymes and key nodes for calibration. Additionally, two tree topologies were tested, evaluating the impact of different scenarios for the emergence of the low-affinity, A-type oxygen reductases. The results of these analyses suggest that HCO enzymes arose prior to the onset of the GOE and plausibly acted as a biological sink for oxygen generated by early Cyanobacteria.

2. Methods

To generate a phylogenetic tree of HCO and NOR protein families amenable to molecular clock analyses, an iterative sampling approach was employed which enabled the selection of sequences representative of the wide extant diversity of the enzymes. This approach is necessitated by the dataset size constraints for molecular clock analyses as trees with more than a few hundred taxa are computationally infeasible to assess using an intensive Bayesian methodology (Lartillot et al., 2015). The majority of the diversity contained in the initial trees constructed in this work is recently branching, and therefore the removal of many of these sequences is not expected to substantially influence the reconstruction of deeply branching, ancient divergences relevant to the early history of HCO enzymes.

35,984 sequences for subunit I of A-type, B-type, and C-type oxygen reductases were initially collected using the BLAST search methods described below. This set was then reduced to 5360 candidate sequences from which the oxygen reductase subtypes could be identified. Finally, this set was further reduced to 386 sequences, which is sufficiently small for a molecular clock analysis to be computationally tractable, while retaining sufficient diversity and sampling depth to recover the crown group nodes of major microbial and eukaryotic clades, including those necessary for molecular clock calibration. Details for each of these steps are provided in **Section 2.1** and **Section 2.2**.

2.1. Sequence curation, alignment, and tree construction

The initial protein sequence set was assembled to contain HCO and NOR enzymes across all domains of life. For the oxygen reductases, subunit I sequences were collected, and for NORs, the homologous NorB subunits were collected. Query sequences (**Supplementary Table 1**) were selected from the National Center for Biotechnology Information's (NCBI) Protein database.

NCBI's non-redundant protein database (Pruitt et al., 2007) was searched using BLASTp (Altschul et al., 1990) with default search settings and maximum sequence targets of 500 or 1000. No sequences were excluded from the search, enabling the collection of uncultured and environmental sequences as well as predicted protein models in order to comprehensively survey the extant diversity within these enzyme subunits. From these searches, 35,984 sequences were collected. An initial multiple sequence alignment was conducted on the sequence set using MAFFT version 7.245 (Katoh and Standley, 2013). The *--auto* strategy was employed, which enables the program to select between progressive alignment algorithms or an interactive refinement method, L-INS-i, based on data size. The algorithm FFT-NS-2, which is a fast, progressive alignment method, was chosen. The resulting alignment was viewed in Jalview 2.11.2.2 (Waterhouse et al., 2009). Following alignment, a phylogenetic tree was constructed using the JTT + CAT model in FastTree version 2.1.8 (Price et al., 2010; Price et al., 2009). The resulting tree was processed with Treemmer version 0.3 (Menardo et al., 2018), a program which systematically reduces redundancy within the tree while maintaining representative diversity. After pruning and inspection, 5414 remaining sequences were re-aligned in MAFFT using the *--auto* setting. Sequences within this smaller set were closely inspected; sequences with large gaps, partial cover, and those without the six invariant histidine residues present in both oxygen reductases and NORs were omitted, leaving 5360 sequences. The sequences within this candidate set were then classified into A-type, B-type, and C-type oxygen reductases and NORs based on the manually classified oxygen reductase and NOR sequences present in the HCO database established by Sousa and colleagues in 2011 (Sousa et al., 2011). Following classification, a phylogenetic tree containing the 5360 sequences was generated using IQTree version 2.1.3 COVID-edition for Linux, 64-bit (Minh et al., 2020) with ModelFinder (Kalyaanamoorthy et al., 2017). ModelFinder identified Q.pfam+R10 (Minh et al., 2021) as the best-fitting model for the resulting phylogeny based on the Bayesian information criterion. Bipartition and branch supports were assessed using ultrafast bootstraps (Hoang et al., 2018) and Shimodaira-Hasegawa (SH) approximate likelihood-ratio tests (Guindon et al., 2010). The resulting unrooted maximum-likelihood tree is shown in Fig. 1, and its groups are coloured in part using the colour scheme of the HCO enzyme superfamily described by Sousa and colleagues in 2011 and 2012 (Sousa et al., 2012; Sousa et al., 2011). The tree was rooted using the Minimal Ancestor Deviation (MAD) algorithm (Tria et al., 2017); this rooting recovers NOR as an outgroup to the A-, B-, and C-type HCO enzymes.

From this tree, each HCO subtype and NORs were subsampled using the Jalview remove redundancy tool; the tool identifies pairs of sequences with high percent identity and discards the shorter sequence. A percent identity threshold of 97 % was applied for each group of enzymes; A-type Cyanobacteria and eukaryotic groups were manually subsampled to preserve nodes necessary for molecular clock calibration. Once downsampled, the resulting sequences were re-aligned in MAFFT using the L-INS-i algorithm, which is an iterative refinement method appropriate for sequences with long gaps and a conserved domain. The resulting alignment was further refined manually in preparation for molecular clock analysis and involved the correction of misalignments, the removal of large gaps and long N- and C-terminal extensions, and the removal of poorly aligned sequences (Fig. 2). This manual curation resulted in a multiple sequence alignment containing 386 sequences with 423 aligned amino acid sites.

This curated alignment was used to generate maximum-likelihood trees for molecular clock analysis using IQTree with ModelFinder, which determines the best-fit model from a variety of basic substitution models (Kalyaanamoorthy et al., 2017), and with GHOST, a complex model for heterotachously evolved sequences (Crotty et al., 2019). The resulting models selected from this survey were compared based on their log-likelihoods, the Akaike information criterion (AIC), and the Bayesian information criterion (BIC) (**Supplementary Table 2**). In each case, the GHOST models with both linked and unlinked parameters and 2, 4, or 16 classes had less favorable AIC and BIC scores as well as comparable or less favorable log-likelihood scores due to the large

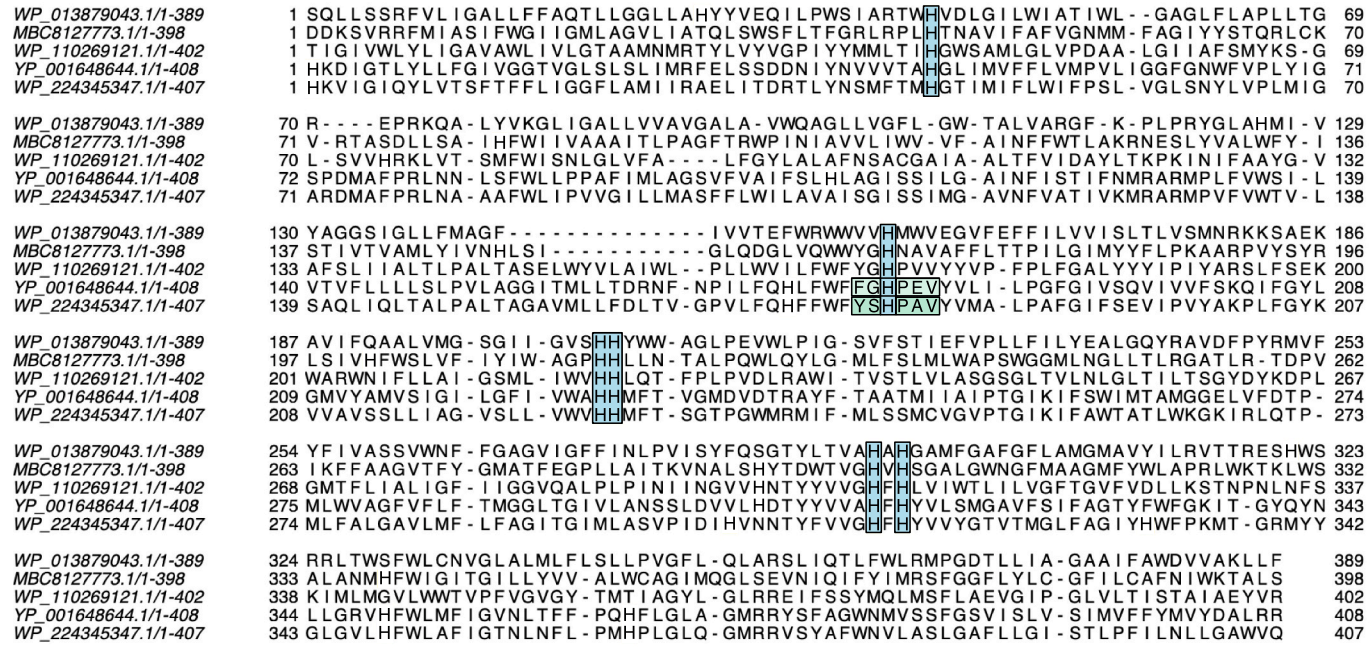


Fig. 2. Multiple sequence alignment of heme-copper oxidase family members. The residues outlined in black and shaded in blue represent the six invariant histidines present in subunit I for all members of the family; the sequence segments outlined in black and shaded in green represent the A1- and A2-type sequence motifs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

number of additional free parameters than the models selected by ModelFinder. As such, these complex models were not considered further for use in this study.

Two maximum-likelihood trees were pursued after selection by ModelFinder. For the first tree, Topology 1, ModelFinder selected the LG + F + I + G4 model, and bipartition and branch supports were assessed with 100 standard bootstraps and approximate likelihood-ratio tests. In this topology, the A-type oxygen reductases descend from within the diversity of B-type oxygen reductases. The second tree, Topology 2, was generated using a guide tree (**Supplementary Fig. 1**) that forces the monophyly of each cytochrome oxidase subtype and NORs as depicted in Fig. 1. For this tree, ModelFinder selected LG + F + R10, and bipartition and branch supports were assessed with 1000 ultrafast bootstraps and Shimodaira-Hasegawa approximate likelihood-ratio tests. Both trees were rooted using the MAD algorithm; the rooting of both trees is consistent and separates the oxygen reductases from the NORs. The tree topologies are compared in more detail in **Results**.

2.2. Molecular clock calibration and analyses

Four nodes on both tree topologies were assigned hard-bound age constraints for molecular clock analyses (Table 2). A hard-bound root age of 4400 Ma to 2750 Ma was imposed on the trees in which the older bound corresponded to the oldest evidence of liquid water and habitability following the moon-forming impact (Wilde et al., 2001), and the younger bound corresponded to the estimated younger-bound age of crown Cyanobacteria in published studies (Fournier et al., 2021), as this clade is present within the tree. Therefore, the root age calibration (prior) is sufficiently broad to encompass all plausible hypotheses for the deep common ancestry of these protein families.

Within the eukaryotic groups represented in the tree, two groups, Archaeplastida and Filozoa, showed phyletic patterns consistent with vertical inheritance and were selected for calibration. The first group, Archaeplastida, was selected due to the recovery of Rhodophyta grouping sister to Viridiplantae (as represented by both Chlorophyta and Streptophyta) and was constrained to 1600 to 1300 Ma, consistent with

Table 2
Age constraints applied to molecular clock analyses in this study. Calibrations were applied as hard bounds to specific nodes in the maximum-likelihood trees generated to estimate the emergences of oxygen reductase subtypes. Ages are specified in millions of years (Ma).

Age Constraint	Older Bound (Ma)	Younger Bound (Ma)	Source(s)
Root	4400	2750	Age estimate for liquid water from detrital zircons (Wilde et al., 2001) corresponding to the earliest potential habitability of Earth; younger bound age of crown Cyanobacteria as estimated in Fournier et al. (2021)
Archaeplastida	1600	1300	Age estimate for the common ancestor of Rhodophyta and Chlorophyta, calibrated by 7 fossil constraints in Parfrey et al. (2011) and 8 fossil constraints in Eme et al. (2014)
Filozoa	1050	875	Age estimate for the common ancestor of Filasterea and Choanoflagellatea based on 6 fossil constraints in Parfrey et al. (2011)
Synechococcus -Prochlorococcus	666	494	Age estimate for the common ancestor of *Synechococcus* and *Prochlorococcus* species as calibrated by 3 cyanobacterial fossils and 4 eukaryotic plastid lineage fossils in Fournier et al. (2021) and as calibrated by 1 diatom and 2 cyanobacterial fossils in Sánchez-Baracaldo (2015)

previously calculated age estimates for Archaeplastida and Hacrobia, which used hard-bound fossil constraints (Eme et al., 2014) and the ages calculated for green algae, red algae, and land plants (Parfrey et al., 2011). The second group, Filozoa, was identified by the close basal grouping of Filasterea and Choanoflagellata to crown Metazoa and was constrained with a narrowed version of the age estimates recovered for Holozoa and Filozoa reported in Parfrey et al. (2011): 1050 Ma to 850 Ma. Within Cyanobacteria, a deep duplication precedes two groups with phylogenies and taxon samplings consistent with crown Cyanobacteria. A clear vertical pattern of inheritance was observed for the Prochlorococcaceae and Synechococcaceae groups in the cyanobacterial tree, leading to their selection for calibration. The common ancestor of these cyanobacterial families was constrained to 666 to 494 Ma as previously reported (Fournier et al., 2021; Sánchez-Baracaldo, 2015).

Following calibration, posterior and prior divergence time estimates for the oxygen reductases and NORs in both fixed tree topologies were calculated using PhyloBayes 4.1c (Lartillot et al., 2015; Lartillot et al., 2007; Lartillot and Philippe, 2006; Lartillot and Philippe, 2004). In total, 72 molecular clocks were run. For both topologies, the following run parameters were applied: an LG amino-acid substitution matrix (Le and Gascuel, 2008), 4 categories for the discrete gamma distribution of rates across sites, and a gamma-distributed root prior with a mean age of 3575 Ma and a standard deviation of 410 Ma, consistent with the hard-bound root calibration. Convergence was assessed every 100 run cycles for clock chains for posterior distributions. Clocks were run with either all 4 calibrations, the root calibration and eukaryotic calibrations, or with the root calibration and cyanobacterial calibrations. Each clock was run with and without birth-death priors on divergence times under the following relaxed clock models: UGAM, or uncorrelated gamma multipliers (Drummond et al., 2006), LN, or log-normal autocorrelated clock (Thorne et al., 1998), and CIR, or the Cox-Ingersoll-Ross process (Lepage et al., 2006). Convergence was monitored during each run by viewing the continuous difference files generated by the PhyloBayes program and was tested using recommended cutoffs (Lartillot et al., 2015). The burn-in for each clock's chains was manually determined by plotting their log-likelihood values; 5 % or less of each run's initial cycles were discarded before the generation of posterior and prior distributions.

3. Results

The phylogenetic and molecular clock analyses of two topologies, which differ mainly in the placement of the low-affinity, A-type HCO enzymes, enabled the recovery of conservative estimates for the Archean emergence of low-affinity oxygen reductases prior to the GOE. Ages recovered for the A-type oxygen reductases are consistent with previously reported age estimates for the emergence of Cyanobacteria (Fournier et al., 2021).

3.1. Heme-copper oxygen reductase phylogenies

The maximum-likelihood tree constructed from 5360 heme-copper oxygen reductase and NOR sequences recovered distinct groups of oxygen reductase subtypes, apart from NORs (Fig. 1). This general, unrooted topology has been recovered in other studies (Ducluzeau et al., 2014; Sousa et al., 2012; Sousa et al., 2011). Rooting by the MAD algorithm placed the root between the HCO enzymes and the NOR group, consistent with an outgroup rooting based on the functional differences between the enzyme groups.

Within the oxygen reductase subtypes, differing distributions of Bacteria, Archaea, and Eukaryota were recovered. In agreement with previous classifications of heme-copper oxygen reductase constituents (Pereira et al., 2001; Sousa et al., 2012; Sousa et al., 2011), bacteria were distributed across all oxygen reductase subtypes as well as NORs, eukaryotes grouped within the A1-type oxygen reductases, and most taxa from Cyanobacteria grouped within the A2-type oxygen reductases and C-type oxygen reductases. Within the C-type oxygen reductases, however, a small group of archaea belonging to the order Methanosarcinales clustered within the cyanobacterial oxygen reductases. Archaea have not been previously reported to have C-type oxygen reductases, and further, members of the Methanosarcinales are not known to conduct aerobic respiration, though a recent study independently detected nitrous oxide/oxygen reductases in class II methanogens (Lyu and Lu, 2018). Cursory searches for cytochrome *c* oxidases within Methanosarcinales in the NCBI Protein database reveal multiple protein sequences annotated as cbb3-type cytochrome *c* oxidase subunit I, possibly revealing that distinguishing characteristics for the C-type oxygen reductases and NORs may fall outside subunit I for Archaea. Given the similarities between the C-type oxygen reductases and NORs, it is possible that these sequences may describe homologous proteins involved in intracellular nitrous oxide or oxygen detoxification schemes or adaptation to more oxidative environments.

31 novel archaeal sequences not reported in previous cytochrome oxidase phylogenies were recovered, grouping basally to the A-type oxygen reductases. Inspection of these archaeal sequences shows the presence of proton conducting channels with unique substitutions potentially impacting their function. For example, the sequence corresponding to accession EQB66259.1 contains a substitution in the position of its A1-type D-channel motif. In the place of the XGHPEV motif, the mutated segment contains WGHPLV, suggesting a substitution of a leucine for a glutamate. A group of archaeal sequences from members of Nitrososphaerota also group basal to the A-type oxygen reductases, and in the positions of their A1-type D-channel motif they contain a TFHPEV segment, potentially representing a substitution of a phenylalanine for a glycine. These sequences may represent transitional states between A- and B- type oxygen reductases or reflect the active loss of a functional D-channel in some archaeal species – possibilities discussed further in **Section 4.0**. These substitutions may make these enzymes more functionally similar to B-type oxygen reductases despite their grouping with an earlier classification as A-type oxygen reductases. During subsampling of the survey tree, 2 of the 31 archaeal sequences were included following the removal of closely related sequences in order to distill the extant diversity of the oxygen reductases.

Topology 1 and Topology 2 were broadly phylogenetically consistent. The principal difference between these topologies was the forced bipartition between the A-type and B-type HCO enzymes in Topology 2; compared to Topology 1, this resulted in only modest changes to the phylogeny within these HCO enzyme subtypes. This forced bipartition resulted in a tree that was favored in terms of log-likelihood, indicating that assuming the monophyly of each of the HCO subtypes in Topology 2 is appropriate for estimating the divergence time of the low affinity, A-type HCO enzymes.

Eukaryotic and cyanobacterial clades present within the A-type oxygen reductases were used to calibrate the molecular clocks run with both topologies. The expanded clades for eukaryotes in Topology 1 (**Supplementary Fig. 2**) and in Topology 2 (**Supplementary Fig. 3**) reveal that while broad diversity within Eukaryota was recovered, the topology of the A1-type oxygen reductases from eukaryotes was incongruent with the overall eukaryotic species tree. However, within eukaryotes, some major groups were congruent with known species-tree relationships, specifically enabling the secondary calibration of nodes corresponding to crown Archaeplastida and crown Filozoa. The phylogeny of Cyanobacteria in Topology 1 (**Supplementary Fig. 4**) and Topology 2 (**Supplementary Fig. 5**) shows a duplication of the A1-type prior to the diversification of crown Cyanobacteria. Both cyanobacterial sub-trees recover similar relationships and taxonomic distributions across major cyanobacterial groups, with the basally branching *Gloeobacteraceae* sister to the other cyanobacteria in each case.

3.2. Molecular dating of the heme-copper oxygen reductases

Molecular clock analyses for both tested topologies recover pre-GOE emergences of crown HCO enzymes. The posterior age estimates generated by the molecular clock calculations vary depending on the calibrations employed and the models used (**Supplementary Tables 3–6**). The posterior age estimates generated by the molecular clocks generally overlapped with their prior age estimates, which were calculated with all parameters being equal minus the contribution of sequence data (**Supplementary Tables 7–10**). Models which applied birth-death priors on divergence times recovered older and often larger age intervals compared to models run with uniform priors. Among the clock models employed, CIR consistently calculated the oldest mean age estimates for the nodes sampled, while UGAM consistently recovered the youngest mean ages. Analysis with jackknifed, or separate groupings of calibrations, revealed that when only root and cyanobacterial calibrations were imposed on the molecular clock, the resulting age estimates were older than those produced when only the root and eukaryotic calibrations were employed. There was only one exception to this finding – the UGAM with birth-death priors model tended to produce distinctly younger posterior age estimates when only the root and cyanobacterial calibrations were used. The recovery of older ages when this clock was run without the birth-death prior likely reflects an exaggerated interaction between the birth-death prior and this configuration of the clock model.

When all calibrations were included, the resulting age estimates reflected the relative impacts each set of calibrations had within that particular run and model. Among the clock models tested, the log-normal autocorrelated clock with the birth-death prior yielded the most consistent age estimates across all the combinations of calibrations tested. These ages and their 95 % credible interval ranges (95 % CI) are discussed below. Chronograms of Topology 1 (Fig. 4) and Topology 2 (Fig. 5) show ages recovered from this model. Ages derived from other models are available in **Supplementary Tables 3–10**.

For both topologies, the age estimates recovered for the common ancestor of A-type and B-type oxygen reductases was older than the age estimates recovered for the common ancestors of C-type oxygen reductases or NORs. In Topology 1, the ancestor to the B-type oxygen reductases, from which the A-type oxygen reductases descend, was assigned a mean age of 3587 Ma (95 % CI 3914 to 3246 Ma), while in Topology 2, the common ancestor of both B-type and A-type oxygen reductases was assigned a younger age of 3399 Ma (95 % CI 3758 to 3033 Ma). For the common ancestor of the A-type oxygen reductases, Topology 1 yielded an age of 3188 Ma (95 % CI 3487 to 2877 Ma), while Topology 2 yielded an age of 3211 Ma (95 % CI 3537 to 2860 Ma). The distributions for these age estimates (Fig. 6) are effectively normal and unaffected by significant skew. While the age estimates for these nodes across both topologies were generally consistent, the slightly older age calculated for the A-type oxygen reductases in Topology 2 likely resulted from the deeply branching, basal archaeal sequences mentioned earlier. The phylogenetic placement of these sequences within the A-type oxygen reductases in both topologies pushed back the ages for the subtype's ancestors by creating a common ancestor node between these archaeal sequences and those that more closely resemble the canonical A-type oxygen reductases. In the absence of these 2 archaeal sequences, the age estimates calculated for the common ancestor of the remaining A-type oxygen reductases was 2897 Ma for Topology 1 (95 % CI 3212 to 2642 Ma) and 2856 Ma for Topology 2 (95 % CI 3164 to 2559 Ma).

The ages for the cyanobacterial A-type oxygen reductases were additionally examined in order to consider the emergence of oxygen reductases in the context of oxygen production via oxygenic photosynthesis (Fig. 7). Within both topologies, the ages recovered for pre-

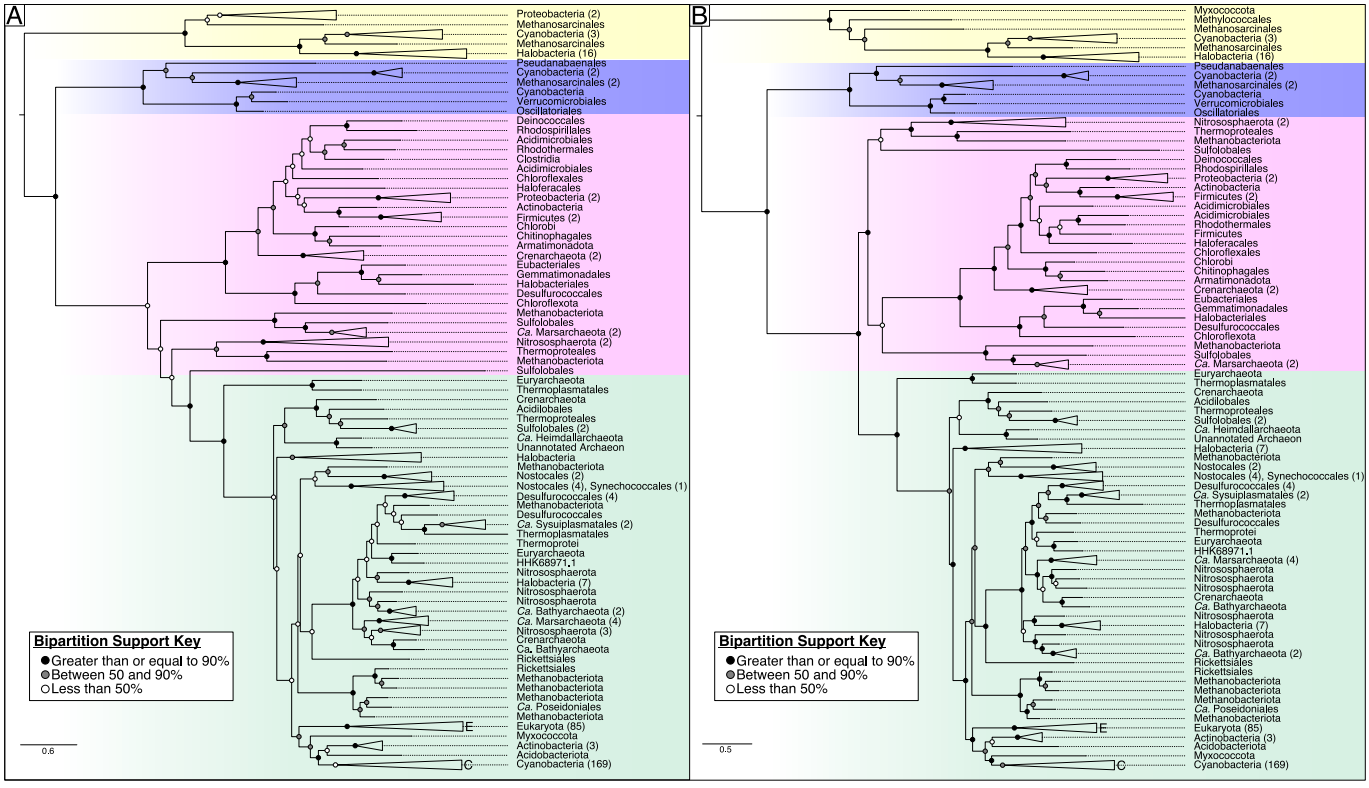


Fig. 3. Maximum-likelihood trees for the heme-copper oxygen reductases. Taxa and collapsed clades are labeled according to NCBI taxonomy annotations at the order level when possible; for taxa without order-level classification, the next highest level of classification was used. Topology 1 is shown in panel (A), and Topology 2 is shown in panel (B). Bipartition support values are coloured according to bootstrap support values, with nodes coloured black having ≥90 % support, nodes coloured gray having between 50 % and 90 % support, and nodes coloured white having <50 % support. Eukaryota and Cyanobacteria clades involved in calibration are collapsed and labeled E and C, respectively.

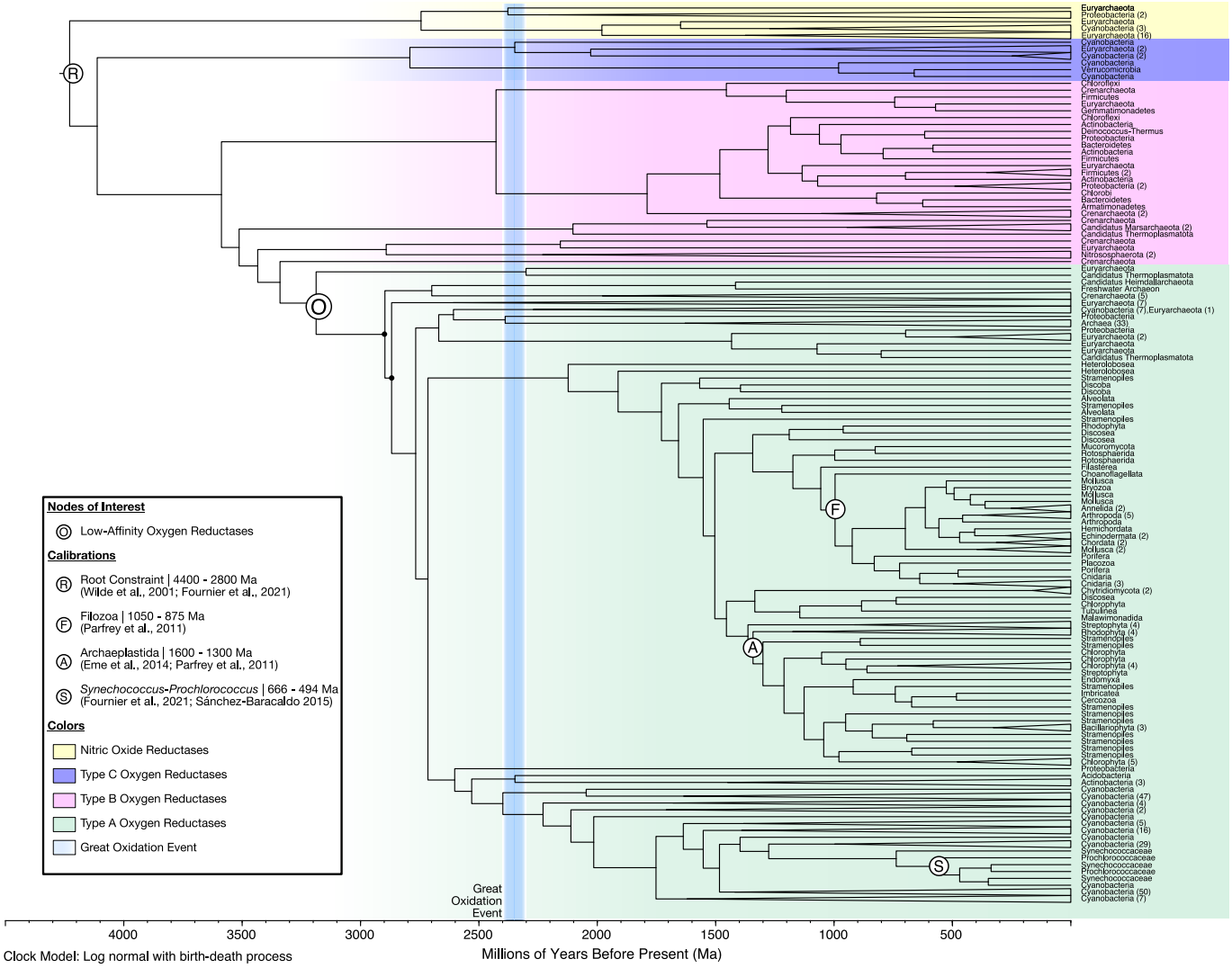


Fig. 4. Chronogram with Heme-Copper Oxygen Reductases using Topology 1. Age estimates are shown for the log-normal model with a birth-death process. Nodes marked with black dots represent the low-affinity oxygen reductases with basal archaeal representatives excluded.

duplication cyanobacterial A-type oxygen reductases push up against the onset of the GOE (2400 to 2300 Ma). In Topology 1, the mean age calculated for cyanobacterial oxygen reductases was 2399 Ma, while in Topology 2, the mean age calculated was 2359 Ma. The 95 % credible intervals for these age estimates extend to 2644 and 2626 Ma, respectively, further suggesting that the topologies independently recover a diversification of cyanobacterial A-type oxygen reductases before the establishment of a persistently oxygenated planet. Both topologies additionally show that the cyanobacterial A-type oxygen reductases undergo major diversification into and after the GOE.

4. Discussion

The analyses presented in this work investigate the evolutionary history of HCO enzymes, the components of aerobic electron transport chains which directly interface with molecular oxygen, as a proxy to reconstruct the evolutionary history of aerobic respiration. Molecular clock analyses of these protein families recover middle Archean age estimates for the advent of low-affinity, A-type oxygen reductases and late Archean age estimates for the cyanobacterial A-type oxygen reductases. Archean aerobic respiration, mediated by HCO enzymes, may have functioned as a biological sink for molecular oxygen that forestalled the establishment of a permanently oxygenated planet.

4.1. Topological Impacts

Significant attention has been directed towards discerning the origins and history of the HCO enzymes, ordering the appearances of the heme-copper oxygen reductase subtypes, as well as determining their relationship with the nitric oxide reductases. Differing views on the appropriate roots for phylogenetic trees of oxygen reductases and NORs, and even whether rooting is possible (Sousa et al., 2012), have led to contrasting views on the antiquity of particular subtypes (Ducluzeau et al., 2014; Gribaldo et al., 2009). The phylogenetic trees considered in this work were rooted by minimal ancestor deviation, independent of the identification of an outgroup. The position of the root selected by this technique suggests that a series of gene duplications, the first of which led to the NORs and the ancestral uroxidase which preceded the development of the current heme-copper oxygen reductase subtypes. It further implies that HCO enzymes share a monophyletic origin, consistent with hypotheses presented in earlier studies (Castresana et al., 1994; Pereira et al., 2001). The second duplication, which took place among the ancestral uroxidases, resulted in the divergence of the C-type oxygen reductases from the more closely grouped B-type and A-type oxygen reductases. This topology is consistent with the classical view of oxygen reductase evolution discussed by Gribaldo and colleagues in 2009, but contrasts with the rooting position proposed in that work in

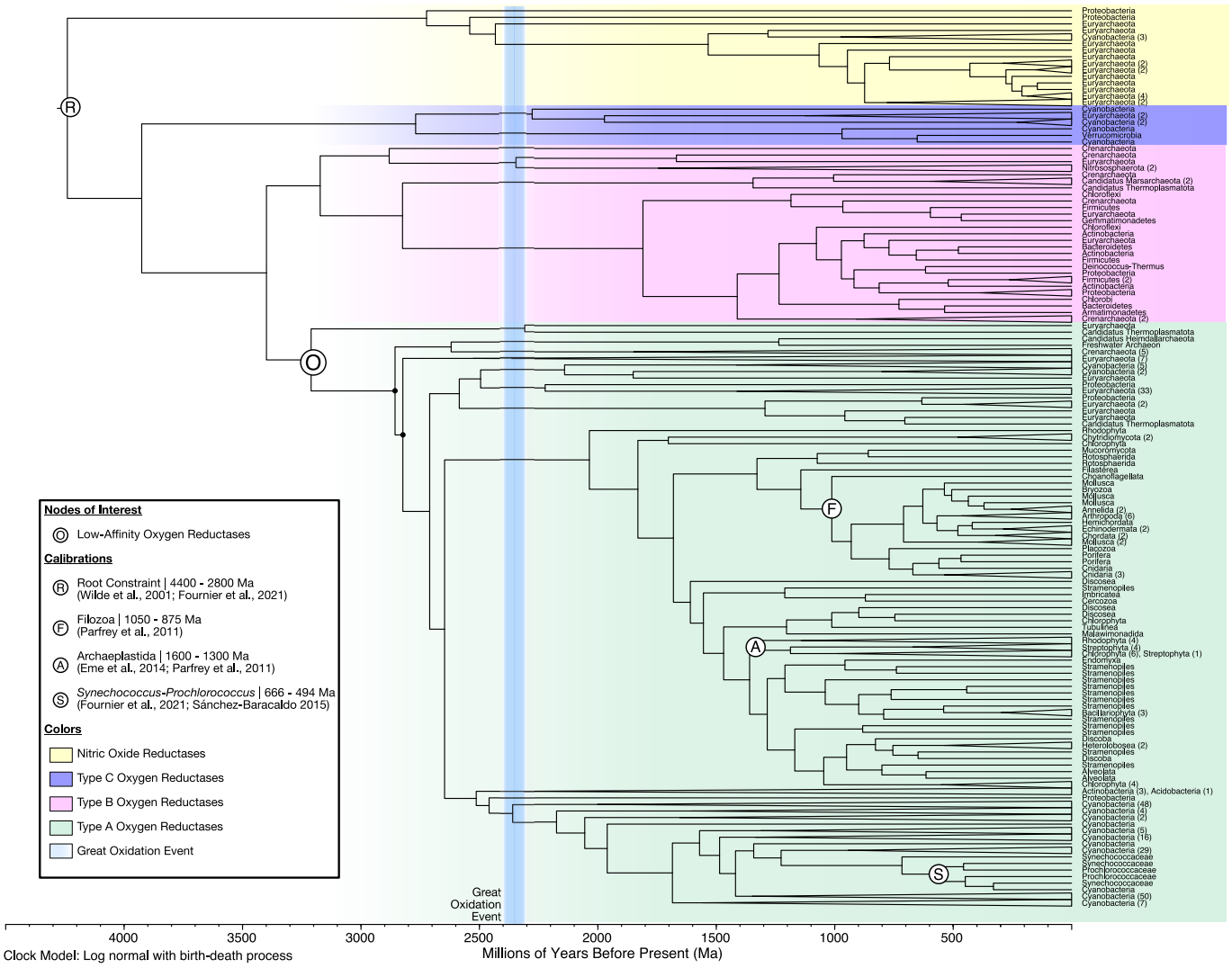


Fig. 5. Chronogram with Heme-Copper Oxygen Reductases using Topology 2. Age estimates are shown for the log-normal model with a birth-death process. Nodes marked with black dots represent the low-affinity oxygen reductases with basal archaeal representatives excluded.

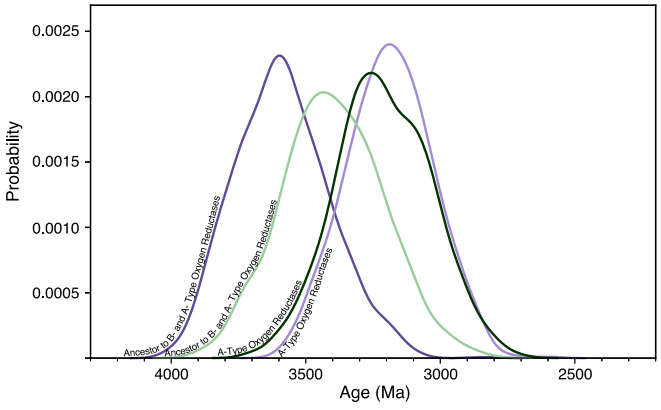


Fig. 6. Divergence time estimates for the A- and B-type oxygen reductases. Ages are shown for Topology 1 (coloured purple) and Topology 2 (coloured green). Probability curves are labeled according to the divergences they represent. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

which the root is placed between the archaeal and bacterial A-type oxygen reductases (Gribaldo et al., 2009).

The two topologies compared in this work diverge following the second proposed gene duplication: in Topology 1, A-type oxygen reductases are monophyletic and nested within paraphyletic B-type oxygen reductase diversity, while in Topology 2, both A-types and B-types are monophyletic, as recovered in previously published phylogenies with much lower taxon sampling densities (Castresana et al., 1994; Pereira et al., 2001; Saraste and Castresana, 1994; Sousa et al., 2012; Sousa et al., 2011). The lack of a distinct separation between the B-type and A-type oxygen reductases in Topology 1 may result from the inclusion of the basal archaeal sequences, which phylogenetically group within the margins of the A-type HCO enzymes. These sequences, which are classified as the catalytic subunits of oxygen reductases through automatic annotation pipelines (HHA13956.1 was assigned protein motif PF00115 and EQB66259.1 was assigned to CDD 440605) and contain the six invariant histidines common to all oxygen reductases, may have a modified D-channel given the residue deviations in their D-channel motifs. Alternatively, they may represent another possible evolutionary intermediate between the A-types and B-type oxygen reductases, as is hypothesized for *Nitrosopumilus maritimus* in Han et al.,

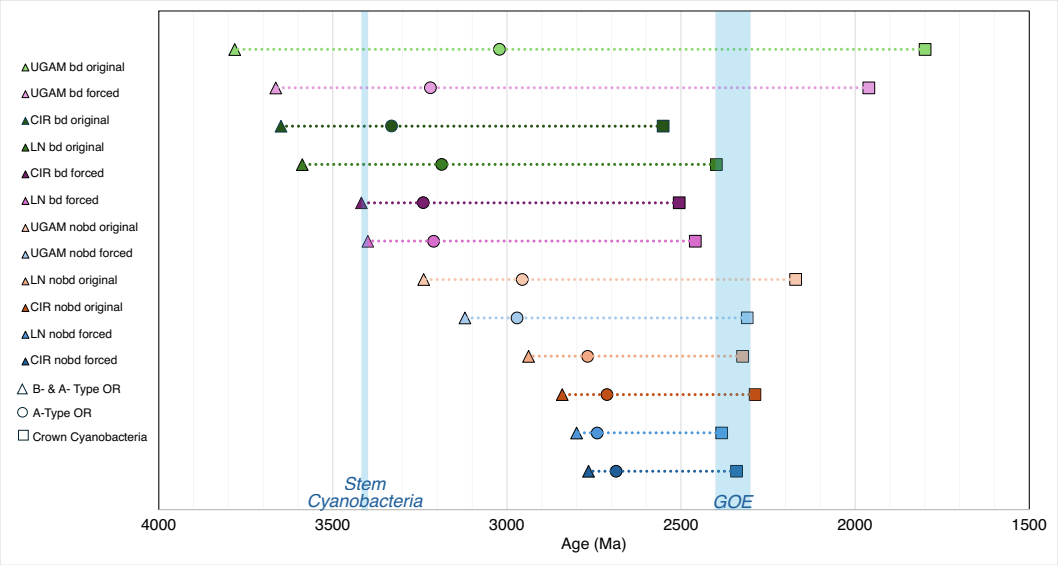


Fig. 7. Distribution of mean ages of A- and B-type oxygen reductases and crown Cyanobacteria. Ages are shown for Topology 1 (labelled original for original bipartition) and Topology 2 (labelled forced for forced bipartition). The mean age estimate for the emergence of stem Cyanobacteria is shown at ~3400 Ma (Fournier et al., 2021) and is coloured in blue; the Great Oxidation Event is shown from 2400 to 2300 Ma and is coloured in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2011, which has a structurally identical D-channel to the A-type oxygen reductases and a K-channel analog resembling that of the B-type oxygen reductases (Han et al., 2011). The existence of HCO enzyme subtype intermediates is consistent with the recovery of the topology in which the A-type oxygen reductases descend from within the B-type oxygen reductase diversity, in that a sharp division between the extant two subtypes does not exist (Ducluzeau et al., 2014). Dedicated examination of the structures and functionality of these archaeal enzymes, as well as those of their other subunits, may yield insights that further elucidate the evolutionary history and functional diversity of these enzymes.

4.2. Oxygen reductases in cyanobacteria

The deep duplication recovered among the A-type containing cyanobacteria in the trees makes it possible to infer that aerobic respiration was ancestral to these microbes and was present in stem Cyanobacteria, and also provides a novel means of constraining the origin of oxygenic photosynthesis within the stem cyanobacterial lineage. Furthermore, this study resulted in relatively young age estimates for crown Cyanobacteria for Topology 1 (95 % CI 2644 to 2172 Ma) and Topology 2 (95 % CI 2626 to 2097 Ma), as it did not explicitly constrain the age of this group. Future studies focused on the emergence of oxygen reductases in cyanobacteria will therefore likely provide a powerful means of further constraining the age of both aerobic respiration and oxygenic photosynthesis in both relative and absolute dating schemas. Nonetheless, the ages recovered here suggest that a major diversification of cyanobacterial HCO enzymes took place during the ecological reorganization of oxygenated and anoxic environments on Earth before or during the GOE, while many other groups of organisms capable of aerobic respiration diversified later.

4.3. Reconciliation of geochemical and geobiological records

Multiple independent records assembled by the examination of Archean rocks track distinct events which suggest that molecular oxygen may have been at least transiently available before the GOE in terrestrial and marine environments. The onset of the GOE around 2400 to 2300 Ma (Catling and Zahnle, 2020) corresponds to the approximate time interval after which signals of mass independent fractionation of sulfur isotopes in sulfur minerals largely disappear in the rock record. Departures from sulfur mass independent fractionation have been modeled to occur at oxygen levels greater than 10^{-6} times present oxygen levels (PAL), meaning that levels of oxygen less than 10^{-6} PAL, if present before the GOE, would not be recorded as mass dependent fractionation in sulfur minerals (Catling and Zahnle, 2020; Pavlov and Kasting, 2002).

Many geochemical paleoredox proxies record instances of "whiffs," or the local, transient accumulation of molecular oxygen prior to persistent oxygenation at the close of the GOE (Anbar et al., 2007) and may represent instances of oxygen accumulation resulting from oxygenic photosynthesis that overwhelmed local geochemical reductants. During oxidative continental weathering, redox sensitive elements are liberated from crustal minerals, enabling their transport to and accumulation in seawater, and, eventually, their deposition. Elevated abundances of molybdenum and rhenium in the late Archean Mt. McRae Shale (~2501 Ma) and of elevated molybdenum at a second site, the coeval Klein Naute Formation, have been associated with oxidative weathering (Anbar et al., 2007; Kendall et al., 2010). Later studies of selenium and osmium isotope distributions at the Mt. McRae Shale recover excursions similarly attributed to the presence of oxygen (Kendall et al., 2015; Stüeken et al., 2015). Additional isotope systems have been employed to infer the presence of molecular oxygen before the GOE, including chromium and molybdenum. These elements undergo characteristic isotopic fractionations when they are sorbed onto manganese oxides (Crowe et al., 2013; Ostrander et al., 2021; Planavsky et al., 2014), minerals which form in the presence of oxygen in terrestrial and marine settings. Paleosols from the Pongola Supergroup, including the Nsuze paleosol, which formed between 2980 Ma and 2960 Ma, and the 2950-Ma-old Sinqeni Formation contain isotopic signatures from chromium (Crowe et al., 2013) and molybdenum (Planavsky et al., 2014), respectively, consistent with the presence of manganese oxides. While recent works have highlighted molecular oxygen independent mechanisms for manganese oxide formation (Daye et al., 2019; Liu et al., 2020), the extent of the operation of these mechanisms in Archean settings remains unclear. Other inorganic geochemical proxies have

recorded changes in the extent of oxygen-mediated weathering during the late Archean (Kendall, 2021; Lyons et al., 2014; Ostrander et al., 2021). Morphological evidence in the form of fossil bubbles associated with cyanobacterial oxygen-production has been recovered in 3220-Ma-old (Homann et al., 2015) and in 2700-Ma-old microbial mats (Bosak et al., 2009). These independent geochemical and morphological records of molecular oxygen availability in a variety of environmental settings during the middle to late Archean are largely compatible with the recovery of a Mesoarchean (3200 to 2800 Ma) emergence and diversification of oxygen-requiring HCO enzymes (Fig. 3 and Fig. 4).

A growing complement of genomic investigations further supports the possibility of pre-GOE interactions between life and accumulated oxygen. The Archean emergence of Cyanobacteria, the microbes capable of oxygenic photosynthesis, has been recovered by fossil- and gene transfer-calibrated molecular clock analyses (Boden et al., 2021; Fournier et al., 2021; Sánchez-Baracaldo, 2015). Superoxide dismutase enzymes with copper and zinc cofactors, which catalyze the removal of superoxide free radicals, have been estimated to have emerged in Cyanobacteria prior to the GOE using fossil-calibrated and geochemically-constrained molecular clocks (Boden et al., 2021). Phylogenetic analyses of high redox potential respiratory quinones, molecules involved in electron transport in aerobic respiration and photosynthesis, revealed a pre-GOE origin for these molecules (Elling et al., 2025). Further, detailed phylogenetic analyses of a collection of oxygen-interfacing enzymes suggest an uptick in diversification at 3100 Ma (Jabłońska and Tawfik, 2021). Additionally, recent analyses, which sought to time-calibrate a tree for Bacteria using the GOE, yielded evidence of some aerobic bacterial lineages arising prior to the GOE (Davín et al., 2025). Therefore, the ages recovered here for the origin and diversification of aerobic respiration metabolisms across microbial groups are broadly consistent with an emerging consensus of an Archean biosphere that included the biological cycling of oxygen.

The diversification of the oxygen reductase subtypes prior to the GOE suggests the presence of persistently oxygenated environments for which geochemical evidence is scarce and still subject to debate. However, a new view is emerging regarding the operational oxygen requirements of HCO enzymes that may aid in the explanation of aerobic respiration in the absence of geochemical evidence for oxygen. Metatranscriptomic surveys of microbial life across oxygen regimes have revealed that while high-affinity oxygen reductases may be more highly expressed in environments with <1 μM O_2/L, low-affinity, A-type oxygen reductases are also still expressed (Berg et al., 2022). Recent examinations of the oxygen requirements for the operation of high and low-affinity oxygen reductases in soil bacteria have shown that low-affinity oxygen reductases can be expressed in oxygen concentrations as low as 1 nM O_2/L without the simultaneous detectable expression of high-affinity oxygen reductases (Trojan et al., 2021). An experiment conducted by Trojan and colleagues, which subjected a low-affinity oxygen reductase containing bacterium to an increasingly diminished oxygen supply showed that the rate of microbial oxygen consumption never exceeded the rate at which oxygen was supplied, maintaining oxygen availability below 0.01 μM O_2/L while other experiments showed that high-affinity oxygen reductases subject to similar conditions would entirely deplete their oxygen supplies, resulting in anoxic conditions (Trojan et al., 2021). These findings suggest that early aerobes with low-affinity oxygen reductases may have inhabited environments with exceedingly low levels of oxygen – simultaneously acting as a biological sink for that oxygen and potentially preventing its accumulation to the extent required to be recorded geochemically. Within this framework, Archean 'whiffs' of oxygen could represent accumulations of oxygen that exceeded the capacity of local geochemical and biological sinks. Aerobic life would have been restricted to environments which supported the growth of Cyanobacteria until higher levels of oxygen accumulated that ultimately initiated the GOE.

CRediT authorship contribution statement

Fatima Husain: Visualization, Methodology, Investigation, Conceptualization, Writing – review & editing, Writing – original draft. **Haitao Shang:** Conceptualization, Writing – review & editing. **Stilianos Louca:** Conceptualization, Writing – review & editing. **Gregory P. Fournier:** Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

FH was supported by the MIT EAPS John H. Carlson Fellowship and the MIT Hugh Hampton Young Memorial Fellowship for the duration of this research. This work was funded by a 2021 Scialog Collaborative Innovation Award (grant no. 2021-3128 awarded to GPF).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.palaeo.2025.113531.

Data availability

All data from this study is available in the main manuscript and in the supplementary material. Sequence and tree files have been deposited in the NASA Astrobiology Habitable Environments Database at the following link: https://doi.org/10.48667/rykp-pn33.

References

Al-Attar, S., De Vries, S., 2015. An electrogenic nitric oxide reductase. FEBS Lett. 589, 2050–2057. https://doi.org/10.1016/j.febslet.2015.06.033.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Blast Local Alignment Search Tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Anbar, A.D., Duan, Y., Lyons, T.W., Arnold, G.L., Kendall, B., Creaser, R.A., Kaufman, A. J., Gordon, G.W., Scott, C., Garvin, J., Buick, R., 2007. A whiff of oxygen before the great oxidation event? Science 317, 1903–1906. https://doi.org/10.1126/science.1140325.

Berg, J.S., Ahmerkamp, S., Pjevac, P., Hausmann, B., Milucka, J., Kuypers, M.M.M., 2022. How low can they go? Aerobic respiration by microorganisms under apparent anoxia. FEMS Microbiol. Rev. 46, fuac006. https://doi.org/10.1093/femsre/fuac006.

Boden, J.S., Konhauser, K.O., Robbins, L.J., Sánchez-Baracaldo, P., 2021. Timing the evolution of antioxidant enzymes in cyanobacteria. Nat. Commun. 12, 1–12. https://doi.org/10.1038/s41467-021-24396-y.

Bosak, T., Liang, B., Sim, M.S., Petroff, A.P., 2009. Morphological record of oxygenic photosynthesis in conical stromatolites. Proc. Natl. Acad. Sci. 106, 10939–10943. https://doi.org/10.1073/pnas.0900885106.

Brochier-Armanet, C., Talla, E., Gribaldo, S., 2009. The Multiple Evolutionary Histories of Dioxygen Reductases: Implications for the Origin and Evolution of Aerobic Respiration. Mol. Biol. Evol. 26, 285–297. https://doi.org/10.1093/molbev/msn246.

Calhoun, M.W., Thomas, J.W., Gennis, R.B., 1994. The cytochrome oxidase superfamily of redox-driven proton pumps. Trends Biochem. Sci. 19, 325–330. https://doi.org/10.1016/0968-0004(94)90071-X.

Castresana, J., Lübben, M., Saraste, M., Higgins, D.G., 1994. Evolution of cytochrome oxidase, an enzyme older than atmospheric oxygen. EMBO J. 13, 2516–2525. https://doi.org/10.1002/j.1460-2075.1994.tb06541.x.

Catling, D.C., Zahnle, K.J., 2020. The Archean atmosphere. Sci. Adv. 6. https://doi.org/10.1126/sciadv.aax1420.

Chen, J., Strous, M., 2013. Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution. Biochim. Biophys. Acta BBA - Bioenerg. 1827, 136–144. https://doi.org/10.1016/j.bbabio.2012.10.002.

Crotty, S.M., Minh, B.Q., Bean, N.G., Holland, B.R., Tuke, J., Jermiin, L.S., Haeseler, A.V., 2019. GHOST: Recovering Historical Signal from Heterotachously Evolved Sequence Alignments. Syst. Biol. syz051. https://doi.org/10.1093/sysbio/syz051.

Crowe, S.A., Døssing, L.N., Beukes, N.J., Bau, M., Kruger, S.J., Frei, R., Canfield, D.E., 2013. Atmospheric oxygenation three billion years ago. Nature 501, 535–538. https://doi.org/10.1038/nature12426.

Davín, A.A., Woodcroft, B.J., Soo, R.M., Morel, B., Murali, R., Schrempf, D., Clark, J.W., Álvarez-Carretero, S., Boussau, B., Moody, E.R.R., Szánthó, L.L., Richy, E., Pisani, D., Hemp, J., Fischer, W.W., Donoghue, P.C.J., Spang, A., Hugenholtz, P., Williams, T. A., Szöllősi, G.J., 2025. A geological timescale for bacterial evolution and oxygen adaptation. Science 388. https://doi.org/10.1126/science.adp1853.

Daye, M., Klepac-Ceraj, V., Pajusalu, M., Rowland, S., Farrell-Sherman, A., Beukes, N., Tamura, N., Fournier, G., Bosak, T., 2019. Light-driven anaerobic microbial oxidation of manganese. Nature 576, 311–314. https://doi.org/10.1038/s41586-019-1804-0.

de Vries, S., Schröder, I., 2002. Comparison between the nitric oxide reductase family and its aerobic relatives, the cytochrome oxidases. Biochem. Soc. Trans. 30, 662–667. https://doi.org/10.1042/bst0300662.

Degli Esposti, M., 2020. On the evolution of cytochrome oxidases consuming oxygen. Biochim. Biophys. Acta Bioenerg. 1861, 148304. https://doi.org/10.1016/j.bbabio.2020.148304.

Degli Esposti, M., Mentel, M., Martin, W., Sousa, F.L., 2019. Oxygen Reductases in Alphaproteobacterial Genomes: Physiological Evolution from Low to High Oxygen Environments. Front. Microbiol. 10, 499. https://doi.org/10.3389/fmicb.2019.00499.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, 699–710. https://doi.org/10.1371/journal.pbio.0040088.

Ducluzeau, A.-L., Ouchane, S., Nitschke, W., 2008. The *cbb*$_3$ Oxidases are an Ancient Innovation of the Domain Bacteria. Mol. Biol. Evol. 25, 1158–1166. https://doi.org/10.1093/molbev/msn062.

Ducluzeau, A.-L., Schoepp-Cothenet, B., Van Lis, R., Baymann, F., Russell, M.J., Nitschke, W., 2014. The evolution of respiratory O_2/NO reductases: An out-of-the-phylogenetic-box perspective. J. R. Soc. Interface 11. https://doi.org/10.1098/rsif.2014.0196.

Elling, F.J., Pierrel, F., Chobert, S.-C., Abby, S.S., Evans, T.W., Reveillard, A., Pelosi, L., Schnoebelen, J., Hemingway, J.D., Boumendjel, A., Becker, K.W., Blom, P., Cordes, J., Nathan, V., Baymann, F., Lücker, S., Spieck, E., Leadbetter, J.R., Hinrichs, K.-U., Summons, R.E., Pearson, A., 2025. A novel quinone biosynthetic pathway illuminates the evolution of aerobic metabolism. Proc. Natl. Acad. Sci. 122. https://doi.org/10.1073/pnas.2421994122.

Eme, L., Sharpe, S.C., Brown, M.W., Roger, A.J., 2014. On the Age of Eukaryotes: Evaluating Evidence from Fossils and Molecular Clocks. Cold Spring Harb. Perspect. Biol. 6. https://doi.org/10.1101/cshperspect.a016139.

Fournier, G.P., Moore, K.R., Rangel, L.T., Payette, J.G., Momper, L., Bosak, T., 2021. The Archean origin of oxygenic photosynthesis and extant cyanobacterial lineages. Proc. R. Soc. B Biol. Sci. 288, 1–10. https://doi.org/10.1098/rspb.2021.0675.

García-Horsman, J.A., Barquera, B., Rumbley, J., Ma, J., Gennis, R.B., 1994. The superfamily of heme-copper respiratory oxidases. J. Bacteriol. 176, 5587–5600. https://doi.org/10.1128/jb.176.18.5587-5600.1994.

Gribaldo, S., Talla, E., Brochier-Armanet, C., 2009. Evolution of the haem copper oxidases superfamily: a rooting tale. Trends Biochem. Sci. 34, 375–381. https://doi.org/10.1016/j.tibs.2009.04.002.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321. https://doi.org/10.1093/sysbio/syq010.

Han, H., Hemp, J., Pace, L.A., Ouyang, H., Ganesan, K., Roh, J.H., Daldal, F., Blanke, S. R., Gennis, R.B., 2011. Adaptation of aerobic respiration to low O_2 environments. Proc. Natl. Acad. Sci. USA 108, 14109–14114. https://doi.org/10.1073/pnas.1018958108.

Hendriks, J., Warne, A., Gohlke, U., Haltia, T., Ludovici, C., Lübben, M., Saraste, M., 1998. The active site of the bacterial nitric oxide reductase is a dinuclear iron center. Biochemistry 37, 13102–13109. https://doi.org/10.1021/bi980943x.

Hendriks, J.H.M., Jasaitis, A., Saraste, M., Verkhovsky, M.I., 2002. Proton and electron pathways in the bacterial nitric oxide reductase. Biochemistry 41, 2331–2340. https://doi.org/10.1021/bi0121050.

Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281.

Homann, M., Heubeck, C., Airo, A., Tice, M.M., 2015. Morphological adaptations of 3.22 Ga-old tufted microbial mats to Archean coastal habitats (Moodies Group, Barberton Greenstone Belt, South Africa). Precambrian Res. 266, 47–64. https://doi.org/10.1016/j.precamres.2015.04.018.

Jabłońska, J., Tawfik, D.S., 2021. The evolution of oxygen-utilizing enzymes suggests early biosphere oxygenation. Nat. Ecol. Evol. 5, 442–448. https://doi.org/10.1038/s41559-020-01386-9.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.

Kendall, B., 2021. Recent advances in geochemical paleo-oxybarometers. Annu. Rev. Earth Planet. Sci. 49, 399–433. https://doi.org/10.1146/annurev-earth-071520-051637.

Kendall, B., Reinhard, C.T., Lyons, T.W., Kaufman, A.J., Poulton, S.W., Anbar, A.D., 2010. Pervasive oxygenation along late Archaean ocean margins. Nat. Geosci. 3, 647–652. https://doi.org/10.1038/ngeo942.

Kendall, B., Creaser, R.A., Reinhard, C.T., Lyons, T.W., Anbar, A.D., 2015. Transient episodes of mild environmental oxygenation and oxidative continental weathering during the late Archean. Sci. Adv. 1, 1–6. https://doi.org/10.1126/sciadv.1500777.

Lartillot, N., Philippe, H., 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol. Biol. Evol. 21, 1095–1109. https://doi.org/10.1093/molbev/msh112.

Lartillot, N., Philippe, H., 2006. Computing Bayes factors using thermodynamic integration. Syst. Biol. 55, 195–207. https://doi.org/10.1080/10635150500433722.

Lartillot, N., Brinkmann, H., Philippe, H., 2007. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. BMC Evol. Biol. 7, 1–26. https://doi.org/10.1186/1471-2148-7-S1-S4.

Lartillot, N., Blanquart, S., Lepage, T., 2015. Phylobayes: A Bayesian Software for Phylogenetic Reconstruction Using Mixture Models. Software Manual.

Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25, 1307–1320. https://doi.org/10.1093/molbev/msn067.

Lepage, T., Lawi, S., Tupper, P., Bryant, D., 2006. Continuous and tractable models for the variation of evolutionary rates. Math. Biosci. 199, 216–233. https://doi.org/10.1016/j.mbs.2005.11.002.

Liu, W., Hao, J., Elzinga, E.J., Piotrowiak, P., Nanda, V., Yee, N., Falkowski, P.G., 2020. Anoxic photogeochemical oxidation of manganese carbonate yields manganese oxide. Proc. Natl. Acad. Sci. 117, 22698–22704. https://doi.org/10.1073/pnas.2002175117.

Lyons, T.W., Reinhard, C.T., Planavsky, N.J., 2014. The rise of oxygen in Earth's early ocean and atmosphere. Nature 506, 307–315. https://doi.org/10.1038/nature13068.

Lyu, Z., Lu, Y., 2018. Metabolic shift at the class level sheds light on adaptation of methanogens to oxidative environments. ISME J. 12, 411–423. https://doi.org/10.1038/ismej.2017.173.

Menardo, F., Loiseau, C., Brites, D., Coscolla, M., Gygli, S.M., Rutaihwa, L.K., Trauner, A., Beisel, C., Borrell, S., Gagneux, S., 2018. Treemmer: a tool to reduce large phylogenetic datasets with minimal loss of diversity. BMC Bioinformatics 19, 164. https://doi.org/10.1186/s12859-018-2164-8.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Minh, B.Q., Dang, C.C., Vinh, L.S., Lanfear, R., 2021. QMaker: Fast and Accurate Method to Estimate Empirical Models of Protein Evolution. Syst. Biol. 70, 1046–1060. https://doi.org/10.1093/sysbio/syab010.

Morris, R.L., Schmidt, T.M., 2013. Shallow breathing: bacterial life at low O_2. Nat. Rev. Microbiol. 11, 205–212. https://doi.org/10.1038/nrmicro2970.

Murali, R., Pace, L.A., Sanford, R.A., Ward, L.M., Lynes, M.M., Hatzenpichler, R., Lingappa, U.F., Fischer, W.W., Gennis, R.B., Hemp, J., 2024. Diversity and Evolution of Nitric Oxide Reduction in Bacteria and Archaea. Proc. Natl. Acad. Sci. 121, 1–8. https://doi.org/10.1073/pnas.2316422121.

Nelson, D.L., Cox, M.M., 2013. Lehninger Principles of Biochemistry.

Ostrander, C.M., Johnson, A.C., Anbar, A.D., 2021. Earth's First Redox Revolution. Annu. Rev. Earth Planet. Sci. 49, 337–366. https://doi.org/10.1146/annurev-earth-072020-055249.

Parfrey, L.W., Lahr, D.J.G., Knoll, A.H., Katz, L.A., 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc. Natl. Acad. Sci. 108, 13624–13629. https://doi.org/10.1073/pnas.1110633108.

Pavlov, A.A., Kasting, J.F., 2002. Mass-independent fractionation of sulfur isotopes in archean sediments: strong evidence for an anoxic archean atmosphere. Astrobiology 2, 27–41. https://doi.org/10.1089/153110702753621321.

Pereira, M.M., Teixeira, M., 2004. Proton pathways, ligand binding and dynamics of the catalytic site in haem-copper oxygen reductases: a comparison between the three families. Biochim. Biophys. Acta BBA - Bioenerg. 1655, 340–346. https://doi.org/10.1016/j.bbabio.2003.06.003.

Pereira, M.M., Santana, M., Teixeira, M., 2001. A novel scenario for the evolution of haem-copper oxygen reductases. Biochim. Biophys. Acta Bioenerg. 1505, 185–208. https://doi.org/10.1016/S0005-2728(01)00169-4.

Pitcher, R.S., Watmough, N.J., 2004. The bacterial cytochrome cbb3 oxidases. Biochim. Biophys. Acta Bioenerg. 1655, 388–399. https://doi.org/10.1016/j.bbabio.2003.09.017.

Planavsky, N.J., Asael, D., Hofmann, A., Reinhard, C.T., Lalonde, S.V., Knudsen, A., Wang, X., Ossa Ossa, F., Pecoits, E., Smith, A.J.B., Beukes, N.J., Bekker, A., Johnson, T.M., Konhauser, K.O., Lyons, T.W., Rouxel, O.J., 2014. Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event. Nat. Geosci. 7, 283–286. https://doi.org/10.1038/ngeo2122.

Price, M.N., Dehal, P.S., Arkin, A.P., 2009. FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. Mol. Biol. Evol. 26, 1641–1650. https://doi.org/10.1093/molbev/msp077.

Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for large Alignments. PLoS One 5, e9490. https://doi.org/10.1371/journal.pone.0009490.

Pruitt, K.D., Tatusova, T., Maglott, D.R., 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. 35, D61–D65. https://doi.org/10.1093/nar/gkl842.

Sánchez-Baracaldo, P., 2015. Origin of marine planktonic cyanobacteria. Sci. Rep. 5, 17418. https://doi.org/10.1038/srep17418.

Saraste, M., 1990. Structural features of cytochrome oxidase. Q. Rev. Biophys. 23, 331–366. https://doi.org/10.1017/S0033583500005588.

Saraste, M., Castresana, J., 1994. Cytochrome oxidase evolved by tinkering with denitrification enzymes. FEBS Lett. 341, 1–4. https://doi.org/10.1016/0014-5793(94)80228-9.

Saraste, M., Holm, L., Lemieux, L., Lübben, M., Van Der Oost, J., 1991. The happy family of cytochrome oxidases. Biochem. Soc. Trans. 19, 608–612. https://doi.org/10.1042/bst0190608.

Sharma, V., Wikström, M., 2014. A structural and functional perspective on the evolution of the heme–copper oxidases. FEBS Lett. 588, 3787–3792. https://doi.org/10.1016/j.febslet.2014.09.020.

Sousa, F.L., Alves, R.J., Pereira-Leal, J.B., Teixeira, M., Pereira, M.M., 2011. A Bioinformatics Classifier and Database for Heme-Copper Oxygen Reductases. PLoS One 6, e19117. https://doi.org/10.1371/journal.pone.0019117.

Sousa, F.L., Alves, R.J., Ribeiro, M.A., Pereira-Leal, J.B., Teixeira, M., Pereira, M.M., 2012. The superfamily of heme–copper oxygen reductases: Types and evolutionary considerations. Biochim. Biophys. Acta BBA - Bioenerg. 1817, 629–637. https://doi.org/10.1016/j.bbabio.2011.09.020.

Stüeken, E.E., Buick, R., Anbar, A.D., 2015. Selenium isotopes support free O_2 in the latest Archean. Geology 43, 259–262. https://doi.org/10.1130/G36218.1.

Thorne, J.L., Kishino, H., Painter, I.S., 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15, 1647–1657. https://doi.org/10.1093/oxfordjournals.molbev.a025892.

Tria, F.D.K., Landan, G., Dagan, T., 2017. Phylogenetic rooting using minimal ancestor deviation. Nat. Ecol. Evol. 1. https://doi.org/10.1038/s41559-017-0193.

Trojan, D., Garcia-Robledo, E., Meier, D.V., Hausmann, B., Revsbech, N.P., Eichorst, S.A., Woebken, D., 2021. Microaerobic Lifestyle at Nanomolar O 2 Concentrations Mediated by Low-Affinity Terminal Oxidases in Abundant Soil Bacteria. mSystems 6. https://doi.org/10.1128/msystems.00250-21.

van der Oost, J., de Boer, A.P.N., de Gier, J.W.L., Zumft, W.G., Stouthamer, A.H., van Spanning, R.J.M., 1994. The heme-copper oxidase family consists of three distinct types of terminal oxidases and is related to nitric oxide reductase. FEMS Microbiol. Lett. 121, 1–9. https://doi.org/10.1016/0378-1097(94)90137-6.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., Barton, G.J., 2009. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

Wikström, M., 1989. Identification of the electron transfers in cytochrome oxidase that are coupled to proton-pumping. Nature 338, 776–778. https://doi.org/10.1038/338776a0.

Wilde, S.A., Valley, J.W., Peck, W.H., Graham, C.M., 2001. Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. Nature 409, 175–178. https://doi.org/10.1038/35051550.

Supplementary Material for **Molecular Clock Evidence for an Archean Diversification of Heme-Copper Oxygen Reductase Enzymes**

Fatima Husain[a,*], Haitao Shang[b,c], Stilianos Louca[b], Gregory P. Fournier[a]

[a]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, 55-101, Cambridge, Massachusetts 02139, United States of America
[b]Department of Biology, University of Oregon, 77 Klamath Hall, 1210 University of Oregon, Eugene, Oregon 97403, United States of America
[c]Department of Earth, Environmental and Resource Sciences, The University of Texas at El Paso, 500 W University, El Paso, Texas 79902, United States of America
*corresponding author

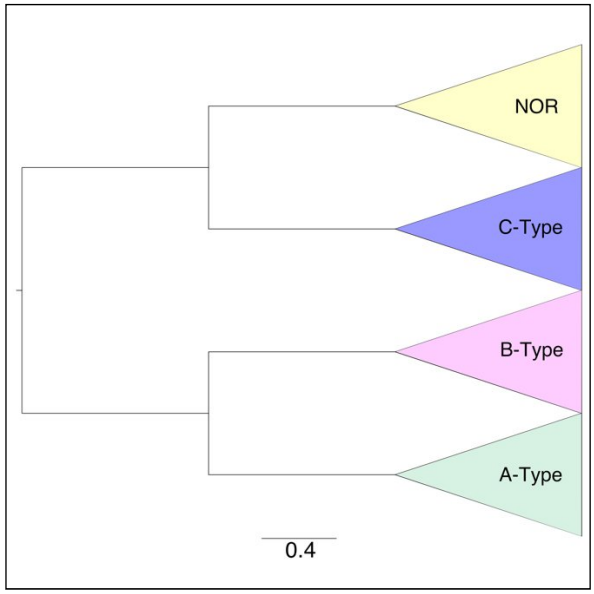
Email Addresses:
fhusain@mit.edu
hshang@utep.edu
g4nier@mit.edu

Supplementary Material contains:

- Accession numbers, organism, annotation, subunit classification, and domains of query sequences
- Guide tree for the heme-copper oxygen reductase family and the nitric oxide reductases
- Maximum-likelihood trees of eukaryotes and Cyanobacteria
- Posterior and prior age estimates from heme-copper oxygen reductase molecular clock analyses

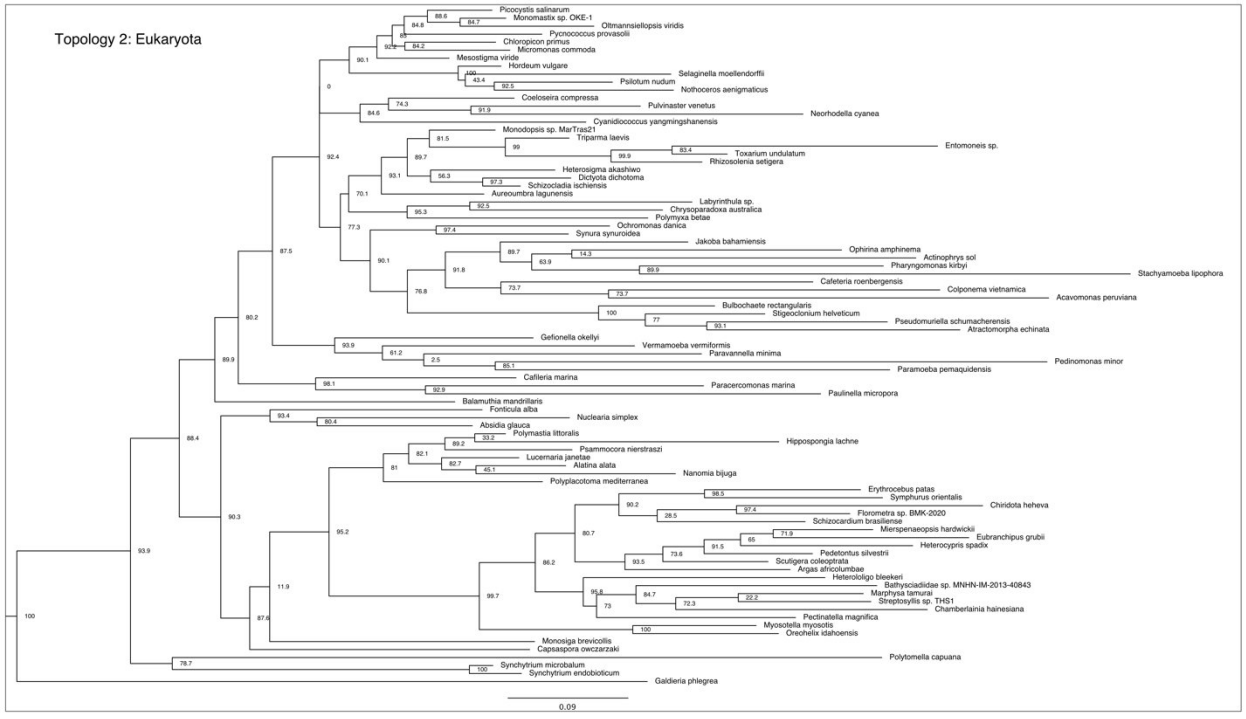
Sequence Accession	Organism	Annotation	Subunit	Domain
WP_045981596.1	*Paracoccus*	cytochrome *c* oxidase subunit 1	I	Bacteria
WP_011173204.1	*Thermus*	ba_3-type cytochrome c oxidase subunit I	I	Bacteria
WP_120510824.1	*Photobacterium salinisoli*	cytochrome-c oxidase, cbb_3-type subunit I	I	Bacteria
EEX11521.1	*Ruegeria lacuscaerulensis*	nitric oxide reductase subunit B	NorB	Bacteria
QYU70609.1	*Leptolyngbya sp. 15MV*	cytochrome *c* oxidase subunit I	I	Bacteria
MBL1178935.1	*Pantanalinema sp. GBBB05*	cytochrome *c* oxidase subunit I	I	Bacteria
NJM11088.1	*Synechococcaceae cyanobacterium*	cytochrome-*c* oxidase, cbb_3-type subunit I	I	Bacteria
NJM13844.1	*Synechococcaceae cyanobacterium*	nitric-oxide reductase large subunit	NorB	Bacteria
MBK69941.1	*Methanobacteriota*	cytochrome *c* oxidase subunit I	I	Archaea
WP_162414133.1	*Haloarcula salina*	$b(o/a)_3$-type cytochrome-*c* oxidase subunit 1	I	Archaea
WP_096203922.1	*Ca. Methanoperedens nitroreducens*	cbb_3-type cytochrome *c* oxidase subunit I*	I	Archaea
CDK41162.1	*Halorubrum sp. AJ67*	cytochrome *b* subunit of nitric oxide reductase	NorB	Archaea
P11947.1	*Tetrahymena pyriformis*	cytochrome *c* oxidase subunit 1	I	Eukaryota
BBD14158.1	*Ophirina amphinema*	cytochrome *c* oxidase subunit 1	I	Eukaryota
YP_001315139.1	*Chlorokybus atmophyticus*	cytochrome *c* oxidase subunit 1	I	Eukaryota
ACX50286.1	*Bigelowiella natans*	cytochrome oxidase subunit 1	I	Eukaryota
Q07434.1	*Physarum polycephalum*	cytochrome *c* oxidase subunit 1	I	Eukaryota

Supplementary Table 1. Accession numbers, organism, annotation, subunit classification, and domain of the query sequences used in this study. Annotations listed are provided in the NCBI sequence listing. *cbb_3-type oxygen reductases have not been found in Archaea previously[8,24]; annotations for subunit I may have varying degrees of accuracy.

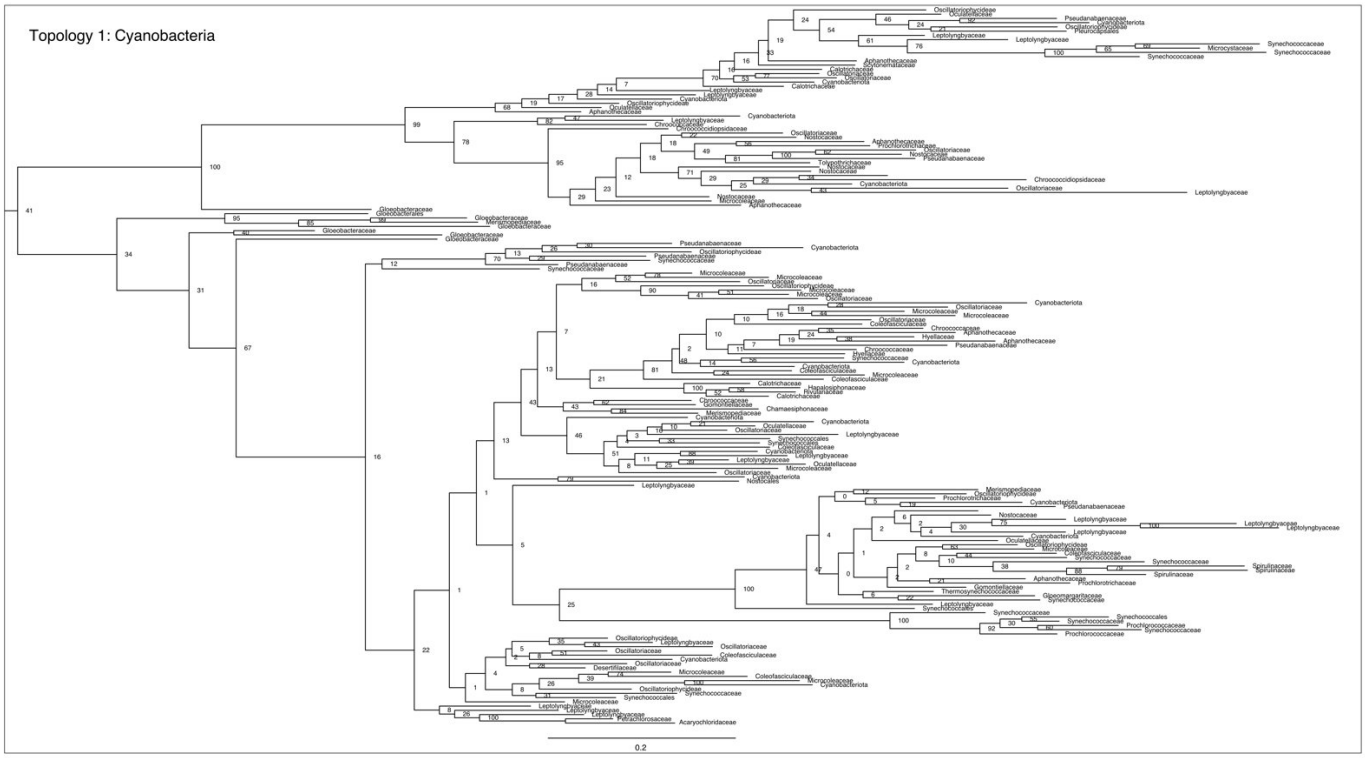


Supplementary Figure 1. Guide tree for the heme-copper oxygen reductase family and the nitric oxide reductases. Sequences in the collapsed clades of this quartet were classified using the heme-copper oxidase classifier and database outlined in Sousa et al., 2011.

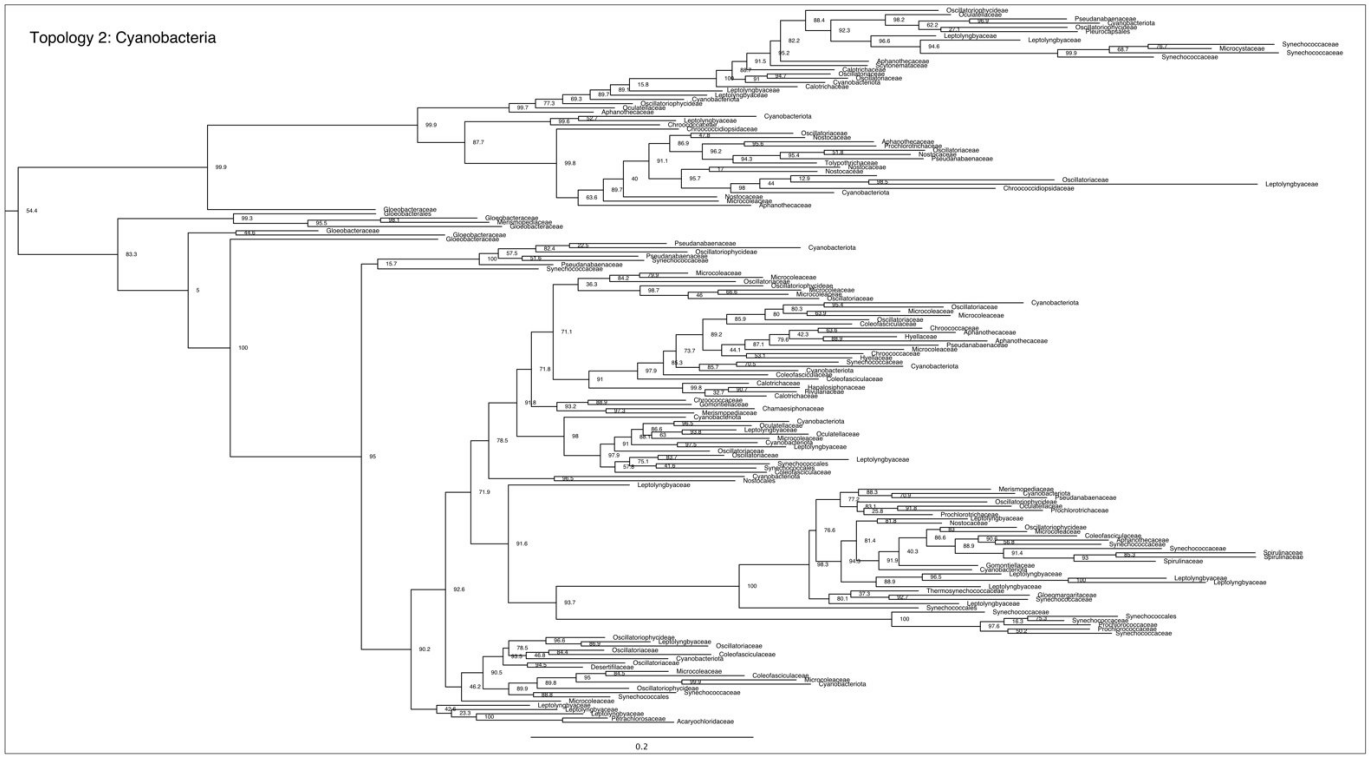
Supplementary Figure 2. Maximum-likelihood tree of eukaryotes from Figure 3, Panel A. Taxa were identified at the species level when possible. Bootstrap support values are located adjacent to nodes.



Supplementary Figure 3. Maximum-likelihood tree of eukaryotes from Figure 3, Panel B. Taxa were identified at the species level when possible. Bootstrap support values are located adjacent to nodes.



Supplementary Figure 4. Maximum-likelihood tree of Cyanobacteria from Figure 3, Panel A. Taxa were identified at the family level when possible. Bootstrap support values are located adjacent to nodes.



Supplementary Figure 5. Maximum-likelihood tree of Cyanobacteria from Figure 3, Panel B. Taxa were identified at the family level when possible. Bootstrap support values are located adjacent to nodes.

Posterior Distributions -bd	CIR_all	CIR_cyanos	CIR_euks	LN_all	LN_cyanos	LN_euks	UGAM_all	UGAM_cyanos	UGAM_euks
All Cytochrome Oxidases	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**
Mean	3988	4192	3885	4112	4194	4143	4247	3993	4212
Median	4014	4210	3906	4137	4240	4178	4266	4017	4252
95% HDI	3585	3963	3374	3808	3879	3826	4047	3588	3949
	4366	4392	4316	4381	4394	4393	4374	4351	4396
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
C-Types	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**
Mean	3263	3454	3063	2792	2826	2719	2673	2325	2732
Median	3280	3468	3074	2792	2884	2762	2667	2304	2724
95% HDI	2724	3016	2530	2140	2030	2015	2103	1706	2037
	3680	3805	3648	3451	3479	3263	3388	3102	3357
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
B-Types	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**
Mean	3649	3863	3502	3587	3727	3605	3781	3375	3801
Median	3669	3871	3517	3595	3754	3620	3793	3383	3811
95% HDI	3267	3578	3065	3246	3334	3284	3504	2956	3480
	4022	4121	3937	3914	4133	3915	4020	3779	4043
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**
Mean	3332	3557	3159	3188	3343	3208	3021	2502	3057
Median	3337	3557	3174	3188	3364	3220	3029	2500	3067
95% HDI	2980	3289	2739	2877	2913	2902	2735	2166	2768
	3704	3847	3514	3487	3757	3489	3314	2886	3377
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 1 node	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**
Mean	3050	3282	2865	2897	3053	2909	2695	2162	2709
Median	3055	3282	2869	2898	3070	2925	2694	2156	2705
95% HDI	2723	3012	2502	2642	2643	2635	2405	1848	2451
	3388	3534	3182	3212	3472	3161	2964	2468	3031
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 2 nodes	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**
Mean	3020	3254	2838	2869	3025	2875	2629	2107	2651
Median	3023	3255	2839	2871	3042	2882	2629	2101	2647
95% HDI	2698	3009	2484	2579	2581	2621	2342	1835	2360
	3358	3532	3155	3135	3428	3140	2881	2433	2928
Posterior Distributions -bd	**CIR_all**			**LN_all**			**UGAM_all**		
Cyanobacteria	**603**			**603**			**603**		
Mean	2552			2399			1799		
Median	2549			2395			1794		
95% HDI	2264			2172			1532		
	2811			2644			2070		

Supplementary Table 2. Posterior age estimates from UGAM, LN, and CIR models run with birth-death priors for Topology 1. Mean, median, and 95% high density intervals (HDI) are shown in millions of years. UGAM is short for the uncorrelated gamma multipliers model; LN is short for the log-normal autocorrelated clock model; CIR is short for the Cox-Ingersoll-Ross process model.

Posterior Distributions -nobd	CIR_all	CIR_cyanos	CIR_euks	LN_all	LN_cyanos	LN_euks	UGAM_all	UGAM_cyanos	UGAM_euks
All Cytochrome Oxidases	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**

Mean	2930	3650	2902	3047	3709	2980	3393	3675	3383
Median	2909	3644	2876	3021	3717	2951	3377	3666	3369
95% HDI	2782	3060	2789	2794	3121	2790	2948	3048	2924
	3139	4325	3090	3357	4346	3276	3902	4299	3860
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
C-Types	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**
Mean	2592	3253	2587	2545	3126	2470	2598	2844	2580
Median	2582	3253	2574	2530	3140	2458	2597	2838	2584
95% HDI	2363	2657	2409	2182	2506	2175	2077	2222	2103
	2839	3841	2808	2916	3747	2851	3151	3497	3093
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
B-Types	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**
Mean	2841	3538	2812	2938	3578	2887	3239	3504	3234
Median	2815	3527	2786	2913	3590	2854	3225	3497	3226
95% HDI	2670	2923	2678	2673	3034	2689	2771	2968	2791
	3047	4146	2992	3231	4218	3166	3669	4160	3694
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**
Mean	2712	3374	2678	2768	3373	2724	2956	3211	2945
Median	2687	3366	2656	2749	3388	2695	2946	3212	2931
95% HDI	2559	2818	2544	2518	2849	2518	2546	2616	2581
	2931	3978	2863	3064	3963	3000	3336	3760	3396
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 1 node	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**
Mean	2581	3218	2546	2640	3216	2590	2779	3018	2766
Median	2558	3218	2528	2620	3227	2563	2771	3014	2757
95% HDI	2428	2669	2418	2404	2713	2373	2426	2502	2381
	2767	3780	2731	2928	3785	2857	3149	3602	3130
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 2 nodes	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**
Mean	2567	3202	2532	2625	3200	2576	2745	2985	2733
Median	2544	3201	2513	2606	3216	2549	2734	2981	2722
95% HDI	2423	2660	2388	2370	2663	2367	2413	2483	2356
	2763	3769	2697	2904	3730	2848	3138	3575	3101
Posterior Distributions -nobd	**CIR_all**			**LN_all**			**UGAM_all**		
Cyanobacteria	**603**			**603**			603		
Mean	2288			2323			2171		
Median	2274			2311			2165		
95% HDI	2135			2113			1853		
	2473			2609			2495		

Supplementary Table 3. Posterior age estimates from UGAM, LN, and CIR models run without birth-death priors for Topology 1. Mean, median, and 95% high density intervals (HDI) are shown in millions of years. UGAM is short for the uncorrelated gamma multipliers model; LN is short for the log-normal autocorrelated clock model; CIR is short for the Cox-Ingersoll-Ross process model.

Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
All Cytochrome Oxidases	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**
Mean	3893	3987	3890	3926	4103	4046	4203	4001	4185

Median	3918	3979	3925	3949	4141	4066	4238	4036	4253
95% HDI	3409	3675	3395	3494	3743	3680	3933	3567	3775
	4322	4350	4332	4303	4375	4373	4391	4337	4397
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
C-Types	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**
Mean	3075	3258	2982	2769	2957	2856	2696	2309	2774
Median	3079	3291	2991	2788	2996	2881	2689	2276	2774
95% HDI	2478	2747	2416	2113	2314	2220	2139	1686	2091
	3706	3770	3514	3402	3582	3578	3409	3018	3465
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
Ancestor to B- and A- Types	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**
Mean	3418	3548	3376	3399	3622	3456	3664	3297	3685
Median	3415	3549	3384	3403	3630	3462	3674	3300	3717
95% HDI	2968	3176	2970	3033	3272	3079	3343	2897	3288
	3860	3937	3783	3758	3941	3810	3972	3708	3997
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
B-Types	**418**	**418**	**418**	**418**	**418**	**418**	**418**	**418**	**418**
Mean	3284	3410	3234	3172	3387	3245	3482	3111	3519
Median	3278	3406	3234	3184	3398	3253	3492	3112	3548
95% HDI	2855	3013	2802	2741	3004	2835	3087	2679	3103
	3741	3790	3611	3615	3800	3703	3802	3499	3904
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**
Mean	3240	3396	3192	3211	3441	3255	3219	2786	3275
Median	3237	3395	3196	3216	3454	3256	3225	2773	3291
95% HDI	2808	3042	2818	2860	3083	2895	2833	2408	2866
	3683	3762	3576	3537	3739	3650	3541	3245	3655
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 1 node	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**
Mean	2891	3069	2817	2856	3091	2861	2749	2263	2800
Median	2884	3058	2814	2853	3102	2860	2745	2259	2800
95% HDI	2508	2769	2476	2559	2774	2483	2435	1936	2457
	3283	3471	3122	3164	3369	3184	3069	2568	3149
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 2 nodes	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**
Mean	2857	3035	2779	2823	3052	2821	2665	2180	2708
Median	2846	3024	2775	2817	3062	2822	2666	2173	2707
95% HDI	2480	2698	2485	2512	2745	2505	2356	1892	2350
	3246	3415	3116	3123	3331	3168	2986	2497	3026
Posterior Distributions -bd	**CIR_all**			**LN_all**			**UGAM_all**		
Cyanobacteria	**603**			**603**			**603**		
Mean	2397			2359			1761		
Median	2391			2353			1755		
95% HDI	2066			2097			1477		
	2706			2626			2052		

Supplementary Table 4. Posterior age estimates from UGAM, LN, and CIR models run with birth-death priors for Topology 2. Mean, median, and 95% high density intervals (HDI) are shown in millions of years. UGAM is short for the uncorrelated gamma multipliers model; LN is short for the log-normal autocorrelated clock model; CIR is short for the Cox-Ingersoll-Ross process model.

Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
All Cytochrome Oxidases	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**
Mean	2885	3627	2877	2909	3684	2884	3291	3676	3240
Median	2861	3632	2859	2884	3684	2855	3288	3672	3222
95% HDI	2771	3029	2783	2784	3103	2767	2805	3063	2830

	3076	4287	3052	3096	4352	3074	3721	4293	3691
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
C-Types	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**
Mean	2565	3233	2544	2487	3143	2448	2545	2843	2499
Median	2557	3240	2538	2486	3141	2441	2539	2830	2496
95% HDI	2383	2723	2380	2255	2535	2152	2023	2150	2001
	2830	3861	2752	2749	3709	2708	3016	3473	2986
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
Ancestor to B- and A- Types	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**
Mean	2766	3488	2754	2800	3556	2769	3121	3491	3082
Median	2743	3489	2737	2782	3556	2743	3114	3489	3065
95% HDI	2640	2854	2634	2654	2995	2638	2676	2970	2682
	2959	4065	2921	2985	4209	2955	3528	4144	3494
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
B-Types	**418**	**418**	**418**	**418**	**418**	**418**	**418**	**418**	**418**
Mean	2730	3442	2717	2741	3487	2719	3043	3414	3002
Median	2708	3444	2703	2725	3487	2695	3035	3419	2989
95% HDI	2585	2811	2586	2583	2940	2584	2601	2896	2606
	2935	4012	2882	2936	4135	2913	3459	4056	3436
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**
Mean	2687	3400	2678	2741	3482	2704	2971	3332	2933
Median	2666	3401	2663	2720	3487	2682	2962	3329	2921
95% HDI	2537	2779	2551	2597	2923	2585	2566	2784	2556
	2864	3960	2849	2936	4116	2903	3371	3902	3326
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 1 node	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**
Mean	2546	3227	2529	2609	3326	2570	2780	3121	2743
Median	2532	3230	2515	2590	3332	2552	2772	3117	2727
95% HDI	2395	2663	2404	2452	2824	2426	2409	2573	2397
	2725	3788	2692	2785	3977	2742	3167	3639	3141
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 2 nodes	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**
Mean	2530	3208	2513	2593	3308	2555	2738	3081	2707
Median	2515	3208	2499	2575	3313	2538	2731	3076	2697
95% HDI	2406	2650	2392	2427	2732	2421	2371	2573	2332
	2735	3765	2679	2762	3879	2737	3109	3634	3064
Posterior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
Cyanobacteria	**603**			**603**			**603**		
Mean	2263			2311			2167		
Median	2251			2300			2161		
95% HDI	2126			2143			1819		
	2444			2498			2459		

Supplementary Table 5. Posterior age estimates from UGAM, LN, and CIR models run without birth-death priors for Topology 2. Mean, median, and 95% high density intervals (HDI) are shown in millions of years. UGAM is short for the uncorrelated gamma multipliers model; LN is short for the log-normal autocorrelated clock model; CIR is short for the Cox-Ingersoll-Ross process model.

Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
All Cytochrome Oxidases	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**
Mean	3429	3382	3394	3457	3354	3440	3387	3363	3458
Median	3420	3375	3374	3475	3378	3436	3387	3376	3446
95% HDI	2705	2404	2639	2698	2225	2715	2511	2387	2775
	4312	4376	4359	4333	4365	4307	4352	4384	4320
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**

C-Types	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**
Mean	1907	2035	1896	1873	2061	1917	1861	1982	1923
Median	1887	1977	1848	1825	2047	1861	1793	1958	1880
95% HDI	475	409	455	389	433	427	390	401	527
	3441	3661	3400	3346	3680	3414	3345	3560	3492
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
B-Types	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**
Mean	3326	3269	3280	3333	3247	3334	3264	3257	3353
Median	3305	3264	3255	3345	3286	3336	3267	3285	3340
95% HDI	2582	2170	2389	2472	1979	2548	2272	2100	2622
	4285	4328	4248	4264	4324	4280	4242	4309	4247
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**
Mean	3014	2971	2950	2981	2959	3018	2930	2960	3034
Median	2984	3005	2928	2974	3015	3013	2929	3004	3010
95% HDI	2123	1697	2053	2033	1558	2066	1913	1657	2181
	3927	4122	3934	3908	4171	3899	3871	4127	3934
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 1 node	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**
Mean	2947	2910	2883	2913	2906	2954	2864	2902	2971
Median	2917	2941	2863	2898	2963	2950	2864	2939	2954
95% HDI	2046	1599	2002	1929	1515	2055	1825	1569	2116
	3839	4055	3842	3789	4150	3862	3762	4074	3855
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 2 nodes	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**
Mean	2886	2856	2818	2847	2854	2892	2802	2847	2910
Median	2859	2890	2800	2836	2900	2888	2805	2886	2888
95% HDI	2060	1551	1966	1896	1445	1974	1782	1536	2024
	3833	4003	3798	3734	4113	3764	3684	4042	3764
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
Cyanobacteria	**603**			**603**			**603**		
Mean	2498			2447			2414		
Median	2468			2435			2396		
95% HDI	1714			1639			1585		
	3327			3258			3239		

Supplementary Table 6. Prior age estimates from UGAM, LN, and CIR models run with birth-death priors for Topology 1. Mean, median, and 95% high density intervals (HDI) are shown in millions of years. UGAM is short for the uncorrelated gamma multipliers model; LN is short for the log-normal autocorrelated clock model; CIR is short for the Cox-Ingersoll-Ross process model.

Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
All Cytochrome Oxidases	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**
Mean	2858	3438	2864	2858	3432	2863	2856	3415	2863
Median	2838	3423	2842	2838	3409	2842	2838	3383	2842
95% HDI	2772	2800	2776	2773	2797	2776	2777	2788	2776
	2994	4061	3017	2998	4052	3012	2993	4024	3013
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
C-Types	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**

Mean	2493	3003	2498	2501	3014	2502	2507	2985	2512
Median	2574	3029	2578	2588	3024	2590	2599	3010	2597
95% HDI	1813	2070	1815	1815	2080	1740	1791	1967	1816
	2965	3988	2987	2969	3950	2959	2957	3876	2981
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
B-Types	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**
Mean	2849	3428	2854	2848	3422	2853	2846	3405	2853
Median	2830	3414	2833	2829	3400	2833	2829	3376	2833
95% HDI	2756	2794	2762	2754	2790	2762	2760	2784	2759
	2985	4055	3017	2991	4043	3007	2983	4025	3005
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**
Mean	2807	3386	2813	2807	3380	2812	2805	3364	2811
Median	2791	3374	2795	2791	3359	2795	2790	3334	2794
95% HDI	2702	2759	2693	2702	2753	2703	2704	2770	2703
	2953	3997	2971	2964	3998	2968	2946	4013	2971
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 1 node	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**
Mean	2797	3375	2803	2796	3369	2802	2794	3353	2801
Median	2782	3363	2785	2781	3346	2785	2781	3323	2784
95% HDI	2679	2767	2683	2686	2740	2685	2689	2750	2681
	2935	4006	2963	2949	3980	2955	2937	3983	2951
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 2 nodes	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**
Mean	2786	3365	2792	2785	3358	2791	2783	3343	2790
Median	2772	3353	2775	2770	3335	2776	2771	3312	2774
95% HDI	2670	2717	2673	2660	2741	2678	2677	2734	2680
	2927	3954	2962	2928	3979	2951	2926	3962	2956
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
Cyanobacteria	**603**			**603**			**603**		
Mean	2712			2711			2709		
Median	2703			2701			2701		
95% HDI	2574			2575			2575		
	2867			2871			2859		

Supplementary Table 7. Prior age estimates from UGAM, LN, and CIR models run without birth-death priors for Topology 1. Mean, median, and 95% high density intervals (HDI) are shown in millions of years. UGAM is short for the uncorrelated gamma multipliers model; LN is short for the log-normal autocorrelated clock model; CIR is short for the Cox-Ingersoll-Ross process model.

Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
All Cytochrome Oxidases	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**
Mean	3381	3360	3422	3400	3402	3401	3379	3404	3409
Median	3354	3379	3411	3384	3416	3384	3364	3398	3383
95% HDI	2611	2311	2650	2660	2490	2704	2547	2630	2698
	4333	4382	4336	4345	4364	4347	4322	4365	4334
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
C-Types	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**
Mean	1794	1968	1789	1830	2050	1859	1847	2083	1804
Median	1728	1920	1720	1756	2020	1809	1799	2061	1742
95% HDI	474	407	484	405	426	466	476	378	475

	3382	3576	3429	3393	3643	3375	3425	3624	3420
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
Ancestor to B- and A- Types	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**
Mean	3252	3236	3292	3276	3293	3279	3257	3300	3293
Median	3231	3262	3282	3253	3316	3251	3247	3298	3262
95% HDI	2339	2092	2451	2361	2227	2432	2374	2380	2563
	4232	4374	4294	4212	4347	4230	4283	4318	4318
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
B-Types	**418**	**418**	**418**	**418**	**418**	**418**	**418**	**418**	**418**
Mean	2608	2636	2604	2644	2764	2655	2622	2739	2660
Median	2611	2700	2614	2655	2809	2665	2632	2782	2666
95% HDI	1450	1194	1376	1515	1374	1529	1434	1412	1569
	3807	3954	3762	3881	4127	3868	3851	4055	3853
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**
Mean	3146	3134	3185	3165	3204	3175	3156	3210	3189
Median	3123	3176	3167	3144	3227	3149	3149	3204	3162
95% HDI	2233	1854	2237	2226	1993	2292	2170	2222	2322
	4160	4272	4174	4123	4248	4143	4144	4270	4180
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 1 node	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**
Mean	3052	3041	3087	3074	3125	3086	3064	3135	3102
Median	3039	3090	3061	3055	3142	3061	3065	3133	3075
95% HDI	2082	1699	2143	2198	1904	2167	2025	2117	2217
	3998	4187	4097	4093	4228	4046	4006	4251	4095
Prior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 2 nodes	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**
Mean	2971	2965	3003	2989	3054	3005	2984	3064	3019
Median	2951	2998	2981	2967	3078	2975	2989	3059	2993
95% HDI	2002	1665	2023	2090	1803	2154	1888	2014	2126
	3909	4182	3962	3971	4175	4027	3854	4200	3986
Posterior Distributions -bd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
Cyanobacteria	**603**			**603**			**603**		
Mean	2503			2531			2526		
Median	2473			2510			2522		
95% HDI	1696			1647			1611		
	3412			3364			3432		

Supplementary Table 8. Prior age estimates from UGAM, LN, and CIR models run with birth-death priors for Topology 2. Mean, median, and 95% high density intervals (HDI) are shown in millions of years. UGAM is short for the uncorrelated gamma multipliers model; LN is short for the log-normal autocorrelated clock model; CIR is short for the Cox-Ingersoll-Ross process model.

Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
All Cytochrome Oxidases	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**	**409**
Mean	2886	3414	2895	2887	3421	2899	2886	3427	2900
Median	2858	3381	2865	2860	3391	2870	2859	3404	2870
95% HDI	2773	2795	2778	2771	2801	2776	2777	2798	2776
	3080	4030	3105	3081	4054	3120	3074	4061	3124
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
C-Types	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**	**410**
Mean	2516	2995	2534	2528	2987	2536	2523	2998	2531
Median	2595	3018	2613	2606	3004	2617	2599	3011	2611
95% HDI	1787	2039	1841	1808	1978	1841	1754	1920	1756
	3015	3893	3064	3016	3910	3112	2980	3893	3047
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**

Ancestor to B- and A- Types	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**	**417**
Mean	2877	3405	2886	2878	3412	2890	2877	3417	2891
Median	2849	3370	2856	2851	3384	2861	2849	3393	2861
95% HDI	2759	2790	2763	2757	2795	2764	2766	2791	2758
	3074	4028	3101	3072	4046	3120	3067	4050	3113
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
B-Types	**418**	**418**	**418**	**418**	**418**	**418**	**418**	**418**	**418**
Mean	2787	3299	2797	2788	3304	2801	2786	3308	2797
Median	2784	3276	2791	2788	3279	2794	2785	3277	2791
95% HDI	2529	2700	2555	2532	2721	2528	2528	2711	2514
	3051	3985	3089	3068	4019	3086	3051	4035	3069
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**	**449**
Mean	2867	3394	2876	2868	3401	2879	2867	3406	2881
Median	2839	3360	2846	2841	3374	2850	2840	3382	2851
95% HDI	2739	2785	2739	2744	2782	2745	2745	2785	2750
	3063	4014	3074	3067	4026	3110	3055	4041	3111
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 1 node	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**	**451**
Mean	2856	3383	2866	2858	3391	2870	2857	3396	2871
Median	2829	3350	2839	2832	3364	2841	2831	3367	2843
95% HDI	2729	2777	2726	2730	2783	2726	2736	2785	2724
	3057	3998	3068	3059	4025	3092	3055	4035	3091
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
A-Types - inwards 2 nodes	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**	**458**
Mean	2846	3372	2855	2847	3380	2859	2846	3385	2861
Median	2818	3341	2828	2821	3351	2831	2821	3356	2832
95% HDI	2717	2763	2718	2712	2774	2720	2721	2777	2719
	3051	3983	3064	3051	4015	3090	3049	4023	3088
Prior Distributions -nobd	**CIR_all**	**CIR_cyanos**	**CIR_euks**	**LN_all**	**LN_cyanos**	**LN_euks**	**UGAM_all**	**UGAM_cyanos**	**UGAM_euks**
Cyanobacteria	**603**			**603**			**603**		
Mean	2772			2773			2772		
Median	2750			2752			2753		
95% HDI	2625			2609			2598		
	2993			2983			2963		

Supplementary Table 9. Prior age estimates from UGAM, LN, and CIR models run without birth-death priors for Topology 2. Mean, median, and 95% high density intervals (HDI) are shown in millions of years. UGAM is short for the uncorrelated gamma multipliers model; LN is short for the log-normal autocorrelated clock model; CIR is short for the Cox-Ingersoll-Ross process model.