Timing the Evolution of Cyanobacterial Antioxidants: Superoxide Dismutases

Joanne Boden
School of Geographical Sciences https://orcid.org/0000-0003-0412-3668
Kurt Konhauser
University of Alberta
Leslie Robbins
University of Regina https://orcid.org/0000-0002-6931-5743
Patricia Sánchez-Baracaldo (p.sanchez-baracaldo@bristol.ac.uk)
University of Bristol

Article

Keywords: cyanobacteria, phylogenetics, superoxide dismutase enzymes (SODs)

Posted Date: January 28th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-149571/v1

Version of Record: A version of this preprint was published at Nature Communications on August 6th, 2021. See the published version at https://doi.org/10.1038/s41467-021-24396-y.

Timing the Evolution of Cyanobacterial Antioxidants: Superoxide Dismutases

Joanne S. Boden[1], Kurt O. Konhauser[2], Leslie J. Robbins[3,4], and Patricia Sánchez-Baracaldo[1*]

[1]School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, United Kingdom

[2]Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, AB T6G 2E3, Canada

[3]Department of Earth and Planetary Sciences, Yale University, New Haven, CT, USA

[4]Department of Geology, University of Regina, Regina, SK S4S 0A2, Canada

*Correspondence: Dr Patricia Sanchez-Baracaldo, School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, United Kingdom

Email: p.sanchez-baracaldo@bristol.ac.uk

https://orcid.org/0000-0003-0412-3668

https://orcid.org/0000-0001-7722-7068

https://orcid.org/0000-0002-6931-5743

https://orcid.org/0000-0002-5216-2664

Key words

Cyanobacteria, superoxide dismutases, oxidative damage, molecular clocks

Abstract

The ancestors of cyanobacteria generated Earth's first biogenic molecular oxygen but how they dealt with its toxicity remains unconstrained. Here we investigated when superoxide dismutase enzymes (SODs) capable of removing superoxide free radicals evolved. We found phylogenetic evidence that ancestral cyanobacteria used SODs with copper and zinc cofactors (CuZnSOD) during the Archaean. By the Paleoproterozoic, they became genetically capable of using iron, nickel, and manganese as cofactors (FeSOD, NiSOD, and MnSOD respectively). The evolution of NiSOD is particularly intriguing because it has been previously hypothesized that declining seawater Ni concentrations at the end of the Archaean caused a fundamental shift in the marine biosphere away from methanogenesis towards oxygenic photosynthesis. Our novel analyses of enzymes dealing with O_2 toxicity now demonstrate that the beneficiaries of this chemical change - marine planktonic cyanobacteria - were able to utilize the remaining Ni from seawater 0.9-0.8 Ga to supplement their existing metabolic capabilities.

Introduction

Oxygen is essential for complex life forms as it is used during aerobic respiration to create more energy per mol of substrate than other available electron acceptors [1]. While today the Earth's atmosphere contains ~ 21% oxygen (O_2), it was at least 100,000 times lower in the Archaean (4.0 to 2.5 Ga) [2,3]. Just how and when O_2 first appeared as a byproduct of biological evolution – oxygenic photosynthesis - remains controversial, with estimates ranging from 3.8 billion years ago (Ga) [4] to immediately preceding the Great Oxidation Event (GOE) [5], which is estimated to have begun by 2.45 Ga [6]. Since the O_2 produced was novel and highly reactive, early cyanobacteria – the first producers of O_2 - would have found exposure to it highly toxic. Environmental pressures likely provided impetus for the evolution of protective enzymes that prevented oxidative damage from reactive oxygen species (ROS) [7-10]. While it is unclear when defense mechanisms against oxidative damage first evolved, it is reasonable

to assume that O_2-generating-organisms, such as cyanobacteria, would have co-evolved protective mechanisms as water oxidation proteins evolved.

Cyanobacteria remove ROS using carotenoids, α-tocopherol and antioxidant enzymes, including peroxidases, catalases, superoxide reductases (SORs), and superoxide dismutases (SODs)[11], and their evolutionary history can be elucidated by phylogenetic methodologies. Whilst peroxidases and catalases protect against oxidative damage by enhancing the rate of removal of peroxides (such as H_2O_2 and R-O-O-H) [12], SORs and SODs remove superoxide free radicals ($O_2^{\cdot-}$) [13]. These $O_2^{\cdot-}$ are produced as a biproduct of photosynthetic and respirational electron transport chains [11]. If left to accumulate, they react with solvent-exposed [4Fe-4S] $^{2+}$ clusters in proteins, including those required for amino acid biosynthesis [14] and photosynthesis [15], generating reactants of the fenton reaction which can ultimately lead to extensive DNA damage [11]. It may be for this reason that SODs and SORs have been found in all three domains of life - Eukarya, Archaea, and Bacteria [13].

Phylogenetic methodologies coupled with protein structural analyses have revealed that three different isoforms of SOD enzymes evolved independently of one another to protect cells against oxidative stress [16]. Each has a unique 3D structure, amino acid sequence and metal cofactor; either manganese (MnSOD), nickel (NiSOD) or a combination of copper and zinc (CuZnSOD)[16,17]. A fourth SOD enzyme shares its evolutionary heritage with MnSOD, but utilises iron as its cofactor (FeSOD)[18,19].

All four SODs are found within cyanobacteria [19,20], but their locations within the cell differ. FeSOD and NiSOD are cytoplasmic [19,21-24], whereas CuZnSODs from *Synechocystis* and MnSODs from *Anabaena, Plectolyngbya* and *Synechococcus* are tethered to membranes [22,24-27]. The choice of which SOD isoform(s) and associated metal cofactor(s) are used by a given species could reflect its evolutionary heritage, function, and environmental history.

Early phylogenetic studies on the evolution of SODs in cyanobacteria were limited by the availability of genomes, but found that NiSODs are present in planktonic marine strains, such as

67 *Prochlorococcus* spp., whereas MnSOD and FeSOD are widespread amongst a variety of ancestral and
68 derived lineages [19,20,28]. CuZnSODs, while rare, are not restricted to any particular group [19,20]. The 3D
69 protein structures required for all copper metalloproteins have previously been predicted to have arisen
70 after Fe- and Mn- utilising proteins in the Proterozoic (~2.4-0.5 Ga), following the GOE [29]. Therefore, it
71 would follow that cyanobacteria used FeSODs and MnSODs before CuZnSODs. While CuZnSODs and
72 FeSODs/MnSODs are widely distributed in bacteria and eukaryotes, NiSODs are restricted only to
73 Bacteria and may have appeared later in the evolutionary history of life [16]. Previous approaches aimed
74 at determining metal usage in SODs have not considered the habitats where strains live or evolved. This
75 is problematic because metal availability differs across oceanic and terrestrial environments, especially
76 in past geological eras [30,31].

77 Cyanobacteria have a long evolutionary record spanning at least 1.88 Ga (the oldest undisputed
78 colonial cyanobacterial microfossils; [32]), but perhaps as old, or older, than 3.22 Ga (based on stromatolite
79 fabric [33] and molecular clock analyses of PSII [34]). They also lived in a variety of marine, freshwater, and
80 terrestrial habitats [35,36]. Here, we analysed when cyanobacteria acquired the genes encoding for NiSOD
81 (*sodN*), CuZnSOD (*sodC*), FeSOD (*sodB*) and MnSOD (*sodA*) in the context of their habitat and metal
82 cofactor availability through geological time. To do this, we constructed phylogenies of all SODs encoded
83 in the genomes of 15,899 bacteria. We have also implemented a phylogenomic approach and multiple
84 molecular clock analyses to estimate when the crown group of Cyanobacteria diverged from their closest
85 non-photosynthetic relatives, the Vampirovibrionia [37,38]. By studying the timing of the divergence of
86 Cyanobacteria from their closest relatives, as well as their acquisition of SODs, we have established
87 points in the Proterozoic and Archaean when SODs were present in this Phylum.

Results

Bacterial SOD Diversity

To elucidate which transition metals cyanobacteria first used to protect themselves against the oxidative stress caused by $O_2^{\cdot-}$, the evolutionary history of SOD genes was modelled and mapped onto an updated time-calibrated phylogeny. In order to do this, we began by screening 15,899 bacterial genomes for genes encoding NiSOD, CuZnSOD, and MnSOD. BLASTP was unable distinguish between genes encoding FeSOD and MnSOD, so they are considered as one hereafter. Results reveal that the *sodN* gene encoding NiSOD is less common in bacteria than *sodC* encoding CuZnSOD. The *sodN* gene was found in 1464 different species, including Gammaproteobacteria, Alphaproteobacteria, Planctomycetes, Bacteroidetes, Actinobacteria, Cyanobacteria, Verrucomicrobia and Chloroflexi (SI Appendix, Fig. S1). By contrast, *sodC* was absent from the latter two phyla, but present in more genomes (5723) of Betaproteobacteria, Firmicutes and the six previous (SI Appendix, Fig. S1). Together, *sodA* and *sodB* are more widespread than *sodC* and *sodN* combined, being present in 13,748 bacterial species across all ten phyla mentioned previously, as well as Fibrobacteres, Chlorobi, and Vampirovibrionia (SI Appendix, Fig. S2).

Cyanobacterial SOD Diversity

A more specific search amongst cyanobacteria revealed that most strains with *sodN* (51 of 55 strains) live in saltwater habitats (Fig. 1). They include representatives from all major clades of marine taxa (Fig. 2). Ten lack a gene (named *sodX*) encoding NiSOD's maturation protease (SI Appendix, Table S1). Of these, six contain genes encoding other SOD isoforms, while the remaining four are all picocyanobacteria, i.e., all *Prochlorococcus* and some *Synechococcus* species (Fig. 2).

Many cyanobacterial strains contain multiple SOD isoforms. These include 32 strains with *sodN*, every strain with *sodC*, and 45 of 115 with *sodA* or *sodB*. The remaining 23 strains which only use NiSOD (including three without *sodX*) are all marine with small genomes. They include picocyanobacteria and endosymbionts living in larger marine tunicates, algae, and sponges (e.g., *Prochloron* spp., UCYNA and two strains of *Synechococcus spongiarum*). The 70 cyanobacteria which only use FeSOD and MnSOD live in a variety of marine, terrestrial, and freshwater habitats (Fig. 2). Only three percent of cyanobacteria (five of 149 strains discounting plastids) have genes encoding every SOD isoform as well as the NiSOD maturation protease (Fig. 2). Paralogues of *sodC* were found in three genomes and paralogues of *sodA* and/or *sodB* were found in at least 13 genomes (SI Appendix, Table S1).

The resources needed to make each SOD isoform vary. NiSOD is composed from a mode of 157 amino acids (range 145-166), whereas CuZnSOD is made from 177 (range 103-236) and FeSOD/MnSOD from 200 (range 197-280) (SI Appendix, Fig. S3). SODs from the family which utilize Fe and Mn are present in all major clades of nitrogen-fixing cyanobacteria (defined by [39], including all Nostocales, one *Trichodesmium* spp., seven unicellular diazotrophs, *Chroococcidiopsis thermalis* and *Leptolyngbya* sp. PCC7375 (Fig. 2).

Transfer of SODs Across Phyla

To investigate whether cyanobacteria obtained their SOD genes from other bacterial phyla, Maximum Likelihood (ML) phylogenies were constructed using the NiSODs, CuZnSODs and combined Mn- and Fe-SODs from bacteria. If, in each case, all cyanobacterial proteins form a monophyletic group ('cluster together'), this would indicate a single origin. Surprisingly, these phylogenetic analyses revealed that cyanobacterial NiSODs, CuZnSODs, MnSODs and FeSODs are polyphyletic (SI Appendix, Figs. S1-2), suggesting several lateral gene transfer events or other modes of reticulated evolution. Furthermore, the SODs with Mn or Fe cofactors found in Vampirovibrionia are not related to those in their sister phyla, Cyanobacteria [37](SI Appendix, Table S2 and Fig. S2).

Molecular phylogenies identified several horizontal gene transfers (HGTs) of each SOD isoform between cyanobacteria and other bacterial phyla: two of *sodN* (NiSODs), five of *sodC* (SI Appendix, Fig. S1), and three to eight of Mn- and Fe-utilising SODs (SI Appendix, Fig. S2). A variety of different phyla are involved in these HGTs. For example, most cyanobacterial NiSODs have diversified from a protein resembling that of benthic marine Deltaproteobacteria (namely *Geobacter electrodiphilus*, UFBoot 88 and *Plesiocystis pacifica*, UFBoot 89, SI Appendix, Fig. S4), whereas other cyanobacterial NiSODs (present in *Synechococcus spongiarum*) diversified from a protein resembling those of Alphaproteobacteria (UFBoot 81, SI Appendix, Fig. S5).

Genes encoding CuZnSOD may have been present in the shared common ancestor of all extant cyanobacteria. This ancestor gave rise to basal lineages before diverging into macrocyanobacteria and microcyanobacteria (Fig. 2). Although CuZnSODs are rare in macro- and micro-cyanobacteria, they are present in most free-living basal lineages (Fig. 2). Two in particular (*Pseudanabaena* spp. and *Gloeobacter* spp.) share *sodC* genes which are monophyletic (PP 1) and closely related in the same way as the species are to one another (SI Appendix, Fig. S6). This suggests they have been vertically inherited from the common ancestor of the cyanobacteria crown group.

As cyanobacteria diversified to occupy new ecological niches and habitats [35], the *sodC* genes which initially allowed crown cyanobacteria to use copper and zinc to protect against oxidative stress were likely lost. Later HGTs likely occurred between non-cyanobacterial phyla and picocyanobacteria, diazotrophs, sponge symbionts, *Acaryochloris* spp. and *Microcoleus chthonoplastes* (SI Appendix, Fig. S1), resulting in the distribution found today.

Bayesian Relaxed Molecular Clock Analyses

Divergence times were estimated in Phylobayes 4.1 [40] using SSU and LSU ribosomal RNA from 164 cyanobacteria and seven vampirovibrionia. The molecular clock's topology was constrained using a ML tree constructed in IQ-TREE v1.6.1 [41] from these same ribosomal RNAs and 136 core cyanobacterial

159 proteins with similar evolutionary trajectories [42]. They have a range of functions in metabolism, cellular
160 processes, and information handling [35,36,43]. All molecular clock analyses implemented six calibration
161 points (Table 1) as previously described [43]. Divergence times were estimated using uncorrelated gamma
162 multipliers (UGAM)[44]. All analyses, regardless of calibration points and models, indicate that
163 Cyanobacteria diverged from Vampirovibrionia between 3.3 and 3.6 Ga in the Archaean eon (Table 2).
164 Estimates that include 95% credibility intervals allow for a range between 2.8 and 4.3 Ga, suggesting
165 that Cyanobacteria diverged from their sister phyla at the latest some 300 Myrs before the GOE (Table
166 2).

167

168 ## Discussion

169

170 ### Early Oxygen Availability

171 The advent of oxygenic photosynthesis had profound impacts on Earth's climate, chemistry, and
172 biota. It destroyed a methane greenhouse which kept the world warm [45], enhanced the supply of sulfate
173 and redox active metals to the oceans [46,47], marginalized the growth of anoxygenic phototrophs [48], and
174 facilitated the chemical oxidation of dissolved Fe(II) in seawater to form banded iron formations - BIF [49].
175 Organisms would have responded by tailoring their metal requirements and evolving more efficient
176 enzymes, such as SODs, to deal with the expansion of O_2 into new habitats [35].

177 Age estimates for the emergence of oxygenic photosynthesis are therefore fundamental to
178 reconciling the biochemistry and use of trace metals. Previous estimates place the divergence of
179 Cyanobacteria and their closest relatives, Vampirovibrionia, in the Neoarchaean between 2.5-2.6 Ga [50],
180 but our molecular clock analyses point to an earlier origin where the cyanobacterial lineage emerges in
181 the Paleoarchaean between 3.3-3.5 Ga (Table 2). This is consistent with geochemical studies showing
182 that oxygen was accumulating in shallow coastal habitats ca. 3.0-3.2 Ga [33,51,52] and molecular clock

183 analyses of the PSII gene family which predict that photosystems capable of splitting water arose 3.3-
184 4.1 Ga [34].

185 ## Timing of SOD Isoforms in Cyanobacteria

186 Photosystems are one of many sources of $O_2^{\cdot-}$ (others include transmembrane NADPH oxidases
187 and flavins)[13] and as a result, the membrane-bound SOD isoforms - CuZnSOD and MnSOD - have been
188 found to protect photosystem repair mechanisms under bright light [22,53]. Some key parts of the
189 photosystem repair mechanism are membrane-bound FtsH proteases, which degrade damaged
190 photosystem components. Previous research suggests that they evolved with the advent of oxygenic
191 photosynthesis in the Archaean and utilize zinc cofactors [54]. Our phylogenetic analyses suggest for the
192 first time that cyanobacteria may also have been genetically capable of removing ROS (specifically $O_2^{\cdot-}$
193) using CuZnSOD in the Archaean between 2.9-2.6 Ga (Table 2).

194 This conclusion is based on the close relationship of CuZnSODs from two basal lineages:
195 *Pseudanabaena* spp. and *Gloeobacter* spp. (SI Appendix, Fig. S6, Fig. S1). The most recent common
196 ancestor (MRCA) of these two genera appeared ~3.45 Ga, and diversified ~2.6 Ga, giving rise to all
197 extant cyanobacteria (Fig. 2). If the ancestor of the crown cyanobacterium had *sodC*, it may have been
198 inherited by its descendants, resulting in a unique phylogenetic signal whereby the CuZnSODs of
199 *Gloeobacter* spp. and *Pseudanabaena* spp. are sisters. Because of this phylogenetic pattern (SI
200 Appendix, Fig. S6), it is, therefore, likely that crown cyanobacteria could have used CuZnSOD. An
201 alternative scenario could result in the same pattern if *sodC* was transferred laterally from the MRCA of
202 *Pseudanabaena* spp. to the MRCA of *Gloeobacter* spp. This would put an upper age constraint on the
203 origin of CuZnSOD ~950 mya (Fig. 2).

204 Timing the origin of Cu-based metalloenzymes is a difficult task. Characteristic protein-folds
205 required to bind Cu are predicted to have evolved during, or following, the GOE [29]. However, geological
206 evidence suggests that metabolisms which rely on copper metalloenzymes (eg. nitrification and methane

207 oxidation), were present in the Archaean, 2.7 Ga [55]. If proteins were capable of using Cu cofactors in the
208 Archaean, then CuZnSODs could have been present at the root of cyanobacteria as a mechanism to
209 deal with O_2 generated by photosynthesis, before it accumulated in the global atmosphere (at the GOE).
210 This would have been particularly important for Mesoarchaean and Neoarchaean individuals living inside
211 benthic mats [56,57], where there was a smaller diffusion gradient to pull O_2 out of the cells. As a result,
212 higher concentrations of O_2 accumulated, raising the potential for superoxide radical generation,
213 particularly in the periplasm [58].

214 Why are there only a few cyanobacterial strains with the gene encoding CuZnSOD today?
215 Perhaps it is related to physiological considerations resulting in the replacement of CuZnSOD for
216 MnSODs. The thylakoid membranes of *Nostoc* sp. PCC7120 lack CuZnSOD (Fig. 2) and instead contain
217 MnSOD [27]. When exposed to intense sunlight these MnSODs protect the cell from photoinhibition [53] in a
218 similar manner to CuZnSOD [22]. Furthermore, the gene encoding MnSOD can be post-translationally
219 processed to produce three proteins of different sizes which vary in concentration between heterocysts
220 and vegetative cells [26]. Under nitrogen-supplemented conditions, the smallest cytosolic 24 kDa protein
221 predominates the slightly larger 27 kDa protein, but under nitrogen-limiting conditions, both proteins are
222 present in near equal proportions [26]. No such flexibility has been documented in CuZnSODs. Therefore,
223 the ability to modify the size and localisation of MnSOD proteins in response to changes in the
224 environment could explain why the *sodA* gene encoding MnSOD is widespread in all major clades of
225 nitrogen-fixing cyanobacteria [39](Fig. 2) and many phyla of non-photosynthetic bacteria with diazotrophic
226 representatives (e.g. Firmicutes, Actinobacteria and Proteobacteria, SI Appendix, Fig. S2)[59].
227 Phylogenetic evidence for the utilization of SODs which incorporate Fe and Mn cofactors appears
228 relatively recently, in the middle of the Proterozoic (Table 2, Fig. 3). Intriguingly, modern cyanobacteria
229 might be less sensitive to oxidative stress caused by $O_2^{\cdot-}$ because the proteins with $[4Fe\text{-}4S]^{2+}$ clusters
230 have evolved to protect themselves by positioning them in less solvent-exposed parts of the molecule.

231 It may be that with this method that picocyanobacteria are able to thrive with only a cytoplasmic NiSOD
232 (Fig. 2).

233 The relatively late appearance of FeSOD and MnSOD in cyanobacteria is surprising as previous
234 studies have postulated an Archaean origin in bacteria [16,60,61]. What caused this delay? Earth system
235 models suggest that photoferrotrophs outcompeted cyanobacteria for upwelling nutrients in aquatic
236 habitats prior to the GOE [48]. Perhaps they also limited the soluble Fe^{2+} available for cyanobacteria, thus
237 facilitating a selective advantage to lineages which used alternative metals for relieving oxidative stress.
238 As the atmosphere became more oxygenated, photoferrotrophs were marginalized to shrinking pools of
239 Fe^{2+}. Our analyses suggest that the Neoproterozoic oxygenation increased cyanobacteria's requirement
240 for ROS defense mechanisms so much that lineages began using FeSOD or MnSOD regardless of the
241 waning global concentrations of Fe and Mn (Fig. 3). Further study, however, will be needed to assess
242 how effectively FeSODs, MnSODs and CuZnSODs protect against oxidative stress under different
243 ambient concentrations of oxygen and transition metals.

244 Phylogenetic evidence also indicates that cyanobacteria started using cytosolic NiSOD isoforms
245 in the MRCA of *Leptolyngbya* sp. PCC 7375 and *Nodosilinea nodulosa* PCC 7104 which diversified in
246 marine Neoproterozoic (Table 2, Fig. 3) habitats (SI Appendix, Fig. S7.SI Appendix, Fig. S8). During the
247 early Palaeoproterozoic and Mesoproterozoic, phylogenomic evidence suggests that cyanobacteria
248 were living in terrestrial, coastal brackish and marine benthic environments [35,62]. Therefore, NiSOD usage
249 also corresponds with the first appearance of marine planktonic unicellular nitrogen-fixing cyanobacteria
250 and non-nitrogen-fixing picocyanobacteria (*Synechococcus* and *Prochlorococcus*) in the open ocean at
251 the end of the Precambrian [36,43,63,64]. Although only speculative at this time, it is possible that the
252 acquisition of NiSOD and its associated maturation protease assisted the invasion of cyanobacteria into
253 pelagic marine habitats. Pelagic planktonic cyanobacteria are typically limited by P, Fe, and N (in cases
254 of non-nitrogen fixers) due to their distance from sources of riverine discharge [65]. As NiSOD is composed
255 of fewer amino acids than other SOD isoforms (SI Appendix, Fig. S3), and does not require Fe, its

utilization by benthic marine cyanobacteria in the Neoproterozoic may have imparted an evolutionary advantage and increased their resilience to nutrient limitation in the open ocean.

By studying only extant taxa, it is inherently difficult to estimate whether extinct lineages of cyanobacteria used NiSOD, MnSOD, FeSOD or CuZnSOD prior to the estimations provided in Fig. 2, Fig. 3 and Table 2. It is reassuring, however, that our estimates of all SOD isoforms predate the Cretaceous-Paleogene mass extinction of non-avian dinosaurs from terrestrial and aquatic environments 66 Mya [66], the Permian-Triassic mass extinction of 56% of all marine animal genera 252 Mya [67], and the Snowball Earth glaciations 720-635 Mya, which subjected life to extreme climate fluctuations [68]. Therefore, our estimations of timing are robust across several past extinction events. Any future discovery of novel lineages of basal cyanobacteria may change our estimated order of SOD appearance.

SODs and the Chemical Evolution of Seawater

It has been proposed that trace metal inventories of ancient marine sediments should reflect trace metal availability and correspond to the emergence of novel metabolisms and metalloenzymes [29]. In light of this suggestion, the past two decades have seen a number of papers purporting the use of ancient marine chemical sediments (i.e., banded iron formations - BIF, shales, pyrite) to track seawater composition through time (see [31] for a review). With regards to the metal cofactors contained in SOD, their temporal trajectories have been reconstructed based on both thermodynamic considerations (e.g. [69]) and sedimentary records (e.g. [31,70,71]). However, there is a poor record of trace metal inventories outside of the marine realm, so it can be more difficult to establish links between metal availability and diversification of cyanobacteria living in terrestrial and freshwater habitats.

Our ancestral state reconstructions predict that a non-marine cyanobacterium could have acquired CuZnSOD (SI Appendix, Fig. S7) in the Archaean (Table 2). In the Archaean to Proterozoic ocean, Cu and Zn were supposed to have been present at exceedingly low abundances given the low solubility of sulfide phases and the belief that they would be sequestered into euxinic sediments after the

GOE. However, BIF and shale records indicate near static reservoirs for Zn [70,71] and Cu [72]. This may hint at either a strong biological control on water column cycling being established very early on or the possible complexation of these trace metals by organic ligands. While the possible release of Cu though the oxidative weathering of sulfides in the lead up to, and following the GOE [72], may explain the emergence of CuZnSOD in the Mesoarchaean to Neoarchaean, at this point there is a dearth of geochemical evidence to support this hypothesis, at least from a global first-order perspective. Despite this, existing chemical sediment compilations cannot preclude the possibility of local or transient increases in Cu, creating oases of bioavailability that stimulated the emergence of CuZnSODs.

Past reconstructions of Ni in seawater based on BIF and syn-sedimentary to early diagenetic pyrite [73-75] show a dramatic, and unidirectional, decline that immediately preceded the Great Oxidation Event. This drop has been proposed to reflect a decline in the flux of nickel to the oceans due to waning komatiite weathering and would have marginalized methanogenic microbial communities from the water column to the sediment pile, corresponding to seawater concentrations decreasing from as high as 400 nM down to effectively near modern levels of Ni (~9 nM) [73,74]. In turn, this would have created environmentally-favourable conditions for a rise in oxygenic photosynthesis by cyanobacteria [74]. This Neoarchaean to Paleoproterozoic decline in Ni is also reflected in a compilation of shale data, normalized to evolving upper continental crust (Fig. 4). Interestingly, however, there appears to be a slight uptick in Ni in the terminal Neoproterozoic to Phanerozoic (Fig. 4) that occurs shortly after the appearance of NiSOD in the Tonian period (Table 2), which is not captured in the BIF record [73,74]. This discrepancy in sedimentary records is likely due to the declining concentration of Fe in seawater through time as Earth's oceans moved from a dominantly anoxic and ferruginous state in the Archaean to the modern oxic oceans we have today; a trend perhaps best reflected in the disappearance of large Superior-type BIF following the Paleoproterozoic (e.g., [76]).

An increase in Ni during the Neoproterozoic presumably reflects a relatively greater Ni flux to the oceans from either increased oxidative weathering, similar to other sulfide bound metals such as

molybdenum or vanadium at this time [47,77], or through the weathering of freshly emplaced large igneous provinces as has been proposed for phosphorous [78]. At present, the exact mechanisms and timing of higher Ni concentrations in Neoproterozoic shale require further investigation as it lags behind the observed increases in molybdenum, vanadium, and phosphorus [47,77-79].

What Controls SOD Isoforms in Cyanobacteria?

Given the histories of these trace metals, one may ask whether it is availability driving the assimilation of Cu, Zn, Fe, Mn and Ni into SOD metalloenzymes, or an inherent competition with other extant lineages at any given time. For instance, metalloenzyme fold superfamilies that incorporate zinc proliferate late and have been proposed to contribute to the delay in eukaryotic evolution [29]. Alternatively, this has been proposed to reflect an intrinsic biological property of eukaryotic evolution rather than a sudden shift in zinc availability [70]. In the latter view, the early incorporation of copper and zinc into SODs would reflect a lack of competition from the yet unevolved eukaryotes in the Archaean. Then in the late Mesoproterozoic, as Eukaryotes began to become more abundant and dominate marine and terrestrial environments, they may have outcompeted cyanobacteria for copper and zinc, providing an impetus for the emergence of other membrane-bound SODs, such as MnSOD – effectively providing a strategy for alleviating limitation imparted by competition with emergent Eukaryotes. Although Eukaryotes also utilize Mn, protein structures required to bind Mn have not been preferentially retained in their genomes [80] and cyanobacteria, having evolved earlier than photosynthetic Eukaryotes [81], had plenty of time to develop an efficient Mn uptake system.

Later, in Neoproterozoic marine habitats, cyanobacteria began supplementing their ROS defense mechanisms with a new and smaller SOD isoform (NiSOD) which requires less phosphorus and nitrogen to manufacture. This emergence of NiSOD coincides with Earth's second planetary oxygenation [3] and the emergence of planktonic marine cyanobacteria [36]. Therefore, we suggest that cyanobacteria

328 began assimilating nickel into SODs as it was being under-utilized in Neoproterozoic planktonic
329 communities when Neoproterozoic oxygen levels rose.
330
331 ## Methods
332
333 ### Acquisition of Superoxide Dismutase Sequences
334 Bacterial SOD diversity was assessed by downloading 15,899 bacterial genomes from the NCBI
335 RefSeq database (https://www.ncbi.nlm.nih.gov/refseq/) in 2018. They were searched for presence of
336 genes encoding NiSOD, CuZn-SOD, FeSOD and MnSOD using the basic local alignment search tool
337 (BLAST) for proteins (https://blast.ncbi.nlm.nih.gov/Blast.cgi) with a word size of six, gap opening penalty
338 of 11, gap extension penalty of 1 and e value cut off of <1x10^-5. Query sequences are detailed in SI
339 Appendix, Table S3. They were aligned to relevant subject sequences using the BLOSUM62 substitution
340 matrix. Each genome was treated as a separate subject (functionally equivalent to a unique database).
341 The functions of each hit were verified by choosing a representative subsample of hits with distant
342 evolutionary relationships and noting functional domains predicted by the NCBI conserved domain
343 search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi)(SI Appendix, Table S4 and Fig. S9). If
344 possible, query sequences with experimentally-defined functions were utilised.

345 Cyanobacterial SOD diversity was assessed by sampling 153 additional genomes (strains are
346 listed in SI Appendix, Table S1 and Table S5), including representatives from all major clades of salt-
347 tolerant and freshwater cyanobacteria [36]. These were supplemented with five eukaryotic plastid
348 genomes [81] and eight vampirovibrionial genomes representing all four sub-classes (gastranaerophilus,
349 vampirovibionales, caenarcaniphilales and obscuribacterales)[37].

350 ### SOD Gene Phylogenies

Amino acid sequences for all SOD isoforms (including paralogues found in cyanobacteria) were aligned using MAFFT [82] and gaps present in ≥80% sequences removed. The evolutionary history of each isoform was estimated by constructing nine ML phylogenies with IQ-TREE version 1.6.1 [41]. Substitution models were identified by ModelFinder [83] and node supports measured using ultrafast bootstrap two approximations [84]

Species Phylogeny of Cyanobacteria

An evolutionary tree of cyanobacteria was generated from a data set including SSU rRNA (1687 nucleotides), LSU rRNA (3387 nucleotides), and 136 core proteins (52,227 aa). Each protein is encoded by an orthologous gene which is conserved among cyanobacteria and involved in a key cellular function, such as information processing, metabolism or cellular processes. For a full description of the genes, see previous papers [35,36,63]. Single proteins and ribosomal RNA were aligned using MAFFT [82] and concatenated using an alignment viewer (http://sdsssdfd.altervista.org/arklumpus/AlignmentViewer/AlignmentViewer.html). Gaps were removed if present at the same position in ≥80% sequences and ML methodology implemented in IQ-TREE v1.6.1 [41] to estimate the cyanobacterial phylogeny. Each protein and rRNA was characterised under a unique partition with -SPP to account for heterotachy and appropriate substitution models identified by ModelFinder [83]. Ultrafast bootstrap values were calculated to measure support for branching relationships by resampling sites within partitions 1000 times [84].

Divergence Time Estimation

Bayesian relaxed molecular clocks were implemented using the topology described above. This topology was fixed and ages estimated based on predicted mutation rates for ribosomal RNA (SSU and LSU, 5074 aligned positions). Substitutions were modelled using a flexible general time reversible model inferred from the alignment and a Dirichlet process prior [85](GTR+CAT) to account for different rates of evolution between distant sites of the molecule. Divergence times were estimated in phylobayes 4.1 [40]

using uncorrelated gamma multipliers [40,44], a birth-death prior on divergence times and root prior chosen from a gamma probability distribution with mean 3060 and standard deviation 404. This ensured that the ancestor of all Vampirovibrionia and Cyanobacteria originated after the end of the late heavy bombardment 3.9 Ga [86,87].

We also implemented microfossil calibrations as follows: filamentous cyanobacteria more than 1.9 Ga [88,89], akinete-forming cyanobacteria 1.6 to 1.888 Ga [88,90,91] and apical cells of cyanobacteria 1.7 to 1.888 Ga [88,90,92]. These fossil constraints were supplemented with evidence dating the appearance of eukaryotic hosts of endosymbiotic cyanobacteria. For example, *Richelia species* diversify before their diatom host, named *Hemiaulus* [93], appeared 110 Mya [94] and UCYNA species diversify before their prymnesiophyte host, named *Braarudosphaera begelowii*, appeared 91 Mya [95]. Geochemical evidence were used to constrain cyanobacteria first diversifying before the Great Oxygenation Event of 2.32 Ga [96], but after either; a) 3 Ga when molybdenum oxides document the first 'whiffs' of atmospheric oxygen [51]; or b) 2.7 Ga based on the earliest fossilized evidence of cyanobacterial stromatolites [97]. Soft bounds were applied throughout to allow 5% of the prior probability density to fall outside of specified minimum and maximum ages.

Models were considered complete when four replicate independent chains converged. This was tested by estimating effective sample sizes and relative differences using Tracecomp (in Phylobayes with effsize >50, reldiff <0.3). Final branch supports represent 95% HPD intervals calculated for a single representative chain using Readdiv (in Phylobayes) with the same burn-in (25% of mean chain length) and sampling frequency (1 in every 10 points) as Tracecomp.

Chronology of SOD Isoforms

The order in which SOD isoforms appeared in cyanobacteria, was estimated using a method described as 'topological comparison' [98]. First, a Bayesian protein phylogeny was created for each SOD isoform using only cyanobacterial sequences (SI Appendix, Fig. S6 and Fig. S8). Bayesian phylogenetic

reconstructions of MnSODs and FeSODs had not converged after 2 weeks, so the sequences were separated into 7 groups based on their position in the ML bacterial phylogenies (SI Appendix, Fig. S2) to conduct alignments and phylogenetic reconstructions on more similar sequences (SI Appendix, Fig. S10). The resulting protein phylogenies were then compared to the species phylogeny of cyanobacteria (Fig. 2) to identify monophyletic groups whose NiSOD, CuZnSOD or MnSOD/FeSOD had evolved as expected by vertical inheritance (without HGT). The last common ancestor of each of these clades was assumed to have been capable of using the corresponding SOD.

Ancestral State Reconstruction of Habitat Preference

To find out which habitats ancestral cyanobacteria lived in when SOD isoforms appeared, Bayesian stochastic character mapping [99] was implemented with SIMMAP v1.5 [100] in the phytools package [101] of R using our time-calibrated trees. Prior distribution on the root node of the tree was estimated based on the data, 1000 simulations were performed and the *all rates different model* was utilised to allow transition rates between marine and non-marine habitats to vary based the data (as implemented in [81]). Character states were coded as either 'marine' or 'non-marine' (SI Appendix, Table S1).

Compilation of Ni data

From a database of >4,000 literature shale analyses spanning the Archaean to modern, 1,584 had both Ni and Ti data available and are used here to reconstruct the trajectory of Ni in seawater. Molar Ni/Ti ratios were normalized to evolving continental crust [102], then time-binned based on age (Archean, Paleoproterozoic, Mesoproterozoic, Neoproterozoic, Phanerozoic). Mean values for each time bin were bootstrap resampled (n=10,000; Fig. 4) in Matlab® 2019b using the *bootstrp* function.

References

421 1 Wheelis, M. *Principles Of Modern Microbiology*. 140-167 (Jones and Bartlett Publishers,
422 London, 2008).

423 2 Kump, L. R. The rise of atmospheric oxygen. *Nature* **451,** 277-278 (2008).

424 3 Lyons, T. W., Reinhard, C. T. & Planavsky, N. J. The rise of oxygen in Earth's early ocean and
425 atmosphere. *Nature* **506,** 307-315 (2014).

426 4 Rosing, M. T. & Frei, R. U-rich Archaean sea-floor sediments from Greenland - indications of >
427 3700 Ma oxygenic photosynthesis. *Earth Planet. Sci. Lett.* **217,** 237-244 (2004).

428 5 Fischer, W. W., Hemp, J. & Johnson, J. E. Evolution of oxygenic photosynthesis. *Annu. Rev.*
429 *Earth Planet. Sci*. **44,** 647-683 (2016).

430 6 Konhauser, K. O. et al. Aerobic bacterial pyrite oxidation and acid rock drainage during the Great
431 Oxidation Event. *Nature* **478,** 369-374 (2011).

432 7 Passardi, F. et al. Phylogenetic distribution of catalase-peroxidases: Are there patches of order
433 in chaos? *Gene* **397,** 101-113 (2007).

434 8 Slesak, I., Kula, M., Slesak, H., Miszalski, Z. & Strzalka, K. How to define obligatory
435 anaerobiosis? An evolutionary view on the antioxidant response system and the early stages of the
436 evolution of life on Earth. *Free Radic. Biol. Med.* **140,** 61-73 (2019).

437 9 Slesak, I., Slesak, H., Zimak-Piekarczyk, P. & Rozpadek, P. Enzymatic antioxidant systems in
438 early anaerobes: Theoretical considerations. *Astrobiology* **16,** 348-358 (2016).

439 10 Zamocky, M., Gasselhuber, B., Furtmuller, P. G. & Obinger, C. Molecular evolution of hydrogen
440 peroxide degrading enzymes. *Arch. Biochem. Biophys.* **525,** 131-144 (2012).

441 11 Latifi, A., Ruiz, M. & Zhang, C. C. Oxidative stress in Cyanobacteria. *FEMS Microbiol. Rev.* **33,**
442 258-278 (2009).

443 12 Bernroitner, M., Zamocky, M., Furtmuller, P. G., Peschek, G. A. & Obinger, C. Occurrence,
444 phylogeny, structure, and function of catalases and peroxidases in Cyanobacteria. *J. Exp. Bot.* **60,** 423-
445 440 (2009).

446 13 Sheng, Y. et al. Superoxide dismutases and superoxide reductases. *Chem. Rev.* **114,** 3854-
447 3918 (2014).

448 14 Wallace, M. A. et al. Superoxide inhibits 4Fe-4S cluster enzymes involved in amino acid
449 biosynthesis. Cross-compartment protection by CuZn-superoxide dismutase. *J. Biol. Chem.* **279,** 32055-
450 32062 (2004).

451 15 Vassiliev, I. R., Ronan, M. T., Hauska, G. & Golbeck, J. H. The bound electron acceptors in
452 green sulfur bacteria: Resolution of the g-tensor for the F(X) iron-sulfur cluster in *Chlorobium tepidum*.
453 *Biophys. J.* **78,** 3160-3169 (2000).

454 16 Miller, A. F. Superoxide dismutases: Ancient enzymes and new insights. *FEBS Lett.* **586,** 585-
455 595 (2012).

456 17 Perry, J. J., Shin, D. S., Getzoff, E. D. & Tainer, J. A. The structural biochemistry of the
457 superoxide dismutases. *BBA Proteins and Proteomics* **1804,** 245-262 (2010).

458 18 Wolfe-Simon, F., Grzebyk, D., Schofield, O. & Falkowski, P. G. The role and evolution of
459 superoxide dismutases in algae. *J. Phycol.* **41,** 453-465 (2005).

460 19 Priya, B. et al. Comparative analysis of cyanobacterial superoxide dismutases to discriminate
461 canonical forms. *BMC Genomics* **8,** 435 (2007).

462 20 Dupont, C. L., Neupane, K., Shearer, J. & Palenik, B. Diversity, function and evolution of genes
463 coding for putative Ni-containing superoxide dismutases. *Environ. Microbiol.* **10,** 1831-1843 (2008).

464 21 Caiola, M. G., Canini, A. & Ocampo-Friedmann, R. Iron superoxide dismutase (Fe-SOD)
465 localization in *Chroococcidiopsis* sp (Chroococcales, Cyanobacteria). *Phycologia* **35,** 90-94 (1996).

466 22 Ke, W. T., Dai, G. Z., Jiang, H. B., Zhang, R. & Qiu, B. S. Essential roles of iron superoxide
467 dismutase in photoautotrophic growth of *Synechocystis* sp. PCC 6803 and heterogeneous expression
468 of marine *Synechococcus* sp. CC9311 copper/zinc superoxide dismutase within its *sodB* knockdown
469 mutant. *Microbiol.* **160,** 228-241 (2014).

470 23 Regelsberger, G. et al. The iron superoxide dismutase from the filamentous cyanobacterium
471 *Nostoc* PCC 7120. Localization, overexpression, and biochemical characterization. *J. Biol. Chem.* **279,**
472 44384-44393 (2004).

473 24 Okada, S., Kanematsu, S. & Asada, K. Intracellular distribution of manganese and ferric
474 superoxide dismutases in blue-green algae. *FEBS Lett.* **103,** 106-110 (1979).

475 25 Regelsberger, G. et al. Biochemical characterization of a membrane-bound manganese-
476 containing superoxide dismutase from the cyanobacterium *Anabaena* PCC 7120. *J. Biol. Chem.* **277,**
477 43615-43622 (2002).

478 26 Raghavan, P. S., Rajaram, H. & Apte, S. K. Membrane targeting of MnSOD is essential for
479 oxidative stress tolerance of nitrogen-fixing cultures of *Anabaena* sp. strain PCC7120. *Plant Mol. Biol.*
480 **88,** 503-514 (2015).

481 27 Li, T. et al. Differential expression and localization of Mn and Fe superoxide dismutases in the
482 heterocystous cyanobacterium *Anabaena* sp strain PCC 7120. *J. Bacteriol.* **184,** 5096-5103 (2002).

483 28 Schmidt, A., Gube, M., Schmidt, A. & Kothe, E. In silico analysis of nickel containing superoxide
484 dismutase evolution and regulation. *J. Basic Microbiol.* **49,** 109-118 (2009).

29 Dupont, C. L., Butcher, A., Valas, R. E., Bourne, P. E. & Caetano-Anolles, G. History of biological metal utilization inferred through phylogenomic analysis of protein structures. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 10567-10572 (2010).

30 Sigel, H. & Sigel, R. *Metal ions in biological systems*. 1st edn, Vol. 43 (Taylor & Francis Group, Boca Raton, 2005).

31 Robbins, L. J. et al. Trace elements at the intersection of marine biological and geochemical evolution. *Earth-Sci. Rev.* **163,** 323-348 (2016).

32 Javaux, E. J. Challenges in evidencing the earliest traces of life. *Nature* **572,** 451-460 (2019).

33 Homann, M. Earliest life on Earth: Evidence from the Barberton Greenstone Belt, South Africa. *Earth-Sci. Rev.* **196,** 102888 (2019).

34 Cardona, T., Sanchez-Baracaldo, P., Rutherford, A. W. & Larkum, A. W. Early Archean origin of Photosystem II. *Geobiology* **17,** 127-150 (2019).

35 Blank, C. E. & Sanchez-Baracaldo, P. Timing of morphological and ecological innovations in the cyanobacteria - a key to understanding the rise in atmospheric oxygen. *Geobiology* **8,** 1-23 (2010).

36 Sanchez-Baracaldo, P. Origin of marine planktonic cyanobacteria. *Sci. Rep.* **5,** 17418 (2015).

37 Soo, R. M. et al. An expanded genomic representation of the phylum cyanobacteria. *Genome Biol. Evol.* **6,** 1031-1045 (2014).

38 Di Rienzi, S. C. et al. The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *Elife* **2** (2013).

39 Sanchez-Baracaldo, P., Hayes, P. K. & Blank, C. E. Morphological and habitat evolution in the cyanobacteria using a compartmentalization approach. *Geobiology* **3,** 145-165 (2005).

506 40 Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: a Bayesian software package for
507 phylogenetic reconstruction and molecular dating. *Bioinformatics* **25,** 2286-2288 (2009).

508 41 Trifinopoulos, J., Nguyen, L. T., von Haeseler, A. & Minh, B. Q. W-IQ-TREE: A fast online
509 phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **44,** W232-235 (2016).

510 42 Shi, T. & Falkowski, P. G. Genome evolution in Cyanobacteria: The stable core and the variable
511 shell. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 2510-2515 (2008).

512 43 Sanchez-Baracaldo, P., Bianchini, G., Di Cesare, A., Callieri, C. & Chrismas, N. A. M. Insights
513 into the evolution of picocyanobacteria and phycoerythrin genes (mpeBA and cpeBA). *Front. Microbiol.*
514 **10,** (2019).

515 44 Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating
516 with confidence. *PLoS Biol.* **4,** 699-710 (2006).

517 45 Kopp, R. E., Kirschvink, J. L., Hilburn, I. A. & Nash, C. Z. The Paleoproterozoic snowball Earth:
518 A climate disaster triggered by the evolution of oxygenic photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.*
519 **102,** 11131-11136 (2005).

520 46 Reinhard, C. T., Raiswell, R., Scott, C., Anbar, A. D. & Lyons, T. W. A late Archean sulfidic sea
521 stimulated by early oxidative weathering of the continents. *Science* **326,** 713-716 (2009).

522 47 Scott, C. et al. Tracing the stepwise oxygenation of the Proterozoic ocean. *Nature* **452,** 456-459
523 (2008).

524 48 Ozaki, K., Thompson, K. J., Simister, R. L., Crowe, S. A. & Reinhard, C. T. Anoxygenic
525 photosynthesis and the delayed oxygenation of Earth's atmosphere. *Nat. Commun.* **10,** 3026 (2019).

526 49 Schad, M., Konhauser, K. O., Sanchez-Baracaldo, P., Kappler, A. & Bryce, C. How did the
527 evolution of oxygenic photosynthesis influence the temporal and spatial development of the microbial
528 iron cycle on ancient earth? *Free Radic. Biol. Med.* **140,** 154-166 (2019).

529 50 Shih, P. M., Hemp, J., Ward, L. M., Matzke, N. J. & Fischer, W. W. Crown group
530 Oxyphotobacteria postdate the rise of oxygen. *Geobiology* **15,** 19-29 (2017).

531 51 Planavsky, N. J. et al. Evidence for oxygenic photosynthesis half a billiion years before the Great
532 Oxidation Event. *Nat. Geosci.* **7,** 283-286 (2014).

533 52 Satkoski, A. M., Beukes, N. J., Weiqiang, L., Beard, B. L. & Johnson, C. M. A redox-stratified
534 ocean 3.2 billion years ago. *Earth Planet. Sci. Lett.* **430,** 43-53 (2015).

535 53 Zhao, W. X., Guo, Q. X. & Zhao, J. D. A membrane-associated Mn-superoxide dismutase
536 protects the photosynthetic apparatus and nitrogenase from oxidative damage in the cyanobacterium
537 *Anabaena* sp PCC 7120. *Plant Cell Physiol.* **48,** 563-572 (2007).

538 54 Shao, S., Cardona, T. & Nixon, P. J. Early emergence of the FtsH proteases involved in
539 photosystem II repair. *Photosynthetica* **56,** 163-177 (2018).

540 55 Moore, E. K., Jelen, B. I., Giovannelli, D., Raanan, H. & Falkowski, P. G. Metal availability and
541 the expanding network of microbial metabolisms in the Archaean eon. *Nat. Geosci.* **10,** 629-636 (2017).

542 56 Lalonde, S. V. & Konhauser, K. O. Benthic perspective on Earth's oldest evidence for oxygenic
543 photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 995-1000 (2015).

544 57 Wilmeth, D. T. et al. Neoarchean (2.7 Ga) lacustrine stromatolite deposits in the Hartbeesfontein
545 Basin, Ventersdorp Supergroup, South Africa: Implications for oxygen oases. *Precambrian Res.* **320,**
546 291-302 (2019).

547 58 Korshunov, S. & Imlay, J. A. Detection and quantification of superoxide formed within the
548 periplasm of *Escherichia coli*. *J. Bacteriol.* **188,** 6326-6334 (2006).

549 59 Hartmann, L. S. & Barnum, S. R. Inferring the evolutionary history of Mo-dependent nitrogen
550 fixation from phylogenetic studies of *nifK* and *nifDK*. *J. Mol. Evol.* **71,** 70-85 (2010).

551 60 Inupakutika, M. A., Sengupta, S., Devireddy, A. R., Azad, R. K. & Mittler, R. The evolution of
552 reactive oxygen species metabolism. *J. Exp. Bot.* **67,** 5933-5943 (2016).

553 61 Kirschvink, J. L. et al. Paleoproterozoic snowball earth: Extreme climatic and geochemical global
554 change and its biological consequences. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 1400-1405 (2000).

555 62 Raven, J. A., Beardall, J., Larkum, A. W. D. & Sanchez-Baracaldo, P. Interactions of
556 photosynthesis with genome size and function. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **368,** (2013).

557 63 Sanchez-Baracaldo, P. & Cardona, T. On the origin of oxygenic photosynthesis and
558 Cyanobacteria. *New Phytol.* **225,** 1440-1446 (2020).

559 64 Sanchez-Baracaldo, P., Ridgwell, A. & Raven, J. A. A Neoproterozoic transition in the marine
560 nitrogen cycle. *Curr. Biol.* **24,** 652-657 (2014).

561 65 Bristow, L. A., Mohr, W., Ahmerkamp, S. & Kuypers, M. M. M. Nutrients that limit growth in the
562 ocean. *Curr. Biol.* **27,** R474-R478 (2017).

563 66 Brusatte, S. L. et al. The extinction of the dinosaurs. *Biol. Rev.* **90,** 628-642 (2015).

564 67 Bambach, R. K. Phanerozoic biodiversity mass extinctions. *Annu. Rev. Earth Planet. Sci.* **34,**
565 127-155 (2006).

566 68 Hoffman, P. F. et al. Snowball Earth climate dynamics and Cryogenian geology-geobiology. *Sci.*
567 *Adv.* **3,** e1600983 (2017).

568 69 Saito, M. A., Sigman, D. M. & Morel, F. M. M. The bioinorganic chemistry of the ancient ocean:
569 The co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean-
570 Proterozoic boundary? *Inorganica Chim. Acta* **356,** 308-318 (2003).

571 70 Robbins, L. J. et al. Authigenic iron oxide proxies for marine zinc over geological time and
572 implications for eukaryotic metallome evolution. *Geobiology* **11,** 295-306 (2013).

573 71 Scott, C. et al. Bioavailability of zinc in marine systems through time. *Nat. Geosci.* **6,** 125-128
574 (2013).

575 72 Fru, E. C. et al. Cu isotopes in marine black shales record the Great Oxidation Event. *Proc. Natl.*
576 *Acad. Sci. U.S.A.* **113,** 4941-4946 (2016).

577 73 Konhauser, K. O. et al. Oceanic nickel depletion and a methanogen famine before the Great
578 Oxidation Event. *Nature* **458,** 750-U785 (2009).

579 74 Konhauser, K. O. et al. The Archean nickel famine revisited. *Astrobiology* **15,** 804-815 (2015).

580 75 Large, R. R. et al. Trace element content of sedimentary pyrite as a new proxy for deep-time
581 ocean–atmosphere evolution. *Earth Planet. Sci. Lett.* **389,** 209-220 (2014).

582 76 Konhauser, K. O. et al. Iron formations: A global record of Neoarchaean to Palaeoproterozoic
583 environmental history. *Earth-Sci. Rev.* **172,** 140-177 (2017).

584 77 Sahoo, S. K. et al. Ocean oxygenation in the wake of the Marinoan glaciation. *Nature* **489,** 546-
585 549 (2012).

586 78 Horton, F. Did phosphorus derived from the weathering of large igneous provinces fertilize the
587 Neoproterozoic ocean? *Geochem. Geophy. Geosy.* **16,** 1723-1738 (2015).

588 79 Reinhard, C. T. et al. Evolution of the global phosphorus cycle. *Nature* **541,** 386-389 (2017).

589 80 Dupont, C. L., Yang, S., Palenik, B. & Bourne, P. E. Modern proteomes contain putative imprints
590 of ancient shifts in trace metal geochemistry. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 17822-17827 (2006).

591 81 Sanchez-Baracaldo, P., Raven, J. A., Pisani, D. & Knoll, A. H. Early photosynthetic eukaryotes
592 inhabited low-salinity habitats. *Proc. Natl. Acad. Sci. U.S.A.* **114,** E7737-E7745 (2017).

593 82 Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:
594 Improvements in performance and usability. *Mol. Biol. Evol.* **30,** 772-780 (2013).

595 83 Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder:
596 Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14,** 587-591 (2017).

597 84 Hoang, D. T. C., O. Haeseler,A.V. Minh,B.Q. Vinh,L.S. UFBoot2: Improving the ultrafast
598 bootstrap approximation. *Mol. Biol. Evol.* **35,** 518-522 (2017).

599 85 Lartillot, N. & Philippe, H. A Bayesian mixture model for across-site heterogeneities in the amino-
600 acid replacement process. *Mol. Biol. Evol.* **21,** 1095-1109 (2004).

601 86 Sleep, N. H., Zahnle, K. J., Kasting, J. F. & Morowitz, H. J. Annihilation of ecosystems by large
602 asteroid impacts on the early Earth. *Nature* **342,** 139-142 (1989).

603 87 Zahnle, K. et al. Emergence of a habitable planet. *Space Sci. Rev.* **129,** 35-78 (2007).

604 88 Golubic, S. & Lee, S. J. Early cyanobacterial fossil record: Preservation, palaeoenvironments
605 and identification. *Eur. J. Phycol.* **34,** 339-348 (1999).

606 89 Hofman, H. J. Precambrian microflora, Belcher Islands, Canada: Significance and systematics.
607 *J. Paleontol.* **50,** 1040-1073 (1976).

608 90 Sergeev, V. N., Gerasimenko, L. M. & Zavarzin, G. A. The Proterozoic history and present state
609 of Cyanobacteria. *Microbiology* **71,** 623-637 (2002).

610 91 Golubic, S., Sergeev, V. N. & Knoll, A. H. Mesoproterozoic Archaeoellipsoides: Akinetes of
611 heterocystous cyanobacteria. *Lethaia* **28,** 285-298 (1995).

612 92 Zhang, Y. & Golubic, S. Endolithic microfossils (cyanophyta) from early Proterozoic
613 stromatolites, Hebei, China. *Acta Micropaleontol. Sin.* **4,** 1-12 (1987).

614 93 Foster, R. A. et al. Nitrogen fixation and transfer in open ocean diatom-cyanobacterial
615 symbioses. *ISME J.* **5,** 1484-1493 (2011).

616 94 Sims, A. P., Mann, D. G. & Medlin, L. K. Evolution of the diatoms: Insights from fossil, biological
617 and molecular data. *Phycologia* **45,** 361-402 (2006).

618 95 Cornejo-Castillo, F. M. et al. Cyanobacterial symbionts diverged in the late Cretaceous towards
619 lineage-specific nitrogen fixation factories in single-celled phytoplankton. *Nat. Commun*. **7,** 11071 (2016).

620 96 Bekker, A. et al. Dating the rise of atmospheric oxygen. *Geochim. Cosmochim. Acta* **68,** A780-
621 A780 (2004).

622 97 Bosak, T., Liang, B., Sim, M. S. & Petroff, A. P. Morphological record of oxygenic photosynthesis
623 in conical stromatolites. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 10939-10943 (2009).

624 98 Chan, C. X., Beiko, R. G. & Ragan, M. A. Scaling up the phylogenetic detection of lateral gene
625 transfer events. *Methods Mol. Biol.* **1525,** 421-432 (2017).

626 99 Huelsenbeck, J. P., Nielsen, R. & Bollback, J. P. Stochastic mapping of morphological
627 characters. *Syst. Biol.* **52,** 131-158 (2003).

628 100 Bollback, J. P. SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC*
629 *Bioinformatics* **7,** 88 (2006).

630 101 Revell, L. J. Phytools: An R package for phylogenetic comparative biology (and other things).
631 *Methods Ecol. Evol.* **3,** 217-223 (2012).

102 Condie, K. C. Chemical-composition and evolution of the upper continental-crust - contrasting results from surface samples and shales. *Chem. Geol.* **104,** 1-37 (1993).

Acknowledgments

We thank Andrew H. Knoll, Paul J. Valdes, John Raven and Paul Falkowski for helpful comments on the manuscript. We also thank Giorgio Bianchini for technical help designing scripts for automated BLASTP and leaf colouring as well as developing a trimming tool for alignments. All phylogenetic analyses were performed at the high-performance computing facility (BlueCrystal 3 and 4) at the University of Bristol. Funding support for this work came from a Royal Society University Research Fellowship to P.S.-B. L.J.R. gratefully acknowledges the support of a Donnelley Environmental Postdoctoral Fellowship from the Yale Institute for Biospheric Studies.

Author Contributions

J.S.B. and P.S-B. conceived the project; J.S.B. and P.S-B. designed the phylogenetic and molecular clocks analyses, J.S.B. performed analyses, J.S.B. and P.S-B. interpreted evolutionary results; K.O.K and L.J.R. interpreted the Ni shale records; J.S.B., K.O.K, L.J.R. and P.S-B. wrote the paper.

Competing Interests

The authors declare no conflict of interest.

Figure legends

Figure 1: Habitat distribution of SOD genes. Panels represent the distribution of SOD isoforms in Cyanobacteria from (A) marine, (B) non-marine, (C) freshwater, (D) geothermal springs, and (E) terrestrial habitats.

Figure 2: Time calibrated cyanobacterial tree of life and superoxide dismutases. Genes encoding NiSOD (orange circles), NiSOD maturation protease (orange stars), CuZnSOD (red squares), and SODs with Mn- or Fe- cofactors (blue left-facing triangles) are highlighted next to leaf labels. The earliest node predicted to contain each SOD isoform is highlighted with a coloured circle and label (orange for NiSOD, red for CuZnSOD, and blue for the ancestor of FeSOD and MnSOD; see Table 2 for posterior age probabilities and SI Appendix, Fig. S11 for age distribution of these events). The phylogenetic tree was estimated from SSU and LSU ribosomal RNA and 136 core cyanobacterial proteins from 167 different taxa using maximum likelihood methodology implemented in IQ-TREE v1.6.1 [41]. Node labels represent ultrafast bootstrap approximations less than 100 [84]. Ages were estimated using a Bayesian relaxed molecular clock with Uncorrelated Gamma Multipliers [44] for ribosomal RNA. Black circles represent calibration points (Table 1). The first divergence of Cyanobacteria was constrained to occur between 2.7 Ga and 2.32 Ga [96,97]. GOE, Great Oxidation Event. Uni. diaz. refers to unicellular diazotrophs and Phan. refers to the phanerozoic eon.

Figure 3: Emergence of Cyanobacterial SOD isoforms inferred through geological time. Coloured bands represent the distribution of age posterior estimates (Table 2) for key nodes in a Bayesian molecular clock parameterised with uncorrelated Gamma Multipliers [44] and calibrated with the first divergence of Cyanobacteria between 2.32 and 2.7 Ga. Vertical black lines represent the mean in each instance. A description of each node is presented beneath the age distribution. Any given node is the

677 earliest predicted to have contained a gene encoding the SOD isoform stated above it based on
678 topological comparisons of gene trees and species trees. Phan. Refers to the phanerozoic eon; Meso.
679 Refers to the Mesoproterozoic era and Neo. Refers to the Neoproterozoic era. NOE, Neoproterozoic
680 oxygenation event.
681

682 **Figure 4: Bootstrap resampled mean values from time binned shale samples**. Each group
683 represents a time binned subset of the shale database with mean values of molar Ni/Ti, normalized to
684 evolving crust [102], being bootstrap resampled (n=10,000 for each bin).

685 **Table 1. Calibration points used for molecular clock analyses**. The most recent common ancestor
686 of cyanobacteria was constrained to diversify before widespread atmospheric oxygenation 2.32 Ga [96]
687 and after either early whiffs of oxygen dated 3 Ga [51] or formation of the first fossilized cyanobacterial
688 stromatolites 2.7 Ga [97].

689

Calibration	Minimum Age / mya	References	Maximum Age / mya	References	Diversification between…
1st Cyanobacterial diversification	2320	96	2700	97	*Gloeobacter.violaceus.PCC7421* *Acaryochloris.sp.MBIC11017*
	2320	96	3000	51	
Filamentous Cyanobacteria	1900	89	n/a	n/a	*Pseudanabaena.biceps.PCC7429* *Leptolyngbya.sp.PCC7376*
Akinete-forming cyanobacteria	1600	91	1888	88,90	*Calothrix.sp.PCC7103* *Nostoc.azollae.0708*
Apical cells of Cyanobacteria	1700	92	1888	88,90	*Pleurocapsa.sp.PCC7327* *Pleurocapsa.sp.PCC7319*
Endosymbionts of *Hemiaulus* spp.	110	94	n/a	n/a	*Richelia.intracellularis.HH01* *Richelia.intracellularis.HM01*
Endosymbionts of *B. begelowii*	91	95	n/a	n/a	*UCYNA2* *UCYNA*

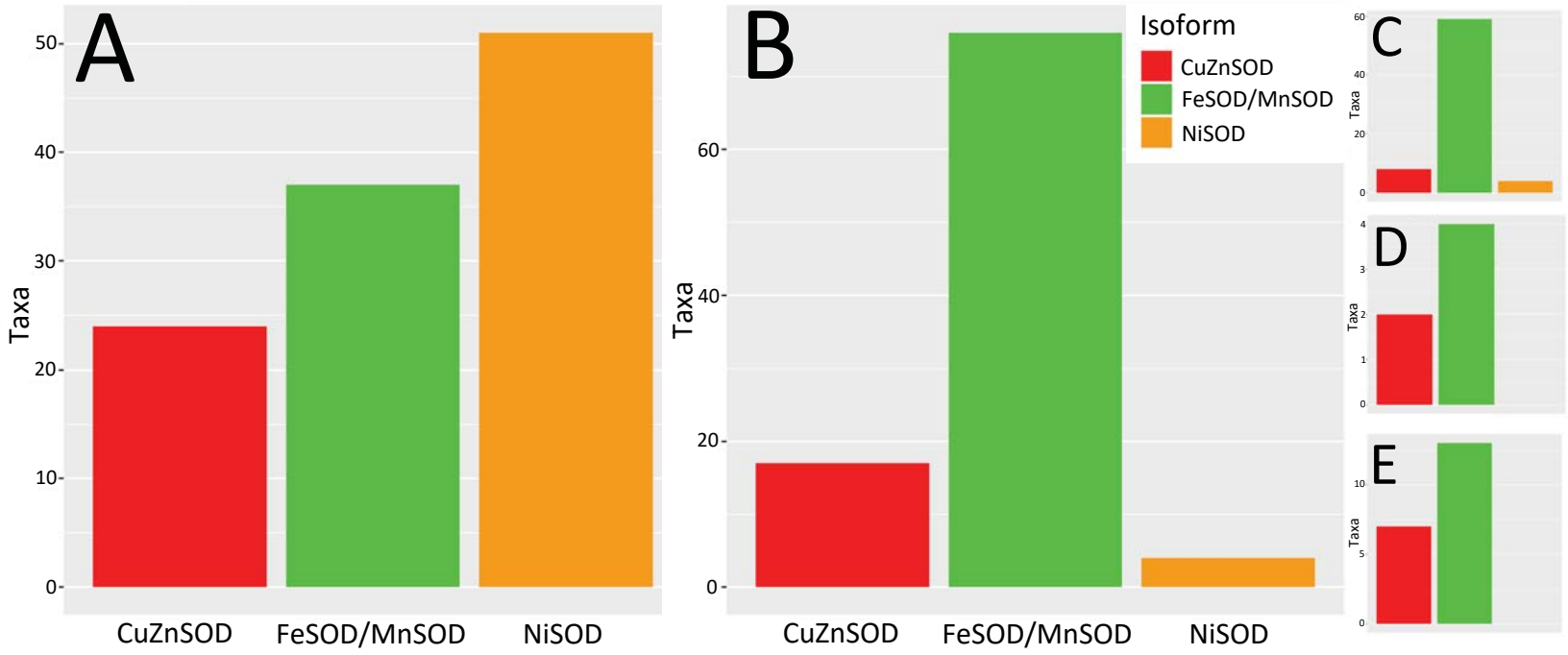
690 **Table 2. Statistics summarizing divergence times predicted by molecular clock analyses.** Mean
691 divergence times are presented in millions of years alongside 95% HPD confidence intervals in brackets.

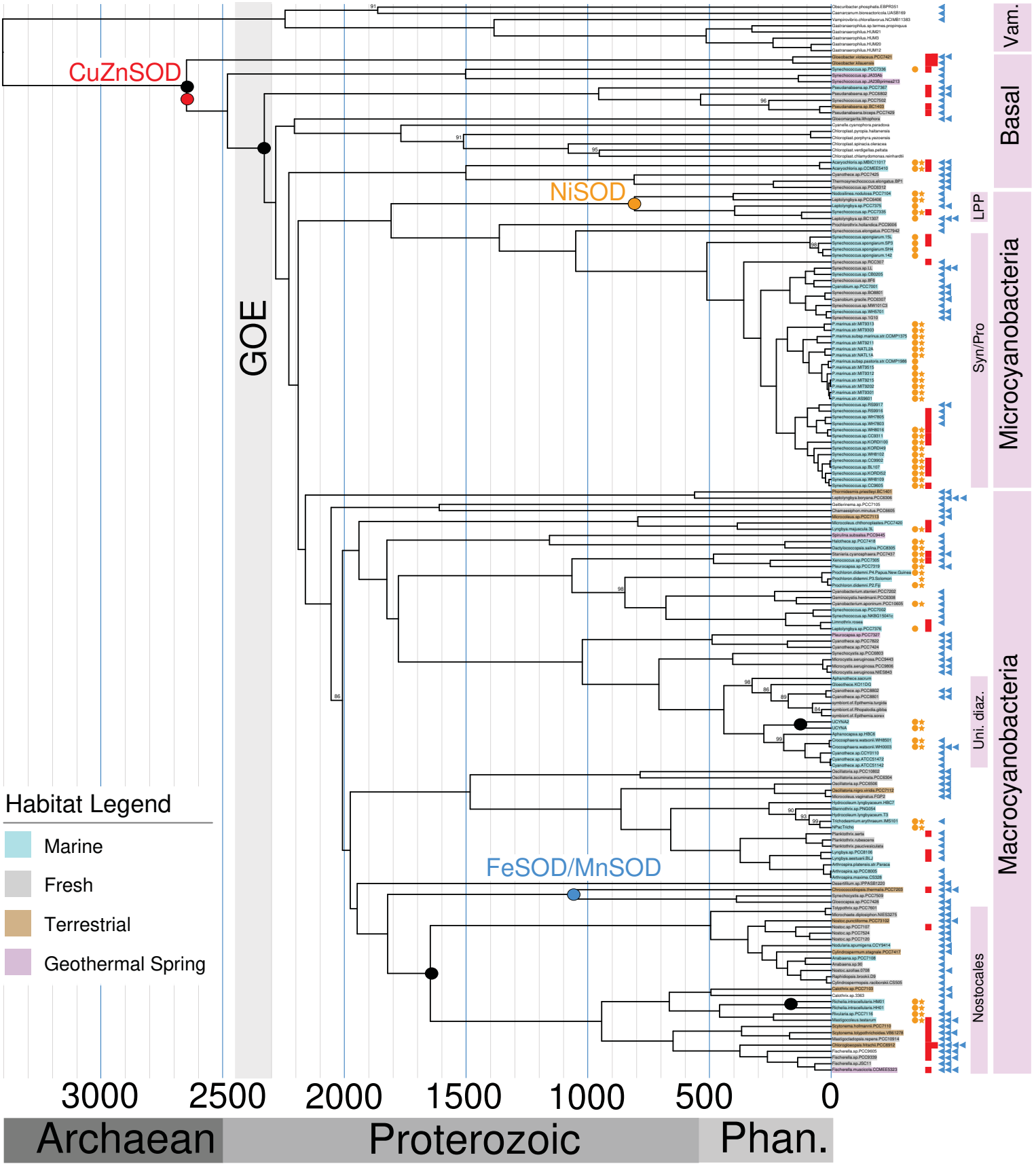
	First Divergence of Cyanobacteria	
	3 to 2.32 Ga	2.7 to 2.32 Ga
Divergence of Cyanobacteria and Vampirovibrionia	3544 (4235-3001)	3374 (4058-2807)
NiSOD	912 (1982-281)	806 (1898-226)
CuZnSOD	2926 (3043-2692)	2649 (2720-2490)
MnSOD/FeSOD	1140 (1941-207)	1045 (1893-171)

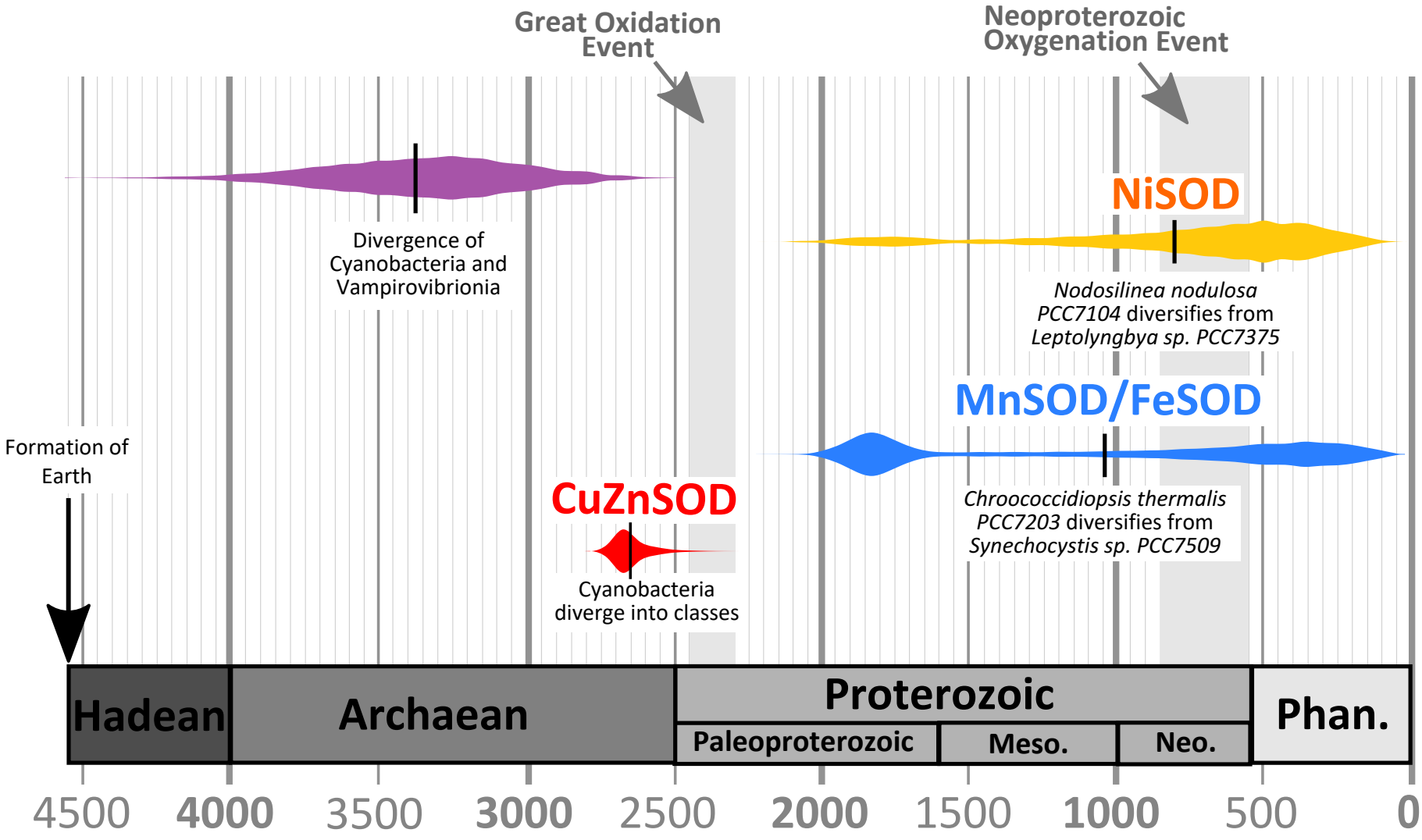
692

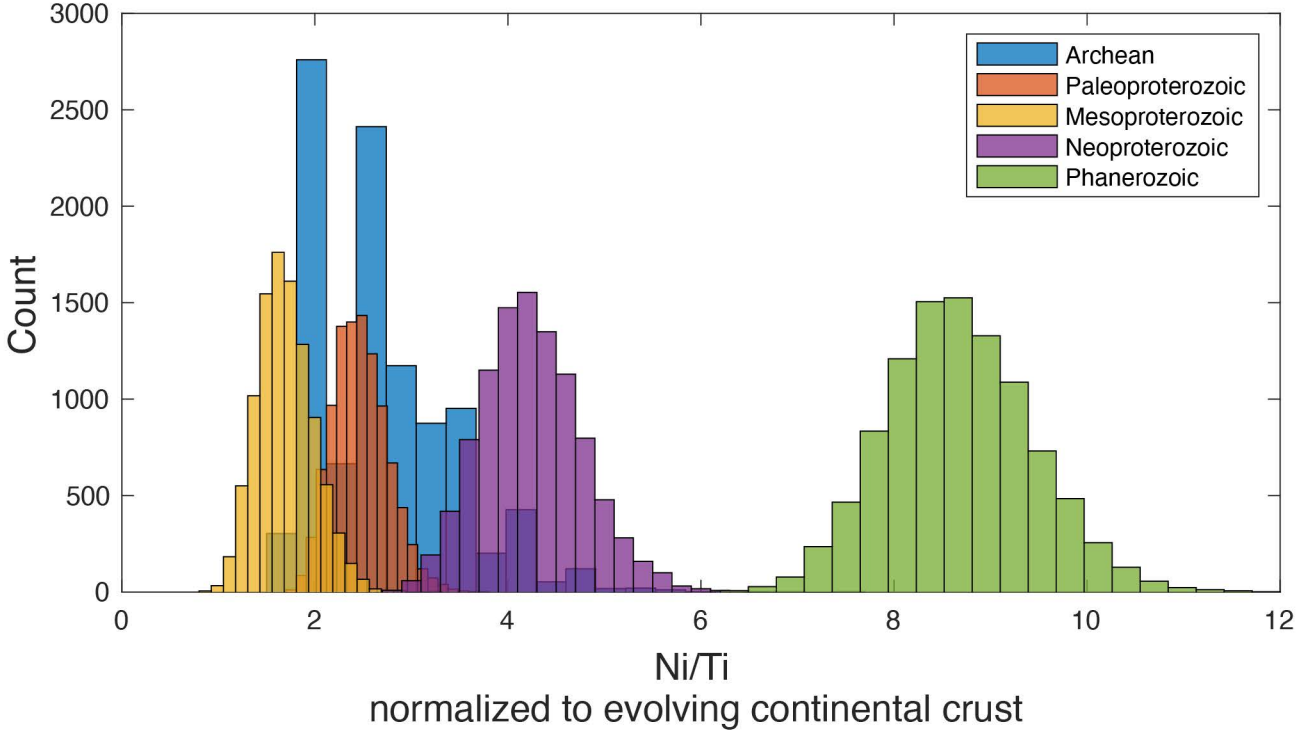
693

694


A
B
C
D
E
Taxa
CuZnSOD
FeSOD/MnSOD
NiSOD
Isoform
CuZnSOD
FeSOD/MnSOD
NiSOD


CuZnSOD
NiSOD
FeSOD/MnSOD
GOE
Habitat Legend
Marine
Fresh
Terrestrial
Geothermal Spring
Vam.
Basal
LPP
Microcyanobacteria
Syn/Pro
Uni. diaz.
Macrocyanobacteria
Nostocales
3000
2500
2000
1500
1000
500
0
Archaean
Proterozoic
Phan.


Great Oxidation Event
Neoproterozoic Oxygenation Event
Divergence of Cyanobacteria and Vampirovibrionia
NiSOD
Nodosilinea nodulosa PCC7104 diversifies from Leptolyngbya sp. PCC7375
MnSOD/FeSOD
Chroococcidiopsis thermalis PCC7203 diversifies from Synechocystis sp. PCC7509
Formation of Earth
CuZnSOD
Cyanobacteria diverge into classes
Hadean
Archaean
Proterozoic
Paleoproterozoic
Meso.
Neo.
Phan.
4500
4000
3500
3000
2500
2000
1500
1000
500
0


Archean
Paleoproterozoic
Mesoproterozoic
Neoproterozoic
Phanerozoic
Count
3000
2500
2000
1500
1000
500
0
0
2
4
6
8
10
12
Ni/Ti
normalized to evolving continental crust

Figures

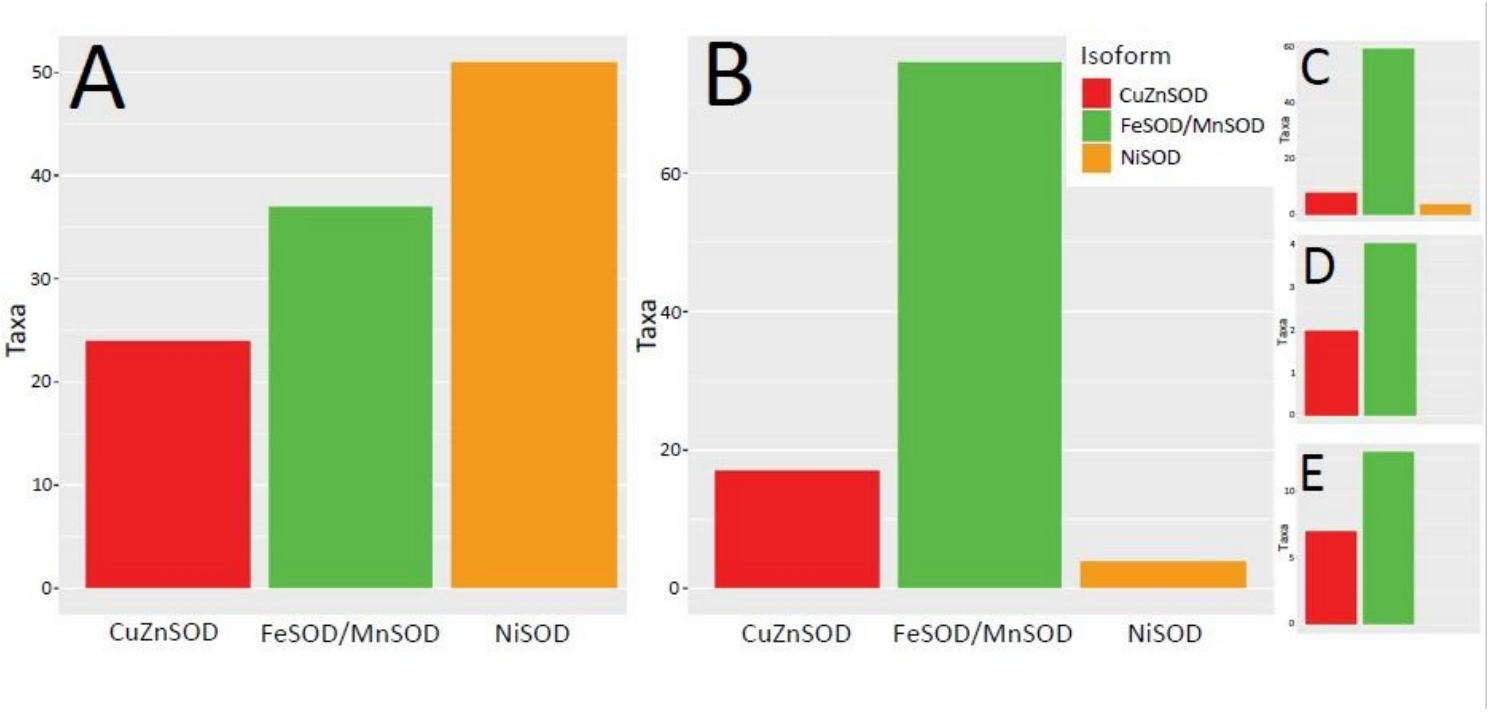


Figure 1

Habitat distribution of SOD genes. Panels represent the distribution of SOD isoforms in Cyanobacteria from (A) marine, (B) non-marine, (C) freshwater, (D) geothermal springs, and (E) terrestrial habitats.

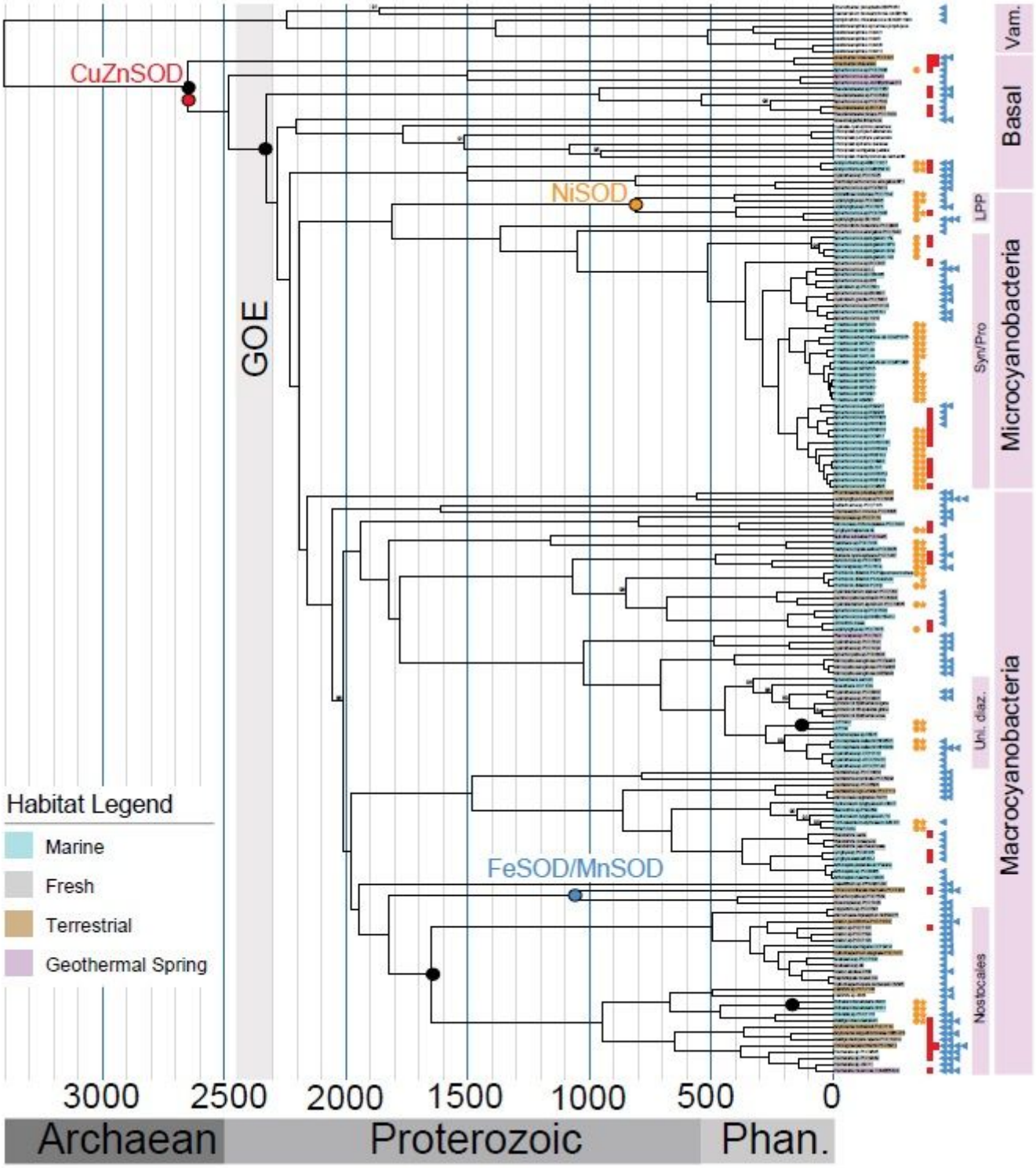


Figure 2

Time calibrated cyanobacterial tree of life and superoxide dismutases. Genes encoding NiSOD (orange circles), NiSOD maturation protease (orange stars), CuZnSOD (red squares), and SODs with Mn- or Fe-cofactors (blue left-facing triangles) are highlighted next to leaf labels. The earliest node predicted to contain each SOD isoform is highlighted with a coloured circle and label (orange for NiSOD, red for CuZnSOD, and blue for the ancestor of FeSOD and MnSOD; see Table 2 for posterior age probabilities

and SI Appendix, Fig. S11 for age distribution of these events). The phylogenetic tree was estimated from SSU and LSU ribosomal RNA and 136 core cyanobacterial proteins from 167 different taxa using maximum likelihood methodology implemented in IQ-TREE v1.6.1 41. Node labels represent ultrafast bootstrap approximations less than 100 84. Ages were estimated using a Bayesian relaxed molecular clock with Uncorrelated Gamma Multipliers 44 for ribosomal RNA. Black circles represent calibration points (Table 1). The first divergence of Cyanobacteria was constrained to occur between 2.7 Ga and 2.32 Ga 96,97. GOE, Great Oxidation Event. Uni. diaz. refers to unicellular diazotrophs and Phan. refers to the phanerozoic eon.

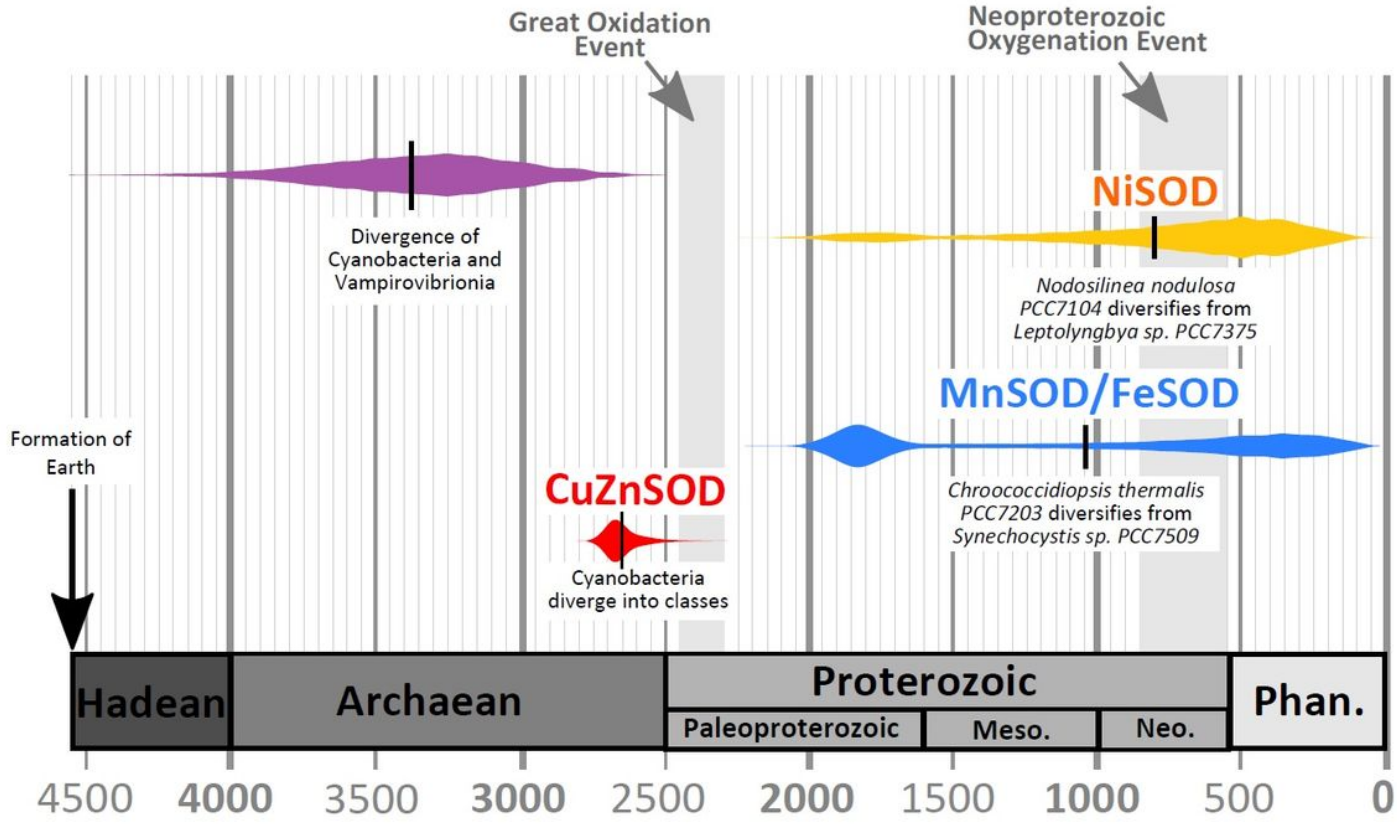


Figure 3

Emergence of Cyanobacterial SOD isoforms inferred through geological time. Coloured bands represent the distribution of age posterior estimates (Table 2) for key nodes in a Bayesian molecular clock parameterised with uncorrelated Gamma Multipliers 44 and calibrated with the first divergence of Cyanobacteria between 2.32 and 2.7 Ga. Vertical black lines represent the mean in each instance. A description of each node is presented beneath the age distribution. Any given node is the earliest predicted to have contained a gene encoding the 677 SOD isoform stated above it based on topological comparisons of gene trees and species trees. Phan. Refers to the phanerozoic eon; Meso. Refers to the Mesoproterozoic era and Neo. Refers to the Neoproterozoic era. NOE, Neoproterozoic oxygenation event.

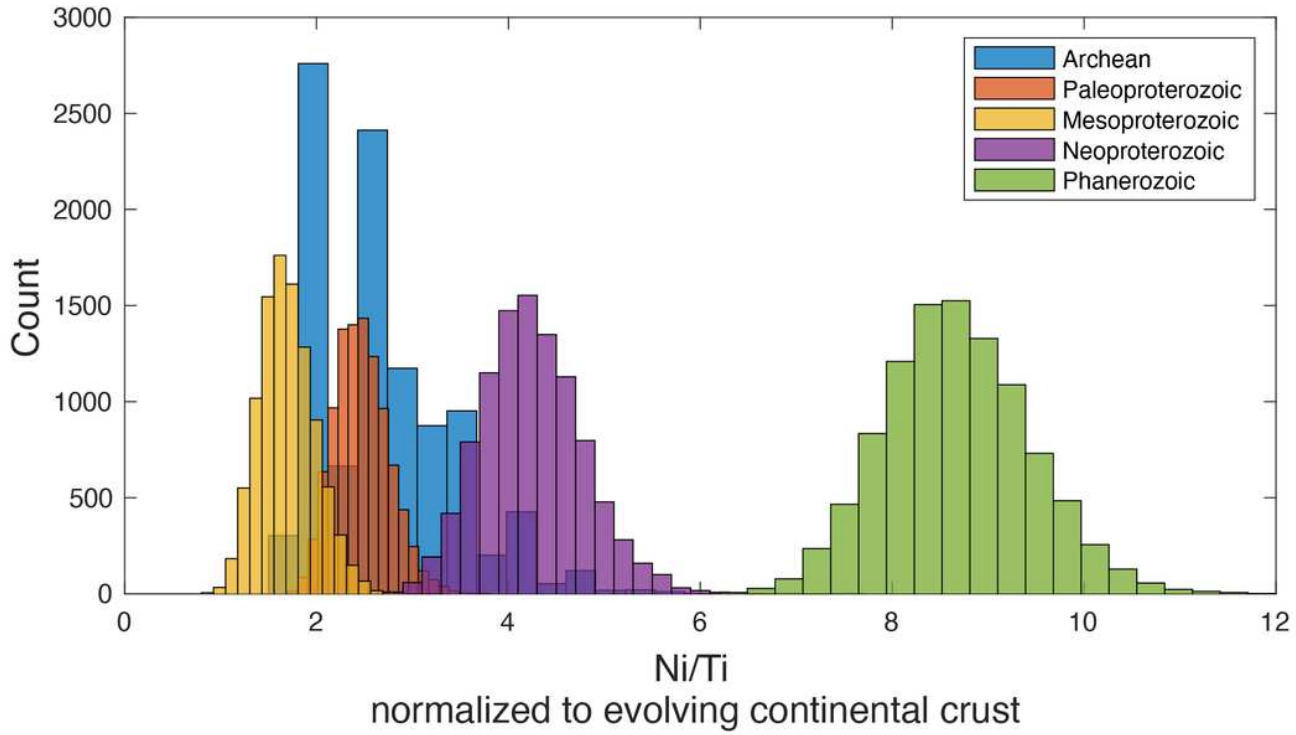


Figure 4

Bootstrap resampled mean values from time binned shale samples. Each group represents a time binned subset of the shale database with mean values of molar Ni/Ti, normalized to evolving crust 102, being bootstrap resampled (n=10,000 for each bin).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformationFINAL.pdf