ORIGINAL ARTICLE

WILEY gebiology

# Crown group Oxyphotobacteria postdate the rise of oxygen

P. M. Shih[1,2] | J. Hemp[3] | L. M. Ward[3] | N. J. Matzke[4] | W. W. Fischer[3]

[1]Joint BioEnergy Institute, Emeryville, CA, USA

[2]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

[3]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

[4]National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, TN, USA

**Correspondence**
Patrick M. Shih, Joint BioEnergy Institute, Emeryville, CA, USA.
Email: pmshih@gmail.com
and
Woodward W. Fischer, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA.
Email: wfischer@caltech.edu

## Abstract

The rise of oxygen ca. 2.3 billion years ago (Ga) is the most distinct environmental transition in Earth history. This event was enabled by the evolution of oxygenic photosynthesis in the ancestors of Cyanobacteria. However, long-standing questions concern the evolutionary timing of this metabolism, with conflicting answers spanning more than one billion years. Recently, knowledge of the Cyanobacteria phylum has expanded with the discovery of non-photosynthetic members, including a closely related sister group termed Melainabacteria, with the known oxygenic phototrophs restricted to a clade recently designated Oxyphotobacteria. By integrating genomic data from the Melainabacteria, cross-calibrated Bayesian relaxed molecular clock analyses show that crown group Oxyphotobacteria evolved ca. 2.0 billion years ago (Ga), well after the rise of atmospheric dioxygen. We further estimate the divergence between Oxyphotobacteria and Melainabacteria ca. 2.5–2.6 Ga, which—if oxygenic photosynthesis is an evolutionary synapomorphy of the Oxyphotobacteria—marks an upper limit for the origin of oxygenic photosynthesis. Together, these results are consistent with the hypothesis that oxygenic photosynthesis evolved relatively close in time to the rise of oxygen.

## 1 | INTRODUCTION

Oxygenic photosynthesis was responsible for the most profound environmental shift in Earth history: the rise of oxygen. It was long recognized that this metabolism evolved in the Cyanobacteria phylum and that this unique ability was a necessary precondition for the rise of oxygen at ca. 2.35 Ga (Luo et al., 2016). However, the evolutionary origins of Cyanobacteria remain uncertain, due to many conflicting lines of evidence ranging from microfossils (Schopf & Packer, 1987), geochemical data (Crowe et al., 2013; Johnson, Gerpheide, Lamb, & Fischer, 2014; Kopp, Kirschvink, Hilburn, & Nash, 2005; Rosing & Frei, 2004), biomarkers (Brocks, Logan, Buick, & Summons, 1999), and geochemical models (Holland, 2009; Kump & Barley, 2007). These different proxies for Cyanobacteria provide divergence estimates that span more than one billion years of Earth history. This quandary leaves a major gap in knowledge regarding the $O_2$ cycle; consequently, it remains unclear whether the rise of oxygen was directly related to the evolution of oxygenic photosynthesis (Kopp et al., 2005; Ward, Kirschvink, & Fischer, 2015), or alternatively driven by a change in Earth's geophysical processes (Holland, 2009; Kump & Barley, 2007).

The origins of Cyanobacteria have been intensely debated over the last half-century. Historically, hypotheses have been grounded in observations from the fossil and sedimentary rock records. However, the validity of some of the landmark studies, which have become core to many fundamental assumptions of the field, has been called into question, specifically concerning the biogenicity and taxonomic affinity of microfossils (Brasier et al., 2002; Butterfield, 2015; Knoll, 2003) and the recent reevaluation of Archean molecular fossils as sample contaminants (French et al., 2015). Because of the paucity of unequivocal evidence from early Precambrian sedimentary rocks, phylogenetics and comparative genomics provide an independent approach to constrain the evolutionary timing of Cyanobacteria, offering a useful point of comparison with the geological and fossil records (Shih, 2015).

A natural way to reconstruct and evaluate the evolution of oxygenic photosynthesis by Cyanobacteria from comparative biology is to identify their closest living relatives, which until recently remained a mystery. The notion that all Cyanobacteria evolved from a common photosynthetic ancestor hinged upon limited sampling of their extant diversity. Adding to the uncertainty, previous phylogenetic efforts to distinguish bacterial phylum-level relationships and identify the closest relative

to Cyanobacteria have yielded conflicting results (Daubin, Gouy, & Perrière, 2002; Wolf, Rogozin, Grishin, Tatusov, & Koonin, 2001; Wu et al., 2009; Zhaxybayeva & Gogarten, 2002). However, rapidly growing genomic and metagenomic datasets have substantially added to our understanding of the Cyanobacteria phylum. 16S rDNA surveys first described a substantial diversity of microbes related to oxygenic Cyanobacteria from aphotic environments (Ley et al., 2005). Recently, genomes from a wide range of aphotic environments, such as gut microbiomes and groundwater samples, were analyzed that revealed the presence of a close sister group to oxygenic Cyanobacteria, named Melainabacteria, filling in an important gap in the diversity and evolution of the Cyanobacteria phylum (Di Rienzi et al., 2013; Johnson et al., 2013a; van der Lelie et al., 2012; Ley et al., 2005; Soo et al., 2014). The close evolutionary relationships of Melainabacteria and previously known Cyanobacteria fit with widely applied guidelines using 16S rDNA sequence identity for inclusion in microbial phyla (Soo et al., 2014); this is a similar degree of sequence identity, for example, as that observed between *Escherichia coli* and *Pseudomonas aeruginosa*—two common and well-studied members of the class γ-Proteobacteria in the phylum Proteobacteria (Hugenholtz, Goebel, & Pace, 1998). The placement of Melainabacteria as a close sister group to all known phototrophic Cyanobacteria is objective and reproducible on the basis a myriad of different gene and protein comparisons (Di Rienzi et al., 2013; Fischer, Hemp, & Johnson, 2016; Fischer, Hemp, & Valentine, 2016; Johnson et al., 2013a; Soo et al., 2014). Additionally, the known oxygenic Cyanobacteria are much more closely related to the Melainabacteria than they are to any of the currently known members of other phyla capable of anoxygenic phototrophy, which are placed at far greater evolutionary distances, with numerous non-phototrophic groups interspersed between them (Fischer et al., 2016). These data illustrated that the Cyanobacteria phylum hosts a greater degree of physiological diversity than previously recognized—a condition similar to all the other known phototrophic phyla, which contain both phototrophic and non-phototrophic members (Fischer et al., 2016).

With the discovery of a substantial diversity of close-living relatives to the oxygenic Cyanobacteria, it has been proposed to update the systematics of the Cyanobacteria phylum by adding a Melainabacteria class and relegating the oxygenic Cyanobacteria to the Oxyphotobacteria class (Soo et al., 2014). For ease of discussion, we refer to these two sister clades using the following nomenclature: Oxyphotobacteria (class *Oxyphotobacteria* containing all known oxygenic Cyanobacteria) and Melainabacteria, wherein the Melainabacteria class to date includes four orders *Gastranaerophilales* (Di Rienzi et al., 2013; Ley et al., 2005), *Obscuribacterales*, *Vampirovibrionales*, and *Caenarcaniphilales* (Soo et al., 2014; Fig. 1). Interestingly, out of all the Melainabacteria genomes sequenced, none contain any genes associated with photosynthesis, and many lack genes necessary for aerobic or anaerobic respiration (Di Rienzi et al., 2013; Soo et al., 2014). Despite substantial environmental and genomic diversity, due to the current lack of observed photosynthetic basal lineages, it is important to consider the hypothesis that oxygenic photosynthesis is a derived, and perhaps relatively recent, feature of the Cyanobacteria phylum (Fig. 1). Thus, the newly identified Melainabacteria can add useful

phylogenetic information in the sparsely covered regions closer to the base of the Cyanobacterial phylum.

To estimate the origin of Oxyphotobacteria, molecular clock studies have typically used either molecular fossils (Battistuzzi, Feijao, & Hedges, 2004; Battistuzzi & Hedges, 2009) or cyanobacterial-like microfossils (Sánchez-Baracaldo, Ridgwell, & Raven, 2014; Schirrmeister, de Vos, Antonelli, & Bagheri, 2013; Tomitani, Knoll, Cavanaugh, & Ohno, 2006) as calibration constraints. It was once common to interpret microfossils as specific living lineages on the basis of morphological traits; however, phylogenetic analyses have revealed multiple independent acquisitions and widespread convergences of many of the classical cyanobacterial morphotypes, such as baeocystous and filamentous cells (Shih et al., 2013; Turner, Pryer, Miao, & Palmer, 1999), highlighting classic challenges in assigning microfossils to extant clades (Knoll & Golubic, 1992). Moreover, 2-methylhopane molecular fossils, once thought to have been specific to the oxygenic Cyanobacteria, appear to have evolved in other phyla (Ricci, Michel, & Newman, 2015) and may not offer unique calibration constraints, concerns about syngeneity aside (French et al., 2015). A number of studies have used the rise of oxygen as a calibration point for the minimal age of Oxyphotobacteria (Battistuzzi et al., 2004; Falcon, Magallon, & Castillo, 2010; Sánchez-Baracaldo et al., 2014; Schirrmeister et al., 2013)—this placement implicitly assumes that crown group Oxyphotobacteria were responsible for the rise of oxygen. However, it is equally plausible that extinct lineages existing before the most recent common ancestor of Oxyphotobacteria (i.e., stem lineages) sourced the $O_2$ fluxes connected with the rise of oxygen. In order to test these assumptions, it is useful to relax these constraints. Instead, valuable evolutionary insights into the origins of Oxyphotobacteria come from endosymbiosis, which lends constraints from the fossil records from plants and algae—characteristics that most other bacterial phyla do not share. Additionally, by incorporating new molecular data from Melainabacteria taxa, molecular clock analyses can now obtain a tighter estimate of the divergence time of Oxyphotobacteria.

Here, we revisit this evolutionary problem by performing cross-calibrated Bayesian relaxed molecular clock analyses with increased taxonomic sampling encompassing Melainabacteria, Oxyphotobacteria, plastids, and mitochondria. Distinct from previous cross-calibration efforts, we expand upon the technique by using a concatenated dataset composed of slowly evolving proteins and genes found in both plastids and mitochondria, which permit calibrations to be used multiple times across a phylogenetic tree (Shih & Matzke, 2013).

## 2 | METHODS

### 2.1 | Generation of concatenated dataset

Sequences from subunits of ATP synthase, the ribosomal large subunit, the ribosomal small subunit, elongation factor Tu, and 16S rDNA were gathered. The protein sequences gathered from ATP synthase machinery consisted of AtpA, AtpB, AtpE, AtpF, AtpH, and AtpI. The ribosomal protein subunits collected for this study were Rpl2, Rpl16,
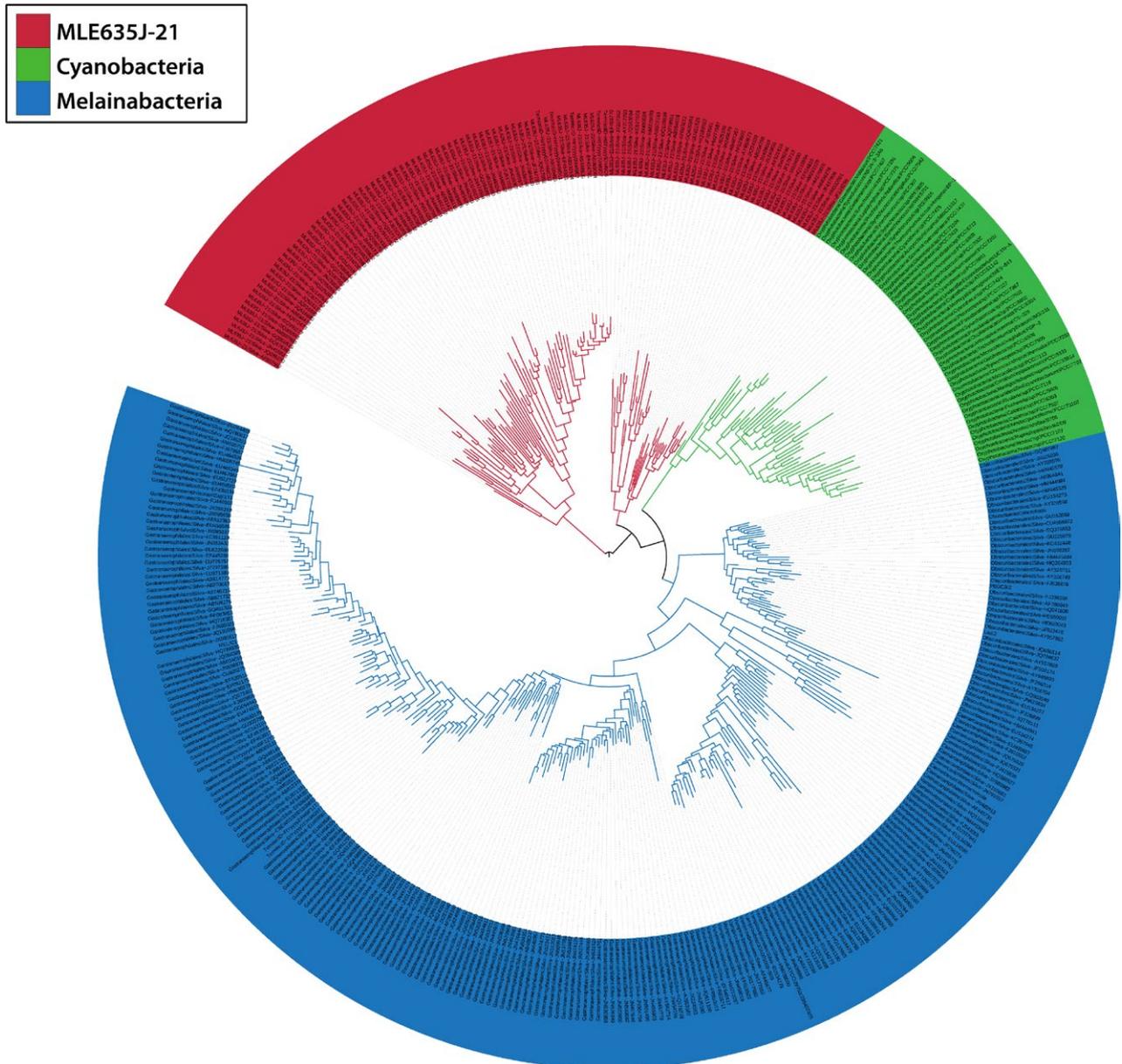
**FIGURE 1** Complete labeled 16S rDNA tree of the Cyanobacterial phylum. Class Oxyphotobacteria in green, class Melainabacteria (orders Gastranaerophilales, Obscuribacterales, Vampirovibrionales, Caenarcaniphilales) in blue, and ML635J-21 clades shown in red. The majority of these Melainabacteria and ML635J-21 sequences were collected from aphotic and anaerobic environments, consistent with the physiologies reconstructed from existing Melainabacteria genomes. This topology warrants consideration of the hypothesis that oxygenic photosynthesis was derived from within the Cyanobacterial phylum

Rps3, and Rps12. All sequences and their corresponding accessions are listed in Table S1. Sequences were collected to maximize the coverage of the Oxyphotobacteria and plastid-bearing eukaryotes. Plant mitochondrial genomes were used along with α-Proteobacteria to serve as both an appropriate outgroup and enable cross-calibration between the corresponding plant mitochondrial and plastid lineages. Alignments for each protein or nucleotide family were performed using the—maxiterate strategy in the alignment program MAFFT (Katoh, Kuma, Toh, & Miyata, 2005), and then concatenated to generate the final dataset. The dataset was partitioned into two parts: concatenated protein and 16S nucleotide sequences, respectively.

## 2.2 | Age calibrations

Dating calibration priors were primarily chosen to avoid biasing analyses with the introduction of controversial microfossil occurrences, instead relying on well-accepted divergence times of plant fossils estimated. The posterior divergence times from the comprehensive molecular clock analysis of land plants by Smith, Beaulieu, and Donoghue (2010) were used as priors in this study. A summary of all constraints used is described in Table 1. Normal distributions of $217 \pm 40$, $327 \pm 30$, $432 \pm 30$, and $477 \pm 70$ Ma were used as divergence time calibration points for Angiospermae, Gymnospermatophyta, Tracheophyta, and

**TABLE 1** Summary of calibration constraints used in this study

| Divergence event | Type of distribution | Constraint in Mya (±SD) |
|---|---|---|
| Monocotyledoneae | Normal | 156 (±14) |
| Angiospermae | Normal | 217 (±40) |
| Gymnospermatophyta | Normal | 327 (±30) |
| Tracheophyta | Normal | 432 (±30) |
| Land plants | Normal | 477 (±70) |
| *Bangiomorpha* | Uniform | 1,174–1,222 |
| "Rise of Oxygen" | Uniform | 2,400–3,000 |
| Last common ancestor | Uniform | 2,400–3,800 |

land plants, respectively. Importantly, the use of land plant divergence events enabled the cross-calibration between divergence events happening simultaneously in lineages that contain both plastids and mitochondria. In BEAST, runs incorporating the fossil *Bangiomorpha pubescens* were constrained using a normal distribution between 1,174 and 1,222 Ma, as done in Yoon, Hackett, Ciniglia, Pinto, and Bhattacharya (2004). Strict geochronologic constraints for the strata from which these fossils were found—the Hunting Formation on Somerset Island, Nunavut, Canada—are between ca. 1,200 and ca. 900 Ma. We examined a constraint associated with the oldest estimate placed between 1,174 and 1,222 Ma. This was chosen in part to provide a conservative (i.e., oldest possible) estimate on the divergence of crown group Oxyphotobacteria (Butterfield, Knoll, & Swett, 1988, 1990). Because the timing of the evolution of oxygenic photosynthesis remains controversial based on geological observations, we used a uniform prior spanning between 2,400 and 3,000 Ma and tested this prior on the two different nodes that represent (i) the divergence between Melainabacteria and Oxyphotobacteria and (ii) the radiation of crown group Oxyphotobacteria. The "Rise of Oxygen" constraint provides a minimum age of 2,400 Ma (Hoffman, 2013), marking the oldest ages hypothesized for the rise of oxygen, and an upper age of 3,000 Ma as suggested by various geological studies (Crowe et al., 2013; Planavsky et al., 2014). A uniform distribution enables the Markov chain Monte Carlo (MCMC) search to agnostically explore with no initial bias before ultimately converging onto a date within the provided hard upper and lower bounds of 2,400–3,000 Ga. Finally, a uniform prior between 3,800 and 2,400 Ma was used as a calibration for the last common ancestor of all taxa used in this study. The large range was used to permit an unbiased and largely unrestricted constraint, which assumes that oxygenic photosynthesis must predate the rise of oxygen and that these taxa had not diverged prior to the Late Heavy Bombardment ca. 3,800 Ma (Cohen, Swindle, & Kring, 2000).

## 2.3 | Molecular clock analysis

Dated phylogenies were estimated using the program BEAST (Drummond & Rambaut, 2007) using the CIPRES Science Gateway v. 3.3 server (Miller, Pfeiffer, & Schwartz, 2010). Cross-calibrated analyses were coded into BEAST XML files as previously described (Shih &

Matzke, 2013). For each macrofossil that provides an age calibration, the same date is applied to every node in the molecular phylogeny that corresponds to the speciation event. Thus, if paralogs of a gene are found in the nucleus, chloroplast, and mitochondrion, one calibration fossil supplies calibration distributions for several nodes in the gene phylogeny. The CpREV model was chosen as the best-fitting amino acid substitution model based on ProtTest analysis (Abascal, Zardoya, & Posada, 2005). The CpREV model was used for the concatenated amino acid sequences and the 16S rDNA sequences used the GTR + G model in accordance with the 16S rDNA molecular clock study using BEAST by Schirrmeister et al. (2013). A variety of BEAST runs were calculated using different combinations of date calibration priors, described below. For all runs, we ran five MCMC chains for the maximum limit allowed on the CIPRES Science Gateway, sampling every 10,000th generation. On average, more than 20 million generations were collected from each MCMC chain, and the initial 5 million generations were discarded as burn-in (for exact numbers of post-burn-in generations reported in Table 2). Maximum clade credibility (MCC) trees were generated using TreeAnnotator v1.7.5. All BEAST runs are summarized in Table 2. In general, the resulting chronograms of runs T64, T65, T68, and T69 were consistent with the overall phylogenomic analyses revealing the sister relationship between Oxyphotobacteria and Melainabacteria described by both Soo et al.

**TABLE 2** Summary of cross-calibrated BEAST runs generated in this study

| Name | Description | Generations |
|---|---|---|
| T64 | "Rise of Oxygen" prior set on Mel/Oxyph divergence. Bangiomorpha prior used | 78.04 million post-burn-in generations across 5 runs (each run had a burn-in of 5 million), 103.04 million total |
| T65 | "Rise of Oxygen" prior set on Mel/Oxyph divergence. Bangiomorpha prior omitted | 79.08 million post-burn-in generations across 5 runs (each run had a burn-in of 5 million), 104.08 million total |
| T68 | "Rise of Oxygen" prior omitted. Bangiomorpha prior used | 99.12 million post-burn-in generations across 5 runs (each run had a burn-in of 5 million), 124.12 million total |
| T69 | "Rise of Oxygen" prior omitted. Bangiomorpha prior omitted | 133.14 million post-burn-in generations across 5 runs (each run had a burn-in of 5 million), 158.14 million total |
| T72 | "Rise of Oxygen" prior set on crown cyano divergence. Bangiomorpha prior used | 81.14 million post-burn-in generations using a burn-in of 5 million, 106.14 million total |
| T73 | "Rise of Oxygen" prior set on crown cyano divergence. Bangiomorpha prior omitted | 81.78 million post-burn-in generations using a burn-in of 5 million, 106.78 million total |

(2014) and Di Rienzi et al. (2013). In contrast, the topology of T72 and T73 placed Melainabacteria sister to α-Proteobacteria.

## 2.4 | Regression analysis of node age uncertainty

The 95% highest posterior density (HPD) widths of node dates from the MCC trees generated by each analysis were extracted and plotted in R. For each pair of analyses, the list of HPD widths was reduced to the set of nodes common between the trees (the trees had mostly common nodes), and the HPD widths were regressed on each other and plotted. A 1:1 line was also plotted for visual comparison. To statistically test for deviations from a 1:1 line (indicating that one analysis had a higher or lower amount of uncertainty in node dates than the other), the expected 1:1 relationship was subtracted and the regression was repeated; the resulting *p*-value on slope (printed on each plot) thus tests for significant deviation from the expected 1:1 relationship.

## 2.5 | Expanded phylogeny of Melainabacteria and Oxyphotobacteria

16S rDNA sequences from members of the Cyanobacteria phylum (including the Oxyphotobacteria, Melainabacteria, and ML635J-21 classes) greater than 1,250 base pairs in length were acquired from the Silva database (Quast et al., 2013). Alignments were performed using the Infernal aligner (Nawrocki, Kolbe, & Eddy, 2009) implemented by the Ribosomal Database Project (Cole et al., 2014). Multiple phyla (Chloroflexi, OP11, Armatimonadetes, BD1-5, OD1, SHA-109, SR1, TM7, WWE3, and WS6) were used as out groups for the 16s rDNA phylogeny. Phylogenetic analyses were performed at the CIPRES Science Gateway using RAxML under the GTR model and default parameters for nucleotides. Trees were imaged using Interactive Tree of Life software (Letunic & Bork, 2011).

## 2.6 | Phylogenetic analysis of $O_2$ reductases

Sequences for the A- and C-family heme–copper $O_2$ reductases were acquired from public genomic and metagenomic databases and aligned using MAFFT. The resulting alignments were manually curated using structural and biochemical data. Phylogenetic analyses were performed at the CIPRES Science Gateway using RAxML (Stamatakis, 2014) under the LG model. The A and C-family $O_2$ reductases were

each midpoint rooted due to uncertainty in the true position of the roots for these related protein families. Trees were imaged using Interactive Tree of Life software (Letunic & Bork, 2011).

## 3 | RESULTS AND DISCUSSION

### 3.1 | Crown group Oxyphotobacteria postdate the rise of oxygen

We have previously shown that cross-calibrated Bayesian relaxed molecular clock analyses improve estimates of timing of evolutionary events and can be useful to date Precambrian divergences (Shih & Matzke, 2013). Here, we built upon this technique by constructing and analyzing a dataset of slowly evolving proteins and genes found in both plastids and mitochondria, which permit fossil calibrations to be used multiple times across a phylogenetic tree. This included proteins from ATP synthase, the ribosomal large subunit, the ribosomal small subunit, and elongation factor Tu, as well as 16S rDNA sequences from Oxyphotobacteria, Melainabacteria, plastids, α-Proteobacteria, and plant mitochondria (Tables S1–S3).

By expanding the cross-calibrated analysis to include Melainabacteria, we estimate the divergence of crown group Oxyphotobacteria at ca. 2.0 Ga, postdating the rise of oxygen by at least 300 million years (Tables 3 and S2). A useful way to test the robustness of this estimate is by examining the results under different combinations of age calibration constraints. Permutations of the two deepest Proterozoic calibrations—the "Rise of Oxygen" constraint and the *Bangiomorpha* fossil constraint—all show a Paleoproterozoic radiation of crown group Oxyphotobacteria, ranging in time between 2.071 and 1.741 Ga (Table 3, Figs 2 and S1–S4). This result contrasts with previous molecular clock studies that explicitly constrained the most recent common ancestor of all Oxyphotobacteria to the rise of oxygen, under the assumption that the radiation of the crown group was responsible for the rise of oxygen (Battistuzzi et al., 2004; Falcon et al., 2010; Sánchez-Baracaldo et al., 2014; Schirrmeister et al., 2013). However, this finding is in agreement with the fossil record, where the first widely accepted fossil Cyanobacteria (*Eoentophysalis* sp.) occur in shallow marine carbonate strata at ca. 1.9 Ga (Hofmann, 1976). In contrast, when the "Rise of Oxygen" constraint was placed on the radiation of crown Oxyphotobacteria, the Melainabacteria clade instead

**TABLE 3** Age estimates for key evolutionary divergences within the oxygenic Cyanobacteria using cross-calibrated methods

| Run | "Rise of Oxygen" constraint | Bangiomorpha constraint | Melainabacteria/Oxyphotobacteria split | Crown Oxyphotobacteria |
|-----|------------------------------|--------------------------|------------------------------------------|------------------------|
| T64 | Yes | Yes | 2,630 (2,400–2,930) | 2,071 (1,805–2,393) |
| T65 | Yes | No | 2,536 (2,400–2,853) | 1,909 (1,556–2,254) |
| T68 | No | Yes | 2,541 (2,090–3,066) | 2,024 (1,723–2,374) |
| T69 | No | No | 2,238 (1,750–2,790) | 1,741 (1,361–2,161) |

Despite substantially different combinations of deep time constraints ("Rise of Oxygen" and "*Bangiomorpha*" constraints), all BEAST runs robustly estimate a Neoarchean divergence between oxygenic Cyanobacteria and Oxyphotobacteria, as well as a Paleoproterozoic radiation of crown group Oxyphotobacteria. Each column identifies the age constraints that were used in different model runs in the study. Dates are in units of millions of years, and parentheses denote the limits of the 95% highest posterior density.
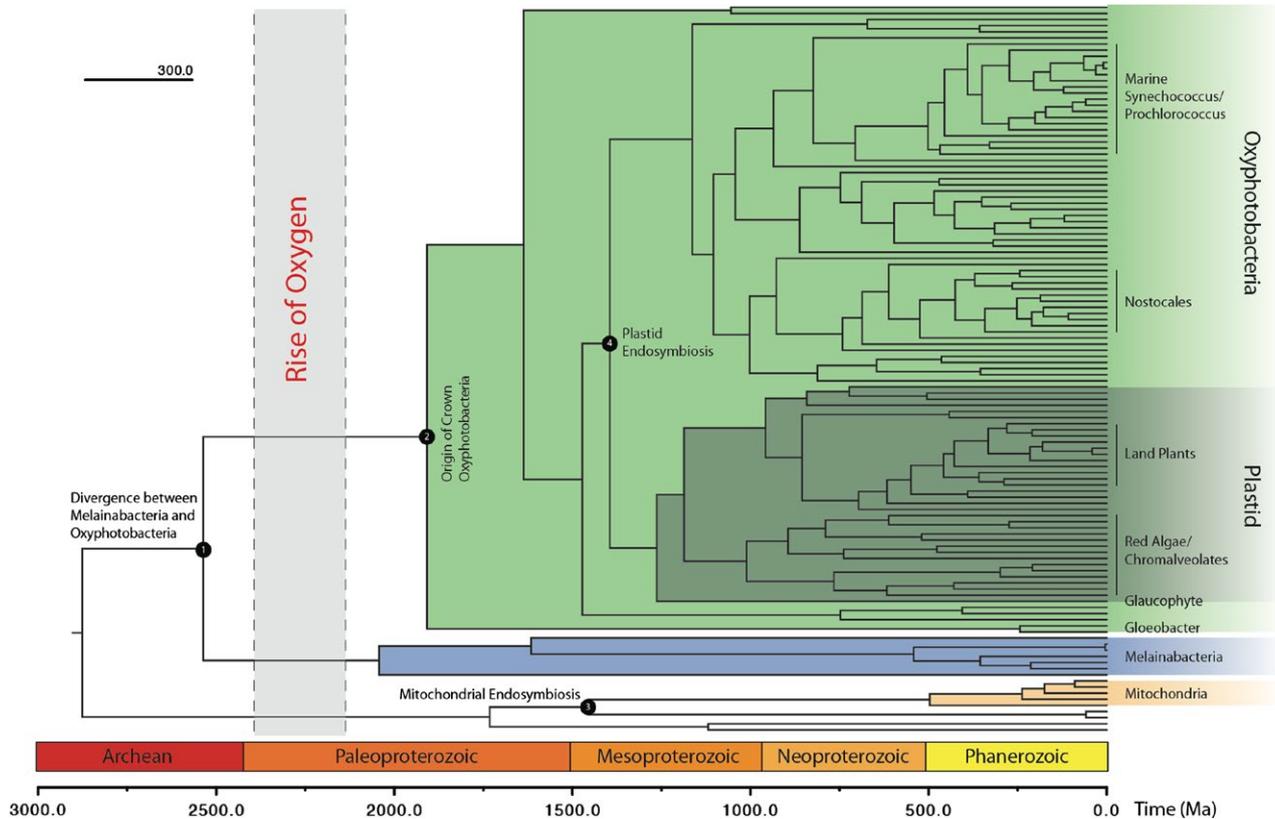
**FIGURE 2** Divergence time estimates for the Cyanobacterial phylum. Dated phylogeny generated from cross-calibrated Bayesian analysis of a concatenated dataset (Run T65). The alignment is composed of conserved proteins found in plastids, mitochondria, and bacteria as well as their 16S rDNA sequences. All analyses illustrate that crown group Oxyphotobacteria postdate the rise of oxygen, which in turn reflects $O_2$ sourced from stem group lineages. Error bars on ages of nodes are shown and summarized in Fig. S2

became sister to the α-Proteobacteria/mitochondria clade—a result in conflict with all phylogenetic studies to date (Di Rienzi et al., 2013; Ley et al., 2005; Soo et al., 2014), which suggested that this calibration set poorly fit the data (Figs S5,S6). Moreover, comparisons of node age uncertainty between BEAST analyses supported placement of the "Rise of Oxygen" constraint on the divergence of Melainabacteria and Oxyphotobacteria over the radiation of crown Oxyphototbacteria (Fig. 3).

## 3.2 | Dating the divergence between Oxyphotobacteria and Melainabacteria

Next, we examined the hypothesis that Oxyphotobacteria and Melainabacteria diverged prior to the rise of oxygen. With a sister group relationship from current genomic data, there is a degree of ambiguity regarding metabolic reconstruction of the last common ancestor of these clades. Two reasonable evolutionary scenarios exist (Fig. 4). (i) The common ancestor to both Oxyphotobacteria and Melainabacteria was aphototrophic and oxygenic photosynthesis evolved after their divergence in the Oxyphotobacteria lineage. If correct, the divergence of these clades would mark an upper age limit on the evolution of oxygenic photosynthesis. (ii) Oxygenic photosynthesis evolved prior to the divergence of Oxyphotobacteria and Melainabacteria and was subsequently lost from the Melainabacteria

clade. The BEAST results can, in principle, help select between these possibilities. For example, if we observe that they diverged after the rise of oxygen, that result would support a loss of photosynthesis from the Melainabacteria (scenario 2).

Our cross-calibrated analyses estimate the divergence of Oxyphotobacteria and Melainabacteria to have occurred ca. 2.5–2.6 Ga (Table 3). This result is consistent with the hypothesis that stem group Oxyphotobacteria evolved oxygenic photosynthesis after their divergence from the Melainabacteria, relatively close in time to the rise of oxygen (Fischer, 2008; Shih, 2015; Ward et al., 2015). This is supported by the observation that the Oxyphotobacteria are nested within a diverse clade of organisms from aphotic environments (Fig. 1). Furthermore, no characterized Melainabacteria genomes encode genes involved in photosynthesis, whereas oxygenic photosynthesis is conserved within all extant Oxyphotobacteria except for a few clear cases of losses in obligate symbionts (Nakayama et al., 2014; Tripp et al., 2010). Additional support for scenario 1 is provided by the contrasting evolutionary histories of other aspects of high potential metabolism (e.g., aerobic respiration) between Oxyphotobacteria and Melainabacteria. As mentioned above, no Melainabacteria contain genes for photosynthesis, but all extant Oxyphotobacteria and some Melainabacteria are able to perform oxidative phosphorylation (Soo et al., 2014), leading to the question of when this ability was acquired within the Cyanobacteria phylum. The $F_0F_1$-ATP synthase
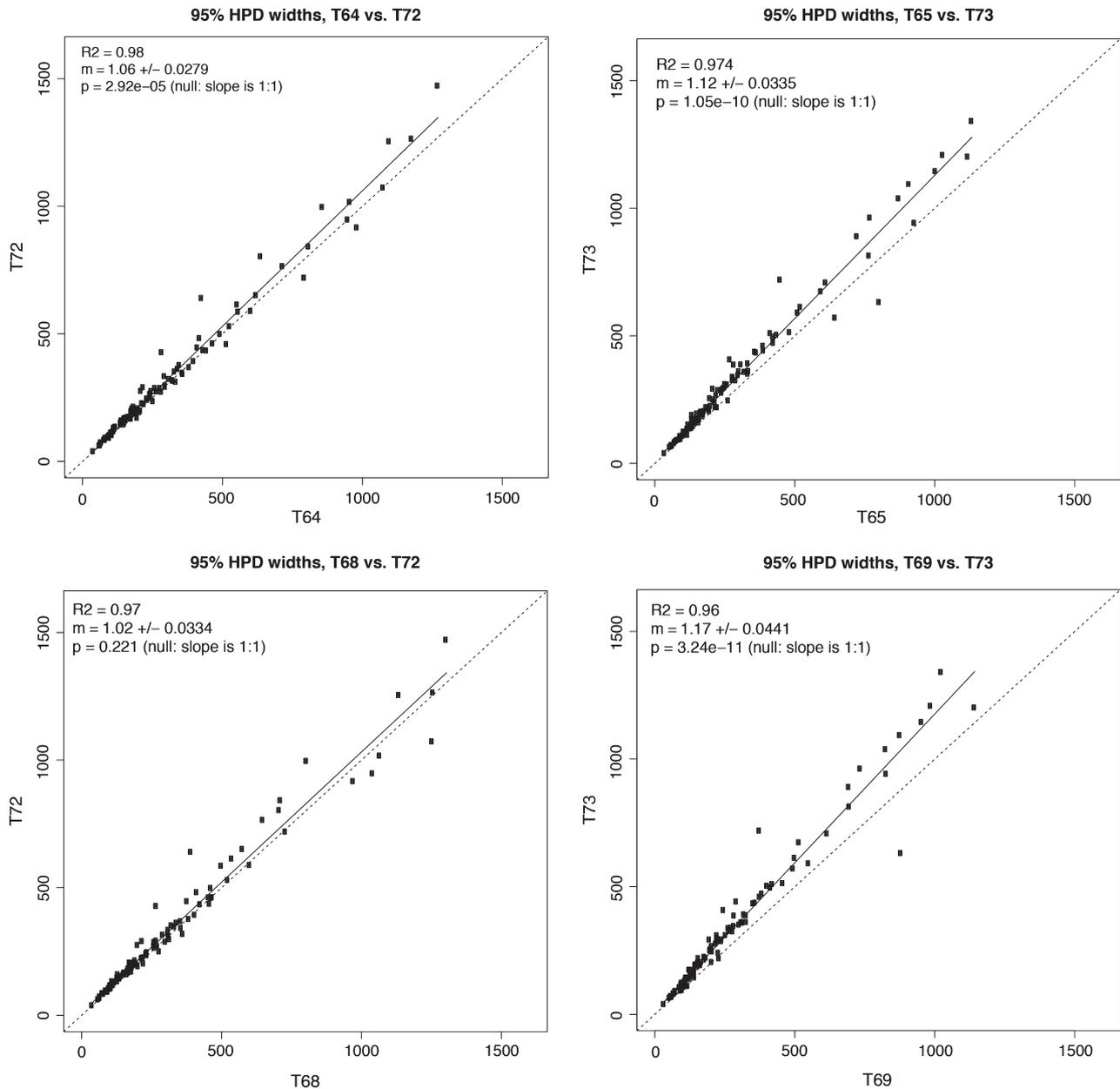
**FIGURE 3** Higher observed age uncertainty when the "Rise of Oxygen" constraint is placed on the radiation of crown group Oxyphotobacteria. Each plot is a comparison of two BEAST analyses, where the width of the 95% HPD represents the amount of dating uncertainty. Each dot represents a corresponding node-date estimate in both trees. All plots are in comparison with one of the two analyses that place the "Rise of Oxygen" constraint on the radiation of crown group Oxyphotobacteria (T72 and T73). In all plots, the observed trend shows that there are higher levels of uncertainty for the T72 and T73 analyses, suggesting that the placement of this constraint is in poorer agreement with the dataset than the combinations of constraints placed with analyses T64, T65, T68, and T69. *p*-Values were calculated using a 1:1 slope as the null hypothesis, assuming that there is no difference in age uncertainty between both analyses

is the only phylogenetically congruent protein shared between the Oxyphotobacteria and Melainabacteria used for oxidative phosphorylation. Phylogenetic analyses of the other respiratory proteins described below illustrate that they were acquired after the divergence of the two clades, suggesting that aerobic respiration and high potential metabolism were acquired independently in the Oxyphotobacteria and Melainabacteria.

For aerobic respiration, Oxyphotobacteria use evolutionarily conserved A-family $O_2$ reductases; these proteins display a characteristically low

affinity for $O_2$ and imply that Oxyphotobacteria acquired their A-family $O_2$ reductases after the evolution of oxygenic photosynthesis and oxygenation of their environment. In contrast, some Melainabacteria in the *Obscuribacterales*, *Vampirovibrionales*, and *Caenarcaniphilales* orders are capable of aerobic respiration using C-family $O_2$ reductases. Notably, the C-family $O_2$ reductases are only very distantly related to the A-family $O_2$ reductases from Oxyphotobacteria (Pereira, Santana, & Teixeira, 2001) and occur within an operon with genes for complex III. This suggests that the Melainabacteria capable of aerobic
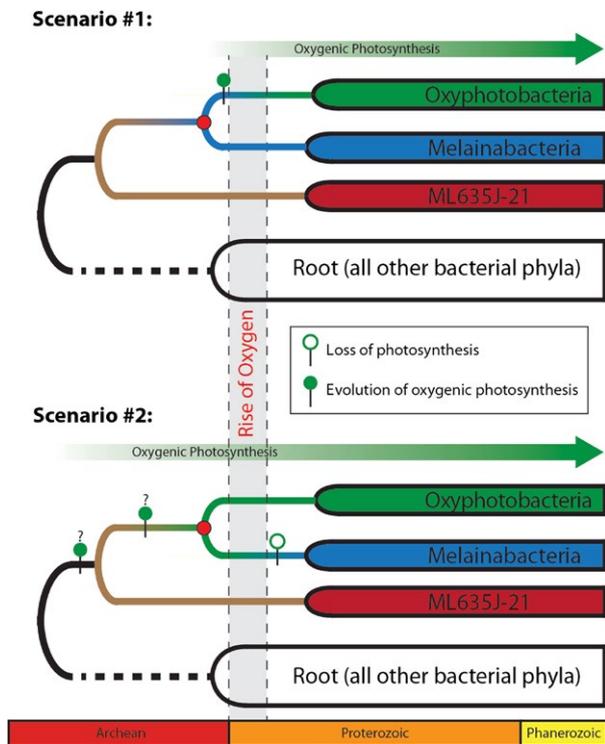
**FIGURE 4** Schematic of two possible scenarios for the timing of oxygenic photosynthesis in relation to the divergence of Oxyphotobacteria and Melainabacteria. Scenario 1. Stem group Oxyphotobacteria diverge from their common ancestor with Melainabacteria (red circle) and evolve oxygenic photosynthesis later. This is the most parsimonious scenario given current data. In this scenario, the divergence of Oxyphotobacteria and Melainabacteria forms a maximum age constraint for oxygenic photosynthesis ca. 2.5–2.6 Ga. and is consistent with a wide range of geological and geochemical data for the rise of oxygen at ca. 2.35 Ga. Scenario 2. Oxygenic photosynthesis evolved prior to the divergence of the Oxyphotobacteria and Melainabacteria, with subsequent loss(es) in the Melainabacteria. Genomic data from members of ML635J-21 clades will help resolve this ambiguity

respiration acquired this ability by lateral gene transfer after the evolution of oxygenic phototrophy (Figs S7,S8). Thus, aerobic respiration was acquired independently at least twice within this phylum (once each in the Oxyphotobacteria and Melainabacteria). Oxyphotobacteria and Melainabacteria also utilize two substantially different complex IIIs for aerobic respiration. The Oxyphotobacteria use a $b_6f$ complex that functions for both phototrophy and aerobic respiration. The $b_6f$ complex has a split cytochrome $b$, an additional cytochrome $c_i$ near the $Q_i$ plastoquinol binding site, and a novel cytochrome $f$. The aerobic Melainabacteria instead have a complex III with a full-length cytochrome $b$ and no cytochrome $c$-containing subunit (Soo et al., 2014). Thus, current genomic data supports a scenario in which aerobic respiration evolved after the origin of oxygenic photosynthesis in this phylum. Again, this is more consistent with scenario 1: Oxygenic photosynthesis evolved after the divergence of Oxyphotobacteria from Melainabacteria, with the independent acquisition of aerobic respiration within these groups after the rise of oxygen. Although several lines of evidence favor scenario 1 over scenario 2, a formal test

of the hypothesis described in scenario 2 will require genomic data from yet more basal members of this bacterial phylum (e.g., members of ML635J-21).

Compared with geological data, a Neoarchean 2.5–2.6 Ga divergence result for Oxyphotobacteria from their closest living relatives is also consistent with observations of a transitional photosystem using Mn prior to water splitting at 2.415 Ga (Johnson et al., 2013b), as well as a number of studies suggesting small photosynthetic fluxes of $O_2$ in Neoarchean environments prior to the rise of oxygen (Anbar et al., 2007; Godfrey & Falkowski, 2009). On the other hand, this result is discordant with hypotheses that stretch Oxyphotobacteria deep into Mesoarchean and older intervals (Crowe et al., 2013; Planavsky et al., 2014; Rosing & Frei, 2004), and raises the possibility that if those geochemical data are interpreted correctly, they may reflect non-cyanobacterial sources of $O_2$ or other oxidants (Ettwig et al., 2010; Fischer et al., 2016; Liang, Hartman, Kopp, Kirschvink, & Yung, 2006).

## 3.3 | Evolution of photosynthesis in Oxyphotobacteria

Oxygenic photosynthesis would eventually become the core engine of the carbon cycle, but this need not have occurred synchronously with the evolution of water splitting. All analyses show several hundred million years between the divergence of Melainabacteria and Oxyphotobacteria, and the radiation of crown group Oxyphotobacteria (Fig. 2). This length of time is important because a large number of evolutionary characters (including a number of multi-subunit complexes) are associated with extant oxygenic photosynthetic organisms (and missing from Melainabacteria), including photosystems I and II, the $b_6f$ complex, RuBisCO and the Calvin cycle, and aerobic respiration—all are synapomorphies of Oxyphotobacteria (Mulkidjanian et al., 2006). It is unlikely that these traits all evolved at the same time, and comparisons of genomic data from Oxyphotobacteria, Melainabacteria, and additional basal clades can help ordinate the relative timing of the appearance of these characters.

The new dates for the divergence of Oxyphotobacteria from Melainabacteria, and the crown group radiation of Oxyphotobacteria, also provide a test of certain models for the evolution of phototrophy. As mentioned above, although all phototrophy is built around the same common molecular machinery (reaction centers, chlorins, and complex III or alternative complex III), it is also clear that the six phyla that are known to contain phototrophic members today are not closely related to one another; and consequently, lateral gene transfer was an important evolutionary vector for the sparse and scattered distribution of phototrophy observed today (Fischer et al., 2016; Raymond, Zhaxybayeva, Gogarten, Gerdes, & Blankenship, 2002). In one class of hypotheses for origin of phototrophy—termed the fusion hypothesis—the type I and type II reaction centers evolved in different lineages (none of which must remain extant), and that ancestral Oxyphotobacteria then ultimately acquired both types of reaction centers by lateral gene transfer (Fischer et al., 2016; Hohmann-Marriott & Blankenship, 2011). If it is correct that oxygenic photosynthesis is derived within the Cyanobacteria phylum, then our estimates

of 2.5–2.6 Ga for the divergence between Oxyphotobacteria and Melainabacteria support the fusion hypothesis, because geological observations highlight that anoxygenic phototrophy was present at ca. 3.4 Ga (Tice & Lowe, 2004).

## 4 | CONCLUSIONS

We presented the first molecular clock estimates for key divergences in the Cyanobacteria phylum that include data from close-living non-phototrophic relatives, but do not employ Archean lipid biomarkers, putative Cyanobacteria microfossils, or the rise of oxygen as calibration constraints. The results from cross-calibrated molecular analyses that include genomic data from the newly discovered Melainabacteria suggest that all known oxygenic Cyanobacteria did not appear until relatively late in Earth history (Fischer et al., 2016; Ward et al., 2015). Estimates of the timing of divergence between Oxyphotobacteria and Melainabacteria are ~2.5–2.6 Ga. If it can be uniquely determined that oxygenic photosynthesis is a synapomorphy of the Oxyphotobacteria—and not simply loss or losses of both phototrophy and respiration from the Melainabacteria—this constraint would correspond to a maximum age of oxygenic photosynthesis (Fig. 4, scenario 1). Current data best support this interpretation, but a formal test of this hypothesis awaits more genomic data from basal members of this phylum.

The results also consistently highlight that crown group oxygenic Cyanobacteria substantially postdate the rise of oxygen—which might instead be attributed to $O_2$ fluxes provided by organisms belonging to stem lineages. The results also consistently highlight that crown group Oxyphotobacteria (and thus all currently known Cyanobacteria capable of oxygenic photosynthesis) substantially postdate the rise of oxygen—which might instead be attributed to $O_2$ fluxes provided by organisms belonging to stem lineages. Many previous studies have placed this crown group divergence on or prior to the rise of oxygen, but based on our results that show this assumption is not well supported; we suggest that it may be useful for future molecular clock studies to examine analyses with and without it. Further comparative genomic analyses reveal that aerobic respiration was acquired at least twice in the Cyanobacteria phylum after the evolution of oxygenic phototrophy.

Ultimately, our understanding of the major metabolic innovations that underpin Earth's biogeochemical cycles is limited to interpretations of the geochemical and fossil data in the context of frameworks built by knowledge of extant biology. This study highlights the opportunity to leverage the exponentially growing amount of genomic and metagenomic data in efforts to update and improve the quality of these frameworks.

## REFERENCES

Abascal, F., Zardoya, R., & Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*, *21*, 2104–2105.

Anbar, A. D., Duan, Y., Lyons, T. W., Arnold, G. L., Kendall, B., Creaser, R. A., … Buick, R. (2007). A whiff of oxygen before the great oxidation event? *Science*, *317*, 1903–1906.

Battistuzzi, F., Feijao, A., & Hedges, S. B. (2004). A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evolutionary Biology*, *4*, 44.

Battistuzzi, F. U., & Hedges, S. B. (2009). A major clade of prokaryotes with ancient adaptations to life on land. *Molecular Biology and Evolution*, *26*, 335–343.

Brasier, M. D., Green, O. R., Jephcoat, A. P., Kleppe, A. K., Van Kranendonk, M. J., Lindsay, J. F., … Grassineau, N. V. (2002). Questioning the evidence for Earth's oldest fossils. *Nature*, *416*, 76–81.

Brocks, J. J., Logan, G. A., Buick, R., & Summons, R. E. (1999). Archean molecular fossils and the early rise of eukaryotes. *Science*, *285*, 1033–1036.

Butterfield, N. J. (2015). Proterozoic photosynthesis – a critical review. *Palaeontology*, *58*, 953–972.

Butterfield, N. J., Knoll, A. H., & Swett, K. (1988). Exceptional preservation of fossils in an Upper Proterozoic shale. *Nature*, *334*, 424–427.

Butterfield, N. J., Knoll, A. H., & Swett, K. (1990). A bangiophyte red alga from the Proterozoic of arctic Canada. *Science*, *250*, 104–107.

Cohen, B. A., Swindle, T. D., & Kring, D. A. (2000). Support for the lunar cataclysm hypothesis from lunar meteorite impact melt ages. *Science*, *290*, 1754–1756.

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., … Tiedje, J. M. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, *42*, D633–D642.

Crowe, S. A., Dossing, 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.

Daubin, V., Gouy, M., & Perrière, G. (2002). A phylogenomic approach to bacterial phylogeny: evidence of a core of genes sharing a common history. *Genome Research*, *12*, 1080–1090.

Di Rienzi, S. C., Sharon, I., Wrighton, K. C., Koren, O., Hug, L. A., Thomas, B. C., … Ley, R. E. (2013). The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *Elife*, *2*, e01102.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*, 214.

Ettwig, K. F., Butler, M. K., Le Paslier, D., Pelletier, E., Mangenot, S., Kuypers, M. M. M., ... Strous, M. (2010). Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature*, *464*, 543–548.

Falcon, L. I., Magallon, S., & Castillo, A. (2010). Dating the cyanobacterial ancestor of the chloroplast. *ISME Journal*, *4*, 777–783.

Fischer, W. W. (2008). Biogeochemistry: Life before the rise of oxygen. *Nature*, *455*, 1051–1052.

Fischer, W. W., Hemp, J., & Johnson, J. E. (2016). Evolution of oxygenic photosynthesis. *Annual Review of Earth and Planetary Sciences*, *44*.

Fischer, W. W., Hemp, J., & Valentine, J. S. (2016). How did life survive Earth's great oxidation? *Current Opinion in Chemical Biology*, *31*, 166–178.

French, K. L., Hallmann, C., Hope, J. M., Schoon, P. L., Zumberge, J. A., Hoshino, Y., ... Summons, R. E. (2015). Reappraisal of hydrocarbon biomarkers in Archean rocks. *Proceedings of the National Academy of Sciences*, *112*, 5915–5920.

Godfrey, L. V., & Falkowski, P. G. (2009). The cycling and redox state of nitrogen in the Archaean ocean. *Nature Geoscience*, *2*, 725–729.

Hoffman, P. F. (2013). The Great Oxidation and a Siderian snowball Earth: MIF-S based correlation of Paleoproterozoic glacial epochs. *Chemical Geology*, *362*, 143–156.

Hofmann, H. J. (1976). Precambrian microflora, Belcher Islands, Canada: significance and systematics. *Journal of Paleontology*, *50*, 1040–1073.

Hohmann-Marriott, M. F., & Blankenship, R. E. (2011). Evolution of photosynthesis. *Annual Review of Plant Biology*, *62*, 515–548.

Holland, H. D. (2009). Why the atmosphere became oxygenated: A proposal. *Geochimica et Cosmochimica Acta*, *73*, 5241–5255.

Hugenholtz, P., Goebel, B. M., & Pace, N. R. (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Journal of Bacteriology*, *180*, 4765–4774.

Johnson, J. E., Gerpheide, A., Lamb, M. P., & Fischer, W. W. (2014). $O_2$ constraints from Paleoproterozoic detrital pyrite and uraninite. *Geological Society of America Bulletin*, *126*, 813–830.

Johnson, J. E., Webb, S. M., Thomas, K., Ono, S., Kirschvink, J. L., & Fischer, W. W. (2013a). Reply to Jones and Crowe: Correcting mistaken views of sedimentary geology, Mn-oxidation rates, and molecular clocks. *Proceedings of the National Academy of Sciences*, *110*, E4119–E4120.

Johnson, J. E., Webb, S. M., Thomas, K., Ono, S., Kirschvink, J. L., & Fischer, W. W. (2013b). Manganese-oxidizing photosynthesis before the rise of cyanobacteria. *Proceedings of the National Academy of Sciences*, *110*, 11238–11243.

Katoh, K., Kuma, K.-I., Toh, H., & Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, *33*, 511–518.

Knoll, A. H. (2003). The geological consequences of evolution. *Geobiology*, *1*, 3–14.

Knoll, A. H., & Golubic, S. (1992). Proterozoic and Living Cyanobacteria. In M. Schidlowski, S. Golubic, M. Kimberley, D. McKirdy Sr, & P. Trudinger (Eds.), *Early organic evolution* (pp. 450–462). Berlin Heidelberg: Springer.

Kopp, R. E., Kirschvink, J. L., Hilburn, I. A., & Nash, C. Z. (2005). The Paleoproterozoic snowball Earth: A climate disaster triggered by the evolution of oxygenic photosynthesis. *Proceedings of the National Academy of Sciences*, *102*, 11131–11136.

Kump, L. R., & Barley, M. E. (2007). Increased subaerial volcanism and the rise of atmospheric oxygen 2.5 billion years ago. *Nature*, *448*, 1033–1036.

van der Lelie, D., Taghavi, S., McCorkle, S. M., Li, L.-L., Malfatti, S. A., Monteleone, D., ... Tringe, S. G. (2012). The metagenome of an anaerobic microbial community decomposing poplar wood chips. *PLoS ONE*, *7*, e36740.

Letunic, I., & Bork, P. (2011). Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Research*, *39*, W475–W478.

Ley, R. E., Bäckhed, F., Turnbaugh, P., Lozupone, C. A., Knight, R. D., & Gordon, J. I. (2005). Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences*, *102*, 11070–11075.

Liang, M.-C., Hartman, H., Kopp, R. E., Kirschvink, J. L., & Yung, Y. L. (2006). Production of hydrogen peroxide in the atmosphere of a Snowball Earth and the origin of oxygenic photosynthesis. *Proceedings of the National Academy of Sciences*, *103*, 18896–18899.

Luo, G., Ono, S., Beukes, N. J., Wang, D. T., Xie, S., & Summons, R. E. (2016). Rapid oxygenation of Earth's atmosphere 2.33 billion years ago. *Science Advances*, *2*, e1600134.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *In Gateway Computing Environments Workshop (GCE)*, 2010, 1–8.

Mulkidjanian, A. Y., Koonin, E. V., Makarova, K. S., Mekhedov, S. L., Sorokin, A., Wolf, Y. I., ... Galperin, M. Y. (2006). The cyanobacterial genome core and the origin of photosynthesis. *Proceedings of the National Academy of Sciences*, *103*, 13126–13131.

Nakayama, T., Kamikawa, R., Tanifuji, G., Kashiyama, Y., Ohkouchi, N., Archibald, J. M., & Inagaki, Y. (2014). Complete genome of a nonphotosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. *Proceedings of the National Academy of Sciences*, *111*, 11407–11412.

Nawrocki, E. P., Kolbe, D. L., & Eddy, S. R. (2009). Infernal 1.0: inference of RNA alignments. *Bioinformatics*, *25*, 1335–1337.

Pereira, M. M., Santana, M., & Teixeira, M. (2001). A novel scenario for the evolution of haem–copper oxygen reductases. *Biochimica et Biophysica Acta*, *1505*, 185–208.

Planavsky, N. J., Asael, D., Hofmann, A., Reinhard, C. T., Lalonde, S. V., Knudsen, A., ... Rouxel, O. J. (2014). Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event. *Nature Geoscience*, *7*, 283–286.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*, D590–D596.

Raymond, J., Zhaxybayeva, O., Gogarten, J. P., Gerdes, S. Y., & Blankenship, R. E. (2002). Whole-genome analysis of photosynthetic prokaryotes. *Science*, *298*, 1616–1620.

Ricci, J. N., Michel, A. J., & Newman, D. K. (2015). Phylogenetic analysis of HpnP reveals the origin of 2-methylhopanoid production in Alphaproteobacteria. *Geobiology*, *13*, 267–277

Rosing, M. T., & Frei, R. (2004). U-rich Archaean seafloor sediments from Greenland—indications of >3700 Ma oxygenic photosynthesis. *Earth and Planetary Science Letters*, *217*, 237–244.

Sánchez-Baracaldo, P., Ridgwell, A., & Raven, J. A. (2014). A neoproterozoic transition in the marine nitrogen cycle. *Current Biology*, *24*, 652–657.

Schirrmeister, B. E., de Vos, J. M., Antonelli, A., & Bagheri, H. C. (2013). Evolution of multicellularity coincided with increased diversification of cyanobacteria and the Great Oxidation Event. *Proceedings of the National Academy of Sciences*, *110*, 1791–1796.

Schopf, J. W., & Packer, B. M. (1987). Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia. *Science*, *237*, 70–73.

Shih, P. M. (2015). Photosynthesis and early Earth. *Current Biology*, *25*, R855–R859.

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

Shih, P. M. (2015). Cyanobacterial evolution: fresh insight into ancient questions. *Current Biology*, *25*, R192–R193.

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., ... Kerfeld, C. A. (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences*, *110*, 1053–1058.

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

Soo, R. M., Skennerton, C. T., Sekiguchi, Y., Imelfort, M., Paech, S. J., Dennis, P. G., … Hugenholtz, P. (2014). An expanded genomic representation of the phylum Cyanobacteria. *Genome Biology and Evolution*, *6*, 1031–1045.

Stamatakis, A. (2014). RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313.

Tice, M. M., & Lowe, D. R. (2004). Photosynthetic microbial mats in the 3,416-Myr-old ocean. *Nature*, *431*, 549–552.

Tomitani, A., Knoll, A. H., Cavanaugh, C. M., & Ohno, T. (2006). The evolutionary diversification of cyanobacteria: Molecular–phylogenetic and paleontological perspectives. *Proceedings of the National Academy of Sciences*, *103*, 5442–5447.

Tripp, H. J., Bench, S. R., Turk, K. A., Foster, R. A., Desany, B. A., Niazi, F., … Zehr, J. P. (2010). Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. *Nature*, *464*, 90–94.

Turner, S., Pryer, K. M., Miao, V. P. W., & Palmer, J. D. (1999). Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *Journal of Eukaryotic Microbiology*, *46*, 327–338.

Ward, L., Kirschvink, J., & Fischer, W. (2015). Timescales of oxygenation following the evolution of oxygenic photosynthesis. *Origins of Life and Evolution of the Biosphere*, *46*, 51–65.

Wolf, Y., Rogozin, I., Grishin, N., Tatusov, R., & Koonin, E. (2001). Genome trees constructed using five different approaches suggest new major bacterial clades. *BMC Evolutionary Biology*, *1*, 8.

Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N. N., … Eisen, J. A. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature*, *462*, 1056–1060.

Yoon, H. S., Hackett, J. D., Ciniglia, C., Pinto, G., & Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. *Molecular Biology and Evolution*, *21*, 809–818.

Zhaxybayeva, O., & Gogarten, J. P. (2002). Bootstrap, Bayesian probability and maximum likelihood mapping: exploring new tools for comparative genome analyses. *BMC Genomics*, *3*, 4.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.