

Granick Revisited: Synthesizing Evolutionary and Ecological Evidence for the Late Origin of Bacteriochlorophyll via Ghost Lineages and Horizontal Gene Transfer

Authors: Lewis M. Ward¹ and Patrick M. Shih^{2,3,4,5}

Affiliations:

1. Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, United States
2. Department of Plant Biology, University of California, Davis, CA, USA.
3. Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.
4. Feedstocks Division, Joint BioEnergy Institute, Emeryville, CA, USA.
5. Genome Center, University of California, Davis, Davis, CA, USA.

Abstract

Photosynthesis—both oxygenic and more ancient anoxygenic forms—has fueled the bulk of primary productivity on Earth since it first evolved more than 3.4 billion years ago. However, the early evolutionary history of photosynthesis has been challenging to interpret due to the sparse, scattered distribution of metabolic pathways associated with photosynthesis, long timescales of evolution, and poor sampling of the true environmental diversity of photosynthetic bacteria. Here, we reconsider longstanding hypotheses for the evolutionary history of phototrophy by leveraging recent advances in metagenomic sequencing and phylogenetics to analyze relationships among phototrophic organisms and components of their photosynthesis pathways, including reaction centers and individual proteins and complexes involved in the multi-step synthesis of (bacterio)-chlorophyll pigments. We demonstrate that components of the photosynthetic apparatus have undergone extensive, independent histories of horizontal gene transfer. This suggests an evolutionary mode by which modular components of phototrophy are exchanged between diverse taxa in a piecemeal process that has led to biochemical innovation. We hypothesize that the evolution of extant anoxygenic photosynthetic bacteria has been spurred by ecological competition and restricted niches following the evolution of oxygenic Cyanobacteria and the accumulation of O₂ in the atmosphere, leading to the relatively late evolution of bacteriochlorophyll pigments and the radiation of diverse crown group anoxygenic phototrophs. This hypothesis expands on the classic “Granick hypothesis” for the stepwise evolution of biochemical pathways, synthesizing recent expansion in our understanding of the diversity of phototrophic organisms as well as their evolving ecological context through Earth history.

Intro:

Earth’s biosphere today is incredibly productive, with >99% of the organic carbon fixed per year fueled by a single metabolism—oxygenic photosynthesis (Ward and Shih 2019, Raven 2009). This metabolism is an evolutionary singularity, having evolved via serial coupling of two photosystems in ancestors of crown group oxygenic Cyanobacteria, and is uniquely capable of using sunlight to power water splitting, yielding O₂ as a byproduct as well as providing electrons for carbon fixation (Fischer et al. 2016). Oxygenic photosynthesis evolved from more ancient forms of anoxygenic

photosynthesis, which supported the biosphere early in its history (Kharecha et al. 2005, Canfield et al. 2006, Ward et al. 2019a, Ozaki et al. 2019, Hao et al. 2020). However, the origin and early evolution of anoxygenic photosynthesis is not well understood due to the antiquity of these events and the lack of clear signatures in the rock record (Fischer et al. 2016, Ward and Shih 2019). The earliest widely accepted evidence for anoxygenic photosynthesis in the rock record is in the form of depth-dependent organic carbon production in preserved microbial mats from ~3.4 Ga (Tice and Lowe 2004, Tice and Lowe 2006), and it is clear that both forms of reaction center had evolved in time to be brought back together in stem group Cyanobacteria to invent oxygenic photosynthesis and trigger the Great Oxygenation Event ~2.3 Ga (Shih et al. 2017a). However, this leaves over a billion years in which photosynthesis was likely present on Earth and actively evolving, and it is unclear how long ago the reaction centers diverged or how long the stem lineage persisted. While molecular clocks and other studies of evolution in deep time often assume that crown group synapomorphies are acquired either at the base of the crown group (e.g. Battistuzzi et al. 2004, Schirmer et al. 2013) or the base of the total group (e.g. Magnabosco et al. 2018), there is no *a priori* way of determining where along a stem lineage these traits were acquired. It is therefore essential to recognize uncertainty that is inherited with long stem lineages and to make use of all available basal lineages and sister groups in phylogenetic analyses to break up long branches and reduce uncertainty in timing of acquisition of important traits (e.g. Shih et al. 2017b).

While simple forms of photoheterotrophy can be supported by ion-pumping rhodopsins, there are no known examples of this metabolism driving carbon fixation (Larkum et al. 2018); true photosynthesis is only known to be supported by a more complicated pathway involving multiple components including an electron transport chain (including Complex III or Alternative Complex III), phototrophic reaction centers or photosystems, synthesis and modification of chlorophyllide pigments for light harvesting, and optionally carbon fixation to enable photoautotrophy. The capacity for reaction center-based phototrophy is scattered across the tree of Bacteria, with a handful of phototrophic lineages separated by many nonphototrophic groups (Figure 1). While early hypotheses for the evolutionary history of photosynthesis invoked vertical inheritance and extensive loss in most lineages (e.g. Woese 1987), more recent work is more consistent with this distribution being driven by horizontal gene transfer (HGT) of phototrophy (Zeng et al. 2014, Ward et al. 2018a, Brinkmann et al. 2018, Ward et al. 2019b).

While previous attempts to understand the history of HGT of phototrophy have investigated relatively recent transfer events of complete phototrophy pathways (Igarashi et al. 2001, Nagashima and Nagashima 2013, Zeng et al. 2014, Brinkmann et al. 2018, Ward et al. 2018a, Ward et al. 2019b), it is possible that individual components of the pathway may be transferred independently of one another, with modern phototrophs encoding chimeric pathways with components acquired from multiple sources via multiple HGT events. An example of this is indications that the (bacterio)chlorophyll synthesis pathway may have a distinct evolutionary history from the phototrophic reaction center, such as HGT of (bacterio)chlorophyll synthesis in ancestors of phototrophic Chloroflexi, Chlorobi, and WPS2 (Eremiobacterota) (Bryant et al. 2012, Sousa et al. 2013, Ward et al. 2019b).

The extant distribution of chlorophyll versus bacteriochlorophyll synthesis is an additional puzzle. Bacteriochlorophyll is biochemically more complex to synthesize and uses lower quality light (Chew and Bryant 2007) yet is found today in bacteria using more ancient anoxygenic phototrophy (Table 1). Oxygenic Cyanobacteria, in contrast, use chlorophyll, which harvests higher quality light and is biochemically more straightforward to synthesize. This has led to hypotheses for the evolution of (bacterio)chlorophylls ranging from the stepwise evolution of progressively more complex pathways, with chlorophyll being more evolutionarily ancient than bacteriochlorophyll (i.e. the Granick hypothesis, Granick 1965), or hypotheses invoking the complete bacteriochlorophyll synthesis pathway in the last common ancestor of extant phototrophs followed by secondary simplification of the pathway in Cyanobacteria (Xiong et al. 2000). The synthesis of bacteriochlorophyll and chlorophyll both proceeds through chlorophyllide a, and so the two pathways share a “backbone” of shared steps, while bacteriochlorophyll synthesis has additional later steps, including those performed by the BchXYZ complex, which is homologous to the BchLNB complex involved in the synthesis of chlorophyllide a from protochlorophyllide a (Supplemental Table 1, Supplemental Figure 1). This homology provides an opportunity for querying the relative evolutionary histories of bacteriochlorophyll and chlorophyll: if the two complexes have congruent phylogenies with the exception of the absence of BchXYZ in Cyanobacteria, this would argue for a secondary simplification of the pathway as suggested by Xiong et al (2000). However, incongruent phylogenies would indicate an independent history of HGT, such as might be expected if bacteriochlorophyll synthesis was secondarily acquired in anoxygenic phototrophs relatively later in their evolutionary history after the divergence of crown group lineages.

Here, we perform phylogenetic analyses of various components of the phototrophy pathway, including individual steps in (bacterio)chlorophyll synthesis, and demonstrate that phylogenetic trees of these components have incongruent topologies, reflecting independent histories of horizontal gene transfer. This suggests a mix-and-match style of modular evolution by which bacteria acquire separate components of phototrophy from separate sources at different times in their evolutionary histories. Moreover, we provide support for an early diverging “ghost lineage” of extinct or undiscovered phototrophs responsible for evolutionary divergence of RC2 and BchXYZ. The genes encoding these complexes were later transferred into crown group phototrophs, leading to evolutionary novelty such as coupled photosystems and bacteriochlorophyll synthesis. We then propose an ecological model for the evolution of phototrophy in deep time, whereby competition with Cyanobacteria and environmental partitioning by O₂ tolerance has led to innovation, diversification, and specialization by crown group anoxygenic phototrophs, leading to diverse bacteriochlorophyll-synthesizing anoxygenic phototrophs today that radiated after the GOE, supplanting ecologically and genetically distinct chlorophyll-synthesizing anoxygenic phototrophs which fueled primary productivity in Archean time.

Results and Discussion:

Phylogenetic analyses of reaction center and (bacterio)chlorophyll synthesis proteins

We have constructed phylogenies of essential proteins for phototrophy including reaction centers and (bacterio)chlorophyll synthesis (listed in Supplemental Table 1) using representatives from all known clades of phototrophic bacteria (Table 1). Consensus trees for phototrophy proteins are shown in Figure 2, with individual phylogenies available as Supplemental Information (Supplemental Figures 2-12).

While these trees have difficulty recovering deep evolutionary relationships due to long evolutionary distances relative to protein lengths, a problem previously recognized for interpreting the evolutionary history of these proteins (e.g. Lockhart et al. 1996, Bryant et al. 2012, Fischer et al. 2016), they robustly recover overall topologies and sister-group relationships that differ between different proteins. Breaking up long branches by adding newly discovered phototrophic taxa (e.g. Chloracidobacteria and Baltobacterales) improves resolution of deep branches relative to previous attempts with more limited datasets (e.g. Xiong et al. 2000). The topologies of phylogenetic trees constructed here are largely incongruent (e.g. organismal tree, RC tree, backbone chlorophyll synthesis, and bacteriochlorophyll synthesis, Figure 2), indicating that components of the phototrophic apparatus have undergone independent horizontal gene transfer events over the course of their evolutionary history (Doolittle 1986).

While branch support within diverse, shallow radiations such as the Cyanobacteria and the Proteobacteria were consistently poor (e.g. Supplemental Figures 3-16), and well-supported roots were not recovered for all proteins investigated (e.g. long branches between BchL/N/B and BchX/Y/Z made root placement inconsistent, Supplemental Figures 13-16, the consistent topology between consensus BchLNB and BchHDIM trees clustered at higher taxonomic levels (i.e. order through phylum) and robust rooting of BchH and BchM via characterized homolog outgroups allows extrapolation of branching order from these trees to the consensus backbone (bacterio)chlorophyll synthesis tree. While uncertainty in early branching order may affect hypotheses of the directionality of HGT invoked in hypotheses for evolutionary history of phototrophy, major sister-group relationships and incongruent topologies between different complexes are robustly recovered, supporting overall evolutionary trends even if the exact number, timing, and directionality of HGT events are uncertain. As a result, interpretations of the role of HGT in driving the evolution of phototrophy are considered robust here, even if the scenario depicted in Figure 3 is only a hypothesis.

Backbone (bacterio)chlorophyll synthesis genes have a somewhat different evolutionary history than downstream bacteriochlorophyll-specific genes, which have a distinct evolutionary history from phototrophic reaction centers, which have independent histories from carbon fixation pathways—all of which are incongruent with organismal phylogenies. Taken all together, the reticulated and piecemeal nature of this incongruencies suggest that the evolution of phototrophy, like many other metabolic pathways, is modular (Chistoserdova 2011, Chen and Strous 2013, Shih et al. 2017b, Fischer et al. 2016, Ward et al. 2018a, Ward et al. 2019c).

Evolutionary history of photosynthesis-associated proteins

The incongruent topologies of trees built with various phototrophy-related proteins (Figure 2) suggests that these proteins may not share a unified evolutionary history but may instead have undergone independent horizontal gene transfer events. By overlaying components of the phototrophy apparatus onto the backbone tree made from

the consensus topology of proteins involved in shared early steps of (bacterio)chlorophyll synthesis, we can reconstruct a hypothetical evolutionary history of the genes, even if a largely unconstrained cooccurring history of organismal transfer must also be occurring throughout stem lineages (Figure 3). For example, the node representing the last common ancestor of (bacterio)chlorophyll synthesis genes in Chlorobi and Chloroflexi almost certainly did not occur in crown group members of either of those phyla or their last common ancestor, but instead independent HGT events introduced phototrophy to those groups following the divergence of their (bacterio)chlorophyll pathways in other host lineages. In some cases, we see transfer of the complete phototrophy pathway (e.g. into Gemmatimonadetes from Proteobacteria, Zeng et al. 2014), while in others it appears that some components were transferred independently of others (e.g. bacteriochlorophyll synthesis and reaction centers into Chlorobi and Chloroflexi, respectively). In many cases, we see transfer of phototrophy but not carbon fixation (e.g. Anaerolineae, Gemmatimonadetes, Ward et al. 2020, Ward et al. 2018a, Zeng et al. 2014) though we also see cases of gain, loss, and reacquisition of carbon fixation (e.g. within the Chloroflexia, Shih et al. 2017b).

Below, we describe a hypothesis for the evolutionary history of phototrophy genes that is consistent with all of the available data, from the “ur-phototroph” (with the prefix “ur-“ denoting the first or proto-version of a thing, here referencing the first phototrophic lineage to evolve, i.e. the first member of stem and total group phototrophs) until the radiation of extant clades of phototrophs (the crown group). The history presented here relies on undiscovered, likely extinct, “ghost lineages” which served as the source of genes to crown group phototrophs during ancient transfer events. The concept of “ghost lineages” is carried over from traditional fossil record-based phylogenetics, where it refers to a lineage which is inferred to exist but which has no known fossil record (Norell et al. 1992, Norell 1993). Importantly, this history is considered from the perspective of phototrophy genes, not of the organisms which harbor them.

The ur-phototroph likely utilized a single ancestral reaction center (almost certainly a homodimer that formed a stem lineage prior to the RC1/RC2 divergence, and which may have been biochemically more similar to heliobacterial RC1 than other extant types, made up of a relatively simple and inefficient homodimer which loosely bound mobile quinones to drive cyclic electron flow, which was adapted in the various reaction center and photosystem lineages to optimize reactions and eventually to adapt to oxygen, Orf et al. 2018) and chlorophyll a, synthesized using a DPOR complex ancestral to both BchLNB and BchXYZ. Eventually, two lineages of phototrophs diverged, either due to speciation of a single organismal lineage or due to HGT of phototrophy genes into a second host organism. One lineage (the “ghost lineage”) possessed a reaction center that evolved into RC2 and a BchLNB-like complex that eventually evolved into BchXYZ but which functioned as a DPOR complex. The other lineage possessed an ancestral RC1 and BchLNB in order to synthesize chlorophyll. The RC1 lineage diversified, with RC1 and chlorophyll synthesis in several lineages. Eventually, HGT of RC2 from the ghost lineage into stem group Cyanobacteria led to the evolution of oxygenic photosynthesis. This triggered ecological restructuring and widespread evolutionary adaptation, including further HGT from the ghost lineage into other lineages of anoxygenic phototrophs. This included HGT of the BchXYZ complex from the ghost lineage, perhaps into stem group Proteobacteria, leading to the coupling of BchLNB and BchXYZ in series to lead to

bacteriochlorophyll synthesis and the ability of anoxygenic phototrophs to better compete in deeper, lower oxygen regions of microbial mats and water columns (see below section “Ecological perspectives on the evolution of photosynthesis”). Further HGT of RC2 from the ghost lineage into other anoxygenic phototroph lineages led to further adaptation and specialization of anoxygenic phototroph lineages to specialized environments. The long branches between homologous proteins in the BchLNB and BchXYZ complexes is consistent with their early divergence sometime during Archean time before the radiation of crown group phototrophs; the congruence between BchLNB and other backbone chlorophyll synthesis genes suggests that these genes have largely been inherited together into extant phototrophs, while the differing topology of BchXYZ indicates an independent history of HGT (Supplemental Figure 2), in this scenario driven by HGT of BchXYZ from the ghost lineage into crown group anoxygenic phototrophic clades, allowing them to produce bacteriochlorophyll.

For simplicity one ghost lineage is described here, but it is likely there were many lineages that are not represented in characterized phototrophs today, both diverging from the stem of the phototroph tree as well as from throughout the crown group. For example, the ghost lineage invoked here is hypothesized to use a type 2 reaction center (though this may have been a more ancient homodimeric form, as heterodimerization postdated the divergence of RC2 and PSII, Cardona 2013), but there are also suggestions from the carbon isotope record suggesting ancient Wood-Ljungdahl utilizing Type 1 phototrophs (Ward and Shih 2019), indicating that there are multiple undiscovered or extinct phototroph lineages.

While some phototrophic Chloroflexi in the Anaerolineae class lack BchLNB (Klatt et al. 2011, Ward et al. 2018a, Ward et al. 2020), this appears to be a derived trait based on their placement in BchHDI and BchXYZ trees. The genomes of these organisms are derived from metagenomic data and so are somewhat incomplete, but the probability that the organisms contain these genes but they weren't recovered in the MAG is incredibly low (estimated by MetaPOAP as $\sim 10^{-12}$, Ward et al. 2018b). Instead, this appears to be a case of secondary loss potentially coupled with bifunctionalization of BchXYZ to perform the reduction of both the C17/C18 double bond normally reduced by BchLNB as well as the C7/C8 double bond normally reduced by this complex. This seems feasible, as chimeras of other homologs of these genes have been demonstrated to be functionally exchangeable (e.g., Cheng et al., 2005; Wätzlich et al., 2009). Notably, Cheng et al. demonstrated the enzymatic flexibility of the complex with ChlN functionally replacing the NifH subunit in the more distant homologous nitrogenase complex (NifHDK) (Cheng et al, 2005). Following the evolutionary history proposed in Figure 3, the BchXYZ complex is descended from a BchLNB-like complex for the reduction of the C17/C18 double bond that was later coopted to instead reduce the C7/C8 double bond to enable bacteriochlorophyll synthesis. However, isolation of phototrophic Anaerolineae and biochemical characterization of their BchXYZ complexes will be necessary to test this hypothesis.

Ecological perspectives on the evolution of photosynthesis

Overlaid onto the history of HGT-driven evolution of phototrophy is also a history of adaptation to changing environmental conditions, particularly the rise of atmospheric oxygen. The earliest evidence for photosynthesis on Earth is found in rocks

of Paleoproterozoic age (Tice and Lowe 2004, Tice and Lowe 2006), but many of the extant groups of anoxygenic phototrophs are thought to represent relatively recent radiations (i.e. post-GOE) (Fischer et al. 2016, Shih et al. 2017b, Ward et al. 2018a, Ward et al. 2019b). Thus, many modern anoxygenic phototrophs are therefore not relicts of the Archean Earth, but instead have undergone billions of years of evolution, leading to extant phototrophs that are a palimpsest of genetic innovation and horizontal gene transfer, with over a billion years of evolution hidden in stem lineages. Straightforward comparative genomic and phylogenetic techniques extrapolate only from extant diversity and therefore may overlook nuance and complexity in the early evolution of pathways before the emergence of the last common ancestors of extant clades. A complementary alternative approach is to develop theories for the ecological drivers of the evolution of phototrophy, which can produce hypotheses which are testable with the biological record, and which can be used to choose between competing hypotheses which are equally supported by the biological record but of which some may be more ecologically viable.

While in principle the innovation and expansion of bacteriochlorophyll synthesis could have occurred before the evolution of oxygenic photosynthesis in Cyanobacteria, we propose that it was instead the expansion of oxygenic phototrophs and the rise of oxygen that likely led to the rampant HGT of the capacity for bacteriochlorophyll synthesis and overall success of this strategy among taxonomically and ecologically diverse anoxygenic phototrophs (Figure 4).

Before the rise of oxygen ~2.3 Gya, surface environments were anoxic (Farquhar et al. 2000, Pavlov and Kasting 2002, Guo et al. 2009, Johnson et al. 2013, Johnson et al. 2014, Gumsley et al. 2017), and electron donors for anoxygenic photosynthesis would have been reasonably abundant (Kharecha et al. 2005, Canfield et al. 2006, Ward et al. 2016, Ward et al. 2019a). As a result, these environments would have been permissive to anoxygenic phototrophs, allowing them to thrive in shallow water environments and as the major primary producers in microbial mats (Ward and Shih 2019). At this time, anoxygenic phototrophs would have been free to exploit abundant, high-quality light conditions to which chlorophyll a is well adapted. The earliest anoxygenic phototrophs therefore would have had little to no evolutionary pressure to evolve more biochemically complex bacteriochlorophyll pigments. Anoxygenic phototrophs may have specialized to particular environments where electron donor compounds or other nutrients were especially abundant (e.g. volcanic environments or near shallow hydrothermal vents), but otherwise would have had little barrier to dispersal and colonization of widespread environments.

By 2.3 Gya, early Cyanobacteria had evolved oxygenic photosynthesis, leading to the accumulation of atmospheric O₂ and a sharp increase in primary productivity associated with this metabolism (Sleep and Bird 2007, Shih et al. 2017a, Crockford et al. 2018, Ward et al. 2019a, Ward and Shih 2019, Hodgskiss et al. 2019, Hao et al. 2020). As Cyanobacteria are quite oxygen tolerant relative to other phototrophs and are not dependent on limited electron donors, these organisms would have been capable of colonizing all shallow aquatic environments, displacing anoxygenic phototrophic communities and leading to a distribution of phototroph types more similar to that seen today. Oxygenic phototrophs are able to grow in well-oxygenated surface environments with largely unobstructed sunlight, and so have had little evolutionary pressure to evolve away from using ancestral chlorophyll pigments. Anoxygenic phototrophs, in contrast,

would have been unable to compete with Cyanobacteria in surface environments where they would experience oxygen toxicity as well as limited availability of electron donors due to biological or abiotic oxidation driven by O₂. As a result, anoxygenic phototrophs would be restricted to environments deeper in water columns or microbial mats where oxygen concentrations are lower and electron donors are more available. As these environments typically underlie cyanobacterial populations (e.g., microbial mats), anoxygenic phototrophs are light-limited in the wavelengths absorbed by cyanobacterial pigments—including chlorophyll (e.g. Pierson et al. 1987). These organisms therefore have undergone significant evolutionary pressure to evolve light-harvesting pigments that are shifted to different wavelengths that reach these deep layers. Bacteriochlorophylls fill this role well: bacteriochlorophyll pigments have peak absorbances at longer wavelengths than chlorophyll, resulting in less energy absorbed per photon but allowing them to work deeper in the water column/microbial mat, once more light has been absorbed (both by the medium and by Cyanobacteria with chlorophylls) (Fenchel and Staarup 1971, Larkum et al. 2018). This scenario is consistent with the ancestral use of chlorophyll in anoxygenic lineages followed by multiple HGT-enabled acquisitions of BchXYZ and therefore bacteriochlorophyll synthesis in anoxygenic lineages after the GOE. Ecological exclusion of anoxygenic phototrophs from high-light surface environments by competition with Cyanobacteria and the distribution of O₂ and electron donors likely therefore provided the evolutionary pressure for the initial evolution of bacteriochlorophyll pigments as well as to drive horizontal gene transfer leading to the widespread adoption of bacteriochlorophyll pigments by diverse anoxygenic phototrophs.

As anoxygenic phototrophic lineages are typically restricted to particular niche environments, island biogeography-like evolutionary radiations have likely led to the diversification of anoxygenic phototrophs seen today, including the diversification of their bacteriochlorophyll pigments. Today, oxygenic phototrophs (Cyanobacteria, algae, and plants) are essentially ubiquitous in habitable Earth surface environments. Cyanobacteria therefore have a largely cosmopolitan distribution, with related species found in diverse environments around the world. However, anoxygenic phototrophs are typically restricted to more limited environments such as stratified water columns (Imhoff 2014), hot springs (e.g. Hallenbeck et al. 2016, Bryant et al. 2007, Klatt et al. 2011, Ward et al. 2017), and geothermal soils (e.g. Kimble et al. 1995) where the presence of sulfide and/or high temperature inhibit Cyanobacteria, though a few lineages of phototrophic Proteobacteria and Eremiobacteria are adapted to more widespread photosynthetic or photoheterotrophic niches in association with plants (e.g. Atamna-Ismaeel et al. 2012, Ward et al. 2019b). These niches are highly localized, discrete environments with specialized geochemical conditions, limiting the dispersal of anoxygenic phototrophs between sites and inhibiting the ability of anoxygenic lineages to colonize new environments.

Further evolution has of course occurred over the billions of years since the rise of Cyanobacteria, including further HGT (e.g. the transfer of many genes, including BchLNB and the alpha-carboxysome from Proteobacteria into *Synechococcus* and *Prochlorococcus*, Bryant et al. 2012), adaptation of anoxygenic phototrophs to their preferred niches (including diversification of bacteriochlorophylls), and the substitution of ancestral oxygen-sensitive enzymes with O₂-tolerant versions, particularly in the Cyanobacteria and Proteobacteria. For example, the innovation of the POR enzyme

complex as an alternative to the oxygen-sensitive BchLNB complex in oxygenic Cyanobacteria. DPOR (the BchLNB complex) is oxygen sensitive, and doesn't work at O₂ concentrations higher than saturation under ~3% present atmospheric levels (Yamazaki et al. 2006), most likely leading to the evolution of POR as an oxygen-tolerant alternative in stem-group Cyanobacteria prior to the radiation of the crown group ~2 Ga (Shih et al. 2017a). POR has also been occasionally acquired by Proteobacteria (Kaschner et al. 2014), reflecting continuing HGT of phototrophy proteins subsequent to the diversification of crown group lineages.

Conclusions:

The origin and early evolution of phototrophy is not clearly revealed by the rock record, and care must be taken in reading the biological record; in particular, the application of comparative biology to investigating the early evolution of traits in stem lineages before the last common ancestor of extant members of a clade is inherently challenging (Ward and Shih 2019). However, large-scale comparative phylogenetics of organismal relationships and phototrophy-related proteins can provide some insight into the early evolution of photosynthesis, especially given that the independent histories of genes involved with phototrophy allow some insight into the nature of stem groups and extinct lineages in instances where genes from these organisms have been horizontally transferred and then inherited into extant organisms. These data are particularly useful once coupled to consideration of the ecological context of the organisms in question, which can help to choose between otherwise equally viable hypotheses. As we've shown here, phylogenetic relationships provide abundant evidence of horizontal gene transfer of phototrophy-related proteins, though the directionality of transfers and the relationships among the deepest branches remain ambiguous in many cases. These relationships are consistent with an early evolution of chlorophyll synthesis followed by a post-GOE radiation of bacteriochlorophyll-synthesizing anoxygenic phototrophs driven by ecological competition and niche partitioning by Cyanobacteria.

The evolution of photosynthesis is modular, involving not only independent HGT of carbon fixation and Reaction Centers but also of separate components of (bacterio)chlorophyll synthesis. The relationships among characterized extant phototrophs suggests one or more "ghost lineages", some of which likely diverged prior to the radiation of the crown group. Chlorophyll synthesis appears to be more ancient than bacteriochlorophyll synthesis, consistent with the Granick hypothesis (Granick 1965). This also results in an evolutionary scenario for the origin of bacteriochlorophyll synthesis analogous to the origin of coupled photosystems for oxygenic photosynthesis, whereby divergence of paralogs followed by reintroduction into a single host organism allows biochemical innovation (e.g. Swithers et al. 2012). The early divergence between RC1 and RC2, and between BchLNB and BchXYZ, suggested by this history is consistent with the long branches between these sets of homologs relative to divergence within each.

The modular nature of the evolution of phototrophy is similar to broader trends in the evolution of pathways including denitrification (Chen and Strous 2013), methanotrophy (Chistoserdova 2011, Ward et al 2019c), and high-potential metabolism in general (Fischer et al. 2016). The ability of microbes to exchange components of

preexisting metabolisms and recombine them into new and innovative pathways appears to be a major driver of innovation.

The innovation of bacteriochlorophyll by coupling in series of the orthologous BchLNB/BchXYZ complexes is analogous to the innovation of coupling divergent photosystems to drive oxygenic photosynthesis. Although the early evolution of phototrophy involved extensive shuffling of phylogenetic relationships via HGT and a role for extinct stem lineages, we can understand this history by careful comparative phylogenetics as long as we have sufficient sampling of extant diversity. Coupling this understanding to analysis of the rock record, potentially supported by molecular clock analysis, may allow us to tie absolute ages to these phylogenies casting events in relative time. For example, the organic biomarker and molecular clock evidence for the evolution of phototrophy in the Chlorobi by 1.6 Ga (Brocks et al. 2005), molecular clock estimates for the origin of phototrophic Chloroflexia around 1 Ga (Shih et al. 2017b), and the rise of atmospheric oxygen due to total group oxygenic Cyanobacteria around 2.3 Ga can provide some constraints. The fact that each of the clades of anoxygenic phototrophs appears to have acquired bacteriochlorophyll synthesis in stem lineages before the radiation of individual crown groups indicates that the radiations of extant anoxygenic phototroph clades occurred relatively late in Earth history after the origin of oxygenic photosynthesis in Cyanobacteria. This is consistent with estimates for increasing diversity of bacterial lineages through geologic time (Louca et al. 2018). The taxonomic affinity and phenotypes of ancient phototrophs responsible for fueling Archean productivity remain unknown, perhaps due to the extinction of these bacterial lineage following the rise of oxygen.

The ecological partitioning of anoxygenic phototrophs into localized environments with lower oxygen concentrations and lower light energy may have led to an island biogeography-like adaptive diversification of different phototrophic lineages to distinct environments. Discrete anoxygenic phototrophic lineages have adapted to particular preferred environments, such as hot springs for Chloroflexi and Chloracidobacteria (Bryant et al. 2007, Ward et al. 2018a, Ward et al. 2017), thermal soils for Heliobacteria (Madigan and Ormerod 1995), stratified water columns for Chlorobi, and acidic plant-associated niches for Baltobacterales (Ward et al. 2019b) (though exceptions do occur, such as Chloroflexi in carbonate tidal flats, Trembath-Reichert et al. 2016, Ward et al. 2020, and Heliobacteria in hot springs, Kimble et al. 1995). Adaptations to these specialized environments may have been driven by island biogeography-like evolutionary trends in which particular phototroph lineages have been isolated and evolved not in specific locations but in particular geochemical settings

Consistent with other analyses of (bacterio)chlorophyll synthesis protein phylogenies (e.g. Xiong et al. 2000, Bryant et al. 2012), we were unable to recover robust, consistent branching order of deep divergences in many protein families (e.g. BchLNB and BchXYZ), though sister group relationships appear robust for each protein. The inconsistent branching order of deep divergences in rooted phylogenies of individual subunits of the BchLNB and BchXYZ complexes may be an artifact of long branch attraction and saturation of variable sequences in relatively small soluble proteins over billions of years of evolution, or this difference in branching order may reflect actual differences in evolutionary history, whereby subunits of these complexes underwent independent horizontal gene transfer events early in their histories. It is difficult (if not

impossible) to distinguish between these possibilities given the limited sequence data and long evolutionary timescales with which we are left. Nonetheless, broad evolutionary trends and shallower sister-group relationships were robustly recovered, and these clearly indicate that different components of the complete phototrophy pathway have independent evolutionary histories.

These limitations to interpreting the early evolutionary history of (bacterio)chlorophyll synthesis reflect a larger problem that must be confronted in phylogenetic analysis over long geological timescales—sufficient information may not be left in the biological record to answer all questions (e.g. Meyer et al. 1986). The mutational saturation rate of protein sequences, particularly for short poorly conserved soluble proteins, limits the evolutionary timescale over which sequence-based phylogenies remain meaningful. The amount of evolutionary time represented by the diversity of (bacterio)chlorophyll synthesis proteins is on the edge of the range at which phylogenetic relationships are meaningfully recoverable, and the amount of evolutionary distance involved in outgroups necessary for rooting further complicate the recovery of meaningful evolutionary histories of these proteins (e.g. Lockhart et al. 1996, Fischer et al. 2016). We therefore find particular value in integrating ecological scenarios for the evolution of phototrophy and (bacterio)chlorophyll synthesis as an independent means of supplementing the limited resolution available from extant sequence data.

This model predicts several potentially testable hypotheses, including the potential existence of a previously undiscovered early-diverging chlorophyll a-synthesizing anoxygenic phototrophs including the “ghost lineage” depicted in Figure 3 which utilizes a basal (perhaps still homodimeric) form of RC2 along with a BchXYZ-like complex to produce chlorophylls. Environmental metagenomic sequencing has proven incredibly powerful for recovering novel phototrophic lineages (e.g. Bryant et al. 2007, Ward et al. 2018a, Holland-Moritz et al. 2018, Ward et al. 2019b), and so it remains conceivable that a larger diversity of phototrophs may be recovered which can test the evolutionary history described here. However, it remains a strong possibility that this lineage has gone extinct, or has lost the capacity for phototrophy.

Methods:

To compare evolutionary relationships among steps in (bacterio)chlorophyll synthesis, phylogenies were constructed for individual proteins in steps involved in the conversion of protoporphyrin IX to protochlorophyllide a (the shared “backbone” (bacterio)chlorophyll synthesis pathway shared in all reaction center-based phototrophs) as well as the subsequent conversion of protoporphyrin IX to specific chlorophyll and bacteriochlorophyll pigments found in only some phototrophs. Congruence of tree topology between phylogenies of different proteins was taken as indicative of shared evolutionary history (i.e. vertically inherited or horizontally transferred together, but not transferred individually), while incongruence was taken as an indication of independent histories of horizontal gene transfer (i.e. HGT of a subset of (b)chl synthesis proteins rather than of the entire pathway).

Protein translations of all bacterial genomes available from Genbank were downloaded on 3 April 2018. This database was supplemented with data from Ward et al. (2018a) and Holland-Moritz et al. (2018) for phototrophic Anaerolineae and Eremiobacterota, respectively. The database was queried with the BLASTP function of

BLAST+ (Camacho et al., 2008) using reference sequences from *Chloroflexus aurantiacus* and an e value cutoff of $1e-10$. Sequences were aligned with MAFFT (Katoh and Standley 2013). Alignments were curated in Jalview (Waterhouse et al. 2009) for sequence completeness and quality of alignment. In cases where MAFFT produced poor alignments, these were reprocessed with MUSCLE (Edgar 2004). Phylogenetic trees were built with RAxML (Stamatakis 2014) on the CIPRES science gateway (Miller et al. 2010). TBE support values were calculated from RAxML bootstraps with BOOSTER (Lemoine et al. 2018). Trees were visualized with the Interactive Tree of Life (Letunic and Bork 2016). Following initial tree construction, iterative trimming of alignments was performed to remove incomplete sequences and functionally divergent homologs not involved in (bacterio)chlorophyll synthesis (e.g. nitrogenase NifD and NifH sequences returned by BchN and BchL searches, respectively). Outgroups were retained for proteins that displayed relatively short evolutionary distances and for which deep branches were robustly recovered (e.g. BchH and CobD), but trees were left unrooted when long evolutionary distances to outgroup sequences led to inconsistent rooting and deep branching order (e.g. between BchL and BchX). Congruent topologies between rooted and unrooted trees in related steps in (bacterio)chlorophyll synthesis (e.g. BchH with BchI and BchD) was considered sufficient to produce rooted consensus trees for steps in the (bacterio)chlorophyll synthesis pathway (i.e. insertion of Mg into protoporphyrin IX for BchH/D/I, the first committed step in (bacterio)chlorophyll synthesis, shared between all reaction center-based phototrophs).

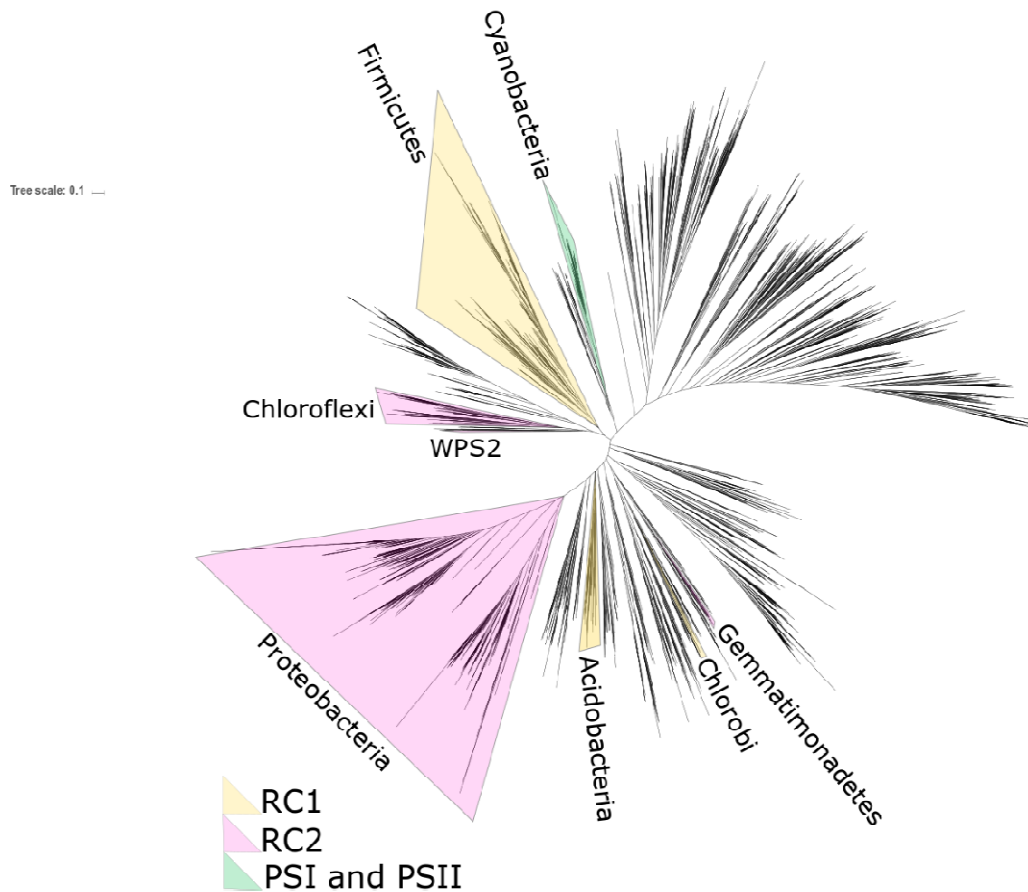
For phototroph lineages only characterized via incomplete metagenome-assembled genomes (e.g. Eremiobacterota, Anaerolineae), the likelihood that missing genes may be present in the source genome was estimated with the False Negative estimate function of MetaPOAP (Ward et al. 2018b).

For steps in (bacterio)chlorophyll synthesis catalyzed by proteins that are more promiscuous and can be recruited into and from distinct pathways (e.g. methyltransferases), and for steps that can be catalyzed by multiple poorly characterized proteins (e.g. C-8 vinyl reductase, Chew and Bryant 2007b), we do not report phylogenies as these were deemed unreliable for recording robust deep evolutionary relationships.

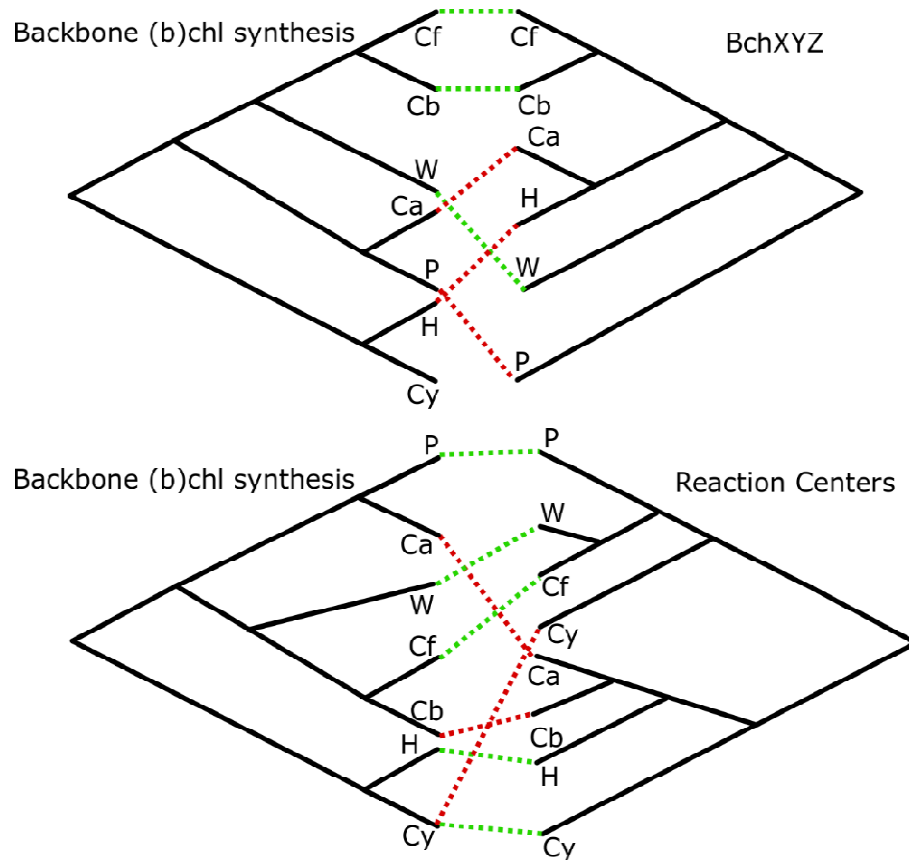
Due to very low support values for deep nodes and extensive artifacts, concatenated protein trees (e.g. BchLNB) were deemed unreliable, likely related to extensive, independent HGT of individual subunits (particularly within the Proteobacteria, e.g. Brinkmann et al. 2018).

Unrooted concatenated ribosomal protein trees of the bacterial domain were constructed following methods from Hug et al. (2016), with eukaryotes and archaea omitted and phototrophic members of Anaerolineae and Eremiobacterota added (Ward et al. 2018a, Holland-Moritz et al. 2018, Ward et al. 2019c). Following addition of new sequences, sequences were realigned with MAFFT (Katoh and Standley 2013) and the tree calculated with RAxML (Stamatakis 2014) on the CIPRES science gateway (Miller et al. 2010) and visualized with iTOL (Letunic and Bork 2016).

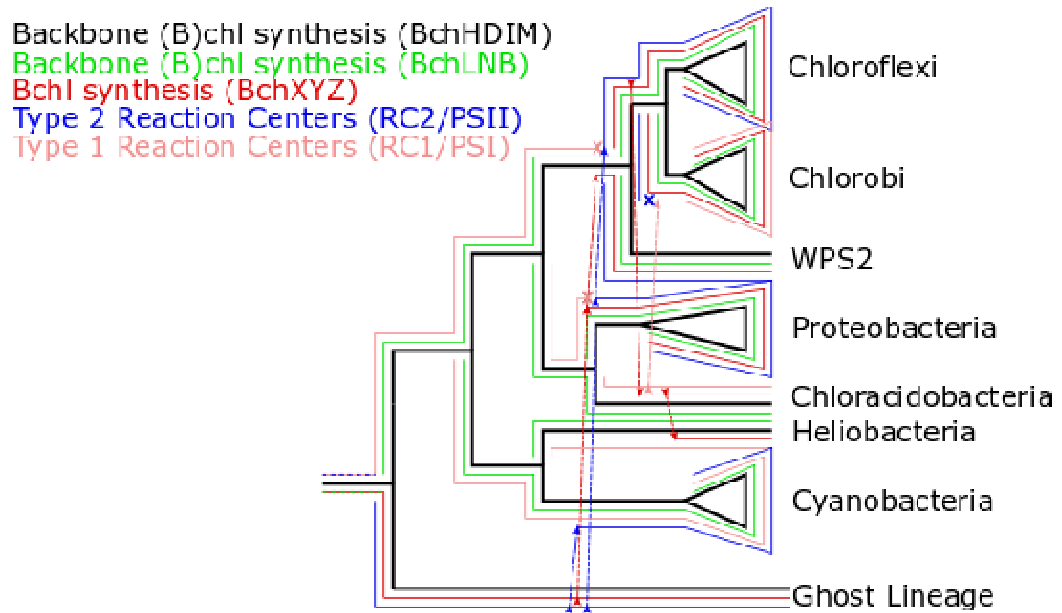
Figures:



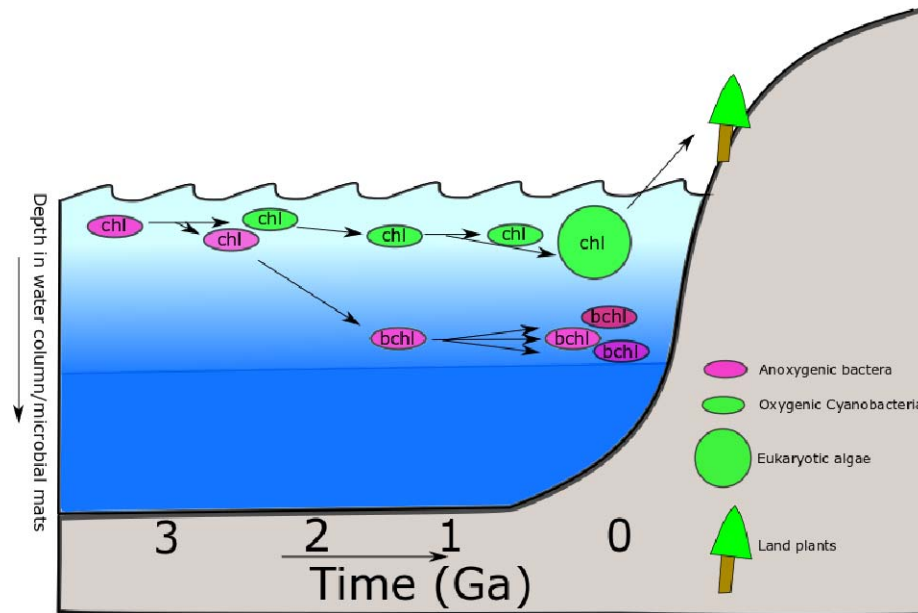
- 1) Unrooted phylogeny of bacteria built with concatenated ribosomal protein sequences, following Hug et al. 2016 with additional data from newly described phototrophs in the phylum Eremiobacterota (Holland-Moritz et al. 2018, Ward et al. 2019b) and the Chloroflexi class Anaerolineae (Ward et al. 2018a). Occurrences of reaction center-based phototrophy within a phylum indicated by shading of the entire phylum, color coded by reaction center type. For clarity, the entire phylum is highlighted even when phototrophy is restricted to only some members (e.g. the phototrophic Heliobacteria within the much broader, predominantly nonphototrophic, Firmicutes phylum).



- 2) Tanglegrams demonstrating phylogenetic incongruities between separate components of phototrophy pathways, including reaction centers, backbone (bacterio)chlorophyll synthesis, and the BchXYZ complex involved in bacteriochlorophyll synthesis; red lines indicate phyla with distinct branching order in phylogenies on the left and right, likely reflecting independent histories of horizontal gene transfer. Taxon abbreviations: Cf, Chloroflexi; Cb, Chlorobi; W, WPS2/Ermiobacterota; Ca, Chloracidobacteria; H, Heliobacteria; P, Proteobacteria (including Gemmatimonadetes); Cy, Cyanobacteria.

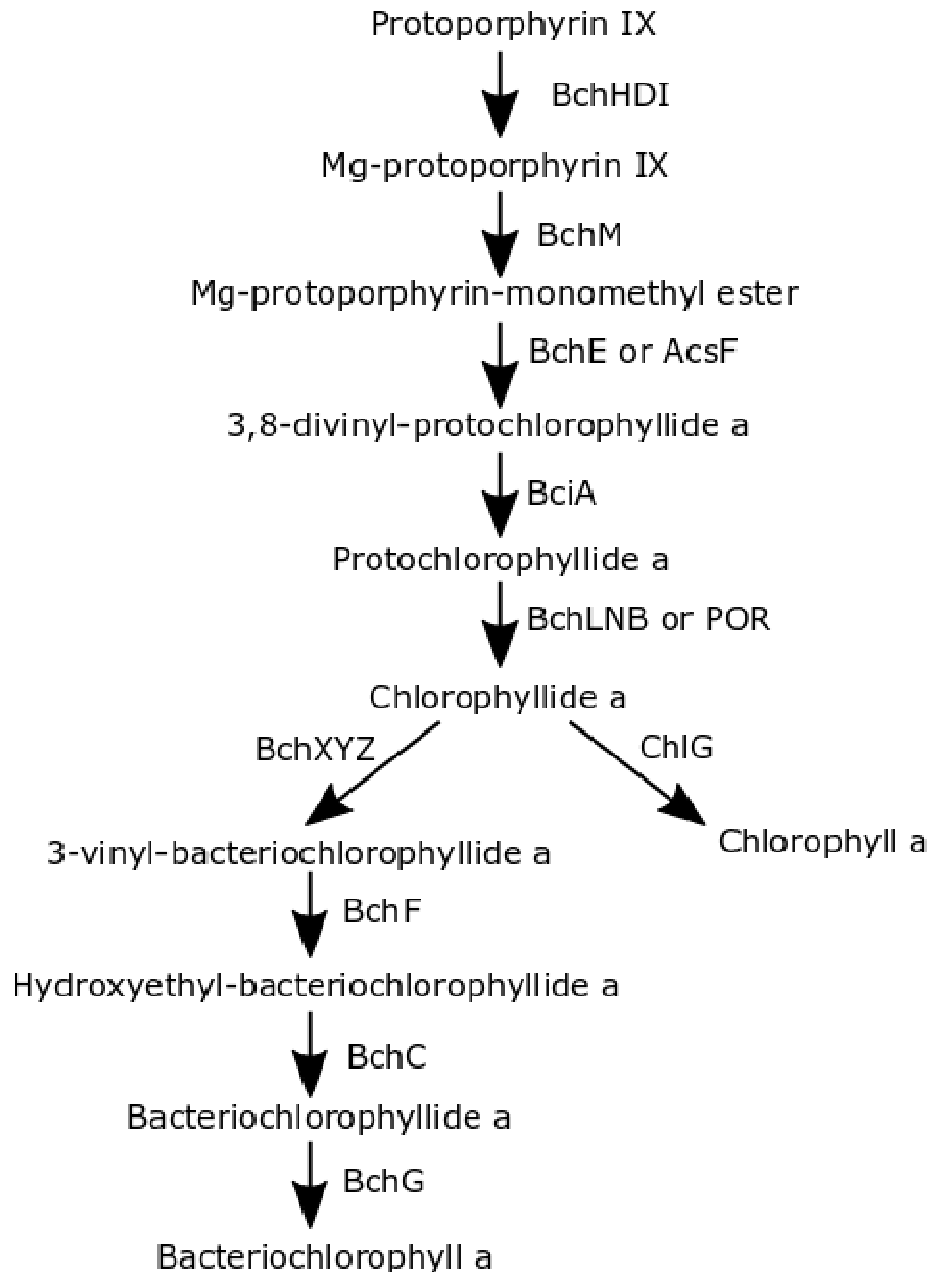


- 3) Cartoon overlaid phylogenies to demonstrate hypothesized history of HGT and ghost lineage. Underlying topology derived from backbone (bacterio)chlorophyll synthesis genes (BchH/D/I/M) (black). BchLNB and BchXYZ are derived from a common ancestor in stem group phototrophs. BchLNB (green) was inherited together with BchHDIM into extant phototrophs; BchXYZ (red) diverged in the ghost lineage before being introduced into extant anoxygenic phototrophs via HGT (a first HGT event introduced it into the stem of the proteobacterial lineage; a second HGT event introduced it into the stem of the WPS2/Chlorobi/Chloroflexi lineage; subsequent HGT introduced it into the Chloracidobacteria and Heliobacteria lineages). Type 1 reaction centers (RC1 and PSI) and Type 2 reaction centers (RC2 and PSII) diverged in stem lineage phototrophs. Type 1 reaction centers (peach) were vertically inherited into extant phototrophic lineages. Type 2 reaction centers (blue) diverged in the same ghost lineage as BchXYZ, and were introduced into extant clades via HGT (first into stem group Cyanobacteria, leading to PSII, then into stem group Proteobacteria and the stem lineage of WPS2, Chloroflexi, and Chlorobi). The most parsimonious history consistent with the data involves a secondary replacement of RC2 with RC1 in Chlorobi. Alternative evolutionary histories are similarly parsimonious, but all involve many events of HGT of individual phototrophy components and most involve secondary loss and replacement in some lineages. The inclusion of one or more ghost lineages improves parsimony and provides a good explanation for long branches between RC1/RC2 and BchLNB/BchXYZ homolog pairs.



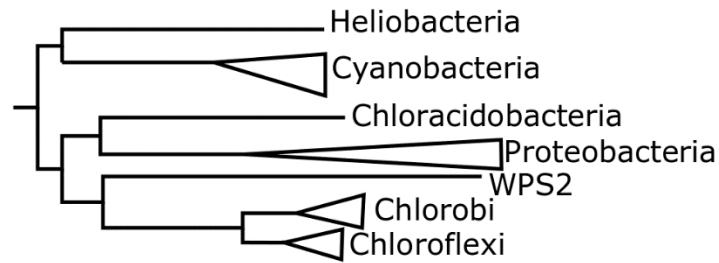
- 4) Cartoon timeline of phototroph evolution as hypothesized here. The “ancestral” phototroph was anoxygenic and utilized chlorophyll pigments, adapted to high light. At least two lineages of phototrophs diverged during Archean time, giving rise to the ancestors of Type 1 Reaction Centers and the BchLNB complex, and the Type 2 Reaction Centers and the BchXYZ complex. Eventually, HGT of a Type 2 RC into an RC1 and BchLNB-containing proto-cyanobacterium enabled the evolution of oxygenic photosynthesis, still using chlorophyll pigments. As oxygenated surface waters and competition with Cyanobacteria forced anoxygenic phototrophs into lower light regions of the water column or microbial mats (where oxygen is lower and electron donors are more abundant), they underwent adaptation to lower quality light. This included the HGT of BchXYZ complexes into BchLNB-containing lineages, allowing the innovation of bacteriochlorophyll pigments, probably first with *bchl a*. Eventually, anoxygenic phototrophs diversified in terms of organisms (including extant groups), pigments (*bchl c-g*), and reaction centers (further HGT of RC2). Oxygenic phototrophy diversified via eukaryotic endosymbiosis (primary and higher order) and colonization of land by plants. While the relative timing of these events can be inferred from comparative biology, absolute timing of many of these events is only poorly constrained if at all, based on the sparse microfossil and biomarker record of early phototrophic microbes and molecular clock estimates for the antiquity of crown group clades.

Supplemental Figures:

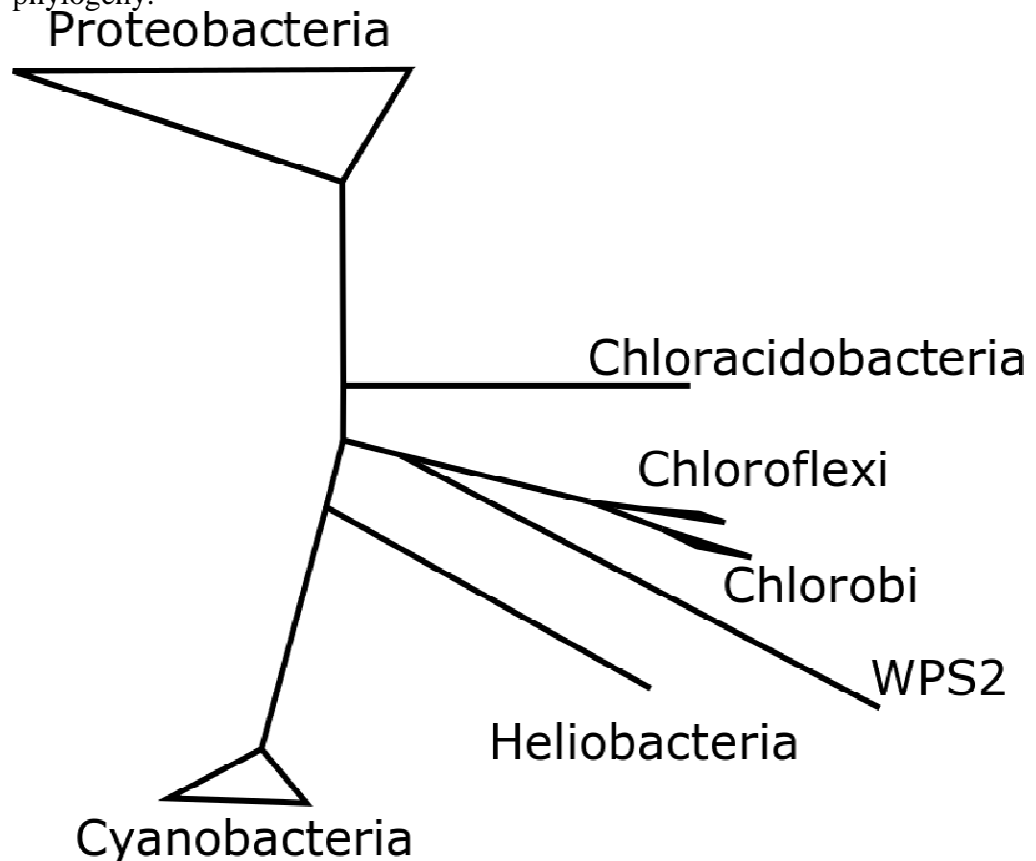


- 1) Simplified diagram of chlorophyll a and bacteriochlorophyll a synthesis from protoporphyrin IX. Steps leading to chlorophyllide a synthesis are shared by all photosynthetic bacteria. From chlorophyllide a, a single enzymatic step can produce chlorophyll a (right branch) as performed in Cyanobacteria, or multiple steps can be taken to produce bacteriochlorophyll a (left branch) as performed in most characterized anoxygenic phototrophs. Not shown are branch points leading to more evolutionarily derived alternative pigments including bchl c, bchl g, chl b, and chl d.
- 2) Consensus phylogenies of steps in (bacterio)chlorophyll synthesis. Major radiations are collapsed at the phylum level. The Chloroflexi clade includes phototrophic Chloroflexia as well as multiple lineages of phototrophic

Anaerolineae. The Proteobacteria clade includes Gemmatimonadetes and in some cases (BchL/N/B) the Synechococcus/Prochlorococcus clade of Cyanobacteria.

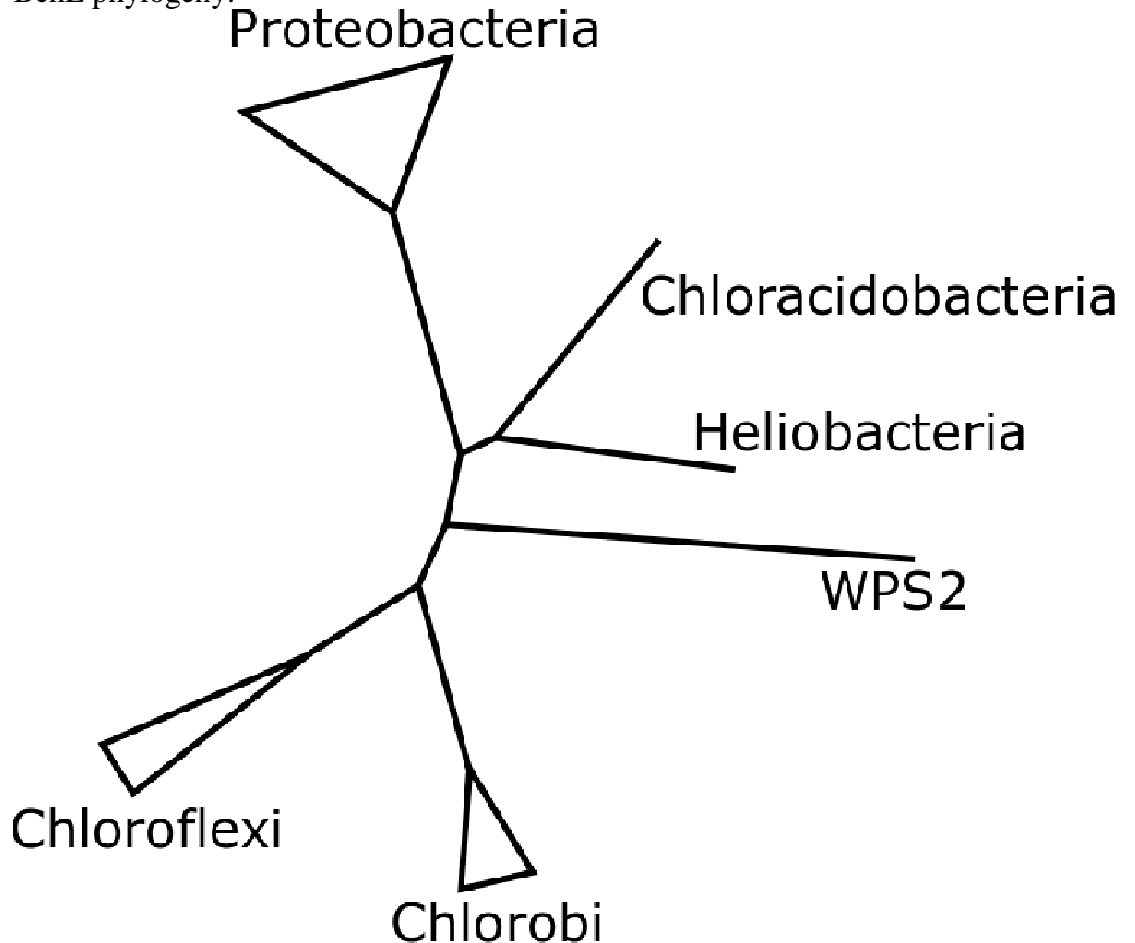


A) consensus tree for earliest dedicated steps in (bacterio)chlorophyll synthesis, including BchH, BchI, and BchD for conversion of protoporphyrin IX to Mg-protoporphyrin IX, and BchM for further conversion to Mg-protoporphyrin monomethyl ester. Consensus tree is rooted based on consistent topology between most branches in all trees, with robust root derived from CobN as an outgroup to BchH. Branch lengths are approximate, derived from the BchM phylogeny.



B) consensus tree for BchL, BchN, and BchB, subunits of the DPOR complex for conversion of protochlorophyllide a to chlorophyllide a, the last step shared in chlorophyll and bacteriochlorophyll synthesis pathways. Long branches between BchL/N/B and closest outgroups (BchX/Y/Z) resulted in poorly supported root placement, so tree is presented unrooted. The topology of the BchL/N/B tree is identical at the phylum level to the BchH/D/I/M tree,

providing support for interpretation of a shared history of the entire “backbone” (bacterio)chlorophyll synthesis pathway and an inferred root for the BchL/N/B tree on the branch between the Cyanobacteria+Heliobacteria clade and the other phyla. Branch lengths are approximate, derived from the BchL phylogeny.



C) consensus tree for BchX, BchY, and BchZ, used for the conversion of chlorophyllide a to 3-vinyl-bacteriochlorophyllide a, the first dedicated step in the synthesis of bacteriochlorophylls a, b, and g, and therefore found in all characterized anoxygenic phototrophs. Long branches between BchX/Y/Z and closest outgroups (BchL/N/B) resulted in poorly supported root placement, so tree is presented unrooted. The topology of this tree is incongruent with those presented in A) and B), suggesting independent histories of HGT of bacteriochlorophyll-specific genes versus shared backbone (bacterio)chlorophyll synthesis genes. Branch lengths are approximate, derived from the BchX phylogeny.

- 3) Protein phylogeny with TBE supports for BchH
- 4) Protein phylogeny with TBE supports for BchD
- 5) Protein phylogeny with TBE supports for BchI
- 6) Protein phylogeny with TBE supports for BchM
- 7) Protein phylogeny with TBE supports for BchL
- 8) Protein phylogeny with TBE supports for BchN
- 9) Protein phylogeny with TBE supports for BchB

- 10) Protein phylogeny with TBE supports for BchX
- 11) Protein phylogeny with TBE supports for BchY
- 12) Protein phylogeny with TBE supports for BchZ
- 13) Protein phylogeny for BchN and BchB
- 14) Protein phylogeny for BchB and BchZ
- 15) Protein phylogeny for BchN and BchY
- 16) Protein phylogeny for BchY and BchZ

Tables:

	Phylum	RC	Pigments	C fixation	Aerobic ?	Distribution of phototrophy
Oxyphotobacteria	Cyanobacteria	PSI, PSII	Chl A	CBB	Yes	Monophyletic, synapomorphic
Chloroflexia	Chloroflexi	RC2	Bchl A, or Bchl C and Bchl A	3HP or CBB	Yes	Monophyletic, derived
Anaerolineae	Chloroflexi	RC2	Bchl A	None	Yes (predicted)	Polyphyletic, derived
Chloracidobacteria	Acidobacteria	RC1	Bchl A and Bchl C	None	Micro	Monophyletic, derived
Heliobacteria	Firmicutes	RC1	Bchl G	None	No	Monophyletic, derived
Chlorobi	Chlorobi	RC1	Bchl A, plus Bchl C, D, E, or F	rTCA or none	No (Yes for one lineage)	Monophyletic, synapomorphic
Alphaproteobacteria	Proteobacteria	RC2	Bchl A	CBB or none	Yes	Polyphyletic, derived
Betaproteobacteria	Proteobacteria	RC2	Bchl A	CBB or none	Yes	Polyphyletic, derived
Gammaproteobacteria	Proteobacteria	RC2	Bchl A	CBB or	Yes	Polyphyletic,

				none		derived
Baltobacterales	Eremiobacterota (WPS-2)	RC2	BchlA	CBB or none	Yes (predicted)	Polyphyletic, derived
Gemmatimonas	Gemmatimonadetes	RC2	BchlA	none	Yes	Monophyletic, derived

Table 1: Traits of extant phototroph lineages

Gene	Substrate	Product	Presence	Evolutionary relatedness	Phylogeny figure
bchH	Protoporphyrin IX	Mg-Protoporphyrin IX	All phototrophs (3 copies in Chlorobi and Chloroflexi)		Supplemental Figure 3
bchD	Protoporphyrin IX	Mg-Protoporphyrin IX	All phototrophs		Supplemental Figure 4
bchI	Protoporphyrin IX	Mg-Protoporphyrin IX	All phototrophs		Supplemental Figure 5
bchM	Mg-Protoporphyrin IX	Mg-protoporphyrinmono methyl ester	All phototrophs except Anaerolineae		Supplemental Figure 6
bchL	Protochlorophyllide a	Chlorophyllide a	All phototrophs except Anaerolineae	Homologous to nifH, bchX	Supplemental Figure 7
bchN	Protochlorophyllide a	Chlorophyllide a	All phototrophs except Anaerolineae	Homologous to nifD bchB, bchY, bchZ	Supplemental Figure 8
bchB	Protochlorophyllide a	Chlorophyllide a	All phototrophs except Anaerolineae	Homologous to nifD bchN, bchY, bchZ	Supplemental Figure 9

bchX	Chlorophyllide a	3-vinyl-bacteriochlorophyllide a	All anoxygenic phototrophs	Homologous to nifH, bchL	Supplemental Figure 10
bchY	Chlorophyllide a	3-vinyl-bacteriochlorophyllide a	All anoxygenic phototrophs	Homologous to nifD bchB, bchN, bchZ	Supplemental Figure 11
bchZ	Chlorophyllide a	3-vinyl-bacteriochlorophyllide a	All anoxygenic phototrophs	Homologous to nifD bchB, bchY, bchN	Supplemental Figure 12

Supplemental Table 1: Proteins involved in (bacterio)-chlorophyll synthesis from protoporphyrin IX.

References:

- 1) Atamna-Ismaeel, N. et al. Bacterial anoxygenic photosynthesis on plant leaf surfaces. *Environ. Microbiol. Rep.* 4, 209-216, doi:10.1111/j.1758-2229.2011.00323.x (2012).
- 2) Battistuzzi, F.U., Feijao, A. and 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(1), p.44.
- 3) Brinkmann, H., Göker, M., Koblížek, M., Wagner-Döbler, I., & Petersen, J. (2018). Horizontal operon transfer, plasmids, and the evolution of photosynthesis in Rhodobacteraceae. *The ISME Journal*, 1.
- 4) Brocks, J. J., Love, G. D., Summons, R. E., Knoll, A. H., Logan, G. A., & Bowden, S. A. (2005). Biomarker evidence for green and purple sulphur bacteria in a stratified Palaeoproterozoic sea. *Nature*, 437(7060), 866.
- 5) Bryant, D.A., Costas, A.M.G., Maresca, J.A., Chew, A.G.M., Klatt, C.G., Bateson, M.M., Tallon, L.J., Hostetler, J., Nelson, W.C., Heidelberg, J.F. and Ward, D.M., 2007. *Candidatus Chloracidobacterium thermophilum*: an aerobic phototrophic acidobacterium. *Science*, 317(5837), pp.523-526.
- 6) Bryant, D., Z. Liu, T. LI, F. Zhao, C. G. Klatt, D. Ward, N. U. Frigaard and J. Overmann (2012). Comparative and functional genomics of anoxygenic green bacteria from the taxa Chlorobi, Chloroflexi, and Acidobacteria. *Functional Genomics and Evolution of Photosynthetic Systems*. R. L. Burnap and W. Vermaas. Dordrecht Springer. 33: 47-102.
- 7) Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer and T. L. Madden (2009). "BLAST+: architecture and applications." *Bmc Bioinformatics* 10: 421. DOI: 10.1186/1471-2105-10-421.
- 8)
- 9) Canfield, D.E., Rosing, M.T. and Bjerrum, C., 2006. Early anaerobic metabolisms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1474), pp.1819-1836.
- 10) Cardona, T. A fresh look at the evolution and diversification of photochemical reaction centers. *Photosynth. Res.* 126, 111-134, doi:10.1007/s11120-014-0065-x (2015).

- 11) Chen, J., & Strous, M. (2013). Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1827(2), 136-144.
- 12) Cheng, Q., Day, A., Dowson-Day, M., Shen, G. F., and Dixon, R. (2005). The *Klebsiella pneumoniae* nitrogenase Fe protein gene (*nifH*) functionally substitutes for the *chlL* gene in *Chlamydomonas reinhardtii*. *Biochem. Biophys. Res. Commun.* 329, 966–975. doi: 10.1016/j.bbrc.2005.02.064
- 13) Chew, A. G. M., & Bryant, D. A. (2007). Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity. *Annu. Rev. Microbiol.*, 61, 113-129.
- 14) Chew, A. G. M., & Bryant, D. A. (2007b). Characterization of a plant-like protochlorophyllide a divinyl reductase in green sulfur bacteria. *Journal of Biological Chemistry*, 282(5), 2967-2975.
- 15) Chistoserdova, L. (2011). Modularity of methylotrophy, revisited. *Environmental Microbiology*, 13(10), 2603-2622.
- 16) Crockford, P.W., Hayles, J.A., Bao, H., Planavsky, N.J., Bekker, A., Fralick, P.W., Halverson, G.P., Bui, T.H., Peng, Y. and Wing, B.A., 2018. Triple oxygen isotope evidence for limited mid-Proterozoic primary productivity. *Nature*, 559(7715), pp.613-616.
- 17) Doolittle RF (1986) Of URFs and ORFs: A primer on how to analyze derived amino acid sequences
- 18) Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- 19) Farquhar, J., Bao, H. and Thieme, M., 2000. Atmospheric influence of Earth's earliest sulfur cycle. *Science*, 289(5480), pp.756-758.
- 20) Fenchel, T. and Straarup, B.J., 1971. Vertical distribution of photosynthetic pigments and the penetration of light in marine sediments. *Oikos*, pp.172-182.
- 21) Fischer, W. W., Hemp, J., & Johnson, J. E. (2016). Evolution of oxygenic photosynthesis. *Annual Review of Earth and Planetary Sciences*, 44, 647-683.
- 22) Granick, S, in *Evolving Genes and Proteins*, V. Bryson and H. J. Vogel, Eds. (Academic Press, New York, 1965), pp. 67–68
- 23) Gumsley, A.P., Chamberlain, K.R., Bleeker, W., Söderlund, U., de Kock, M.O., Larsson, E.R. and Bekker, A., 2017. Timing and tempo of the Great Oxidation Event. *Proceedings of the National Academy of Sciences*, 114(8), pp.1811-1816.
- 24) Guo, Q., Strauss, H., Kaufman, A.J., Schröder, S., Gutzmer, J., Wing, B., Baker, M.A., Bekker, A., Jin, Q., Kim, S.T. and Farquhar, J., 2009. Reconstructing Earth's surface oxidation across the Archean-Proterozoic transition. *Geology*, 37(5), pp.399-402.
- 25) Hallenbeck, P. C., Grogger, M., Mraz, M. & Veverka, D. Draft genome sequence of the photoheterotrophic Chloracidobacterium thermophilum strain OC1 found in a Mat at Ojo Caliente. *Genome Announc.* 4, e01570-01515, doi:10.1128/genomeA.01570-15 (2016).
- 26) Hao, J., Knoll, A.H., Huang, F., Schieber, J., Hazen, R.M. and Daniel, I., 2020. Cycling phosphorus on the Archean Earth: Part II. Phosphorus limitation on primary production in Archean ecosystems. *Geochimica et Cosmochimica Acta*.
- 27) Hodgskiss, M.S., Crockford, P.W., Peng, Y., Wing, B.A. and Horner, T.J., 2019. A productivity collapse to end Earth's Great Oxidation. *Proceedings of the National Academy of Sciences*, 116(35), pp.17207-17212.

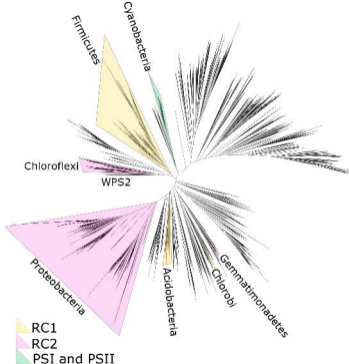
- 28) Holland □ Moritz, H., Stuart, J., Lewis, L.R., Miller, S., Mack, M.C., McDaniel, S.F. and Fierer, N., 2018. Novel bacterial lineages associated with boreal moss species. *Environmental microbiology*, 20(7), pp.2625-2638.
- 29) Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., et al. (2016). A new view of the tree and life's diversity. *Nat. Microbiol.* 1:16048. doi: 10.1038/nmicrobiol.2016.48
- 30) Igarashi, N., Harada, J., Nagashima, S., Matsuura, K., Shimada, K., and Nagashima, K. V. (2001). Horizontal transfer of the photosynthesis gene cluster and operon rearrangement in purple bacteria. *J. Mol. Evol.* 52, 333–341. doi: 10.1007/s002390010163
- 31) Imhoff, J. F. in *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea* (eds Eugene Rosenberg et al.) 501-514 (Springer Berlin Heidelberg, 2014).
- 32) Johnson, J.E., Webb, S.M., Thomas, K., Ono, S., Kirschvink, J.L. and Fischer, W.W., 2013. Manganese-oxidizing photosynthesis before the rise of cyanobacteria. *Proceedings of the National Academy of Sciences*, 110(28), pp.11238-11243.
- 33) Johnson, J.E., Gerpheide, A., Lamb, M.P. and Fischer, W.W., 2014. O₂ constraints from Paleoproterozoic detrital pyrite and uraninite. *Bulletin*, 126(5-6), pp.813-830.
- 34) Kaschner, M., Loeschke, A., Krause, J., Minh, B. Q., Heck, A., Endres, S., ... & Drepper, T. (2014). Discovery of the first light □ dependent protochlorophyllide oxidoreductase in anoxygenic phototrophic bacteria. *Molecular microbiology*, 93(5), 1066-1078.
- 35) Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4), 772-780.
- 36) Kharecha, P., Kasting, J. and Siefert, J., 2005. A coupled atmosphere–ecosystem model of the early Archean Earth. *Geobiology*, 3(2), pp.53-76.
- 37) Kimble, L.K., Mandelco, L., Woese, C.R. and Madigan, M.T., 1995. *Heliobacterium modesticaldum*, sp. nov., a thermophilic heliobacterium of hot springs and volcanic soils. *Archives of microbiology*, 163(4), pp.259-267.
- 38) Klatt, C. G., Wood, J. M., Rusch, D. B., Bateson, M. M., Hamamura, N., Heidelberg, J. F., et al. (2011). Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. *ISME J.* 5, 1262–1278. doi: 10.1038/ismej.2011.73
- 39) Larkum, A.W.D., Ritchie, R.J. and Raven, J.A., 2018. Living off the Sun: chlorophylls, bacteriochlorophylls and rhodopsins. *Photosynthetica*, 56(1), pp.11-43.
- 40) Lemoine, F., Entfellner, J. B. D., Wilkinson, E., Correia, D., Felipe, M. D., Oliveira, T., & Gascuel, O. (2018). Renewing Felsenstein’s phylogenetic bootstrap in the era of big data. *Nature*, 556(7702), 452.
- 41) Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44.W1, W242–W245. doi: 10.1093/nar/gkw290
- 42) Larkum, A. W. D., Ritchie, R. J., & Raven, J. A. (2018). Living off the Sun: chlorophylls, bacteriochlorophylls and rhodopsins. *Photosynthetica*, 1-33.
- 43) Lockhart, P. J., Larkum, A. W., Steel, M., Waddell, P. J., & Penny, D. (1996). Evolution of chlorophyll and bacteriochlorophyll: the problem of invariant sites in sequence analysis. *Proceedings of the National Academy of Sciences*, 93(5), 1930-1934.

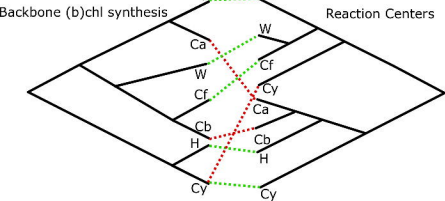
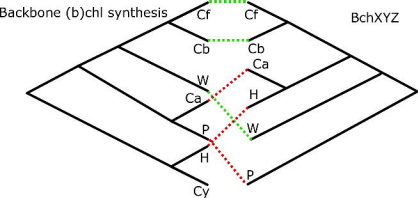
- 44) Louca, S., Shih, P. M., Pennell, M. W., Fischer, W. W., Parfrey, L. W., & Doebeli, M. (2018). Bacterial diversification through geological time. *Nature ecology & evolution*, 2(9), 1458.
- 45) Madigan, M. T. & Ormerod, J. G. in *Anoxygenic Photosynthetic Bacteria* (eds Robert E. Blankenship, Michael T. Madigan, & Carl E. Bauer) 17-30 (Springer Netherlands, 1995).
- 46) Magnabosco, C., Moore, K.R., Wolfe, J.M. and Fournier, G.P., 2018. Dating phototrophic microbial lineages with reticulate gene histories. *Geobiology*, 16(2), pp.179-189.
- 47) Meyer, T.E., Cusanovich, M.A. and Kamen, M.D., 1986. Evidence against use of bacterial amino acid sequence data for construction of all-inclusive phylogenetic trees. *Proceedings of the National Academy of Sciences*, 83(2), pp.217-220.
- 48) Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). "Creating the CIPRES Science Gateway for inference of large phylogenetic trees," in *Proceedings of the Gateway Computing Environments Workshop (GCE)*, (New Orleans, LA).
- 49) Nagashima, S., and Nagashima, K. V. P. (2013). "Comparison of photosynthesis gene clusters retrieved from total genome sequences of purple bacteria," in *Genome Evolution of Photosynthetic Bacteria*, ed T. Beatty (Amsterdam: Academic Press), 151–178.
- 50) Norell, M.A., Novacek, M.J. and Wheeler, Q.D., 1992. Taxic origin and temporal diversity: the effect of phylogeny. *Extinction and phylogeny*. Columbia University Press, New York, pp.89-118.
- 51) Norell, M.A., 1993. Tree-based approaches to understanding history; comments on ranks, rules and the quality of the fossil record. *American Journal of Science*, 293(A), p.407.
- 52) Orf, G.S., Gisriel, C. and Redding, K.E., 2018. Evolution of photosynthetic reaction centers: insights from the structure of the heliobacterial reaction center. *Photosynthesis research*, pp.1-27.
- 53) Ozaki, K., Thompson, K.J., Simister, R.L., Crowe, S.A. and Reinhard, C.T., 2019. Anoxygenic photosynthesis and the delayed oxygenation of Earth's atmosphere. *Nature communications*, 10(1), pp.1-10.
- 54) Pavlov, A.A. and Kasting, J.F., 2002. Mass-independent fractionation of sulfur isotopes in Archean sediments: strong evidence for an anoxic Archean atmosphere. *Astrobiology*, 2(1), pp.27-41.
- 55) Pierson, B., Oesterle, A. and Murphy, G.L., 1987. Pigments, light penetration, and photosynthetic activity in the multi-layered microbial mats of Great Sippewissett Salt Marsh, Massachusetts. *FEMS Microbiology Ecology*, 3(6), pp.365-376.
- 56) Raven, J. A. (2009). Contributions of anoxygenic and oxygenic phototrophy and chemolithotrophy to carbon and oxygen fluxes in aquatic environments. *Aquatic Microbial Ecology*, 56(2-3), 177-192.
- 57) Schirrmeyer, B.E., de Vos, J.M., Antonelli, A. and 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(5), pp.1791-1796.
- 58) Shih, P. M., Hemp, J., Ward, L. M., Matzke, N. J., & Fischer, W. W. (2017a). Crown group Oxyphotobacteria postdate the rise of oxygen. *Geobiology*, 15(1), 19-29.
- 59) Shih, P. M., Ward, L. M., and Fischer, W. W. (2017b). Evolution of the 3-hydroxypropionate bi-cycle and recent transfer of anoxygenic photosynthesis into the

- Chloroflexi. *Proc. Natl. Acad. Sci. U.S.A.* 114, 10749–10754. doi: 10.1073/pnas.1710798114
- 60) Sleep, N.H. and Bird, D.K., 2007. Niches of the pre-photosynthetic biosphere and geologic preservation of Earth's earliest ecology. *Geobiology*, 5(2), pp.101-117.
 - 61) Sousa FL, Shavit-Grievink L, Allen JF, Martin WF. 2013. Chlorophyll biosynthesis gene evolution indicates photosystem gene duplication, not photosystem merger, at the origin of oxygenic photosynthesis. *Genome Biol. Evol.* 5:200–16
 - 62) Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
 - 63) Swithers, K.S., Soucy, S.M. and Gogarten, J.P., 2012. The role of reticulate evolution in creating innovation and complexity. *International journal of evolutionary biology*, 2012.
 - 64) Tice MM, Lowe DR. 2004. Photosynthetic microbial mats in the 3,416-Myr-old ocean. *Nature* 431:549–52
 - 65) Tice MM, Lowe DR. 2006. Hydrogen-based carbon fixation in the earliest known photosynthetic organisms. *Geology* 34:37–40
 - 66) Trembath-Reichert, E., Ward, L.M., Slotznick, S.P., Bachtel, S.L., Kerans, C., Grotzinger, J.P. and Fischer, W.W., 2016. Gene sequencing-based analysis of microbial-mat morphotypes, Caicos Platform, British West Indies. *Journal of Sedimentary Research*, 86(6), pp.629-636.
 - 67) Ward, L.M., Kirschvink, J.L. and Fischer, W.W., 2016. Timescales of oxygenation following the evolution of oxygenic photosynthesis. *Origins of Life and Evolution of Biospheres*, 46(1), pp.51-65.
 - 68) Ward, L. M., McGlynn, S. E. & Fischer, W. W. 2017. Draft genome sequence of *Chloracidobacterium* sp. CP2_5A, a phototrophic member of the phylum Acidobacteria recovered from a Japanese hot spring. *Genome Announc.* 5, e00821-00817, doi:10.1128/genomeA.00821-17 (2017).
 - 69) Ward, L. M., Hemp, J., Shih, P. M., McGlynn, S. E., & Fischer, W. W. (2018a). Evolution of phototrophy in the Chloroflexi phylum driven by horizontal gene transfer. *Frontiers in microbiology*, 9, 260.
 - 70) Ward, LM, PM Shih, and WW Fischer. 2018b. MetaPOAP: Presence or Absence of Metabolic Pathways in Metagenome-Assembled Genomes. *Bioinformatics*, bty510. DOI:10.1093/bioinformatics/bty510
 - 71) Ward, LM*, and PM Shih*. 2019. The evolution of carbon fixation pathways in response to changes in oxygen concentration over geological time. *Free Radical Biology and Medicine*, 140 pp.188-199.
 - 72) Ward, LM, B Rasmussen, and WW Fischer. 2019a. Primary productivity was limited by electron donors prior to the advent of oxygenic photosynthesis. *JGR: Biogeosciences*, 124 (2), 211-226.
 - 73) Ward, LM, T Cardona, and H Holland-Moritz. Evolutionary Implications of Anoxygenic Phototrophy in the Bacterial Phylum Candidatus Eremiobacterota (WPS-2). 2019b. *Frontiers in Microbiology*, 10:1658. doi: 10.3389/fmicb.2019.01658.
 - 74) Ward, L.M., Shih, P.M., Hemp, J., Kakegawa, T., Fischer, W.W. and McGlynn, S.E., 2019c. Genomic Evidence for Phototrophic Oxidation of Small Alkanes in a Member of the Chloroflexi Phylum. *bioRxiv*, p.531582.

- 75) Ward, L.M., Lingappa, U.F., Grotzinger, J.P. and Fischer, W.W., 2020. Microbial mats in the Turks and Caicos Islands reveal diversity and evolution of phototrophy in the Chloroflexota order Aggregatilineales. *Environmental Microbiome*, 15, pp.1-9.
 - 76) Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033
 - 77) Wätzlich, D., Bröcker, M. J., Uliczka, F., Ribbe, M., Virus, S., Jahn, D., et al. (2009). Chimeric Nitrogenase-like enzymes of (bacterio)chlorophyll biosynthesis. *J. Biol. Chem.* 284, 15530–15540. doi: 10.1074/jbc.M901331200
 - 78) Woese, C. R. (1987). Bacterial evolution. *Microbiol. Rev.* 51, 221–271.
 - 79) Xiong, J., Fischer, W. M., Inoue, K., Nakahara, M., and Bauer, C. E. (2000). Molecular evidence for the early evolution of photosynthesis. *Science* 289, 1724–1730. doi: 10.1126/science.289.5485.1724
 - 80) Yamazaki, S., Nomata, J., & Fujita, Y. (2006). Differential operation of dual protochlorophyllide reductases for chlorophyll biosynthesis in response to environmental oxygen levels in the cyanobacterium *Leptolyngbya boryana*. *Plant physiology*, 142(3), 911-922.
 - 81) Zeng, Y., Feng, F., Medová, H., Dean, J., and Koblížek, M. (2014). Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7795–7800. doi: 10.1073/pnas.1400295111
- 82)

Tree scale: 0.1





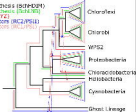
Backbone (Schiff synthesis) (Schiff-D)H)

Backbone (Schiff synthesis) (Schiff-D)H)

Schiff synthesis (Schiff-D)H)

Type 2 Reaction Centers (RC2/PSII)

Type 1 Reaction Centers (RC1/PSI)



Depth in water column/ventral/total mass

