Evolutionary Implications of Anoxygenic Phototrophy in the Bacterial Phylum *Candidatus* Eremiobacterota (WPS-2)

Lewis M. Ward[1]*, Tanai Cardona[2]* and Hannah Holland-Moritz[3,4]*

[1] Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, United States, [2] Department of Life Sciences, Imperial College London, London, United Kingdom, [3] Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, United States, [4] Department of Ecology and Evolutionary Biology, University of Colorado Boulder, Boulder, CO, United States

Genome-resolved environmental metagenomic sequencing has uncovered substantial previously unrecognized microbial diversity relevant for understanding the ecology and evolution of the biosphere, providing a more nuanced view of the distribution and ecological significance of traits including phototrophy across diverse niches. Recently, the capacity for bacteriochlorophyll-based anoxygenic photosynthesis has been proposed in the uncultured bacterial WPS-2 phylum (recently proposed as *Candidatus* Eremiobacterota) that are in close association with boreal moss. Here, we use phylogenomic analysis to investigate the diversity and evolution of phototrophic WPS-2. We demonstrate that phototrophic WPS-2 show significant genetic and metabolic divergence from other phototrophic and non-phototrophic lineages. The genomes of these organisms encode a new family of anoxygenic Type II photochemical reaction centers and other phototrophy-related proteins that are both phylogenetically and structurally distinct from those found in previously described phototrophs. We propose the name *Candidatus* Baltobacterales for the order-level aerobic WPS-2 clade which contains phototrophic lineages, from the Greek for "bog" or "swamp," in reference to the typical habitat of phototrophic members of this clade.

Keywords: phototrophy, anoxygenic, WPS-2, Chloroflexi, reaction center

INTRODUCTION

The vast majority of primary productivity on Earth is fueled by photosynthesis, both today (Raven, 2009) and through most of the history of life (Kharecha et al., 2005; Canfield et al., 2006; Ward et al., 2019b), making the organisms and proteins responsible for driving photosynthesis into critical bases for the carbon cycle. While the quantitatively most significant form of photosynthesis today is oxygenic photosynthesis, which uses water as an electron donor to power carbon fixation, there is also a great diversity of bacteria capable of anoxygenic photosynthesis using compounds such as sulfide, molecular hydrogen, ferrous iron, or arsenic as electron donors (Hohmann-Marriott and Blankenship, 2011; Thiel et al., 2018). Anoxygenic photosynthesis is restricted to members of the bacterial domain, where it has a scattered phylogenetic distribution (**Figure 1**). While the first observations of anoxygenic photosynthesis were made in the late 1800s

(Gest and Blankenship, 2005), the known diversity of bacteria capable of anoxygenic phototrophy has exploded in recent years thanks largely to genome-resolved metagenomics (Bryant et al., 2007; Ward et al., 2018a; **Figure 1**). As our understanding of the physiology and evolution of phototrophy is contingent on adequate sampling of the diversity of phototrophic organisms, continuing discoveries of phylogenetically, ecologically, and biochemically novel phototrophs spurs new insights into phototrophy today and in Earth's history.

Unlike oxygenic photosynthesis which requires the work of Type I and Type II photochemical reaction centers linked in series for water oxidation, anoxygenic phototrophs use exclusively either Type I or Type II reaction centers. Type I and II reaction centers can be differentiated by the nature of their electron acceptor (Hohmann-Marriott and Blankenship, 2011). Type I reaction centers reduce ferredoxin whereas Type II reaction centers reduce quinones. There are only five phyla of bacteria known to have phototrophic representatives that use anoxygenic Type II reaction centers: Proteobacteria, Chloroflexi, and Gemmatimonadetes (Zeng et al., 2014), recent evidence for potential phototrophy in one isolate in the Bacteroidetes (Tahon and Willems, 2017), and the newly discovered WPS-2 (Holland-Moritz et al., 2018; **Figure 1**). The photochemical pigments of anoxygenic Type II reaction centers are bound by two homologous subunits known as L (PufL) and M (PufM). These have traditionally been subdivided into two types, the L and M found in the reaction centers of phototrophic Proteobacteria (PbRC) and the L and M found in those of the phototrophic Chloroflexi (CfRC), with each set making deep-branching monophyletic clades (Beanland, 1990; Cardona, 2015). Therefore, substantial differences exist between the PbRC and the CfRC not only at the level of photochemistry, but also at the level of sequence identity of the core subunits, pigment and subunit composition (Xin et al., 2018; Yu et al., 2018). Phototrophic Gemmatimonadetes and Bacteroidetes, on the other hand, encode Proteobacteria-like reaction centers, suggesting that these organisms acquired phototrophy via horizontal gene transfer from members of the Proteobacteria (Zeng et al., 2014; Tahon and Willems, 2017). In addition, Proteobacterial *pufLM* genes may have also been detected in strains of the genus *Alkalibacterium* of the phylum Firmicutes from perennial springs of the high arctic (Perreault et al., 2008).

Here, we show that the genomes of the newly discovered putative phototrophs of the uncultivated candidate phylum WPS-2 (also known as *Candidatus* Eremiobacterota; Ji et al., 2017) encode a third distinct lineage of anoxygenic Type II reaction centers with novel and unusual characteristics—only the third of this type after the discovery of the Chloroflexi nearly half a century ago (Pierson and Castenholz, 1971). We also show that the WPS-2 clade encoding phototrophy is distinct from basal Eremiobacterota and related phyla in terms of traits including the capacity for aerobic respiration, membrane architecture, and environmental distribution, reflecting a divergent evolutionary history and likely novel ecological roles.

MATERIALS AND METHODS

Metagenome Analyses

Metagenome-assembled genomes (MAGs) of WPS-2 bacteria were downloaded from NCBI WGS and JGI IMG databases. Completeness and contamination of genomes was estimated based on presence and copy number of conserved single-copy proteins by CheckM (Parks et al., 2015). Sequences of ribosomal and metabolic proteins used in analyses (see below) were identified locally with the *tblastn* function of BLAST+ (Camacho et al., 2009), aligned with MUSCLE (Edgar, 2004), and manually curated in Jalview (Waterhouse et al., 2009). Positive BLAST hits were considered to be full length (e.g., >90% the shortest reference sequence from an isolate genome) with *e*-values better than $1e^{-20}$. Genes of interest were screened against outlier (e.g., likely contaminant) contigs as determined by CheckM (Parks et al., 2015) and RefineM (Parks et al., 2017) using tetranucleotide, GC, and coding density content. Presence of metabolic pathways of interest was predicted with MetaPOAP (Ward et al., 2018d) to check for False Positives (contamination) or False Negatives (genes present in source genome but not recovered in metagenome-assembled genomes). Phylogenetic trees were calculated using RAxML (Stamatakis, 2014) on the Cipres science gateway (Miller et al., 2010). Transfer bootstrap support values were calculated by BOOSTER (Lemoine et al., 2018), and trees were visualized with the Interactive Tree of Life viewer (Letunic and Bork, 2016). Taxonomic assignment was confirmed with GTDB-Tk (Parks et al., 2018) and by placement in a concatenated ribosomal protein phylogeny following methods from Hug et al. (2016) Histories of vertical versus horizontal inheritance of metabolic genes was inferred by comparison of organismal and metabolic protein phylogenies to determine topological congruence (Doolittle, 1986; Ward et al., 2018a).
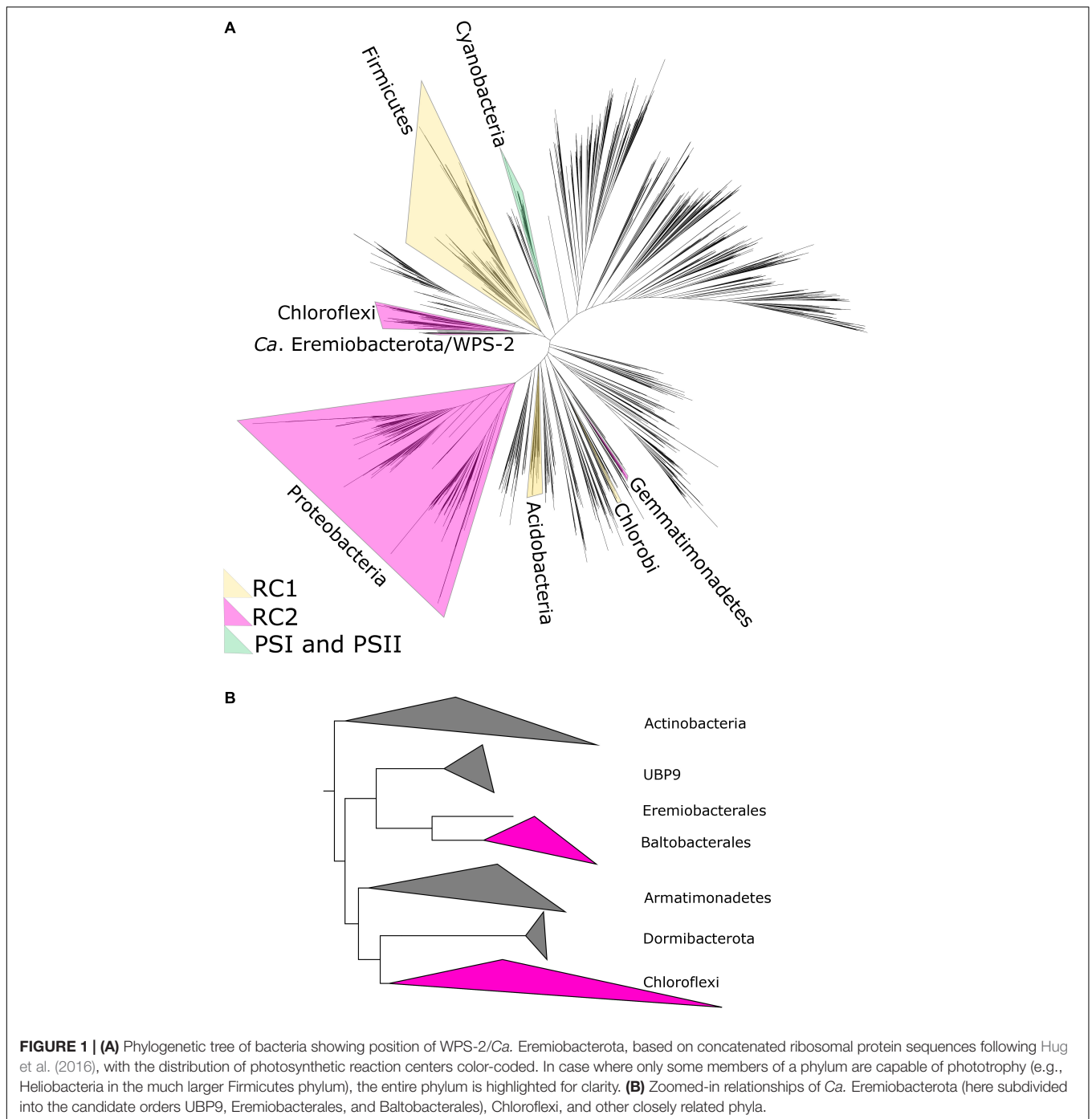
Reaction Center Evolution Analyses

A total of 14 L and 12 M amino acid sequences were collected from the compiled metagenome data. The sequences were added to a dataset of Type II reaction center subunits compiled before (Cardona et al., 2019), which included sequences from Cyanobacteria, Proteobacteria, and Chloroflexi. Sequence alignments were done in Clustal Omega (Sievers et al., 2011) using 10 combined guide trees and Hidden Markov Model iterations. Maximum Likelihood phylogenetic analysis was performed with the PhyML online service[1] (Guindon et al., 2010) using the Smart Model Selection and Bayesian information criterion for the computation of parameters from the dataset (Lefort et al., 2017). Tree search operations were performed with the Nearest Neighbor Interchange approach and the approximate likelihood-ratio test method was chosen for the computation of branch support values. A phylogenetic tree was also constructed using only L and M subunits and excluding cyanobacterial homologs.

Transmembrane helix prediction on WPS-2 L and M subunits was computed using the TMHMM Server, v. 2.0[2] (Krogh et al.,

---

[1]http://www.atgc-montpellier.fr/phyml/
[2]http://www.cbs.dtu.dk/services/TMHMM/

FIGURE 1 | (A) Phylogenetic tree of bacteria showing position of WPS-2/*Ca*. Eremiobacterota, based on concatenated ribosomal protein sequences following Hug et al. (2016), with the distribution of photosynthetic reaction centers color-coded. In case where only some members of a phylum are capable of phototrophy (e.g., Heliobacteria in the much larger Firmicutes phylum), the entire phylum is highlighted for clarity. **(B)** Zoomed-in relationships of *Ca*. Eremiobacterota (here subdivided into the candidate orders UBP9, Eremiobacterales, and Baltobacterales), Chloroflexi, and other closely related phyla.

2001) and the ΔG transmembrane helix prediction tool v. 1.0[3] (Hessa et al., 2007). Structural homology models were carried out with SWISS-MODEL server[4] (Waterhouse et al., 2018) using the 1.9 Å crystal structure of the Type II reaction center from the gammaproteobacterium *Thermochromatium tepidum* as template (Yu et al., 2018) (PDB ID: 5y5s). The WPS-2 L and M subunits annotated as Ga0175859_11240458 and Ga0175859_11402733, respectively, were used for the structural reconstruction.

RESULTS AND DISCUSSION

Ecology and Distribution of WPS-2

Clonal libraries containing 16S rRNA sequences from the candidate WPS-2 (Wittenberg Polluted Soil) phylum were first identified in a 2001 study of soil polluted with polychlorinated

---

[3]http://dgpred.cbr.su.se/

[4]https://swissmodel.expasy.org/

biphenyl (PCB) in Germany (Nogales et al., 2001). Since then, WPS-2 has been found in 16S rRNA amplicon studies from diverse environments including acidic, polluted environments (Nogales et al., 2001; Brantner et al., 2014; Trexler et al., 2014), alpine, high latitude, and Antarctic soils (Costello, 2007; Grasby et al., 2013; Ji et al., 2016, 2017), cryoconite holes in Greenland (Stibal et al., 2015), human/canine oral microbiomes (Camanocha and Dewhirst, 2014), bogs, and peatlands (Serkebaeva et al., 2013; Bragina et al., 2015). More recently, several metagenomic studies have successfully assembled MAGs from WPS-2. These studies include samples from an acidic mine, rich in arsenic (Bertin et al., 2011); MAGs from this study were later identified as WPS-2 by Camanocha and Dewhirst (2014), Arctic (Parks et al., 2017), and Antarctic desert soils (Ji et al., 2017), boreal mosses (Holland-Moritz et al., 2018), and a peatland in Sweden (Woodcroft et al., 2018).

Although these environments are varied, several trends emerge: WPS-2 has a global distribution and is most often found in cool, acidic, and aerobic environments. While not exclusively found in low-temperature environments, most samples containing WPS-2 were collected from areas that are typically cold (i.e., Antarctica, Greenland), or undergo shorter growing seasons (high-latitude and alpine environments). Similarly, WPS-2 is more often found in sites with acidic to moderately acidic pH (between 3 and 6.5). In some cases, this acidity has come from pollution, in others, such as acidic springs, peatlands, bogs, and fir-spruce forest soil, it is a natural part of the habitat. Although WPS-2 has often been associated with polluted environments, it seems more likely that this is due to the acidity accompanying these environments than to any special resistance to the varied pollutants from these studies. Oxygen-rich environments such as mosses, cryoconite holes, and the top few centimeters of soil are typical habitat for WPS-2 and several papers have speculated that members of this phylum are therefore aerobic or microaerobic (Costello, 2007; Ji et al., 2017; Holland-Moritz et al., 2018).

Often WPS-2 makes up a minor part of the bacterial community, however, in some notable cases, the phylum has either dominated the bacterial community (e.g., 23 and 25%) (Grasby et al., 2013; Stibal et al., 2015; Ji et al., 2016) or several phylotypes from WPS-2 have been among the dominant taxa (Bragina et al., 2015; Holland-Moritz et al., 2018). Although most studies identifying WPS-2 as dominant are from specialized or unique environments (i.e., acidic springs, Antarctic soils, and cryoconite holes), an easily accessible and commonplace environment in which phylotypes of WPS-2 are often abundant is boreal mosses. Phylotypes from WPS-2 are the second-most dominant taxa across six common boreal moss species and absent in only one moss of the seven species studied (Holland-Moritz et al., 2018). Interestingly, even when mosses are not the focus of a study, WPS-2 is often more common in environments that contain them (Serkebaeva et al., 2013; Trexler et al., 2014; Bragina et al., 2015; Woodcroft et al., 2018). Currently there are no isolates of WPS-2 to investigate their biochemistry, or to use in comparative genomics, yet given their phototrophic properties and evolutionary context, such an isolate would be useful.

As an accessible and abundant host to phylotypes of WPS-2, boreal mosses provide a logical starting place from which to try to culture WPS-2—an essential step in upgrading this candidate phylum to taxonomically valid status as well as for characterizing its physiology.
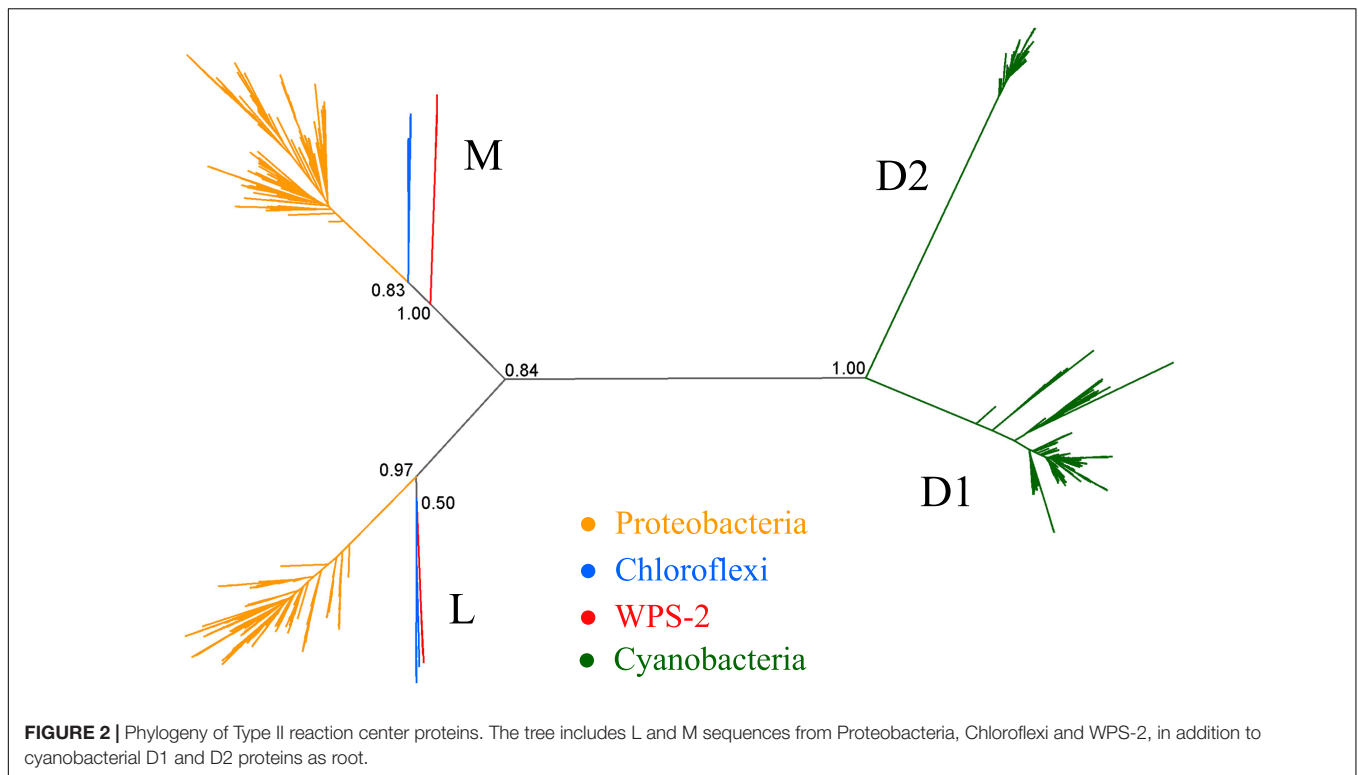
Phylogenetic Context of Phototrophy in the WPS-2 Phylum

The WPS-2 phylum is located on a branch of the bacterial tree of life located near the phyla Chloroflexi, Armatimonadetes, and Dormibacteria (**Figure 1**). This placement has interesting implications for the evolutionary history of a variety of traits (as discussed below) but is particularly relevant for considerations of the evolutionary history of anoxygenic phototrophy.

Phylogenetic analysis of proteins for the Type II reaction center suggests that phototrophy in WPS-2 might be most closely related to that of the Chloroflexi (**Figure 2**, but see below). Given the relatedness of WPS-2 and Chloroflexi phototrophy proteins (**Figure 2** and **Supplementary Figure S1**) and their relative closeness in organismal phylogenies (**Figure 1**), an important outstanding question is whether the last common ancestor of these phyla was phototrophic, implying extensive loss of phototrophy in both phyla and their relatives, or whether WPS-2 and Chloroflexi have independently acquired phototrophy via HGT subsequent to the divergence of these lineages (likely from an earlier branching uncharacterized or extinct group of phototrophic bacteria). While this cannot be conclusively ascertained with available data, it is apparent that aside from bacteriochlorophyll synthesis and phototrophic reaction centers, necessary proteins involved in phototrophic electron transfer (e.g., a *bc* complex or alternative complex III) are not shared, suggesting that members of these phyla independently acquired respiratory electron transfer necessary for phototrophy after their divergence. This is consistent with a relatively recent origin of phototrophy in the Chloroflexi (i.e., within the last ∼1 billion years) after the acquisition of aerobic respiration by the Chloroflexi and other phyla following the Great Oxidation Event (GOE) ∼2.3 billion years ago (Fischer et al., 2016; Shih et al., 2017b). Together with the relatively high protein sequence similarity between reaction center proteins among the phototrophic WPS-2 (discussed below) this may suggest that phototrophy in the WPS-2 has radiated on a similar or even shorter timescale.

Phototrophy in WPS-2 is not restricted to a single lineage, but appears in at least four discrete lineages interspersed with non-phototrophic lineages within the phylum (**Figure 3**). The topology of the organismal phylogeny of the WPS-2 is incongruent with that of phototrophy protein phylogenies (**Figure 3** and **Supplementary Figure S1**), and so this distribution may best be explained by a history of HGT between members of the WPS-2 and not by a deeper ancestry of phototrophy followed by extensive loss. Similar cases of multiple HGT events of complete photosynthesis gene clusters have been demonstrated within the family Rhodobacteraceae of Alphaproteobacteria (Brinkmann et al., 2018), between classes of the phylum Chloroflexi (Ward et al., 2018a), and into the Gemmatimonadetes

FIGURE 2 | Phylogeny of Type II reaction center proteins. The tree includes L and M sequences from Proteobacteria, Chloroflexi and WPS-2, in addition to cyanobacterial D1 and D2 proteins as root.
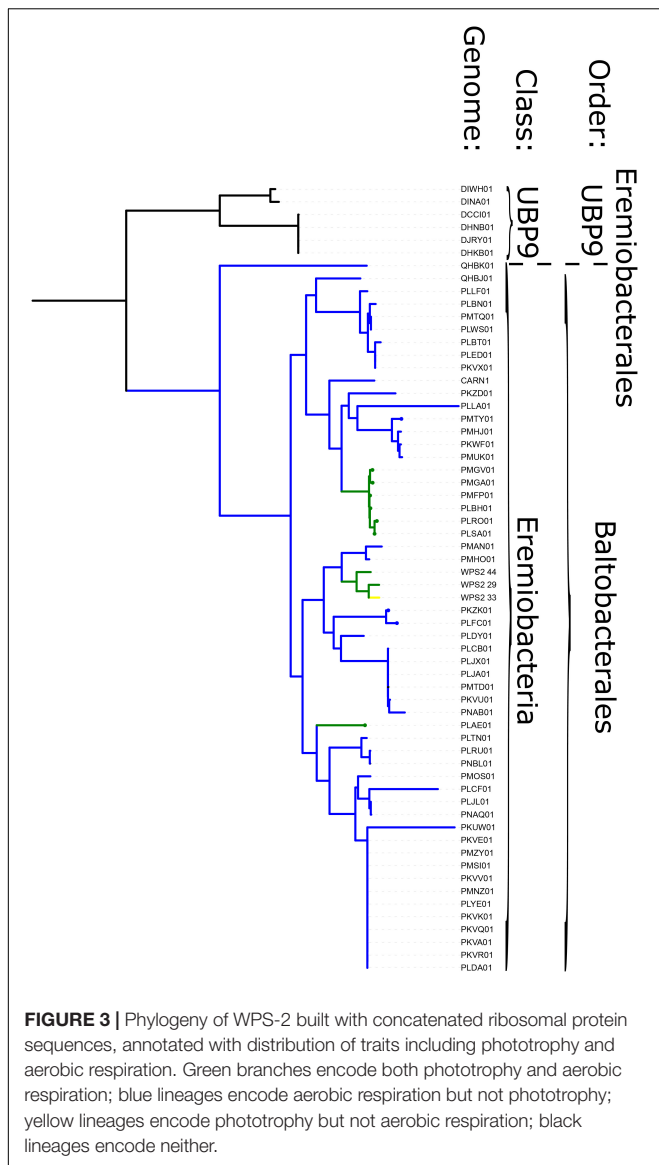
phylum from members of the Proteobacteria (Zeng et al., 2014), which might suggest convergent evolutionary patterns driving the distribution of Type II reaction center-based anoxygenic phototrophy within distant lineages.

Phototrophic WPS-2 appear to encode the synthesis of bacteriochlorophyll *a* (e.g., BchX, BchY, BchZ, BchF, BchC, and BchG) but no other bacteriochlorophylls (e.g., they do not encode proteins such as BchK, BchU, or BchQ for the synthesis of bacteriochlorophyll *c*, *d*, or *e*). These genomes do not appear to encode chlorosome-related proteins such as CsmA, CsmJ, or FmoA. The conversion of Mg-protoporphyrin monomethyl ester to 3,8-divinyl-protochlorophyllide *a* is an essential step in the production of (bacterio) chlorophyll compounds in all phototrophs, and can be encoded by either an aerobic (AcsF) or anaerobic (BchE) protein (Chew and Bryant, 2007). Phototrophic WPS-2 genomes all encode AcsF proteins related to those from phototrophic Chloroflexi and Acidobacteria (with the exception of WPS2_33, which is <80% complete) (**Supplementary Table S1**), consistent with an aerobic lifestyle. Half of the phototrophic WPS-2 genomes also encode BchE, suggesting that these strains may be facultative aerobes which are also capable of growing phototrophically in anaerobic environments.

In genomes in which phototrophy genes were recovered on contigs sufficiently long to assess clustering, bacteriochlorophyll synthesis and reaction center genes appear somewhat clustered but are spread across a larger segment of the genome and with more non-phototrophy genes intermixed than occurs for instance in phototrophic Proteobacteria and Gemmatimonadetes (Zeng et al., 2014). The gene order is fairly consistent across genomes. **Supplementary Figure S7** shows three examples of the

arrangements of bacteriochlorophyll synthesis genes and reaction center genes in genomes PLFC, PLAE, and PMFP. As in Chlorobi, BchC is located between the BchF and BchX in genomes PLAE and PLFC (Chew and Bryant, 2007). The reaction center proteins are within 2000 bp upstream of BchF gene in all three genomes. BchH and a BchH-like homolog are present in PLFC and PMFP as is found in phototrophic Chlorobi and Chloroflexi. In both cases the BchH-like gene is upstream of BchH. However, given the fragmented nature of metagenome assemblies, it is impossible to determine for certain the exact location and orientation of these genes in the genome.

The phylogenetic position of bacteriochlorophyll synthesis proteins including BchL of WPS-2 (**Supplementary Figure S2**) is somewhat in agreement with that of reaction center proteins (**Figure 2** and **Supplementary Figure S1**). However, BchL from the phylum Chlorobi is more closely related to that of the Chloroflexi than either is to the WPS-2. This is consistent with previous studies showing that HGT of protochlorophyllide and chlorophyllide reductase between stem group phototrophic Chlorobi and Chloroflexi had occurred (Bryant et al., 2012; Sousa et al., 2013), although the direction of transfer had remained ambiguous. The new topology of BchL including WPS-2 suggests exchange of bacteriochlorophyll synthesis proteins via HGT from stem group phototrophic Chloroflexi to stem group phototrophic Chlorobi. It is likely that exchange of photosynthetic components between Chloroflexi and Chlorobi had occurred more than once, in either direction, and at different time points and may have involved a swap of reaction center type as suggested before (Bryant et al., 2012) and the acquisition of chlorosomes (Hanada et al., 2002; Garcia Costas et al., 2011).

FIGURE 3 | Phylogeny of WPS-2 built with concatenated ribosomal protein sequences, annotated with distribution of traits including phototrophy and aerobic respiration. Green branches encode both phototrophy and aerobic respiration; blue lineages encode aerobic respiration but not phototrophy; yellow lineages encode phototrophy but not aerobic respiration; black lineages encode neither.

Additional Metabolic Traits in WPS-2 Genomes

Aside from genes encoding phototrophy, the phototroph-containing WPS-2 clade is distinct from its relatives due to a variety of metabolic traits, including proteins used for aerobic respiration and carbon fixation, as well as evidence for an outer membrane.

Most WPS-2 genomes encode proteins consistent with at least a facultative aerobic lifestyle, consistent with their typical recovery from oxic environments and presence of O_2-using proteins such as AcsF as discussed above. The potential capacity for aerobic respiration in most WPS-2 genomes is encoded via an A-family heme-copper oxidoreductase (HCO) and a *bc* complex III. This is in contrast to the Chloroflexi, which typically encode an alternative complex III instead of a *bc* complex. Protein sequences of respiratory proteins from WPS-2 genomes primarily

form single closely related clades with topologies that broadly reflect organismal relationships (**Figure 3** and **Supplementary Figures S3, S4**); these trends suggest vertical inheritance of these genes from the last common ancestor of this clade. The main exceptions to these trends are the divergent WPS-2 genomes in the UBP9 class, which appear to be ancestrally anaerobic, and a handful of WPS-2 genomes that encode additional copies of respiratory proteins, which appear to have been acquired via horizontal gene transfer more recently and have since undergone additional HGT between members of the WPS-2. While the HCO genes in WPS-2 are interpreted here as being associated with aerobic respiration, proteins encoded by these genes are occasionally associated with O_2 detoxification in anaerobic organisms (e.g., Ward et al., 2018c); future isolation and culture-based characterization will be necessary to confidently determine the aerobic phenotype of these organisms.

Like the Chloroflexi, phototrophic WPS-2 typically encode a B-family HCO in addition to the ancestral A-family HCO used for respiration at relatively high O_2 concentrations (Ward et al., 2018a); the functional relationship between the B-family O_2 reductase and phototrophy is unclear, but may relate to oxygen sensitivity of phototrophy-related proteins and the greater efficacy of B-family HCOs at low oxygen concentrations (Han et al., 2011).

While phototrophic Chloroflexi typically perform carbon fixation via the 3-hydroxypropionate bi-cycle (Shih et al., 2017b), this pathway is not encoded in any available WPS-2 genomes. Several phototrophic WPS-2 genomes encode the potential capacity for carbon fixation via the Calvin cycle including a Form I rubisco and phosphoribulokinase (**Supplementary Table S1**). Several other phototrophic WPS-2 genomes recovered genes for either rubisco or phosphoribulokinase but not both, which could reflect an incomplete Calvin cycle or simply the failure to recover both genes in relatively incomplete genomes. Nearly all phototrophic WPS-2 genomes encode closely related rubisco proteins related to that of *Kouleothrix aurantiaca* (Ward et al., 2018a) and *Oscillochloris trichoides* (Kuznetsov et al., 2011), with some also encoding a second copy of rubisco that belongs to the Form IA clade that includes sequences from bacteria including *Nitrosospira*, *Prochlorococcus*, and *Ectothiorhodospira* (**Supplementary Figure S5**; Shih et al., 2016). The phylogeny of rubisco proteins from WPS-2 genomes is not congruent with organismal or phototrophy trees, suggesting that carbon fixation may have undergone an independent history of HGT, consistent with trends in the Chloroflexi (Shih et al., 2017b; Ward et al., 2018a). Further evidence for carbon fixation having an independent history from phototrophy in WPS-2 is the presence of rubisco genes in the PMHO01 and PMHP01 genomes despite these organisms not encoding phototrophy proteins; this rubisco is not closely related to those encoded by phototrophic WPS-2, and so may reflect the independent acquisition of carbon fixation in some non-phototrophic WPS-2 lineages (perhaps to support a lithoautotrophic lifestyle).

The Chloroflexi are thought to possess a modified monoderm membrane (Sutcliffe, 2010, 2011). WPS-2 genomes recovered genes for lipopolysaccharide synthesis and outer membrane proteins such as BamA, indicating that these organisms

are diderm (i.e., possess an outer membrane), reinforcing interpretations of an ancestral diderm membrane architecture followed by secondary loss of the outer membrane in the phylum Chloroflexi (Ward et al., 2017b, 2018a,b,c).

Type II Reaction Centers in WPS-2

The phylogeny of Type II reaction center proteins showed that the L and M subunits of WPS-2 are distant, and clearly distinct, to those found in Proteobacteria and Chloroflexi (**Figure 2**), but the topology of the tree was unstable. It showed varying positions and levels of support for the WPS-2 sequences (**Supplementary Figure S6**). At the sequence identity level, the L and M in WPS-2 share 30–40% sequence identity with those in Proteobacteria and Chloroflexi. In comparison, the level of sequence identity of the Type I reaction center core subunit (PscA) of phototrophic Acidobacteria has about 39% sequence identity to that of the Chlorobi, suggesting substantial divergence. However, the overall structural comparisons do reveal a number of structural parallels to those of the Chloroflexi, in particular of the chlorosome-lacking and relatively early-branching *Roseiflexus* spp. (Xin et al., 2018).

In addition to L and M subunits, the assembled genomes of WPS-2 also contain *pufC*, *puf2A*, and *puf2C* encoding the tetraheme cytochrome direct electron donor to the oxidized special pair, and the light harvesting complex alpha and beta subunits, respectively (Holland-Moritz et al., 2018). No *puhA* was observed indicating that the reaction center of WPS-2 might lack an H subunit in a manner similar to the CfRC.

Within phototrophic WPS-2, the reaction center subunits show considerable diversity, with the level of sequence identity between the two most distant L and two most distant M subunits being about 59%. In comparison, the level of sequence identity between L or M subunits between two relatively distant strains of Chloroflexi, *Roseiflexus* spp., and *Chloroflexus* spp., is about 45% sequence identity. If the obtained sequences are representative of the diversity of phototrophic WPS-2, the greater sequence identity could potentially indicate a relatively more recent origin for their common ancestor in comparison to phototrophic Chloroflexi.

Sequence and structure prediction of the L and M subunits revealed a large number of unique characteristics (**Figure 4**), which is consistent with the level of distinctness seen in the phylogenetic analysis. In Proteobacteria and Cyanobacteria each Type II reaction center protein is made of 5 transmembrane helices each. The recent Cryo-EM structure of the reaction center from *Roseiflexus castenholzii* resolved a 6th helix in the L subunit (Xin et al., 2018). This was located at the N-terminus and cannot be detected through secondary structure prediction methods. In the case of WPS-2 L subunit, transmembrane helix prediction only detected the standard 5 helices; instead a novel N-terminal 6th helix was detected in the M subunit (**Figure 4A**).

Most of the ligands to the special pair and cofactor are conserved, but unique changes that might modulate the energetics of electron transfer can be noted. For example, in the PbRC both "monomeric" bacteriochlorophylls, Bch_L and Bch_M, are coordinated by histidine ligands (**Figures 4B,C**). In the CfRC, a "third" bacteriopheophytin (Pierson et al., 1983;

Xin et al., 2018) occupies the position of Bch_M and accordingly there is no histidine ligand, and instead an isoleucine is found. In WPS-2 a histidine ligand to Bch_M is also absent like in the CfRC, however, instead of isoleucine and strictly conserved aspartate (M-D169) is found, which may suggests the presence of bacteriochlorophyll rather than bacteriopheophytin at this position, but with modified energetics (Heimdal et al., 2007). This M-D169 might play an important role in the inactivation of electron transfer via the M branch.
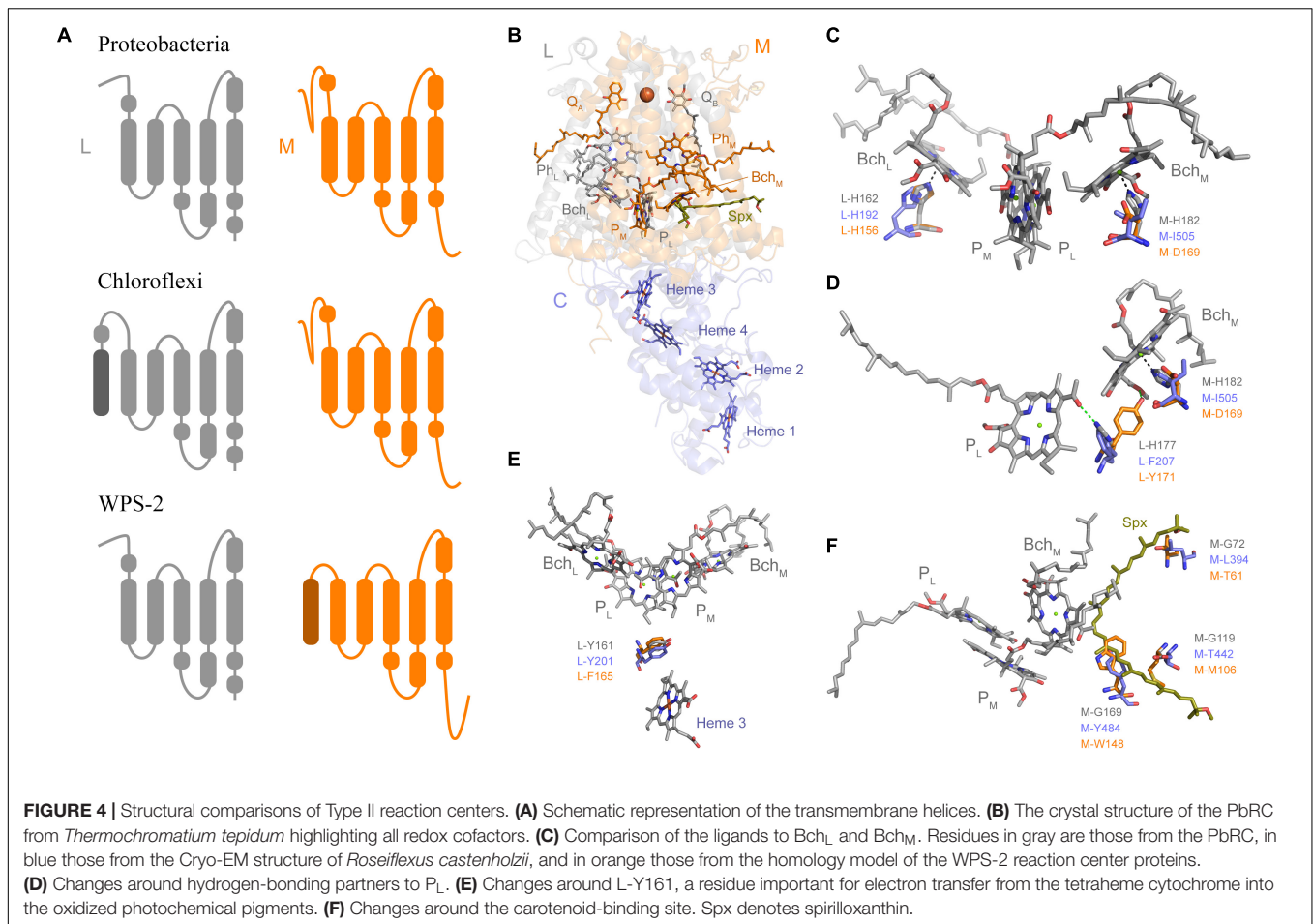
In the PbRC, L-H177 provides a hydrogen bond to the special pair bacteriochlorophyll P_L (**Figure 4D**). Mutagenesis studies have shown that the strength or absence of this bond modulates the midpoint potential of the primary donor pigments P_L and P_M (Lin et al., 1994). In the CfRC, the histidine is substituted by phenylalanine, which should contribute to the lower oxidizing potential of P in comparison to the PbRC. In the CfRC the midpoint potential of the primary donor is +360–390 mV (Bruce et al., 1982; Ivancich et al., 1996; Collins et al., 2011) while in the PbRC is +450–500 mV (Moss et al., 1991; Williams et al., 1992; Woodbury and Allen, 1995; Jones, 2009). In the WPS-2 reaction center, a tyrosine is found at this position and it is potentially oriented toward Bch_M effectively breaking the hydrogen-bond to P_L like in the CfRC. The absence of the hydrogen-bond to P_L could lower the midpoint potential of the primary donor by about −95 mV (Lin et al., 1994), which may suggest that WPS-2 has a primary donor with a potential more similar to that of the CfRC.

In the PbRC, a carotenoid is found in contact with Bch_M and it is thought that this protects the system against the formation of reactive oxygen species by quenching bacteriochlorophyll triplets (Cogdell et al., 2000), see **Figure 4F**. In the CfRC, no carotenoid has been reported at this position and the structure from *Roseiflexus* did not show a bound carotenoid (Xin et al., 2018). In the CfRC and the WPS-2 reaction center three glycine residues, which in the PbRC give space to the carotenoid, are replaced by bulkier residues that should hinder the binding of a potential carotenoid.

In the PbRC and the CfRC a strictly conserved tyrosine (L-Y162) is located in between heme 3 (c-559) of PufC and the special pair (**Figure 4E**). Mutagenesis studies have shown that this tyrosine modulates the midpoint potential of both the heme and the special pair, but it is not required for fast electron transfer (Dohse et al., 1995). While a gene was identified for PufC in WPS-2 (Holland-Moritz et al., 2018), this tyrosine was found to be a phenylalanine in all WPS-2 L sequences. A deprotonation of this conserved tyrosine residue stabilizes the primary charge separation reactions (Wohri et al., 2010), therefore this mechanism is not operational in the reaction center of WPS-2.

PERSPECTIVE AND CONCLUSION

Bacterial genomes previously assigned to WPS-2 have been assigned to a candidate phylum named *Candidatus* Eremiobacterota (Ji et al., 2017). Recent genome-resolved metagenomics has expanded the known genetic diversity of this phylum (Woodcroft et al., 2018) as well as revealing the capacity

FIGURE 4 | Structural comparisons of Type II reaction centers. **(A)** Schematic representation of the transmembrane helices. **(B)** The crystal structure of the PbRC from *Thermochromatium tepidum* highlighting all redox cofactors. **(C)** Comparison of the ligands to Bch$_L$ and Bch$_M$. Residues in gray are those from the PbRC, in blue those from the Cryo-EM structure of *Roseiflexus castenholzii*, and in orange those from the homology model of the WPS-2 reaction center proteins. **(D)** Changes around hydrogen-bonding partners to P$_L$. **(E)** Changes around L-Y161, a residue important for electron transfer from the tetraheme cytochrome into the oxidized photochemical pigments. **(F)** Changes around the carotenoid-binding site. Spx denotes spirilloxanthin.

for anoxygenic phototrophy (Holland-Moritz et al., 2018). Based on the metabolic traits of various *Ca.* Eremiobacterota lineages as discussed above, together with analysis via GTDB-Tk (Parks et al., 2018; **Supplementary Table S1**), we propose for the aerobic WPS-2 clade that includes phototrophic members the designation of a candidate order "Baltobacterales" (from the Greek for "bog" or "swamp," in reference to the typical habitat of these organisms) within the Eremiobacteria class of *Ca.* Eremiobacterota. Following current standards for the taxonomy of uncultured taxa (Chuvochina et al., 2019), we propose the genome WPS2_44 (IMG accession #2734482170) as the type species *Baltobacter phototrophicus* for the candidate ranks Baltobacterales (order) and Baltobacteraceae (family).

Most lineages of anoxygenic phototrophs can be preferentially found in characteristic environments unified only by the presence of light, such as anoxic regions of stratified water columns for Chlorobi (Imhoff, 2014), low-oxygen hot springs for Chloracidobacterium and phototrophic Chloroflexi (Bryant et al., 2007; Klatt et al., 2011; Hallenbeck et al., 2016; Ward et al., 2017a, 2018a), and soils for Heliobacteria (Madigan and Ormerod, 1995), though exceptions do occur e.g., Chloroflexi in carbonate tidal flats (Trembath-Reichert et al., 2016), and Heliobacteria in hot springs (Kimble et al., 1995) or soda lakes (Bryantseva et al., 1999; Asao et al., 2006). In contrast, the

phototrophic Proteobacteria and Gemmatimonadetes appear to have a more cosmopolitan distribution, including freshwater, marine, and soil environments (Zeng et al., 2016; Imhoff, 2017). The preferred environment for phototrophic members of *Ca.* Eremiobacterota appears to be cold, acidic, and aerobic environments with access to sunlight and at least some members of the group forming a close association with plants, particularly mosses. This niche overlaps with that of plant-associated phototrophic Proteobacteria (Atamna-Ismaeel et al., 2012), and may be due to the relatively high oxygen tolerance of these phototrophic lineages as compared to typically more oxygen-sensitive phototrophs such as Heliobacteria and Chlorobiaceae. The apparent close and specific association of phototrophic *Ca.* Eremiobacterota with plants may reflect a long-term evolutionary association, which could further suggest that this group has radiated alongside plants over a timescale of <0.5 billion years, though this hypothesis will require molecular clocks or other analyses to test.

Most members of the class *Ca.* Eremiobacteria (i.e., the candidate orders Baltobacterales and Eremiobacterales) encode aerobic respiration using closely related A-family heme-copper oxidoreductases and *bc* complex III proteins (∼85% of all genomes, and ∼97% of >80% completeness) (**Supplementary Figures S3, S4**). This is in contrast to the basal UBP9 class

which appears to be composed of obligate anaerobes, with no respiration genes encoded in any of the available genomes in this clade, and aerobic Chloroflexi which do not encode closely related complex III or complex IV proteins. This, together with the broad congruence of *Ca.* Eremiobacteria organismal phylogenies with complex III and complex IV protein phylogenies, implies that aerobic respiration was acquired via horizontal gene transfer into stem group *Ca.* Eremiobacteria after their divergence from UBP9 but before the radiation of crown group Eremiobacteria. It is likely that anaerobic ancestors of the Eremiobacteria and UBP9 classes diverged during Archean time, with the radiation of crown group Eremiobacteria occurring after the GOE ∼2.3 billion years ago which led to the expansion of aerobic metabolisms (Fischer et al., 2016). This hypothesis will require molecular clock estimates for the divergence and radiation of Eremiobacteria or other tests to verify, but similar evolutionary trends have been seen in other groups, including the Chloroflexi and Cyanobacteria (Fischer et al., 2016; Shih et al., 2017a,b). Following the acquisition of aerobic respiration by the last common ancestor of crown group Eremiobacteria, aerobic respiration via an A-family HCO and a *bc* complex III appears to have been largely vertically inherited, with additional respiratory proteins (e.g., second copies of *bc* complex and A-family HCOs, and the B-family HCOs associated with phototrophy) acquired via later HGT.

It is well established that all reaction centers originated as homodimers, with electron transfer occurring symmetrically on both side of the reaction center as it is still seen in Type I reaction centers (Hohmann-Marriott and Blankenship, 2008). It is thought that heterodimeric Type II reaction centers evolved from two independent duplication events, one leading to L and M in anoxygenic Type II and another leading to D1 and D2 in cyanobacterial PSII (Beanland, 1990; Sadekar et al., 2006; Cardona, 2015). The heterodimerization process led to the evolution of asymmetric electron transfer, which occurs exclusively via the L branch in anoxygenic Type II and via D1 in PSII, while the M and D2 branches became inactive, respectively. That the three different anoxygenic Type II reaction centers show distinct mechanisms of redox tuning on key positions around Bch_M and other photochemical pigments could indicate that these lineages started to radiate soon after the L and M duplication. Furthermore, given that anoxygenic Type II reaction centers show up to five times faster rates of evolution than cyanobacterial PSII (Cardona et al., 2019) and that the emergence of at least some of major clades of phototrophs with anoxygenic Type II reaction centers postdate the GEO (Shih et al., 2017b), it is possible that the duplication of L and M subunits, and the radiation of the known anoxygenic Type II reaction centers occurred after the origin of oxygenic photosynthesis (Cardona et al., 2019). Therefore, the idea that extant forms of anoxygenic phototrophy powered by heterodimeric Type II reaction centers represent a "primitive" form of photosynthesis is not supported by the available data. Instead, it should be considered a relatively recent, highly specialized, and mobile evolutionary innovation within the larger context of the evolution of anoxygenic photosynthesis, which reaches into Paleoarchean time (Tice and Lowe, 2004, 2006). If so, it would imply

that anoxygenic photosynthesis early in Earth's history may have been driven by now-extinct groups of bacteria using variations on phototropic pathways not yet observed today (Ward and Shih, 2019).

The known phylogenetic diversity of anoxygenic phototrophs has increased substantially in recent years, thanks largely to initial discovery of novel putative phototrophs via genome-resolved metagenomic sequencing of diverse environments (Bryant et al., 2007; Holland-Moritz et al., 2018; Ward et al., 2018a). Better coverage of the extant diversity of phototrophs and their relatives allows us to better query the evolutionary history of both organisms and their metabolisms by providing opportunities for comparative genomic, phylogenetic, and molecular clock analyses. Previous attempts to interpret the early evolution of photosynthesis relied on the narrower set of phototrophic lineages that were known at the time (Xiong et al., 2000; Raymond et al., 2002), whose analysis predated the discovery of phototrophic Chloracidobacterium, Gemmatimonadetes, Thermofonsia, and Eremiobacterota. The discovery of phototrophic Eremiobacterota and other lineages will provide crucial data for future reconsideration of the evolutionary relationships and history of phototrophy, improving our understanding of one of Earth's most important metabolic pathways.

AUTHOR'S NOTE

A version of this manuscript prior to peer review has been released as a preprint by Ward et al. (2019a).

DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

AUTHOR CONTRIBUTIONS

LW, TC, and HH-M designed the study, analyzed the data, and wrote the manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01658/full#supplementary-material

REFERENCES

Asao, M., Jung, D. O., Achenbach, L. A., and Madigan, M. T. (2006). *Heliorestis convoluta* sp. nov., a coiled, alkaliphilic heliobacterium from the wadi el natroun, egypt. *Extremophiles* 10, 403–410. doi: 10.1007/s00792-006-0513-4

Atamna-Ismaeel, N., Finkel, O., Glaser, F., Von Mering, C., Vorholt, J. A., Koblizek, M., et al. (2012). Bacterial anoxygenic photosynthesis on plant leaf surfaces. *Environ. Microbiol. Rep.* 4, 209–216. doi: 10.1111/j.1758-2229.2011.00323.x

Beanland, T. J. (1990). Evolutionary relationships between q-type photosynthetic reaction centers - hypothesis-testing using parsimony. *J. Theor. Biol.* 145, 535–545. doi: 10.1016/S0022-5193(05)80487-4

Bertin, P. N., Heinrich-Salmeron, A., Pelletier, E., Goulhen-Chollet, F., Arsene-Ploetze, F., Gallien, S., et al. (2011). Metabolic diversity among main microorganisms inside an arsenic-rich ecosystem revealed by meta- and proteo-genomics. *ISME J.* 5, 1735–1747. doi: 10.1038/ismej.2011.51

Bragina, A., Berg, C., and Berg, G. (2015). The core microbiome bonds the alpine bog vegetation to a transkingdom metacommunity. *Mol. Ecol.* 24, 4795–4807. doi: 10.1111/mec.13342

Brantner, J. S., Haake, Z. J., Burwick, J. E., Menge, C. M., Hotchkiss, S. T., and Senko, J. M. (2014). Depth-dependent geochemical and microbiological gradients in Fe(III) deposits resulting from coal mine-derived acid mine drainage. *Front. Microbiol.* 5:215. doi: 10.3389/fmicb.2014.00215

Brinkmann, H., Goker, M., Koblizek, M., Wagner-Dobler, I., and Petersen, J. (2018). Horizontal operon transfer, plasmids, and the evolution of photosynthesis in *Rhodobacteraceae*. *ISME J.* 12, 1994–2010. doi: 10.1038/s41396-018-0150-9

Bruce, B. D., Fuller, R. C., and Blankenship, R. E. (1982). Primary photochemistry in the facultatively aerobic green photosynthetic bacterium *Chloroflexus aurantiacus*. *Proc. Natl. Acad. Sci. U.S.A.* 79, 6532–6536. doi: 10.1073/pnas.79.21.6532

Bryant, D., Liu, Z., Li, T., Zhao, F., Klatt, C.G., Ward, D., et al. (2012). "Comparative and functional genomics of anoxygenic green bacteria from the taxa *Chlorobi, Chloroflexi*, and *Acidobacteria*," in *Functional Genomics and Evolution of photosynthetic Systems*, eds R.L. Burnap and W. Vermaas. (Dordrecht: Springer), 47–102. doi: 10.1007/978-94-007-1533-2_3

Bryant, D. A., Garcia Costas, A. M., Maresca, J. A., Chew, A. G. M., Klatt, C. G., Bateson, M. M., et al. (2007). *Candidatus* Chloracidobacterium thermophilum: an aerobic phototrophic acidobacterium. *Science* 317, 523–526. doi: 10.1126/science.1143236

Bryantseva, I. A., Gorlenko, V. M., Kompantseva, E. I., Achenbach, L. A., and Madigan, M. T. (1999). *Heliorestis daurensis*, gen. nov. sp. nov., an alkaliphilic rod-to-coiled-shaped phototrophic heliobacterium from a siberian soda lake. *Arch. Microbiol.* 172, 167–174. doi: 10.1007/s002030050756

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). Blast+: architecture and applications. *Bmc Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Camanocha, A., and Dewhirst, F. E. (2014). Host-associated bacterial taxa from Chlorobi, Chloroflexi, GN02, Synergistetes, SR1, TM7, and WPS-2 phyla/candidate divisions. *J. Oral Microbiol.* 6:25468. doi: 10.3402/jom.v6.25468

Canfield, D. E., Rosing, M. T., and Bjerrum, C. (2006). Early anaerobic metabolisms. *Philos. Trans. R. Soc. B* 361, 1819–1834; discussion 1835–1816. doi: 10.1098/rstb.2006.1906

Cardona, T. (2015). A fresh look at the evolution and diversification of photochemical reaction centers. *Photosynth. Res.* 126, 111–134. doi: 10.1007/s11120-014-0065-x

Cardona, T., Sanchez-Baracaldo, P., Rutherford, A. W., and Larkum, A. W. D. (2019). Early archean origin of Photosystem II. *Geobiology* 17, 127–150. doi: 10.1111/gbi.12322

Chew, A. G., and Bryant, D. A. (2007). Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity. *Annu. Rev. Microbiol.* 61, 113–129. doi: 10.1146/annurev.micro.61.080706.093242

Chuvochina, M., Rinke, C., Parks, D. H., Rappe, M. S., Tyson, G. W., Yilmaz, P., et al. (2019). The importance of designating type material for uncultured taxa. *Syst. Appl. Microbiol.* 42, 15–21. doi: 10.1016/j.syapm.2018.07.003

Cogdell, R. J., Howard, T. D., Bittl, R., Schlodder, E., Geisenheimer, I., and Lubitz, W. (2000). How carotenoids protect bacterial photosynthesis. *Philos. Trans. R. Soc. B* 355, 1345–1349. doi: 10.1098/rstb.2000.0696

Collins, A. M., Kirmaier, C., Holten, D., and Blankenship, R. E. (2011). Kinetics and energetics of electron transfer in reaction centers of the photosynthetic bacterium *Roseiflexus castenholzii. Biochim. Biophys. Acta* 1807, 262–269. doi: 10.1016/j.bbabio.2010.11.011

Costello, E. K. (2007). *Molecular Phylogenetic Characterization of High Altitude soil Microbial Communities and Novel, Uncultivated Bacterial Lineages.* PhD thesis, University of Colorado, Boulder, CO.

Dohse, B., Mathis, P., Wachtveitl, J., Laussermair, E., Iwata, S., Michel, H., et al. (1995). Electron transfer from the tetraheme cytochrome to the special pair in the *Rhodopseudomonas viridis* reaction center: effect of mutations of tyrosine l162. *Biochemistry* 34, 11335–11343. doi: 10.1021/bi00036a006

Doolittle, R. F. (1986). *Of Urfs and Orfs: A Primer on How to Analyze Derived Amino Acid Sequences.* Mill Valley, CA: University Science Books.

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

Fischer, W. W., Hemp, J., and Johnson, J. E. (2016). Evolution of oxygenic photosynthesis. *Annu. Rev. Earth Planet. Sci.* 44, 647–683. doi: 10.1146/annurev-earth-060313-054810

Garcia Costas, A. M., Tsukatani, Y., Romberger, S. P., Oostergetel, G. T., Boekema, E. J., Golbeck, J. H., et al. (2011). Ultrastructural analysis and identification of envelope proteins of "*Candidatus* Chloracidobacterium thermophilum" chlorosomes. *J. Bacteriol.* 193, 6701–6711. doi: 10.1128/Jb.06124-11

Gest, H., and Blankenship, R.E. (2005). "Time line of discoveries: anoxygenic bacterial photosynthesis, in *Discoveries in Photosynthesis*", eds Govindjee, J.T. Beatty, H. Gest J.F. (Dordrecht: Springer), 51–62. doi: 10.1007/1-4020-3324-9_6

Grasby, S. E., Richards, B. C., Sharp, C. E., Brady, A. L., Jones, G. M., and Dunfield, P. F. (2013). The paint pots, kootenay national park, canada - a natural acid spring analogue for mars. *Can. J. Earth Sci.* 50, 94–108. doi: 10.1139/e2012-060

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

Hallenbeck, P. C., Grogger, M., Mraz, M., and Veverka, D. (2016). Draft genome sequence of the photoheterotrophic *Chloracidobacterium thermophilum* strain oc1 found in a mat at ojo caliente. *Genome Announc.* 4:e1570-15. doi: 10.1128/genomeA.01570-15

Han, H., Hemp, J., Pace, L. A., Ouyang, H., Ganesan, K., Roh, J. H., et al. (2011). Adaptation of aerobic respiration to low O_2 environments. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14109–14114. doi: 10.1073/pnas.1018958108

Hanada, S., Takaichi, S., Matsuura, K., and Nakamura, K. (2002). *Roseiflexus castenholzii* gen. nov., sp nov., a thermophilic, filamentous, photosynthetic bacterium that lacks chlorosomes. *Int. J. Syst. Evol. Microbiol.* 52, 187–193.

Heimdal, J., Jensen, K. P., Devarajan, A., and Ryde, U. (2007). The role of axial ligands for the structure and function of chlorophylls. *J. Biol. Inorg. Chem.* 12, 49–61. doi: 10.1007/s00775-006-0164-z

Hessa, T., Meindl-Beinker, N. M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., et al. (2007). Molecular code for transmembrane-helix recognition by the sec61 translocon. *Nature* 450, 1026–1030. doi: 10.1038/Nature06387

Hohmann-Marriott, M. F., and Blankenship, R. E. (2011). Evolution of photosynthesis. *Annu. Rev. Plant Biol.* 62, 515–548. doi: 10.1146/annurev-arplant-042110-103811

Hohmann-Marriott, M.F., and Blankenship, R.E. (2008). "Anoxygenic type-i photosystems and evolution of photosynthetic reaction centers," in *Photosynthetic Protein Complexes*, ed. P. Fromme (Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA), 295–324. doi: 10.1002/9783527623464.ch13

Holland-Moritz, H., Stuart, J., Lewis, L. R., Miller, S., Mack, M. C., Mcdaniel, S. F., et al. (2018). Novel bacterial lineages associated with boreal moss species. *Environ. Microbiol.* 20, 2625–2638. doi: 10.1111/1462-2920.14288

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 of life. *Nat. Microbiol.* 1:16048. doi: 10.1038/nmicrobiol.2016.48

Imhoff, J.F. (2014). "The family *Chlorobiaceae*," in *The prokaryotes: Other major Lineages of Bacteria and the Archaea*, eds E. Rosenberg, E.F. Delong, S. Lory, E. Stackebrandt & F. Thompson (Berlin: Springer), 501–514. doi: 10.1007/978-3-642-38954-2_142

Imhoff, J.F. (2017). "Diversity of anaerobic anoxygenic phototrophic purple bacteria," in *Modern Topics in the Phototrophic Prokaryotes: Environmental and*

Applied Aspects, ed. P.C. Hallenbeck (Cham: Springer International Publishing), 47–85. doi: 10.1007/978-3-319-46261-5_2

Ivancich, A., Feick, R., Ertlmaier, A., and Mattioli, T. A. (1996). Structure and protein binding interactions of the primary donor of the *Chloroflexus aurantiacus* reaction center. *Biochemistry* 35, 6126–6135. doi: 10.1021/Bi952772r

Ji, M., Greening, C., Vanwonterghem, I., Carere, C. R., Bay, S. K., Steen, J. A., et al. (2017). Atmospheric trace gases support primary production in antarctic desert surface soil. *Nature* 552, 400–403. doi: 10.1038/nature25014

Ji, M., Van Dorst, J., Bissett, A., Brown, M. V., Palmer, A. S., Snape, I., et al. (2016). Microbial diversity at Mitchell Peninsula, eastern Antarctica: a potential biodiversity "hotspot". *Polar Biol.* 39, 237–249. doi: 10.1007/s00300-015-1776-y

Jones, M. R. (2009). The petite purple photosynthetic powerpack. *Biochem. Soc. Trans.* 37, 400–407. doi: 10.1042/Bst0370400

Kharecha, P., Kasting, J., and Siefert, J. (2005). A coupled atmosphere-ecosystem model of the early archean earth. *Geobiology* 3, 53–76. doi: 10.1111/j.1472-4669.2005.00049.x

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. *Arch. Microbiol.* 163, 259–267. doi: 10.1007/bf00393378

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

Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Kuznetsov, B. B., Ivanovsky, R. N., Keppen, O. I., Sukhacheva, M. V., Bumazhkin, B. K., Patutina, E. O., et al. (2011). Draft genome sequence of the anoxygenic filamentous phototrophic bacterium *Oscillochloris trichoides* subsp. DG-6. *J. Bacteriol.* 193, 321–322. doi: 10.1128/Jb.00931-10

Lefort, V., Longueville, J. E., and Gascuel, O. (2017). Sms: smart model selection in phyml. *Mol. Biol. Evol.* 34, 2422–2424. doi: 10.1093/molbev/msx149

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

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, W242-W245. doi: 10.1093/nar/gkw290

Lin, X., Murchison, H. A., Nagarajan, V., Parson, W. W., Allen, J. P., and Williams, J. C. (1994). Specific alteration of the oxidation potential of the electron donor in reaction centers from *Rhodobacter sphaeroides*. *Proc. Natl. Acad. Sci. U.S.A.* 91, 10265–10269. doi: 10.1073/pnas.91.22.10265

Madigan, M.T., and Ormerod, J.G. (1995). "Taxonomy, physiology and ecology of heliobacteria," in *Anoxygenic Photosynthetic Bacteria*, eds R.E. Blankenship, M.T. Madigan & C.E. Bauer (Dordrecht: Springer), 17–30. doi: 10.1007/0-306-47954-0_2

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 1–8.

Moss, D. A., Leonhard, M., Bauscher, M., and Mantele, W. (1991). Electrochemical redox titration of cofactors in the reaction center from *Rhodobacter sphaeroides*. *FEBS Lett.* 283, 33–36. doi: 10.1016/0014-5793(91)80547-G

Nogales, B., Moore, E. R. B., Llobet-Brossa, E., Rossello-Mora, R., Amann, R., and Timmis, K. N. (2001). Combined use of 16S ribosomal DNA and 16S rRNA to study the bacterial community of polychlorinated biphenyl-polluted soil. *Appl. Environ. Microbiol.* 67, 1874–1884. doi: 10.1128/Aem.67.4.1874-1884.2001

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* 36, 996–1004 doi: 10.1038/nbt.4229

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

Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P. A., Woodcroft, B. J. Evans, P. N., et al. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* 2, 1533–1542. doi: 10.1038/s41564-017-0012-7

Perreault, N. N., Greer, C. W., Andersen, D. T., Tille, S., Lacrampe-Couloume, G., Lollar, B. S., et al. (2008). Heterotrophic and autotrophic microbial populations in cold perennial springs of the high arctic. *Appl. Environ. Microbiol.* 74, 6898–6907. doi: 10.1128/aem.00359-08

Pierson, B. K., and Castenholz, R. W. (1971). Bacteriochlorophylls in gliding filamentous prokaryotes from hot springs. *Nat. New Biol.* 233, 25–27. doi: 10.1038/newbio233025a0

Pierson, B. K., Thornber, J. P., and Seftor, R. E. B. (1983). Partial purification, subunit structure and thermal stability of the photochemical reaction center of the thermophilic green bacterium *Chloroflexus aurantiacus*. *Biochim. Biophys. Acta* 723, 322–326. doi: 10.1016/0005-2728(83)90132-9

Raven, J. A. (2009). Contributions of anoxygenic and oxygenic phototrophy and chemolithotrophy to carbon and oxygen fluxes in aquatic environments. *Aquat. Microb. Ecol.* 56, 177–192. doi: 10.3354/ame01315

Raymond, J., Zhaxybayeva, O., Gogarten, J. P., Gerdes, S. Y., and Blankenship, R. E. (2002). Whole-genome analysis of photosynthetic prokaryotes. *Science* 298, 1616–1620. doi: 10.1126/science.1075558

Sadekar, S., Raymond, J., and Blankenship, R. E. (2006). Conservation of distantly related membrane proteins: photosynthetic reaction centers share a common structural core. *Mol. Biol. Evol.* 23, 2001–2007. doi: 10.1093/molbev/msl079

Serkebaeva, Y. M., Kim, Y., Liesack, W., and Dedysh, S. N. (2013). Pyrosequencing-based assessment of the bacteria diversity in surface and subsurface peat layers of a northern wetland, with focus on poorly studied phyla and candidate divisions. *PLoS One* 8:e63994. doi: 10.1371/journal.pone.0063994

Shih, P. M., Hemp, J., Ward, L. M., Matzke, N. J., and Fischer, W. W. (2017a). Crown group Oxyphotobacteria postdate the rise of oxygen. *Geobiology* 15, 19–29. doi: 10.1111/gbi.12200

Shih, P. M., Ward, L. M., and Fischer, W. W. (2017b). Evolution of the 3-hydroxypropionate bicycle and recent transfer of anoxygenic photosynthesis into the Chloroflexi. *Proc. Natl. Acad. Sci. U.S.A.* 114, 10749–10754. doi: 10.1073/pnas.1710798114

Shih, P. M., Occhialini, A., Cameron, J. C., Andralojc, P. J., Parry, M. A., and Kerfeld, C. A. (2016). Biochemical characterization of predicted precambrian rubisco. *Nat. Commun.* 7:10382. doi: 10.1038/ncomms10382

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W. Z., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75

Sousa, F. L., Shavit-Grievink, L., Allen, J. F., and Martin, W. F. (2013). Chlorophyll biosynthesis gene evolution indicates photosystem gene duplication, not photosystem merger, at the origin of oxygenic photosynthesis. *Genome Biol. Evol.* 5, 200–216. doi: 10.1093/Gbe/Evs127

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

Stibal, M., Schostag, M., Cameron, K. A., Hansen, L. H., Chandler, D. M., Wadham, J. L., et al. (2015). Different bulk and active bacterial communities in cryoconite from the margin and interior of the greenland ice sheet. *Environ. Microbiol. Rep.* 7, 293–300. doi: 10.1111/1758-2229.12246

Sutcliffe, I. C. (2010). A phylum level perspective on bacterial cell envelope architecture. *Trends Microbiol.* 18, 464–470. doi: 10.1016/j.tim.2010.06.005

Sutcliffe, I. C. (2011). Cell envelope architecture in the *Chloroflexi*: a shifting frontline in a phylogenetic turf war. *Environ. Microbiol.* 13, 279–282. doi: 10.1111/j.1462-2920.2010.02339.x

Tahon, G., and Willems, A. (2017). Isolation and characterization of aerobic anoxygenic phototrophs from exposed soils from the Sør Rondane Mountains, East Antarctica. *Syst. Appl. Microbiol.* 40, 357–369. doi: 10.1016/j.syapm.2017.05.007

Thiel, V., Tank, M., and Bryant, D. A. (2018). Diversity of chlorophototrophic bacteria revealed in the omics era. *Annu. Rev. Plant Biol.* 69, 21–49. doi: 10.1146/annurev-arplant-042817-040500

Tice, M. M., and Lowe, D. R. (2004). Photosynthetic microbial mats in the 3,416-myr-old ocean. *Nature* 431, 549–552. doi: 10.1038/nature02888

Tice, M. M., and Lowe, D. R. (2006). Hydrogen-based carbon fixation in the earliest known photosynthetic organisms. *Geology* 34, 37–40. doi: 10.1130/G22012.1

Trembath-Reichert, E., Ward, L. M., Slotznick, S. P., Bachtel, S. L., Kerans, C., Grotzinger, J. P., et al. (2016). Gene sequencing-based analysis of microbial-mat morphotypes, caicos platform, british west indies. *J. Sediment. Res.* 86, 629–636. doi: 10.2110/jsr.2016.40

Trexler, R., Solomon, C., Brislawn, C. J., Wright, J. R., Rosenberger, A., Mcclure, E. E., et al. (2014). Assessing impacts of unconventional natural gas extraction on microbial communities in headwater stream ecosystems in northwestern pennsylvania. *Front. Microbiol.* 5:522. doi: 10.3389/fmicb.2014.00522

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

Ward, L. M., Mcglynn, S. E., and Fischer, W. W. (2018b). Draft genome sequence of a divergent anaerobic member of the *Chloroflexi* class *Ardenticatenia* from a sulfidic hot spring. *Genome Announc.* 6:e571-18. doi: 10.1128/genomeA. 00571-18

Ward, L. M., Mcglynn, S. E., and Fischer, W. W. (2018c). Draft genome sequences of two basal members of the *Anaerolineae* class of *Chloroflexi* from a sulfidic hot spring. *Genome Announc.* 6:e570-18. doi: 10.1128/genomeA.00570-18

Ward, L. M., Shih, P. M., and Fischer, W. W. (2018d). Metapoap: presence or absence of metabolic pathways in metagenome-assembled genomes. *Bioinformatics* 34, 4284–4289. doi: 10.1093/bioinformatics/bty510

Ward, L. M., Mcglynn, S. E., and Fischer, W. W. (2017a). Draft genome sequence of *Chloracidobacterium* sp. CP2_5A, a phototrophic member of the phylum *Acidobacteria* recovered from a Japanese hot spring. *Genome Announc.* 5:e821-17. doi: 10.1128/genomeA.00821-17

Ward, L. M., Mcglynn, S. E., and Fischer, W. W. (2017b). Draft genome sequences of a novel lineage of *Armatimonadetes* recovered from Japanese hot springs. *Genome Announc.* 5:e820-17. doi: 10.1128/genomeA.00820-17

Ward, L. M., Cardona, T., and Holland-Moritz, H. (2019a). Evolutionary implications of anoxygenic phototrophy in the bacterial phylum candidatus eremiobacterota (WPS-2). *bioRxiv* [Preprint]. doi: 10.1101/534133

Ward, L. M., Rasmussen, B., and Fischer, W. W. (2019b). Primary productivity was limited by electron donors prior to the advent of oxygenic photosynthesis. *J. Geophys. Res. Biogeo.* 124, 211–226. doi: 10.1029/2018jg004679

Ward, L.M., and Shih, P.M. (2019). The evolution and productivity of carbon fixation pathways in response to changes in oxygen concentration over geological time. *Free Radic. Biol. Med.* [Epub ahead of print]

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296-W303. doi: 10.1093/nar/gky427

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., 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

Williams, J. C., Alden, R. G., Murchison, H. A., Peloquin, J. M., Woodbury, N. W., and Allen, J. P. (1992). Effects of mutations near the bacteriochlorophylls in reaction centers from *Rhodobacter sphaeroides*. *Biochemistry* 31, 11029–11037. doi: 10.1021/bi00160a012

Wohri, A. B., Katona, G., Johansson, L. C., Fritz, E., Malmberg, E., Andersson, M., et al. (2010). Light-induced structural changes in a photosynthetic reaction center caught by Laue diffraction. *Science* 328, 630–633. doi: 10.1126/science. 1186159

Woodbury, N.W., and Allen, J.P. (1995). "The pathway, kinetics and thermodynamics of electron transfer in wild type and mutant reaction centers of purple nonsulfur bacteria," in *Anoxygenic photosynthetic bacteria: Advances in photosynthesis*, eds R.E. Blankenship, M.T. Madigan & C.E. Bauer. (Dordrecht: Kluwer Academic Publishers), 527–557. doi: 10.1007/0-306-47954-0_24

Woodcroft, B. J., Singleton, C. M., Boyd, J. A., Evans, P. N., Emerson, J. B., Zayed, A. A. F., et al. (2018). Genome-centric view of carbon processing in thawing permafrost. *Nature* 560, 49–54. doi: 10.1038/s41586-018-0338-1

Xin, Y. Y., Shi, Y., Niu, T. X., Wang, Q. Q., Niu, W. Q., Huang, X. J., et al. (2018). Cryo-EM structure of the RC-LH core complex from an early branching photosynthetic prokaryote. *Nat. Commun.* 9:1568. doi: 10.1038/s41467-018-03881-x

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

Yu, L. J., Suga, M., Wang-Otomo, Z. Y., and Shen, J. R. (2018). Structure of photosynthetic LH1-RC supercomplex at 1.9 Å resolution. *Nature* 556, 209–213. doi: 10.1038/s41586-018-0002-9

Zeng, Y., Baumbach, J., Barbosa, E. G., Azevedo, V., Zhang, C., and Koblizek, M. (2016). Metagenomic evidence for the presence of phototrophic Gemmatimonadetes bacteria in diverse environments. *Environ. Microbiol. Rep.* 8, 139–149. doi: 10.1111/1758-2229.12363

Zeng, Y. H., Feng, F. Y., Medova, H., Dean, J., and Koblizek, 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