

Supplementary Figures for

Photosynthesis is not a Universal Feature of the Phylum Cyanobacteria

Rochelle M. Soo, Connor T. Skennerton, Yuji Sekiguchi, Michael Imelfort, Samuel J. Paech, Paul G. Dennis, Jason A. Steen, Donovan H. Parks, Gene W. Tyson, and Philip Hugenholtz[†]

[†] Correspondence to: Philip Hugenholtz, p.hugenholtz@uq.edu.au

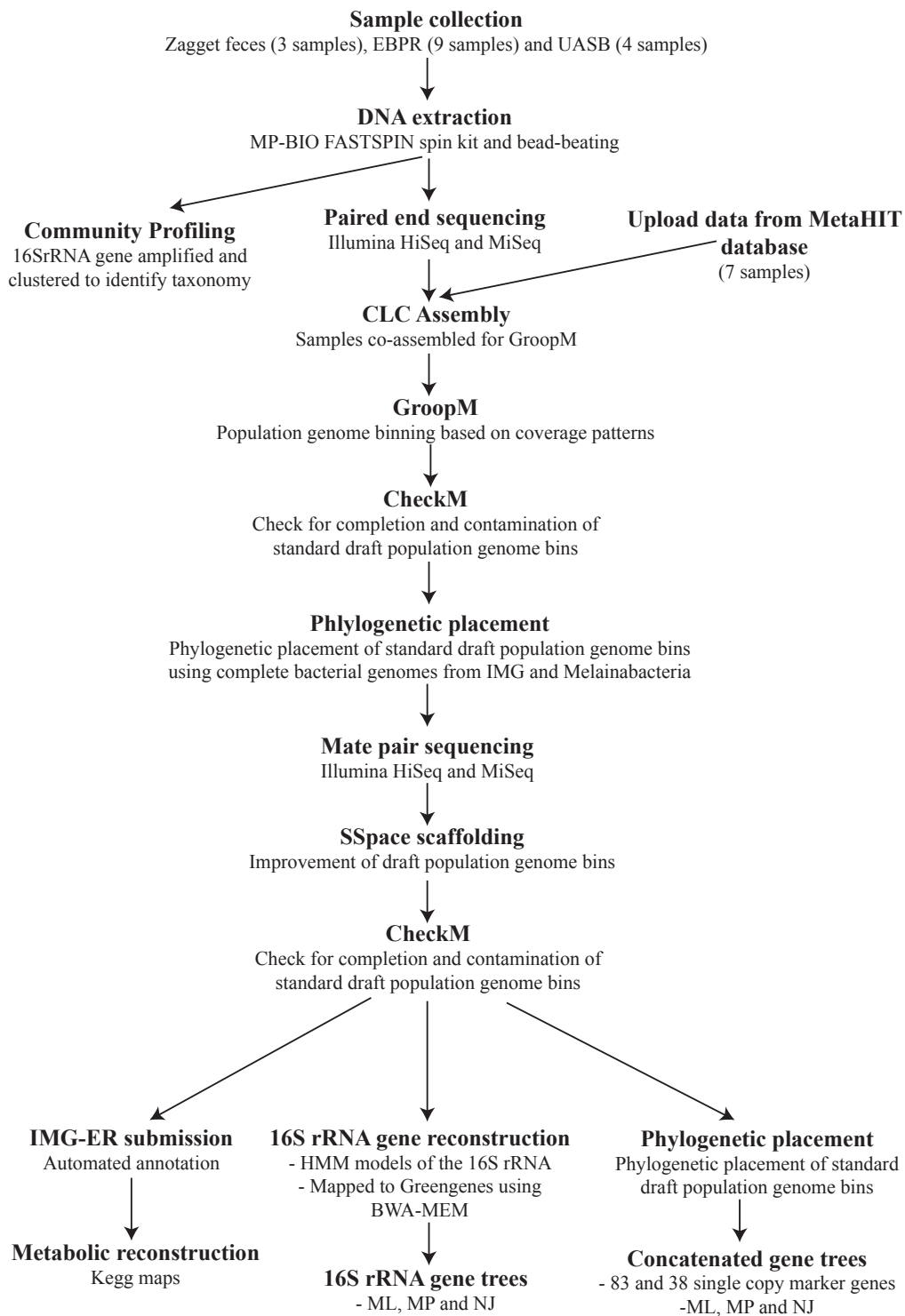
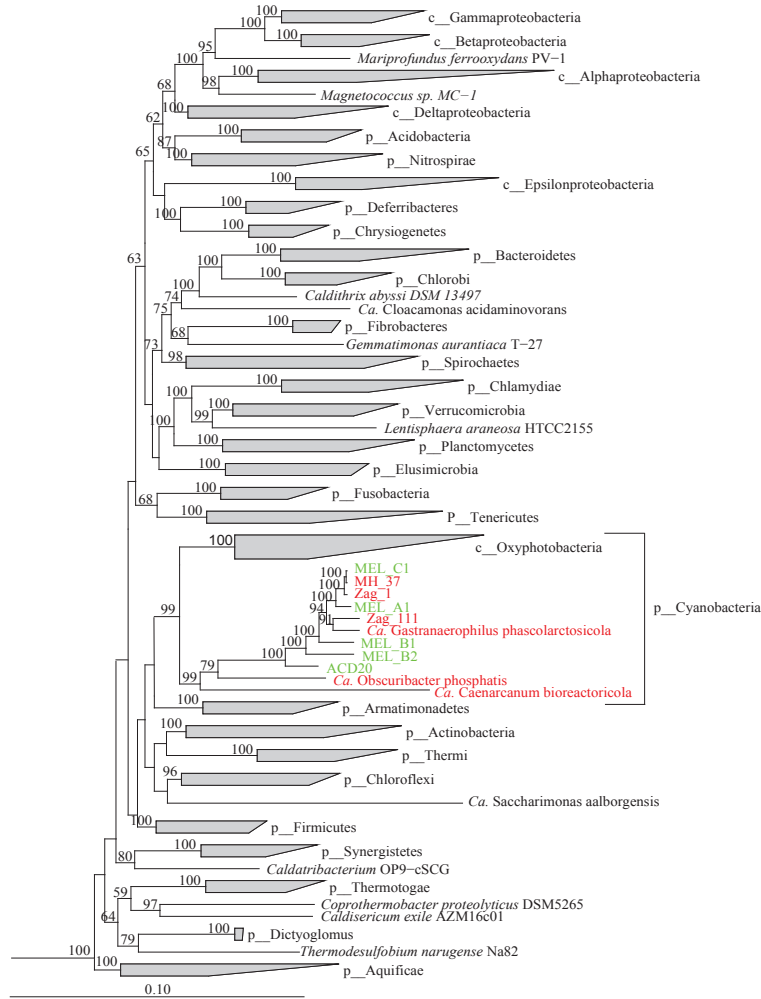
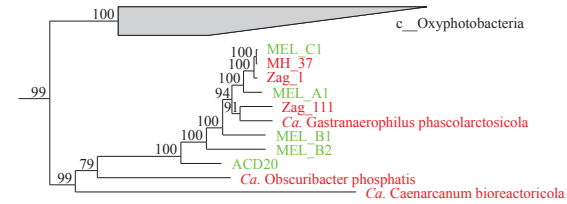


Fig. S1. Flow diagram of the Methods used in this paper

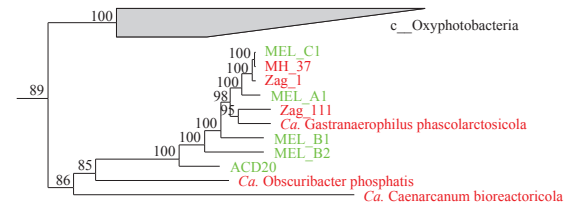
ML (RAxML, JTT+Gamma, 100x bootstrapping) rooted with Archaea



ML (RAxML, JTT+Gamma, 100x bootstrapping)



ML (FastTree, JTT+CAT, 100x bootstrapping) on RAxML tree



MP (PAUP*, 100x bootstrapping, >50% consensus tree)

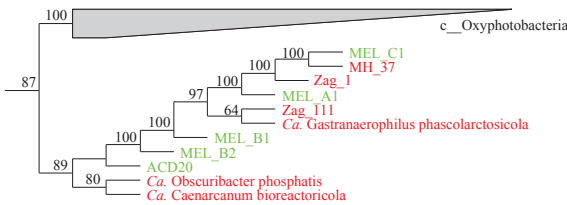
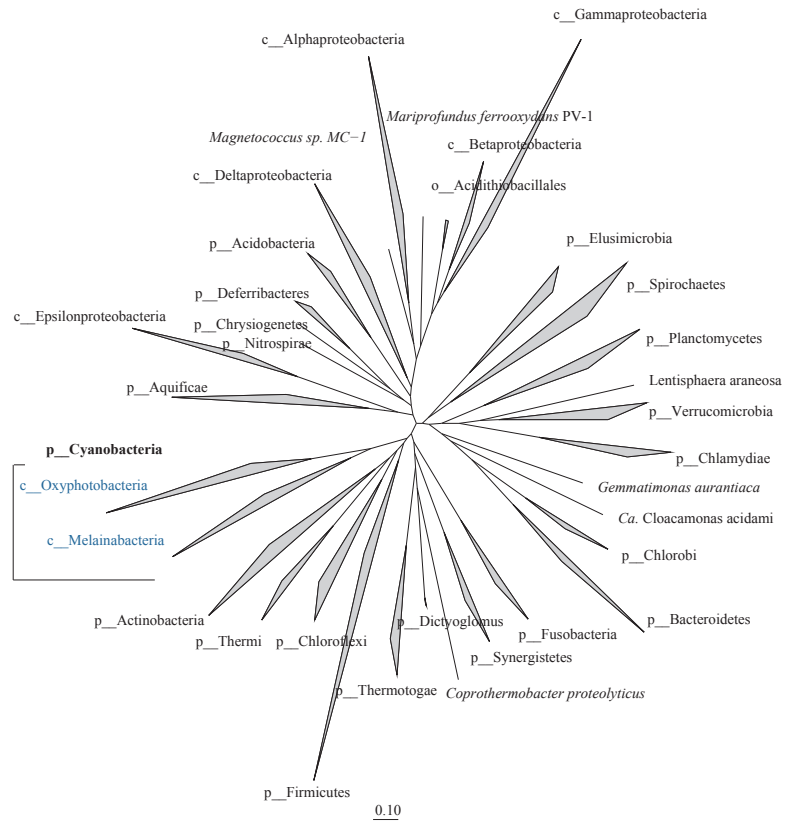


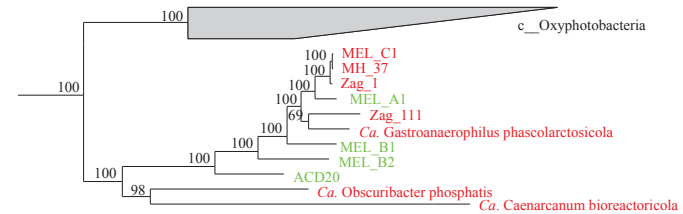
Fig. S2. Phylogeny of Oxyphotobacteria and Melainabacteria genomes among the bacterial phyla based on up to 38 marker genes

Phylogenetic maximum likelihood (RAxML, JTT, G) and maximum parsimony (PAUP*) trees based on a concatenated alignment of up to 38 marker genes, showing the phylogenetic robustness of the phylum Cyanobacteria and the classes Oxyphotobacteria and Melainabacteria. 422 OTUs (operational taxonomic units) from Bacteria and Archaea were used to produce the phylogenetic tree (**Table S4**). Bootstrap analyses (100 times) were performed for the data set with maximum likelihood (RAxML, JTT, G; FastTree, JTT, CAT) and maximum parsimony (PAUP*) methods, and the values obtained are shown in respective trees, except for the values from FastTree, which are shown on the tree generated with RAxML bootstrapping. Genomes in green are representatives from Di Rienzi et al., 2013 and genomes in red are Melainabacteria from this study. Bootstrap values >60% are shown. p_ represents phylum and c_ represents class. *Candidatus* has been abbreviated to *Ca.* for the most complete genomes.

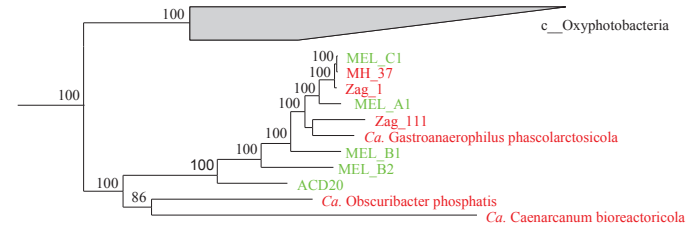
ML (RAxML, JTT+Gamma, 100x bootstrapping) unrooted



ML (RAxML, JTT+Gamma, 100x bootstrapping)



ML (FastTree, JTT+CAT, 100x bootstrapping) on RAxML tree



MP (PAUP*, 100x bootstrapping, >50% consensus tree)

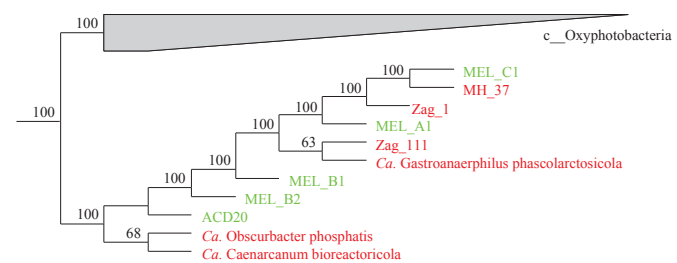
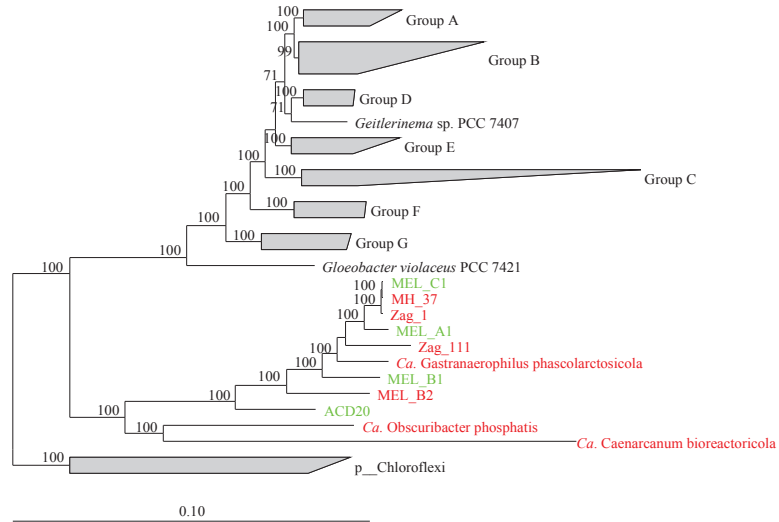


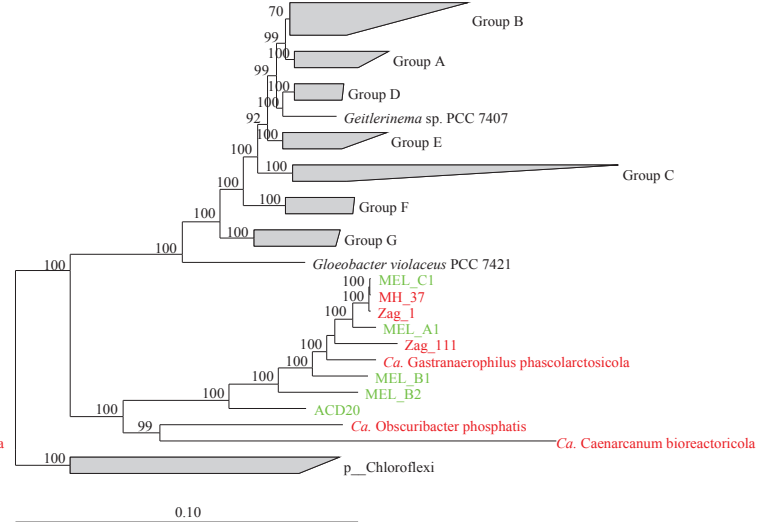
Fig. S3. Phylogeny of Oxyphotobacteria and Melainabacteria among the bacterial phyla based on up to 83 marker genes

Phylogenetic maximum likelihood (RAxML, JTT, G) and maximum parsimony (PAUP*) trees based on a concatenated alignment of up to 83 marker genes, showing the phylogenetic robustness of the phylum Cyanobacteria and the classes Oxyphotobacteria and Melainabacteria. 322 OTUs from Bacteria were used to produce the phylogenetic tree (**Table S4**). Bootstrap analyses (100 times) were performed for the data set with maximum likelihood (RAxML, JTT, G; FastTree, JTT, CAT) and maximum parsimony (PAUP*) methods, and the values obtained are shown in respective trees, except for the values from FastTree, which are shown on the tree generated with RAxML. Genomes in green are representatives from Di Rienzi et al., 2013 and genomes in red are Melainabacteria from this study. Bootstrap values >60% are shown. p_ represents phylum, c_ represents class and o_ represents order. *Candidatus* has been abbreviated to *Ca.* for the most complete genomes.

ML (RAxML, JTT+Gamma, 100x bootstrapping)



ML (FastTree, JTT+CAT, 100x bootstrapping) on RAxML tree



MP (PAUP*, 100x bootstrapping, >50% consensus tree)

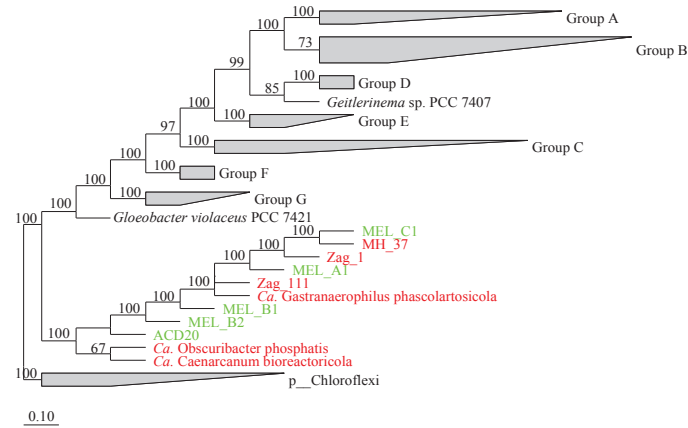


Fig. S4. Phylogeny of Oxyphotobacteria and Melainabacteria in the Cyanobacteria phylum based on up to 83 marker genes

Phylogenetic maximum likelihood (RAxML, JTT, G) and maximum parsimony (PAUP*) trees based on a concatenated alignment of up to 83 marker genes, showing the phylogenetic robustness of the phylum Cyanobacteria and the classes with Chloroflexi used as the outgroup. The Oxyphotobacteria were grouped as according to Shih et al., 2013 (**Table S3**). Bootstrap analyses (100 times) were performed for the data set with maximum likelihood (RAxML, JTT, G; FastTree, JTT, CAT) and maximum parsimony (PAUP*) methods, and the values obtained are shown in respective trees, except for the values from FastTree, which are shown on the tree generated with RAxML. Genomes in green are representatives from Di Rienzi et al., 2013 and genomes in red are Melainabacteria from this study. Bootstrap values >60% are shown. p_ represents phylum, c_ represents class and o_ represents order. *Candidatus* has been abbreviated to *Ca.* for the most complete genomes.

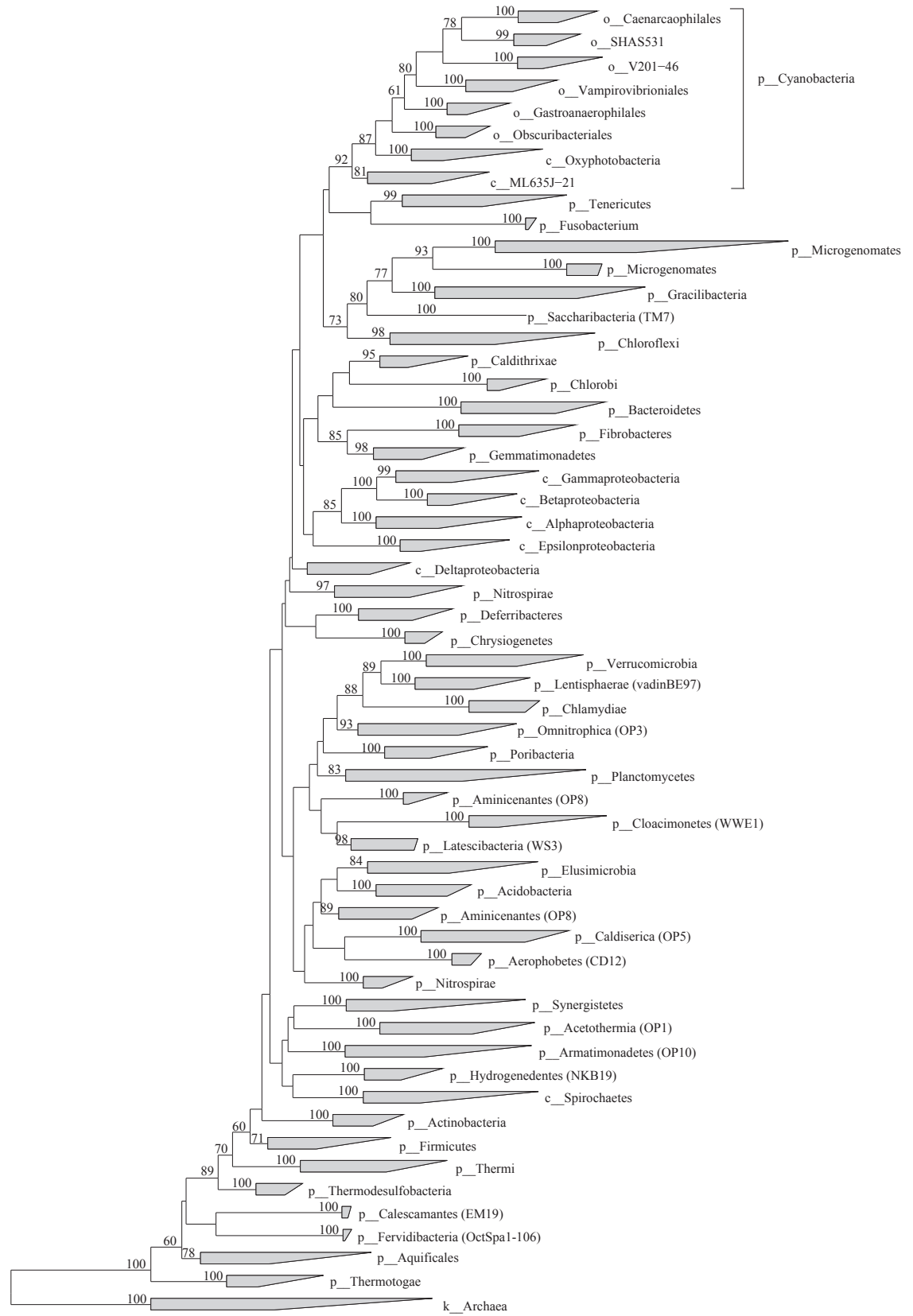


Fig. S5. 16S rRNA gene tree showing the phylogenetic robustness of the Cyanobacteria and its class level lineages in the bacterial domain

Maximum likelihood (RAxML, GTR, G, I; FastTree, JTT, CAT) trees based on the 16S rRNA gene, showing the phylogenetic robustness of the classes Oxyphotobacteria and Melainabacteria. 418 OTUs (operational taxonomic units) from Bacteria and Archaea were used to construct the trees. Bootstrap analyses (100 times) were performed for the data set with maximum likelihood (RAxML, GTR, G, I; FastTree, JTT, CAT). Bootstrap values >60% are shown.

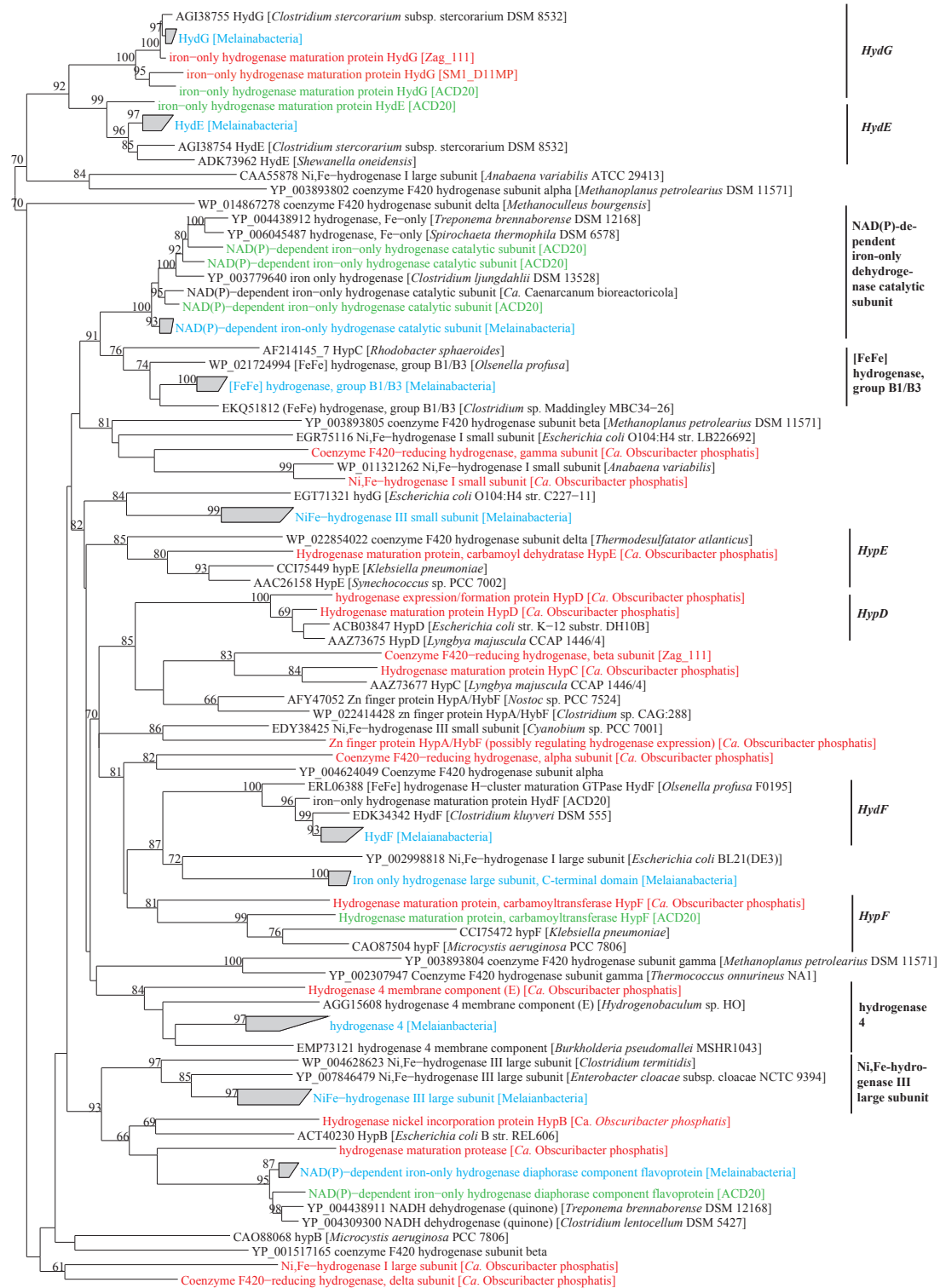


Fig. S6. Gene tree showing the hydrogenase genes found in the Melainabacteria representatives.

A representative collection of hydrogenase genes publicly available from NCBI was combined with genes identified from this study and from Di Rienzi et al., 2013. Hydrogenase genes common to genomes from both studies are colored blue and grouped together. Hydrogenases found only in this study are colored red and genes only identified in the genomes of Di Rienzi et al., 2013 are colored green.

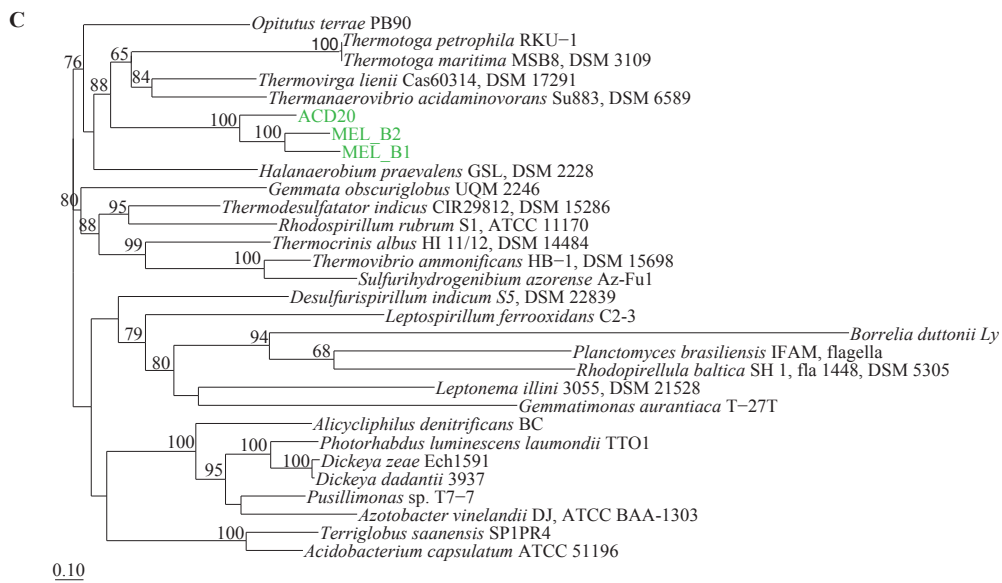
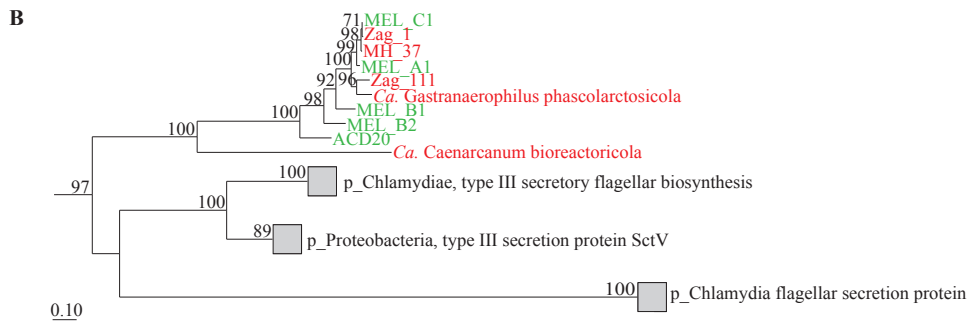
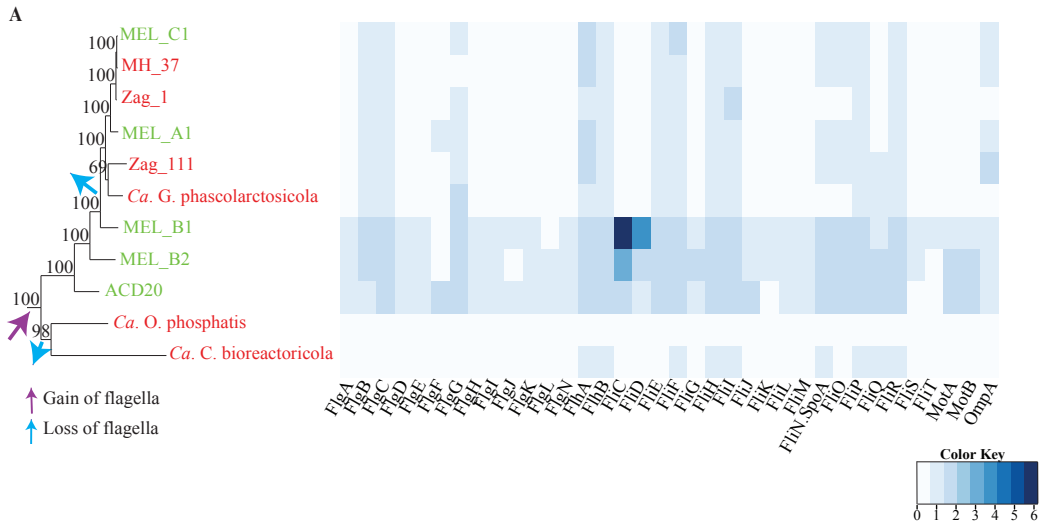


Fig. S7. Heat map showing the presence and absence of flagella genes from the Melainabacteria representatives

(A) A maximum parsimony tree based on a concatenated alignment of up to 83 marker genes. Purple arrows indicate the gain of a functional flagella and blue arrows indicate the loss of a functional flagella. The heat map indicates the square root of the number of flagella genes found in each of the Melainabacteria representatives. (B) Flagella gene trees for FlhA and (C) flagella gene tree for FliM. Genomes used to make the trees are found in **Table S6**. Genomes in green are Melainabacteria representatives from Di Rienzi et al., 2013, genomes in red are Melainabacteria representatives from this study. *Candidatus* has been abbreviated to *Ca.* for the most complete genomes.

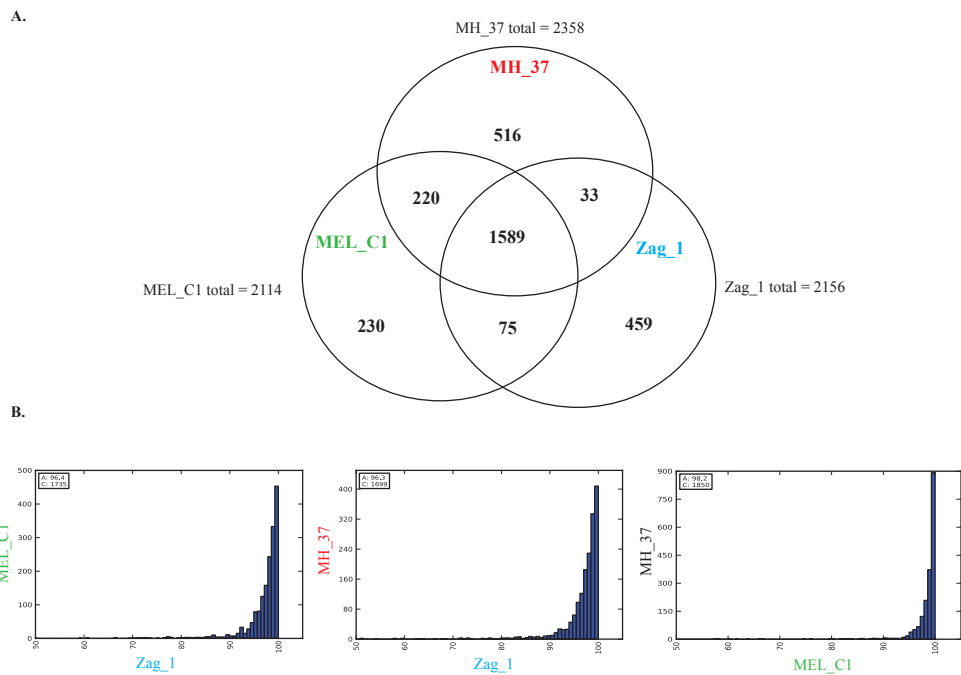


Fig. S8. Venn diagram of the three Gastranaerophilales genomes from the same species

(A) Shared orthologs between MH_37, MEL_C1 and Zag_1 showing the number of genes that are core to the three genomes, and those that are found between two genomes and individual genomes. The total number for each genome does not include paralogs. (B) Average nucleotide identity between MH_37, MEL_C1 and Zag_1, where $\geq 95\%$ ANI is used to define a species.

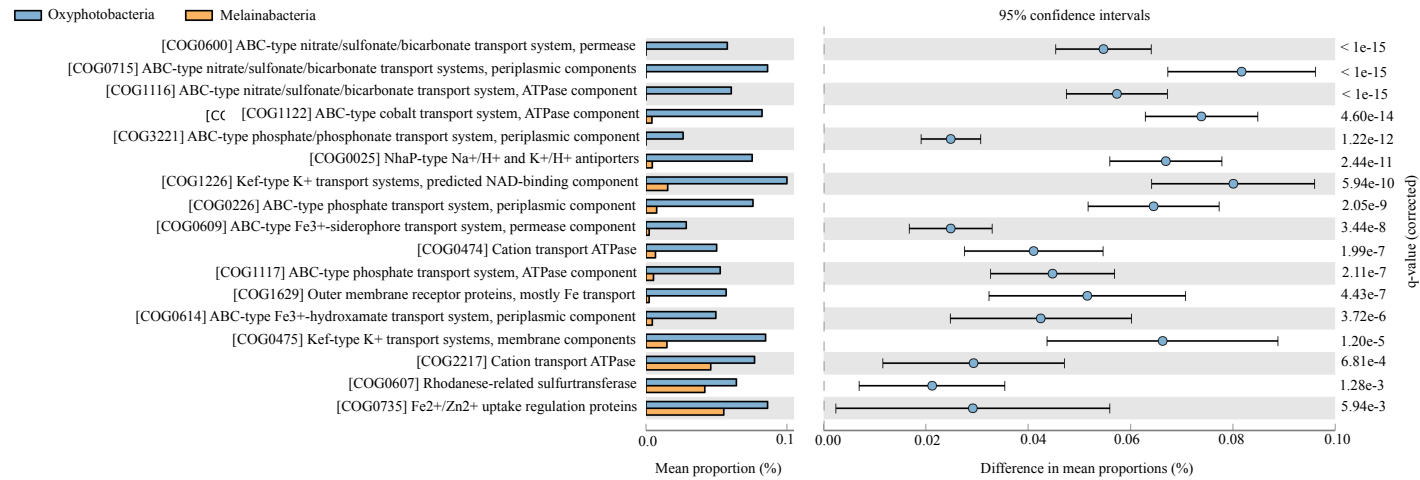


Fig. S9. COG category P (inorganic ion transport and metabolism) for Oxyphotobacteria and Melainabacteria

Shown are all COGs from category P with a difference in mean proportions $\geq 0.02\%$ between Oxyphotobacteria and Melainabacteria and a Storey's q-value of ≤ 0.05 .

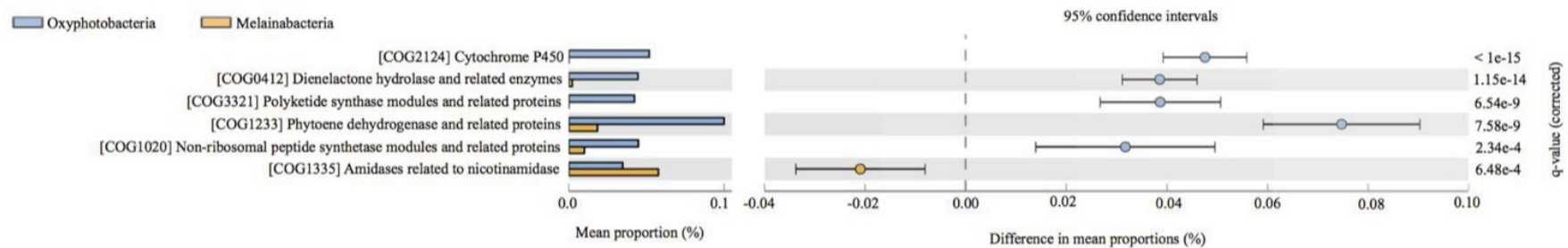


Fig. S10. COG category Q (secondary metabolites and biosynthesis, transport and catabolism) for Oxyphotobacteria and Melainabacteria

Shown are all COGs from category Q with a difference in mean proportions $\geq 0.02\%$ between Oxyphotobacteria and Melainabacteria and a Storey's q-value of ≤ 0.05