Anoxygenic phototrophic *Chloroflexota* member uses a Type I reaction center

J. M. Tsuji[1]*, N. A. Shaw[1], S. Nagashima[2], J. J. Venkiteswaran[1,3],
S. L. Schiff[1], S. Hanada[2], M. Tank[2,4]*, J. D. Neufeld[1]*

[1]University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada, N2L 3G1.

[2]Tokyo Metropolitan University, 1-1 Minami-osawa, Hachioji, Tokyo, Japan, 192-0397.

[3]Wilfrid Laurier University, 75 University Avenue West, Waterloo, Ontario, Canada, N2L 3C5.

[4]Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Inhoffenstrasse 7B, 38124 Braunschweig, Germany.

*Correspondence to: jackson.tsuji@uwaterloo.ca; marcus.tank@dsmz.de; jneufeld@uwaterloo.ca

Abstract:

Phototrophic bacteria within the *Chloroflexota* phylum are puzzling in their evolutionary origin. Previously known phototrophic *Chloroflexota* members use a Type II photosynthetic reaction center for light energy conversion but contain other photosynthesis machinery associated with Type I reaction center-utilizing phototrophs. We sampled an iron-rich boreal lake at the IISD-Experimental Lakes Area and enriched '*Candidatus* Chlorohelix allophototropha', a phototrophic *Chloroflexota* member that uses a Type I reaction center. Phylogenomic evidence suggests that ancestors of '*Ca.* Chx. allophototropha' served as a bridge for historic phototrophy gene exchange within the phylum. The *Chloroflexota* now represents the only bacterial phylum outside the *Cyanobacteria* where both major classes of photosynthetic reaction center occur and can serve as a model system to explore fundamental questions about the evolution of photosynthesis.

Main text:

Known bacterial members of the *Chloroflexota* (*1*) phylum are widely distributed in natural environments and have broad metabolic potential (*2–4*). Anoxygenic chlorophototrophs in the phylum (also called filamentous anoxygenic phototrophs or "green non-sulfur bacteria") were first cultivated in 1974 (*5*) but have presented an evolutionary puzzle. Although these bacteria use a Type II reaction center (RCII) for phototrophy, several phototrophic *Chloroflexota* members also contain chlorosomes, a light-harvesting apparatus that is otherwise unique to Type I reaction center (RCI)-utilizing anoxygenic phototrophs (*6*). Did microorganisms using RCI and RCII for phototrophy interact in the evolutionary past? Some form of genetic interaction between RCI and RCII would have been necessary for oxygenic photosynthesis to arise, given that oxygenic photosynthesis relies on RCII and RCI in tandem as Photosystem II (PSII) and Photosystem I (PSI), respectively. Despite much speculation (*7*, *8*), direct evidence is lacking for such genetic interactions outside the *Cyanobacteria* phylum. Here we report the cultivation of a highly novel phototroph within the *Chloroflexota* phylum that uses RCI and chlorosomes for phototrophy. We describe the physiological and genomic properties of the novel bacterium and its implications for the evolution of phototrophy.

Water was collected from the illuminated and anoxic water column of Lake 227, an iron-rich Boreal Shield lake at the International Institute for Sustainable Development-Experimental Lakes Area (IISD-ELA), to enrich for resident bacterial phototrophs (*9*). After approximately two months of incubation under light, enrichment cultures began showing ferrous iron oxidation activity (Supplementary Text S1) and 16S ribosomal RNA gene amplicon sequencing indicated the presence of a novel bacterium distantly related to known *Chloroflexota* members. Further enrichment in deep agar dilution series eliminated a purple phototrophic bacterium related to the

metabolically versatile *Rhodopseudomonas palustris* (*10*) and a green phototrophic bacterium belonging to the *Chlorobium* genus. Metagenome sequencing and assembly recovered the nearly closed genome of the *Chloroflexota* member (Supplementary Text S2), which represented the only detectable phototroph and comprised 43.4% of raw metagenome sequences (Fig. S1). Given the detection of genes for RCI in the *Chloroflexota* member's genome, we provisionally classify this bacterium as, '*Candidatus* Chlorohelix allophototropha' strain L227-S17 (Chlo.ro.he'lix. Gr. adj. *chloros*, green; Gr. fem. n. *helix*, spiral; N.L. fem. n. *Chlorohelix*, a green spiral. al.lo.pho.to.tro'pha. Gr. adj. *allos*, other, different; Gr. n. *phos*, *-otos*, light; N.L. suff. *-tropha*, feeding; N.L. fem. adj. *allophototropha*, phototrophic in a different way).

Physiologically, '*Ca.* Chx. allophototropha' resembled other cultured phototrophs belonging to the *Chloroflexota* phylum (Fig. 1). Cells grew well in soft agar (Fig. 1A) and developed as long spiraling filaments (Fig. 1B), similar to those observed for *Chloronema giganteum* (*11*), that were composed of cells with dimension of ~2-3 µm x 0.6 µm (Fig. 1C-D). In addition to the presence of light-harvesting chlorosomes (Fig. 1E-F), '*Ca.* Chx. allophototropha' cells were pigmented, with the dominant pigment being bacteriochlorophyll *c*, based on the 433 and 663 nm peaks observed in absorption spectrum data from acetone:methanol pigment extracts (Fig. 1G; see Supplementary Text S3) (*12*).

Despite physiological similarities to cultured *Chloroflexota* members, genome sequencing data revealed that '*Ca.* Chx. allophototropha' encoded a remote homolog of a *pscA*-like RCI gene (Fig. 2) rather than the expected *pufL*/*pufM* RCII genes. A *pscA*-like gene was also detected in a second genome bin, named '*Ca.* Chloroheliaceae bin L227-5C', which we recovered from a second enrichment culture that was subsequently lost (Supplementary Texts S1 and S2). Our detection of RCI genes in '*Ca.* Chx. allophototropha' makes the *Chloroflexota* phylum unique

among all known anoxygenic phototroph-containing microbial phyla, which otherwise contain
75 representatives that use strictly RCI or RCII for phototrophy (Fig. 2A). Based on a maximum-likelihood amino acid sequence phylogeny (Fig. 2B; Fig. S2), the PscA-like predicted protein of '*Ca.* Chx. allophototropha' represents a clearly distinct fourth class of anoxygenic RCI protein. Homology modelling indicated that the '*Ca.* Chx. allophototropha' PscA-like protein contains the expected six N-terminal transmembrane helices for coordinating antenna pigments and five
80 C-terminal transmembrane helices involved in core electron transfer (Fig. 2C) (*13*). The genome did not contain the *pscBCD* photosystem accessory genes found among *Chlorobia* members (*7*), nor a cyanobacterial *psaB*-like paralog, implying that '*Ca.* Chx. allophototropha' uses a homodimeric RCI complex (*14*) like other known anoxygenic phototrophs.

The genome of '*Ca.* Chx. allophototropha' encoded a highly novel homolog of *fmoA*
85 encoding the Fenna-Matthews-Olson (FMO) protein involved in bacteriochlorophyll *a* binding (Fig. 3; Fig. S3) (*15*). This finding supports that '*Ca.* Chx. allophototropha' relies on RCI and chlorosomes for phototrophic energy conversion and makes the *Chloroflexota* the third known phylum with representatives that use the FMO protein for phototrophy (*16*). We also detected a potential homolog of the key chlorosome structural gene *csmA* (Fig. 3) (*6*, *7*). This sequence had
90 a similar predicted tertiary structure to the CsmA protein of *Chlorobium tepidum* (*17*). Although the potential homolog lacked the histidine-25 hypothesized to be required for bacteriochlorophyll *a* binding (Fig. S4) (*18*), no other predicted open reading frames in the genome had close homology to known CsmA sequences, implying that this sequence represents a highly novel CsmA variant. We also detected potential homologs of *csmM* and *csmY* (*7*) in the genome (Fig.
95 3).

Similarly to *Oscillochloris trichoides*, the '*Ca.* Chx. allophototropha' genome encoded genes for the Calvin-Benson-Bassham (CBB) cycle, including a deep-branching Class IC/ID "red type" *cbbL* gene (*19*) representing the large subunit of RuBisCO (Fig. S5) and a *cbbP* gene encoding phosphoribulokinase (Fig. 3). Carbon fixation via the CBB cycle has never before been reported for a RCI-utilizing anoxygenic phototroph (*20*). The novel genome did not encode the potential for carbon fixation via the 3-hydroxypropionate (3HP) bicycle, which is thought to represent a more recent evolutionary innovation within the *Chloroflexales* order (*21*). Additionally, the '*Ca.* Chx. allophototropha' genome encoded the biosynthesis pathway for bacteriochlorophylls *a* and *c* (*7*) and a $nifHI_1I_2BDK$ gene cluster potentially involved in nitrogen fixation (Fig. 3) (*22*). The genome bin of '*Ca.* Chloroheliaceae bin L227-5C' encoded a similar set of phototrophy-related gene homologs as '*Ca.* Chx. allophototropha' (Fig. 3; all gene accessions are summarized in Table S1).

Based on the Genome Taxonomy Database (GTDB) classification system (*23*, *24*), '*Ca.* Chx. allophototropha' represents the first cultivated member of a novel order within the *Chloroflexota* phylum (Fig. 3), which we provisionally name the '*Ca.* Chloroheliales'. This order placed immediately basal to the *Chloroflexales* order, which contains the canonical RCII-utilizing phototroph clade in the phylum that includes the *Chloroflexaceae*, *Roseiflexaceae*, and *Oscillochloridaceae* families (*25*). The *Chloroflexales* order also contains a basal non-phototrophic family, the *Herpetosiphonaceae*, which was formerly placed in its own order prior to reassignment by the GTDB (*23*). Two uncultured genome bins from public databases also placed in the '*Ca.* Chloroheliales' order, although we did not detect the potential for phototrophy in these bins (Fig. 3; Supplementary Text S4).

Previously, the genetic content of RCII-utilizing phototrophic *Chloroflexota* members has been enigmatic because several *Chloroflexota*-associated phototrophy genes appear to be most closely related to genes of RCI-utilizing phototrophs. Our results lead us to hypothesize that '*Ca.* Chx. allophototropha' represents a missing transition form in the evolution of phototrophy among the *Chloroflexota* phylum that resolves this enigma by providing a direct link between RCI- and RCII-utilizing phototrophic taxa. Along with encoding chlorosomes, RCII-utilizing *Chloroflexota* members have been observed to encode bacteriochlorophyll synthesis genes most closely related to those of RCI-utilizing *Chlorobia* members (*7*). Using maximum likelihood phylogenies of the (bacterio)chlorophyll synthesis proteins BchIDH/ChlIDH (Fig. S6) and BchLNB/ChlLNB/BchXYZ (Fig. S7), we observed that '*Ca.* Chloroheliales' sequences placed immediately adjacent to this sister grouping of RCII-utilizing *Chloroflexota* and RCI-utilizing *Chlorobia* members. A phylogeny of the chlorosome structural protein CsmA also indicated closer evolutionary relatedness between '*Ca.* Chloroheliales', RCII-utilizing *Chloroflexota*, and RCI-utilizing *Chlorobia* members compared to RCI-utilizing *Acidobacteriota* members such as *Chloracidobacterium thermophilum* (Fig. S8) (*16*). The phylogenetic clustering of RCI- and RCII-utilizing *Chloroflexota* members with RCI-utilizing *Chlorobia* members suggests a shared genetic history of phototrophy between these groups despite using different core reaction center components. Thus, it is possible that an ancestor of '*Ca.* Chx. allophototropha' received genes for RCI, chlorosomes, bacteriochlorophyll synthesis, and carbon fixation via lateral gene transfer and then provided the photosynthesis accessory genes needed by RCII-containing *Chloroflexales* members via either direct vertical inheritance or a subsequent lateral gene transfer event.

The *Chloroflexota* now represents the only phylum outside of the *Cyanobacteria* where examples of both Type II and Type I photosynthetic reaction centers can be found (Fig. 2A). One

model for the origin of oxygenic photosynthesis suggests that a single organism received quinone and Fe-S type reaction centers via lateral gene transfer and then developed the capacity to use these two reaction center classes in tandem as PSII and PSI (*26*, *27*). This model of evolution can now be tested using the *Chloroflexota* phylum, either through introducing both reaction center classes into a single *Chloroflexota* member or through ancestral state reconstructions to explore the intriguing evolutionary history of RCI- and RCII-utilizing *Chloroflexota* members. Discovering '*Ca.* Chx. allophototropha' demonstrates the possibility for massive photosynthesis gene movements over evolutionary time, which has been required by many evolutionary models previously but has been sparsely demonstrated in nature (*28*, *29*).

References and Notes

1. Formerly the *Chloroflexi* phylum.

2. D. Y. Sorokin, D. Vejmelkova, S. Lücker, G. M. Streshinskaya, W. I. C. Rijpstra, J. S. Sinninghe Damsté, R. Kleerbezem, M. van Loosdrecht, G. Muyzer, H. Daims, *Nitrolancea hollandica* gen. nov., sp. nov., a chemolithoautotrophic nitrite-oxidizing bacterium isolated from a bioreactor belonging to the phylum *Chloroflexi*. *Int J Syst Evol Micr*. **64**, 1859–1865 (2014).

3. K. Wasmund, L. Schreiber, K. G. Lloyd, D. G. Petersen, A. Schramm, R. Stepanauskas, B. B. Jørgensen, L. Adrian, Genome sequencing of a single cell of the widely distributed marine subsurface *Dehalococcoidia*, phylum *Chloroflexi*. *ISME J*. **8**, 383–397 (2014).

4. Y. Okazaki, M. M. Salcher, C. Callieri, S. Nakano, The broad habitat spectrum of the CL500-11 lineage (phylum *Chloroflexi*), a dominant bacterioplankton in oxygenated hypolimnia of deep freshwater lakes. *Front Microbiol*. **9**, 2891 (2018).

5. B. K. Pierson, R. W. Castenholz, A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. *Arch Microbiol*. **100**, 5–24 (1974).

6. N.-U. Frigaard, D. A. Bryant, Seeing green bacteria in a new light: genomics-enabled studies of the photosynthetic apparatus in green sulfur bacteria and filamentous anoxygenic phototrophic bacteria. *Arch Microbiol*. **182**, 265–276 (2004).

7. D. A. Bryant, Z. Liu, T. Li, F. Zhao, A. M. G. Costas, C. G. Klatt, D. M. Ward, N.-U. Frigaard, J. Overmann, in *Functional Genomics and Evolution of Photosynthetic Systems*, R. Burnap, W. Vermaas, Eds. (Springer Netherlands, Dordrecht, 2012; https://doi.org/10.1007/978-94-007-1533-2_3), *Advances in Photosynthesis and Respiration*, pp. 47–102.

8. M. F. Hohmann-Marriott, R. E. Blankenship, Hypothesis on chlorosome biogenesis in green photosynthetic bacteria. *FEBS Lett*. **581**, 800–803 (2007).

9. Materials and methods are available as supplementary materials at the Science website.

10. Y. Jiao, A. Kappler, L. R. Croal, D. K. Newman, Isolation and characterization of a genetically tractable photoautotrophic Fe(II)-oxidizing bacterium, *Rhodopseudomonas palustris* strain TIE-1. *Appl Environ Microbiol*. **71**, 4487–4496 (2005).

11. F. Gich, J. Garcia-Gil, J. Overmann, Previously unknown and phylogenetically diverse members of the green nonsulfur bacteria are indigenous to freshwater lakes. *Arch Microbiol*. **177**, 1–10 (2001).

12. J. Overmann, F. Garcia-Pichel, in *The Prokaryotes: Prokaryotic Communities and Ecophysiology*, E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, F. Thompson, Eds.

(Springer Berlin Heidelberg, 2013; https://doi.org/10.1007/978-3-642-30123-0_51), pp. 203–257.

13. C. Gisriel, I. Sarrou, B. Ferlez, J. H. Golbeck, K. E. Redding, R. Fromme, Structure of a symmetric photosynthetic reaction center–photosystem. *Science*. **357**, 1021–1025 (2017).

14. W. W. Fischer, J. Hemp, J. E. Johnson, Evolution of oxygenic photosynthesis. *Annu Rev Earth Pl Sc*. **44**, 647–683 (2016).

15. J. M. Olson, in *Discoveries in Photosynthesis*, Govindjee, J. T. Beatty, H. Gest, J. F. Allen, Eds. (Springer Netherlands, Dordrecht, 2005; https://doi.org/10.1007/1-4020-3324-9_40), *Advances in Photosynthesis and Respiration*, pp. 421–427.

16. D. A. Bryant, A. M. G. Costas, J. A. Maresca, A. G. M. Chew, C. G. Klatt, M. M. Bateson, L. J. Tallon, J. Hostetler, W. C. Nelson, J. F. Heidelberg, D. M. Ward, *Candidatus* Chloracidobacterium thermophilum: an aerobic phototrophic Acidobacterium. *Science*. **317**, 523–526 (2007).

17. M. Ø. Pedersen, J. Underhaug, J. Dittmer, M. Miller, N. C. Nielsen, The three-dimensional structure of CsmA: A small antenna protein from the green sulfur bacterium *Chlorobium tepidum*. *FEBS Lett*. **582**, 2869–2874 (2008).

18. M. Ø. Pedersen, J. Linnanto, N.-U. Frigaard, N. Chr. Nielsen, M. Miller, A model of the protein–pigment baseplate complex in chlorosomes of photosynthetic green bacteria. *Photosynth Res*. **104**, 233–243 (2010).

19. T. P. Tourova, E. M. Spiridonova, N. V. Slobodova, E. S. Boulygina, O. I. Keppen, B. B. Kuznetsov, R. N. Ivanovsky, Phylogeny of anoxygenic filamentous phototrophic bacteria of the family *Oscillochloridaceae* as inferred from comparative analyses of the *rrs*, *cbbL*, and *nifH* genes. *Microbiology*. **75**, 192–200 (2006).

20. K.-H. Tang, Y. J. Tang, R. E. Blankenship, Carbon metabolic pathways in phototrophic bacteria and their broader evolutionary implications. *Front Microbiol*. **2**, 165 (2011).

21. P. M. Shih, L. M. Ward, W. W. Fischer, Evolution of the 3-hydroxypropionate bicycle and recent transfer of anoxygenic photosynthesis into the *Chloroflexi*. *Proc Natl Acad Sci USA*. **114**, 10749–10754 (2017).

22. J. Raymond, J. L. Siefert, C. R. Staples, R. E. Blankenship, The natural history of nitrogen fixation. *Mol Biol Evol*. **21**, 541–554 (2004).

23. D. H. Parks, M. Chuvochina, D. W. Waite, C. Rinke, A. Skarshewski, P.-A. Chaumeil, P. Hugenholtz, A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol*. **36**, 996–1004 (2018).

24. P.-A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks, GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. **36**, 1925–1927 (2020).

25. S. Hanada, in *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*, E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, F. Thompson, Eds. (Springer Berlin Heidelberg, 2014; https://doi.org/10.1007/978-3-642-38954-2_165), pp. 515–532.

26. M. F. Hohmann-Marriott, R. E. Blankenship, Evolution of photosynthesis. *Annu Rev Plant Biol*. **62**, 515–548 (2011).

27. R. M. Soo, J. Hemp, P. Hugenholtz, Evolution of photosynthesis and aerobic respiration in the cyanobacteria. *Free Radical Bio Med*. **140**, 200–205 (2019).

28. W. F. Martin, D. A. Bryant, J. T. Beatty, A physiological perspective on the origin and evolution of photosynthesis. *FEMS Microbiol Rev*. **42**, 205–231 (2018).

29. Y. Zeng, F. Feng, H. Medová, J. Dean, M. Koblížek, Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum *Gemmatimonadetes*. *Proc Natl Acad Sci USA*. **111**, 7795–7800 (2014).

30. J. M. Tsuji, N. A. Shaw, S. Nagashima, J. J. Venkiteswaran, S. L. Schiff, S. Hanada, M. Tank, J. D. Neufeld, Sequencing data analysis files for "Anoxygenic phototrophic *Chloroflexota* member uses a Type I reaction center." *Zenodo* (2020) (available at https://doi.org/10.5281/zenodo.3934908).

31. S. L. Schiff, J. M. Tsuji, L. Wu, J. J. Venkiteswaran, L. A. Molot, R. J. Elgood, M. J. Paterson, J. D. Neufeld, Millions of Boreal Shield lakes can be used to probe Archaean Ocean biogeochemistry. *Sci Rep*. **7**, 46708 (2017).

32. D. W. Schindler, R. E. Hecky, D. L. Findlay, M. P. Stainton, B. R. Parker, M. J. Paterson, K. G. Beaty, M. Lyng, S. E. M. Kasian, Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. *Proc Natl Acad Sci USA*. **105**, 11254–11258 (2008).

33. F. Hegler, N. R. Posth, J. Jiang, A. Kappler, Physiology of phototrophic iron(II)-oxidizing bacteria: implications for modern and ancient environments. *FEMS Microbiol Ecol*. **66**, 250–260 (2008).

34. J. H. Vandermeulen, N. D. Davis, L. Muscatine, The effect of inhibitors of photosynthesis on zooxanthellae in corals and other marine invertebrates. *Marine Biology*. **16**, 185–191 (1972).

35. L. L. Stookey, Ferrozine - a new spectrophotometric reagent for iron. *Anal Chem*. **42**, 779–781 (1970).

36. N. Pfennig, *Rhodocyclus purpureus* gen. nov. and sp. nov., a ring-shaped, vitamin B12-requiring member of the family *Rhodospirillaceae*. *Int J Syst Evol Microbiol*. **28**, 283–288 (1978).

37. J. F. Imhoff, in *The Prokaryotes*, E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, F. Thompson, Eds. (Springer Berlin Heidelberg, 2014; https://doi.org/10.1007/978-3-642-38922-1_295), pp. 151–178.

38. V. A. Gaisin, E. I. Burganskaya, D. S. Grouzdev, A. A. Ashikhmin, N. A. Kostrikina, I. A. Bryantseva, V. V. Koziaeva, V. M. Gorlenko, '*Candidatus* Viridilinea mediisalina', a novel phototrophic Chloroflexi bacterium from a Siberian soda lake. *FEMS Microbiol Lett*. **366**, fnz043 (2019).

39. J. F. Imhoff, in *The Prokaryotes*, E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, F. Thompson, Eds. (Springer Berlin Heidelberg, 2014; https://doi-org.proxy.lib.uwaterloo.ca/10.1007/978-3-642-38954-2_142), pp. 501–514.

40. K. C. Biswas, N. A. Woodards, H. Xu, L. L. Barton, Reduction of molybdate by sulfate-reducing bacteria. *Biometals*. **22**, 131–139 (2009).

41. A. E. Parada, D. M. Needham, J. A. Fuhrman, Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol*, 1403–1414 (2016).

42. C. Quince, A. Lanzen, R. J. Davenport, P. J. Turnbaugh, Removing noise from pyrosequenced amplicons. *BMC Bioinform*. **12**, 38 (2011).

43. K. Kennedy, M. W. Hall, M. D. J. Lynch, G. Moreno-Hagelsieb, J. D. Neufeld, Evaluating bias of Illumina-based bacterial 16S rRNA gene profiles. *Appl Environ Microbiol*. **80**, 5717–5722 (2014).

44. A. K. Bartram, M. D. Lynch, J. C. Stearns, G. Moreno-Hagelsieb, J. D. Neufeld, Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end Illumina reads. *Appl Environ Microbiol*. **77**, 3846–3852 (2011).

45. M. A. Cavaco, V. L. St. Louis, K. Engel, K. A. St. Pierre, S. L. Schiff, M. Stibal, J. D. Neufeld, Freshwater microbial community diversity in a rapidly changing High Arctic watershed. *FEMS Microbiol Ecol*. **95**, fiz161 (2019).

46. E. Bolyen, J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. K. Cope, R. D. Silva, C. Diener, P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J. M. Gauglitz, S. M. Gibbons, D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. A. Huttley, S. Janssen, A. K. Jarmusch, L. Jiang, B. D. Kaehler, K. B. Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T. Kosciolek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y.-X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. T. Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian,

T. Pearson, S. L. Peoples, D. Petras, M. L. Preuss, E. Pruesse, L. B. Rasmussen, A. Rivers, M. S. Robeson, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R. Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight, J. G. Caporaso, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. **37**, 852–857 (2019).

47. M. D. Lynch, A. P. Masella, M. W. Hall, A. K. Bartram, J. D. Neufeld, AXIOME: automated exploration of microbial diversity. *GigaScience*. **2**, 3–3 (2013).

48. B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, S. P. Holmes, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. **13**, 581–583 (2016).

49. N. A. Bokulich, B. D. Kaehler, J. R. Rideout, M. Dillon, E. Bolyen, R. Knight, G. A. Huttley, J. Gregory Caporaso, Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*. **6**, 90 (2018).

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

51. F. O. Glöckner, P. Yilmaz, C. Quast, J. Gerken, A. Beccati, A. Ciuprina, G. Bruns, P. Yarza, J. Peplies, R. Westram, W. Ludwig, 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J Biotechnol*. **261**, 169–176 (2017).

52. J. Zhou, M. A. Bruns, J. M. Tiedje, DNA recovery from soils of diverse composition. *Appl Environ Microbiol*. **62**, 316–322 (1996).

53. J. Cheng, L. Pinnell, K. Engel, J. D. Neufeld, T. C. Charles, Versatile broad-host-range cosmids for construction of high quality metagenomic libraries. *J Microbiol Methods*. **99**, 27–34 (2014).

54. Z. Chen, L. Pham, T.-C. Wu, G. Mo, Y. Xia, P. L. Chang, D. Porter, T. Phan, H. Che, H. Tran, V. Bansal, J. Shaffer, P. Belda-Ferre, G. Humphrey, R. Knight, P. Pevzner, S. Pham, Y. Wang, M. Lei, Ultra-low input single tube linked-read library method enables short-read second-generation sequencing systems to generate highly accurate and economical long-range sequencing information routinely. *Genome Res*. **30**, 898–909 (2020).

55. S. Kieser, J. Brown, E. M. Zdobnov, M. Trajkovski, L. A. McCue, ATLAS: a Snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data. *BMC Bioinform*. **21**, 257 (2020).

56. Y.-W. Wu, B. A. Simmons, S. W. Singer, MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. **32**, 605–607 (2016).

57. D. D. Kang, F. Li, E. Kirton, A. Thomas, R. Egan, H. An, Z. Wang, MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. **7**, e7359 (2019).

58. D. Hyatt, G.-L. Chen, P. F. LoCascio, M. L. Land, F. W. Larimer, L. J. Hauser, Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform*. **11**, 1–11 (2010).

59. N. A. O'Leary, M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M. Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K. Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill, S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb, W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy, K. D. Pruitt, Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res*. **44**, D733–D745 (2016).

60. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J Mol Biol*. **215**, 403–410 (1990).

61. T. Seemann, Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. **30**, 2068–2069 (2014).

62. M. Kanehisa, Y. Sato, K. Morishima, BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol*. **428**, 726–731 (2016).

63. R. A. White III, J. Brown, S. Colby, C. C. Overall, J.-Y. Lee, J. Zucker, K. R. Glaesemann, C. Jansson, J. K. Jansson, *PeerJ Preprints*, in press, doi:10.7287/peerj.preprints.2843v1.

64. P. Menzel, K. L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun*. **7**, 1–9 (2016).

65. R. G. Singh, A. Tanca, A. Palomba, F. Van der Jeugt, P. Verschaffelt, S. Uzzau, L. Martens, P. Dawyndt, B. Mesuere, Unipept 4.0: functional analysis of metaproteome data. *J Proteome Res*. **18**, 606–615 (2018).

66. P. Petrenko, B. Lobb, D. A. Kurtz, J. D. Neufeld, A. C. Doxey, MetAnnotate: function-specific taxonomic profiling and comparison of metagenomes. *BMC Biol*. **13**, 1–8 (2015).

67. J. A. Fish, B. Chai, Q. Wang, Y. Sun, C. T. Brown, J. M. Tiedje, J. R. Cole, FunGene: the functional gene pipeline and repository. *Front Microbiol*. **4**, 291 (2013).

68. S. R. Eddy, Accelerated profile HMM searches. *PLOS Comput Biol*. **7**, e1002195 (2011).

69. R. D. Finn, P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G. A. Salazar, J. Tate, A. Bateman, The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res*. **44**, D279–D285 (2016).

70. J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: protein structure and function prediction. *Nat Methods*. **12**, 7–8 (2015).

71. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, D. G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. **7**, 539 (2011).

72. L. M. Ward, J. Hemp, P. M. Shih, S. E. McGlynn, W. W. Fischer, Evolution of phototrophy in the *Chloroflexi* phylum driven by horizontal gene transfer. *Front Microbiol*. **9**, 260 (2018).

73. M. Tank, V. Thiel, D. M. Ward, D. A. Bryant, in *Modern Topics in the Phototrophic Prokaryotes: Environmental and Applied Aspects*, P. C. Hallenbeck, Ed. (Springer International Publishing Switzerland, 2017; https://doi.org/10.1007/978-3-319-46261-5_3), pp. 87–137.

74. M. D. Lee, GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics*. **35**, 4162–4164 (2019).

75. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. **32**, 268–274 (2015).

76. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. **14**, 587–589 (2017).

77. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol*. **35**, 518–522 (2018).

78. K.-H. Tang, K. Barry, O. Chertkov, E. Dalin, C. S. Han, L. J. Hauser, B. M. Honchak, L. E. Karbach, M. L. Land, A. Lapidus, F. W. Larimer, N. Mikhailova, S. Pitluck, B. K. Pierson, R. E. Blankenship, Complete genome sequence of the filamentous anoxygenic phototrophic bacterium *Chloroflexus aurantiacus*. *BMC Genomics*. **12**, 334 (2011).

79. O. I. Keppen, O. I. Baulina, E. N. Kondratieva, *Oscillochloris trichoides* neotype strain DG-6. *Photosynth Res*. **41**, 29–33 (1994).

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

81. L. H. Bergstrand, E. Cardenas, J. Holert, J. D. V. Hamme, W. W. Mohn, Delineation of steroid-degrading microorganisms through comparative genomic analysis. *mBio*. **7**, e00166-16 (2016).

82. A. I. Garber, K. H. Nealson, A. Okamoto, S. M. McAllister, C. S. Chan, R. A. Barco, N. Merino, FeGenie: A comprehensive tool for the identification of iron genes and iron gene neighborhoods in genome and metagenome assemblies. *Front Microbiol*. **11**, 37 (2020).

83. T. Cardona, Early Archean origin of heterodimeric Photosystem I. *Heliyon*. **4**, e00548 (2018).

84. F. R. Tabita, T. E. Hanson, S. Satagopan, B. H. Witte, N. E. Kreel, Phylogenetic and evolutionary relationships of RubisCO and the RubisCO-like proteins and the functional lessons provided by diverse molecular forms. *Phil Trans R Soc B*. **363**, 2629–2640 (2008).

85. K. Mendler, H. Chen, D. H. Parks, B. Lobb, L. A. Hug, A. C. Doxey, AnnoTree: visualization and exploration of a functionally annotated microbial tree of life. *Nucleic Acids Res*. **47**, 4442–4448 (2019).

86. G. Talavera, J. Castresana, Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. **56**, 564–577 (2007).

87. S. Hanada, A. Hiraishi, K. Shimada, K. Matsuura, *Chloroflexus aggregans* sp. nov., a filamentous phototrophic bacterium which forms dense cell aggregates by active gliding movement. *Int J Syst Evol Microbiol*. **45**, 676–681 (1995).

88. J. D. Coates, D. J. Ellis, C. V. Gaw, D. R. Lovley, *Geothrix fermentans* gen. nov., sp. nov., a novel Fe(III)-reducing bacterium from a hydrocarbon-contaminated aquifer. *Int J Syst Bacteriol*. **49**, 1615–1622 (1999).

89. C. Leang, M. V. Coppi, D. R. Lovley, OmcB, a c-type polyheme cytochrome, involved in Fe(III) reduction in *Geobacter sulfurreducens*. *J Bacteriol*. **185**, 2096–2103 (2003).

90. R. S. Kantor, A. W. van Zyl, R. P. van Hille, B. C. Thomas, S. T. L. Harrison, J. F. Banfield, Bioreactor microbial ecosystems for thiocyanate and cyanide degradation unravelled with genome-resolved metagenomics. *Environ Microbiol*. **17**, 4929–4941 (2015).

91. M. Ji, C. Greening, I. Vanwonterghem, C. R. Carere, S. K. Bay, J. A. Steen, K. Montgomery, T. Lines, J. Beardall, J. van Dorst, I. Snape, M. B. Stott, P. Hugenholtz, B. C. Ferrari, Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature*. **552**, 400–403 (2017).

Acknowledgments

We thank staff at the IISD-ELA and R. Elgood for providing baseline limnological data; J. Mead, E. Barber, J. Wolfe, K. Thompson, and R. Henderson for assistance with lake sampling; K. Engel, S. Hirose, and T. Chen for assistance with DNA sequencing; Y. Shirotori, M. Saini, and V. Gaisin for assistance and advice in enrichment cultivation; S. Hirose for assistance with pigment analyses; R. Harris and E. Roach for assistance with electron microscopy; Y. Wang and V. Thiel for assistance with sequencing data analyses; and N. Tran and K. Shimada for discussion and advice. **Funding:** This work was supported by a Strategic Partnership Grant for Projects from the National Sciences and Engineering Research Council of Canada (NSERC), Discovery Grants from NSERC, and a research grant from the Institute for Fermentation, Osaka. **Author contributions:** S.L.S., J.D.N., and J.J.V. conceived the study. J.M.T., M.T., and N.A.S. performed enrichment cultivation and characterized the culture. J.M.T. analyzed sequencing data. J.M.T., S.N., M.T., S.H., and J.D.N. interpreted the data. J.M.T. wrote the paper with J.D.N., N.A.S., S.N., and comments from all other authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Raw sequencing data and genome bins are available under the National Center for Biotechnology Information (NCBI) BioProject accession PRJNA640240. Custom scripts and additional raw data files used to analyze sequencing data are available on Zenodo (*30*).

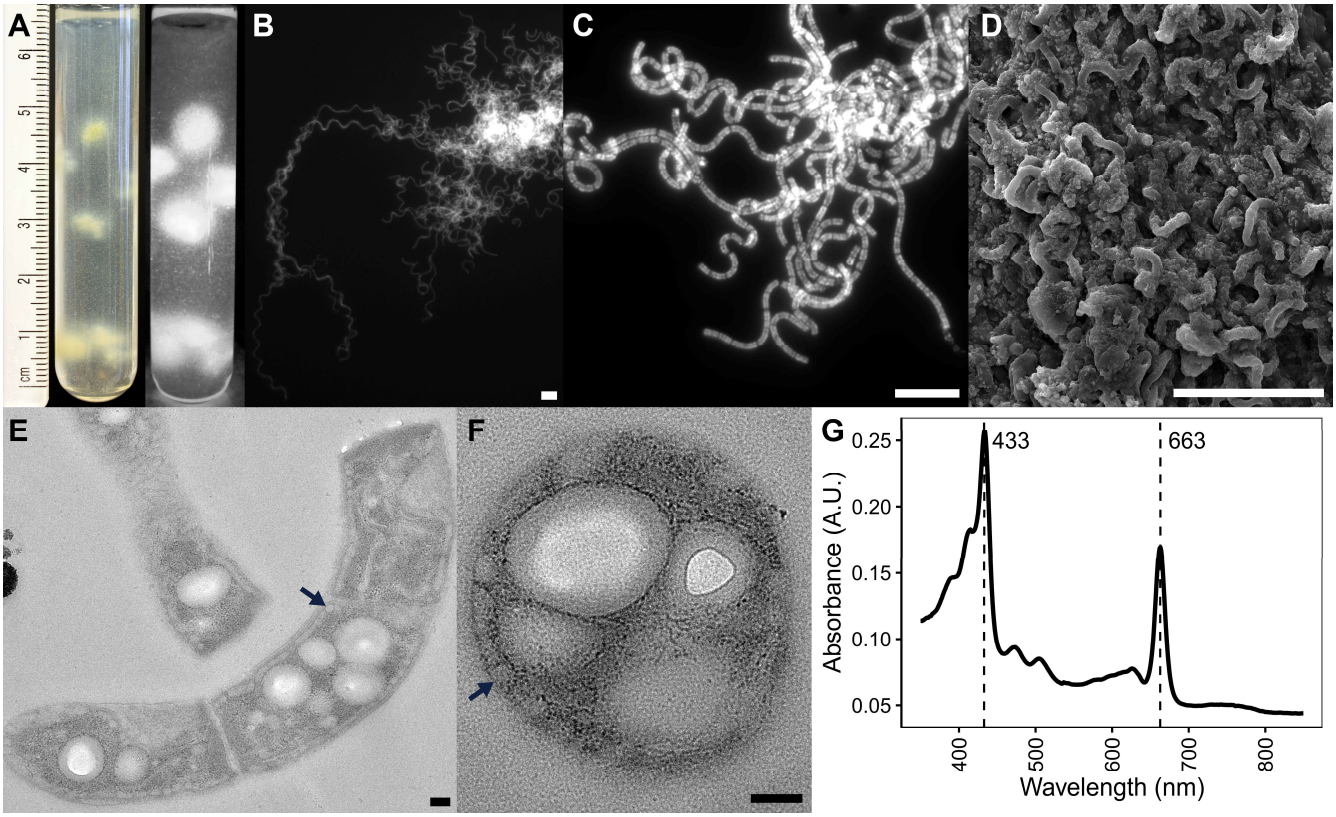

Fig. 1. Physiology of the '*Ca.* Chx. allophototropha culture'. (**A**) Growth of the enrichment culture in soft agar. The left panel is a photograph of an agar tube, and the right panel shows the same tube when viewed using a pigment camera optimized for bacteriochlorophyll *c* (Fig. S9). (**B** and **C**) Epifluorescence microscopy images of the culture. Cell autofluorescence is shown with light excitation and detection wavelengths optimized for bacteriochlorophyll *c*. (**D**) Scanning electron microscopy image of an aggregate of '*Ca.* Chx. allophototropha' cells. (**E** and **F**) Transmission electron microscopy images showing longitudinal and cross sections, respectively, of '*Ca.* Chx. allophototropha' cells. Example chlorosomes are marked with arrows. (**G**) Absorption spectrum of an acetone:methanol extract of pigments from the culture. Scale bars on the top row of images (A, B, C, and D) represent 10 µm, and scale bars on the bottom row of images (E and F) represent 0.1 µm.

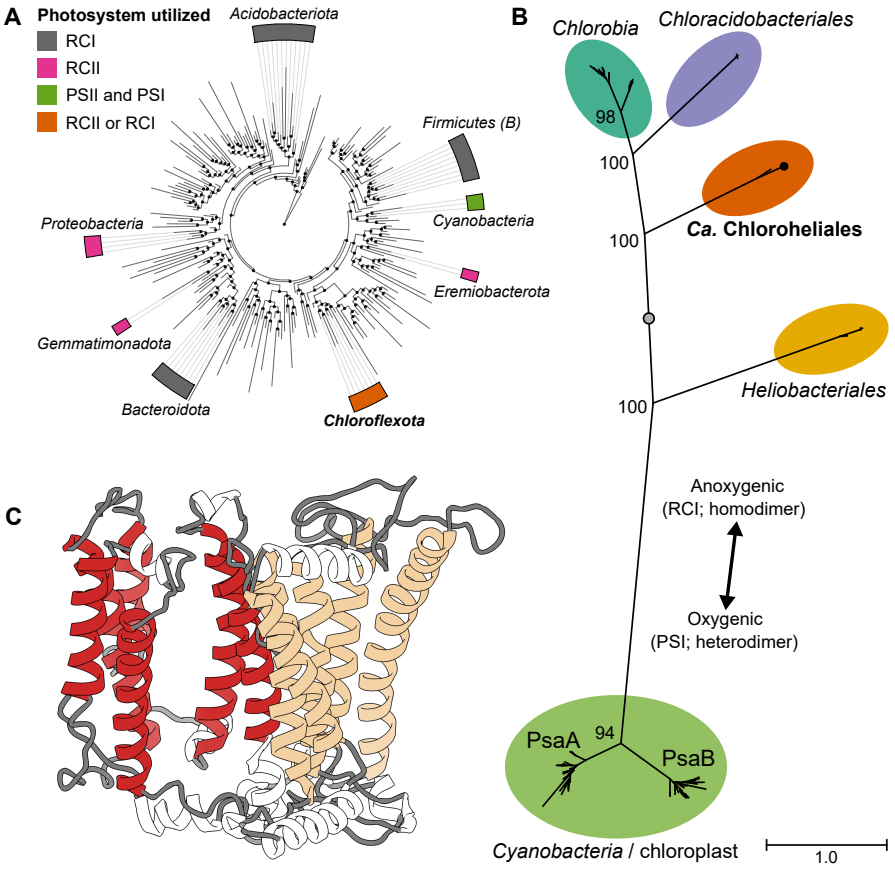

Fig. 2. Functional novelty of the '*Ca.* Chlorohelix allophototropha' photosystem. (**A**) Overview of the eight bacterial phyla containing known chlorophototrophs. (**B**) Maximum likelihood phylogeny of Type I reaction center primary sequences. Ultrafast bootstrap values for major branch points are shown, and the scale bar represents the expected proportion of amino acid change (see Table S2). Chlorophototrophic lineages are summarized by order name. The placement of the PscA-like sequence of '*Ca*. Chx. allophototropha' is indicated by a black dot. All clades above they grey dot utilize chlorosomes. (**C**) Predicted tertiary structure of the '*Ca*. Chlorohelix allophototropha' PscA-like primary sequence. The six N-terminal and five C-terminal transmembrane helices expected for RCI are colored in red and tan, respectively.

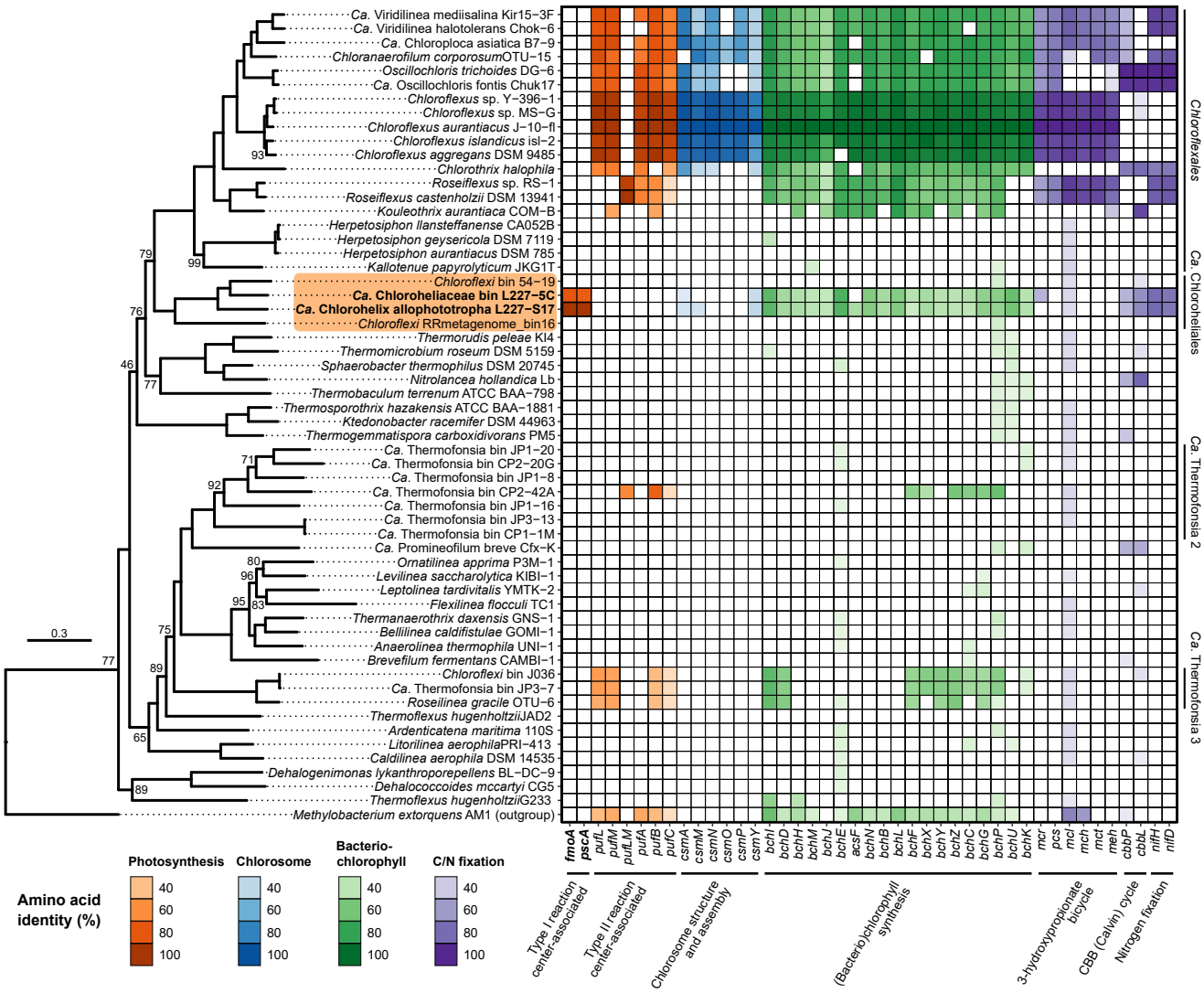

Fig. 3. Genomic potential for phototrophy among members of the *Chloroflexota* phylum. A maximum likelihood phylogeny of *Chloroflexota* members is shown on the left based on 74 concatenated core bacterial proteins. Ultrafast bootstrap values are shown (100/100 omitted for clarity), and the scale bar shows the expected proportion of amino acid change (see Table S2). The '*Ca.* Chloroheliales' clade is highlighted in orange. The heatmap shows the presence/absence of genes involved in photosynthesis or related processes based on bidirectional BLASTP. Intensity of the fill color of each tile corresponds to the percent identity of the BLASTP hit. The query organism used for each gene is indicated by bolded outlines on the heatmap. Orders containing potential phototrophs are labelled on the right of the heatmap.