**Extended Data Table 1 | Microbial community composition of the “*Ca.* Chx. allophototropha” and *Geothrix* sp. L227-G1 enrichment cultures.** Analyzed 16S rRNA gene amplicon data are shown to support spectroscopy, microscopy, and genomic analyses presented in this work. For each culture sample, the subculture generation is indicated (as part of the sample code), along with the figure panel(s) associated with that sample. Data was processed for each sequencing method as described in the methods. Sequences (i.e., ASVs or OTUs) were assigned as “*Ca.* Chx. allophototropha” or *Geothrix* L227-G1 based on >99% nucleotide match. Other ASVs or OTUs are presented in the “Other” category with their genus-level classification. Additional microbial community composition data, based on metagenome reads, are shown in Extended Data Fig. 1e. Abbreviations: bdl, below detection limit.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Culture** | **Code\*** | **Associated figure(s)** | **Sequencing method** | **Relative abundance (%)** | | | **Read count (detection limit, %)\*\*** | **Notes** |
| **“*Ca*. Chx allophoto-tropha”** | **Geothrix L227-G1** | **Other** |
| *Ca.* Chlorohelix | 13.2a-3 | 1d-e | Illumina V4-5 | 96.4 | 2.7 | 0.96 (*Dechloromonas*) | 523 (0.382) | 13.2a-3 was the parent culture of the culture used for SEM.  A parallel culture in the same agar dilution series as 13.2a-3 was used for TEM. |
| *Ca.* Chlorohelix | 19.9 | ED1e, ED3 | Nanopore V1-9 | 100.0 | Bdl | bdl | 28,321 (0.035) | Used for complete genome sequencing. |
| *Ca.* Chlorohelix | 21.2c | 1a-c, ED1f | Illumina V4 | 46.5 | 53.5 | bdl | 49,950 (0.004) | Two sequencing methods were compared for the same sample. |
| Nanopore V1-9 | 2.9 | 97.1 | bdl | 53,805 (0.019) |
| *Ca.* Chlorohelix | 22.2c | 1f-g, ED2e-f | Nanopore V1-9 | 2.8 | 97.2 | bdl | 65,971 (0.015) | Fe(II) + acetate, light |
| *Ca.* Chlorohelix | 22.2d | 1f-g, ED2e-f | Nanopore V1-9 | bdl | 99.9 | 0.07 (*Sphingomonas*) | 63,787 (0.016) | Fe(II) + acetate, dark |
| *Ca.* Chlorohelix | 22.2a | ED2e-f | Nanopore V1-9 | 26.3 | 73.8 | bdl | 50,792 (0.020) | Fe(II), light |
| *Ca.* Chlorohelix | 22.2b | ED2e-f | Nanopore V1-9 | 0.02 | 99.9 | 0.05 (*Microbacterium*), 0.03 (*Acinetobacter*) | 61,971 (0.016) | Fe(II), dark |
| *Geothrix* | 5.1b | ED2x | Nanopore V1-9 | bdl | 100.0 | bdl | 17,398 (0.057) | 5.1b was the parent culture of the culture used for microscopy. |

\*The number before the period (e.g., 12) refers to the subculture generation; subsequent numbers/letters were used internally to distinguish between cultures.

\*\*A sequence was detectable if it generated at least 2 counts (Illumina) or 10 counts (Nanopore).

**Extended Data Table 2 | Genes potentially involved in phototrophy or carbon/nitrogen fixation among genomes of “*Ca.* Chloroheliales” members recovered in this study.** Locus tags are shown each gene. Results correspond to those shown in Fig. 3, except that homologs associated with the incomplete 3-hydroxypropionate bicycle are omitted for clarity, and additional genes involved in bacteriochlorophyll synthesis and the RPP cycle are shown.

|  |  |  |  |
| --- | --- | --- | --- |
| **Protein set** | **Gene** | **“*Ca*. Chx allophototropha”** | **“*Ca*. Chloroheliaceae bin L227-5C”** |
| Type I reaction center-associated | *fmoA* | OZ401\_003705 | HXX20\_00820 |
| *pscA* | OZ401\_000236 | HXX20\_00815 |
| Chlorosome structure and assembly | *csmA* | OZ401\_000329 | HXX20\_02945 |
| *csmM* | OZ401\_000866 | N/A |
| *csmY* | OZ401\_003358 | HXX20\_02040 |
| (Bacterio)chlorophyll synthesis | *bchI* | OZ401\_000920 | HXX20\_00800 |
| *bchD* | OZ401\_002907 | HXX20\_00790 |
| *bchH* | OZ401\_002810 | HXX20\_14895 |
| *bchM* | OZ401\_002809 | HXX20\_07990 |
| *bchJ* | OZ401\_000950 | HXX20\_17620 |
| *bchE* | OZ401\_000949 | HXX20\_17615 |
| *acsF* | OZ401\_004470 | N/A |
| *bchN* | OZ401\_001487 | HXX20\_03120 |
| *bchB* | OZ401\_001486 | HXX20\_03125 |
| *bchL* | OZ401\_001485 | HXX20\_03130 |
| *bciB* | OZ401\_000873 | HXX20\_19450 |
| *chlG* | OZ401\_002374 | HXX20\_02935 |
| *bchF* | OZ401\_001799 | HXX20\_18350 |
| *bchX* | OZ401\_003345 | HXX20\_00700 |
| *bchY* | OZ401\_003346 | HXX20\_00695 |
| *bchZ* | OZ401\_003347 | HXX20\_00690 |
| *bchC* | OZ401\_002223 | HXX20\_23950 |
| *bchG* | OZ401\_001677 | HXX20\_04355 |
| *bchP* | OZ401\_000757 | HXX20\_18140 |
| *bciC* | OZ401\_003683 | HXX20\_08825 |
| *bchR* | OZ401\_000325 | HXX20\_16045 or HXX20\_11195 |
| *bchV* | OZ401\_002358 | HXX20\_00235 |
| *bchU* | OZ401\_002306 | HXX20\_11155 |
| *bchK* | OZ401\_002285 | HXX20\_10895 |
| Reductive pentose phosphate cycle | *prk* | OZ401\_002094 | HXX20\_11340 |
| *rbcL* | OZ401\_002090 | HXX20\_11370 |
| *rbcS* | OZ401\_002100 | HXX20\_11360 |
| Nitrogen fixation | *nifH* | OZ401\_004307 | HXX20\_13555 |
| *nifD* | OZ401\_004303 | HXX20\_13535 |

# Extended Data Table 3 | Summary of physicochemical parameters for sampled Boreal Shield lakes. The table includes the depth and surface area of all nine sampled lakes, along with information on the depths sampled for metagenome/metatranscriptome sequencing between 2016-2018. Summary parameters on the right side of the table show the topmost depth (sampled for metagenome sequencing) where dissolved oxygen was undetectable and the maximum measured concentrations of total dissolved iron, sulfate, and dissolved organic carbon among the collected samples. Full physicochemical data are provided in Supplementary Data 6. Abbreviations: n.d. no data.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Lake** | **Maximum depth (m)** | **Surface area (m2 x 104)** | **Sampling year** | **Sampling month** | **Samples for metagenome sequencing (m)** | **Anoxic zone? (m)** | **Total dissolved iron (max.; µM)** | **Sulfate (max.; µM)** | **Dissolved organic carbon (max.; µM C)** |
| L227 | 10.0 | 5.0 | 2016 | Jun | 6, 8, 10 | Yes (6) | 162.5 | 6.1 | 1092 |
| 2016 | Sep | 6, 8, 10 | Yes (6) | 188.1 | 8.6 | 1014 |
| 2017 | Sep | 1, 3, 4\*, 6, 8, 10 | Yes (6) | 222.9 | 23.3 | 1271 |
| L221 | 5.5 | 9.0 | 2016 | Jun | 5 | Yes (5) | 15.2 | 11.7 | 844 |
| 2018 | Jul | 5\* | Yes (5) | 68.4 | 12.2 | 922 |
| L304 | 6.0 | 3.6 | 2016 | Jun | 6 | Yes (6) | 37.6 | 9.3 | 676 |
| 2018 | Jul | 6\* | Yes (6) | 137.6 | 8.2 | 865 |
| L222 | 6.0 | 16.4 | 2016 | Jun | 5 | No | 5.1 | 12.9 | 681 |
| 2016 | Sep | 5 | Yes (5) | 84.5 | 11.3 | 775 |
| L224 | 25.0 | 25.9 | 2016 | Jun | 25 | No | 0 | 17.8 | 244 |
| 2016 | Sep | 21, 25 | Yes (21) | 102.5 | 8.6 | 304 |
| L373 | 20.0 | 27.3 | 2016 | Sep | 20 | Yes (20) | 207 | 11.7 | 352 |
| L442 | 17.0 | 16.0 | 2016 | Jun | 9, 12, 15, 17 | Yes (15) | 170.7 | 18 | 759 |
| 2016 | Sep | 9, 13, 15, 17 | Yes (9) | 100.7 | 19.8 | 636 |
| L626 | 11.0 | 25.9 | 2016 | Jun | 11 | No | 0 | 12.8 | 361 |
| 2016 | Sep | 11 | Yes (11) | 19.6 | 15.1 | 363 |
| L239 | 30.0 | 54.3 | 2016 | Jun | 10, 20 | No | n.d. | 22.1 | 517 |
| 2016 | Sep | 10, 20 | No | 12.3 | 25.2 | 520 |
| 2018 | Jul | 20 | n.d. | n.d. | n.d. | n.d. |

\*Samples were also collected and used for metatranscriptome sequencing