Microbial life under ice: metagenome diversity and *in situ* activity of Verrucomicrobia in seasonally ice-covered lakes

Patricia Tran[1,2], Arthi Ramachandran[1], Ola Khawasik[1], Beatrix E. Beisner[2,3], Milla Rautio[2,4], Yannick Huot[2,5], David A. Walsh[1,2]

[1] Department of Biology, Concordia University, 7141 Sherbrooke St. West, Montreal, Quebec, H4B 1R6, Canada

[2] Groupe de recherche interuniversitaire en limnologie et environnement aquatique (GRIL), Montréal, Québec, Canada

[3] Département des sciences biologiques, Université du Québec à Montréal, Québec, Canada.

[4] Département des sciences fondamentales, Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada

[5] Département de géomatique appliquée, Université de Sherbrooke, Sherbrooke, Québec, Canada

Corresponding author: David Walsh, Department of Biology, Concordia University, 7141 Sherbrooke St. West Montreal, QC H4B 1R6 Canada

514-848-2424 ext 3477
david.walsh@concordia.ca

Running title: Sub-ice Verrucomicrobia genomes in Quebec lakes

Summary

Northern lakes are ice-covered for a large part of the year, yet our understanding of microbial diversity and activity during winter lags behind that of the ice-free period. In this study, we investigated under-ice diversity and metabolism of Verrucomicrobia in seasonally ice-covered lakes in temperate and boreal regions of Quebec, Canada using 16S rRNA sequencing, metagenomics and metatranscriptomics. Verrucomicrobia, particularly the V1, V3 and V4 subdivisions, were abundant during ice-covered periods. A diversity of Verrucomicrobia genomes were reconstructed from Quebec lake metagenomes. Several genomes were associated with the ice-covered period and were represented in winter metatranscriptomes, supporting the notion that Verrucomicrobia are metabolically active under ice. Verrucomicrobia transcriptome analysis revealed a range of metabolisms potentially occurring under ice, including carbohydrate degradation, glycolate utilization, scavenging of chlorophyll degradation products, and urea use. Genes for aerobic sulfur and hydrogen oxidation were expressed, suggesting chemolithotrophy may be an adaptation to conditions where labile carbon may be limited. The expression of genes for flagella biosynthesis and chemotaxis was detected, suggesting Verrucomicrobia may be actively sensing and responding to winter nutrient pulses, such as phytoplankton blooms. These results increase our understanding on the diversity and metabolic processes occurring under ice in northern lakes ecosystems.

Originality and significance

Reduced ice cover on lakes is occurring worldwide, but there is only limited information on the biogeochemistry and microbiology under ice. This gap in knowledge limits our ability to understand and predict how changes in winter conditions will impact the ecology of lake ecosystems. In this study, we conducted the first meta-omics assessment of bacterial metabolism and gene expression patterns in seasonally ice-covered lakes. Previously uncharacterized lineages within the Verrucomicrobia were strongly associated with ice-covered conditions. Analysis of Verrucomicrobia genomes and gene expression patterns revealed a diversity of functional trait including the use of organic and inorganic energy sources and potential metabolic interactions with winter phytoplankton and zooplankton. The identification of winter-associated lineages and variable metabolic processes indicated that under-ice microbial communities may contribute uniquely to the ecology and nutrient cycling of seasonally ice-covered lakes. This study demonstrates the importance of studying the ice-covered period in the face of climate change and should spur future year-round investigations on microbial community structure and function in ice-covered freshwater ecosystems.

Introduction

Many northern lakes are seasonally ice-covered for over 40% of the year (Weyhenmeyer *et al.*, 2011). During that time, these lakes are typically characterised by low light levels and primary productivity, leading to the traditional idea that they are "dormant" in winter (Bertilsson *et al.*, 2013). However, recent studies describing phytoplankton and bacterial blooms as well as specialized microbial metabolism under ice have challenged this concept (Kankaala *et al.*, 2006; Twiss *et al.*, 2012; Bertilsson *et al.*, 2013; Bižić-Ionescu *et al.*, 2014; Beall *et al.*, 2016; Powers and Hampton, 2016). For example, light penetrating through ice and snow has been shown to fuel substantial primary production and blooms of low light and cold-adapted phytoplankton do occur under the ice (Twiss *et al.*, 2012; Üveges *et al.*, 2012). These phytoplankton blooms can fuel growth of other organisms, including heterotrophic bacteria (Bižić-Ionescu *et al.*, 2014). Moreover, unique niches can develop and persist in ice-covered lakes. Where ice-cover results in lower availability of labile organic substrates from phytoplankton or terrestrial input, volatile compounds such as methane and fermentation end-products produced from anoxic bottom waters or sediments can fuel microbial metabolism (Sundh *et al.*, 2005; Kankaala *et al.*, 2006). Chemolithoautotrophs may also be significant under ice since their sources of energy and carbon acquisition may not be as restricted by ice cover compared to phototrophs and heterotrophs (Auguet *et al.*, 2011). Despite these observations, our understanding of the structure and function of microbial communities under ice lags far behind that of the ice-free period (Bertilsson *et al.*, 2013; Powers and

Hampton, 2016; Hampton *et al.*, 2017). To better understand the contribution of winter communities to lake metabolism and nutrient cycling, year-round investigation of the metabolic traits and activities of freshwater microorganisms is warranted.

Verrucomicrobia are ubiquitous in lakes, yet knowledge of their metabolism and ecology in freshwaters is limited compared to other bacterial groups (Newton *et al.*, 2011). The phylum is comprised of five orders, also referred to as subdivisions (V): Verrucomicrobiae (V1), Spartobacteria (V2), Pedosphaera (V3), Opitutae (V4), and Methylacidiphilum (V6) (Hedlund *et al.*, 1996; Sangwan *et al.*, 2004; Choo *et al.*, 2007; Hou *et al.*, 2008; Kant *et al.*, 2011). 16S rRNA gene surveys have identified all Verrucomicrobia subdivisions in lakes, however, only a few pelagic freshwater isolates exist, all belonging to V1 (Hedlund *et al.*, 1996). In general, while Verrucomicrobia are metabolically diverse, in aquatic ecosystems they are often associated with the degradation of carbohydrates (Martinez-Garcia *et al.*, 2012; Herlemann *et al.*, 2013). For example, analyses of single cell-amplified genomes (SAGs) and metagenome-assembled genomes (MAGs) from coastal environments identified preferences for laminarin, xylan (Martinez-Garcia *et al.*, 2012), cellulose and chitin polymers (Herlemann *et al.*, 2013). Recently, comparative analysis of Verrucomicrobia MAGs from eutrophic Lake Mendota and dystrophic Trout Bog in Wisconsin, USA showed differences in the number and type of glycoside hydrolases (GHs) between the two systems, reflecting adaptations to local environments and carbon substrate availability (He *et al.*, 2017). Additionally, MAGs from freshwater reservoirs in Spain also suggested a preference for polysaccharides, but

also identified genes for the light driven proton pump rhodopsin as well as genes for nitrogen fixation (Cabello-Yeves, Ghai, *et al.*, 2017). Perhaps most relevant from a life under ice perspective, a recent study identified abundant populations of Verrucomicrobia under the ice of Lake Baikal in Siberia, and these lineages also contained multiple polysaccharide degradation pathways in their genomes (Cabello-Yeves, Zemskaya, *et al.*, 2017). Given their widespread distribution across freshwater ecosystems and their prominence under ice, Verrucomicrobia may serve as a suitable model for investigating metabolic adaptations and lifestyle strategies associated with seasonally ice-covered lakes.

Here we present a study on Verrucomicrobia in seasonally ice-covered lakes in Quebec, Canada with a general aim of advancing our understanding of the genomic diversity and metabolic traits of bacterial communities residing under the ice of northern lakes. Using a combination of 16S rRNA gene sequencing, metagenomics, and metatranscriptomics we identified a wide diversity of Verrucomicrobia, including populations and MAGs strongly associated with ice-covered periods. The MAG-associated transcriptomes revealed a range of expressed metabolic genes under ice, including those for the use of phytoplankton and plant-derived organic compounds, chemolithotrophy, as well as motility and chemotaxis. Overall, our study supports the increasing recognition that the winter period represents a dynamic and metabolically important period for lakes despite ice cover.

Results

Verrucomicrobia diversity and distribution in Quebec Lakes

To investigate Verrucomicrobia abundance and diversity during ice-covered and ice-free periods of the year, we analyzed a 3-year time-series of bacterial 16S rRNA gene diversity from three seasonally ice-covered lakes located in the temperate (Lake Croche and Lake Montjoie) and boreal (Lake Simoncouche) regions of Quebec (**Fig. 1A, Table S1**). In total, we generated 16S rRNA data from 143 samples collected from the epilimnion and metalimnion, and corresponding to 6 winter time-points (January and February 2013, 2014, and 2015) and 8 summer time-points (June, July, August 2013 and 2014). The study lakes differed in several environmental characteristics, but were mainly differentiated by nutrient concentrations in a principal component analysis (**Fig. 1B**). The three lakes are distributed along a nutrient gradient, with Croche generally having lower total phosphorus (TP) than Montjoie and Simoncouche (**Fig. S1, Table S2)**.

Verrucomicrobia were common across all lakes, but generally exhibited higher relative abundance of 16S rRNA gene sequences during ice covered period, particularly in Croche and Simoncouche (**Fig. 1C**). The average relative abundance of Verrucomicrobia was about 4-fold higher during ice cover (13%) compared to the ice-free period (4%) in Croche and about 2-fold higher in Simoncouche (10% compared to 5%). Although more abundant under the ice on average, Verrucomicrobia were highly dynamic in time with maximum observed values of 35.6 % and 27.1% in Croche and Simoncouche during winter time points. Verrucomicrobia in Croche were dominated by

V1 and V3, while V4 was the major contributor in Simoncouche. In contrast to Croche and Simoncouche, Verrucomicrobia exhibited a similar mean relative abundance between ice-covered (4.8 %) and ice-free (6.2 %) periods in Montjoie. Maximum values in Montjoie were observed during summer time points and were comprised mostly of V6. Although relative abundance differed between lakes, these results demonstrate Verrucomicrobia are often associated with ice-covered conditions, making them an intriguing group with which to investigate genomic and metabolic adaptations to life under ice.

Verrucomicrobia MAG diversity and gene transcription

Verrucomicrobia genome diversity was investigated in a metagenome co-assembly of 24 samples collected over the 3-year times-series from the three lakes. Following metagenomic binning, we identified 54 Verrucomicrobia MAGs representing all 5 subdivisions (V1 to V4 and V6) previously observed in the 16S rRNA gene survey (**Fig. 2A**). To investigate the distribution of MAGs across lakes and seasons, we performed a Canonical Correspondence Analysis (CCA) of MAG coverage across samples, constrained using 14 environmental variables (see methods for definitions of abbreviations). In the resulting CCA, samples were differentiated along two primary axes (**Fig. 2B**). Croche samples were separated from Montjoie and Simoncouche samples along CCA axis 1, which correlated with nutrient concentrations (*i.e.* TP and TN). Ice-covered samples were generally separated from ice-free samples along CCA axis 2, which differentiated samples from the ice-covered and ice-free period, with the exception

of a single summer Montjoie sample, which clustered with winter samples. A clear separation of MAGs along both axes of the ordination was observed, revealing lake and seasonal preferences (**Fig. 2C**). Interestingly, MAGs within a subdivision, and even those closely related within a subdivision, did not cluster together in the CCA, suggesting niche diversity among members of the same subdivision; taxonomy clearly does not reflect niche preference.

Fifteen MAGs of high completeness and low contamination were selected for further analyses (**Table 1**). These MAGs represented four subdivisions (V1 to V4). The MAGs were distributed across the CCA plot, and are therefore representative of the observed phylogenetic diversity, biogeography, and seasonal associations of Verrucomicrobia MAGs. A concatenated protein phylogeny that included 43 additional MAGs from Wisconsin lakes (He *et al.*, 2017), Lake Baikal (Cabello-Yeves, Zemskaya, *et al.*, 2017) and the Tous and Amadorio reservoirs in Spain (Cabello-Yeves, Ghai, *et al.*, 2017) was generated in order to place the Quebec MAGs in the context of known Verrucomicrobia genomic diversity (**Fig. 3A**). The phylogeny presented in Figure 3A was inferred from a concatenated alignment of four of the five proteins used in He et al., 2017, but exhibited a similar topology to a phylogeny generated based on hundreds of proteins (**Fig. S2**). MAGs were too distantly related to calculate average nucleotide identity (ANI), but average amino acid identity between Quebec MAGs and those from other locations ranged between 91-49% (**Table S3**). For the most part, Quebec MAGs exhibited <70% AAI with other MAGs and therefore represent new lineages for which

genome sequence data is now available. Although common across Quebec lakes, fragment recruitment of metagenomes originating from Wisconsin lakes, Lake Baikal, and the Spanish reservoirs demonstrated that the Verrucomicrobia populations represented by the MAGs are relatively rare in these other freshwater systems (**Table S4**).

As revealed by their wide distribution in the CCA ordination, the MAGs exhibited complex distribution patterns across Quebec lakes (**Fig. 3B**). To investigate gene expression patterns of the Verrucomicrobia MAGs during ice-covered and ice-free periods in Quebec, a temporally overlapping metatranscriptomic times-series was mapped to the MAGs, providing a view of their transcriptional activity across lakes and seasons. A significant number of transcripts were observed for all MAGs, and for the most part reflected MAG distributions across lakes and seasons (**Fig. 3B**). In the following section, we analyzed the distribution, metabolic gene content and gene expression patterns of Verrucomicrobia MAGs with the objective of providing insights into metabolic diversity and activity associated with ice-covered conditions in northern lakes.

Ice cover associated MAGs within V1

A striking association with winter conditions was observed within certain V1 MAGs. Particularly, three MAGs (V1-159, V1-690, and V1-1361) were associated with ice-covered conditions in Croche. These MAGs formed a clade with other freshwater MAGs but were quite unique, sharing between 51-72 % AAI (**Fig 3A, Table S3**). V1-690

and V1-159 were exclusive to winter periods both at the level of the genome and the transcriptome (**Fig. 3B**), while V1-1361 was also identified in the summer. A total of 3,983 orthologs were identified among these three MAGs and 868 were common to all (**Fig. 4A**). Similar to other Verrucomicrobia, the winter-associated V1-159 and V1-690 MAGs contained numerous glycoside hydrolase (GH) genes (**Table S5**). Expression of 33 of 34 GH genes was detected in V1-159. The GHs of highest expression were GH29 (alpha-L-fucosidase) and GH16 (substrate specificity undetermined). In contrast, expression was only detected for 12 of 54 GH genes in V1-690, suggesting other modes of carbon and energy metabolism may be important for this population. Indeed, four proteins annotated as monooxygenases were specific to V1-690 and were relatively highly expressed. Two were annotated as limonene 1,2-monooxygenase and a third as 1 alkane 1-monooxygenase/p-cymene monooxygenase (**Fig. 4B**). Limonene and cymene are plant-derived aromatic compounds, suggesting that these proteins are involved in accessing terrestrial organic matter. In addition, we identified a genomic region encoding the Sox system (*soxXYZABC*) together with two c-type cytochromes (**Fig. 4C**). The Sox system is associated with the use of the reduced sulfur compound thiosulfate as an electron donor in energy metabolism (Ghosh and Dam, 2009). All subunits were expressed during winter except for *soxZ*, suggesting the potential for lithotrophic sulfur oxidation during ice-covered periods.

204 **Ice cover associated MAGs within V4**

205 Specific winter association was observed for V4 MAGs. Two MAGs (Opi-128
206 and Opi-474) were phylogenetically related to MAGs from Trout Bog and the Tous
207 reservoir. Opi-128 exhibited 81 % AAI with TH02519 and TH01800 and Opi-474
208 exhibited 82 % AAI with TH4593, while Opi-128 exhibited 91 % AAI with Tous-
209 C10FEB (**Fig. 3A, Table S3**). Opi-474 and Opi-242 exhibited winter association in
210 Croche and Simoncouche, respectively (**Fig. 3B**). Opi-128 was more commonly
211 associated with summer. A total of 3,776 orthologs were identified among the three
212 genomes and only 11% were shared among all MAGs (**Fig. 4A**), indicating substantial
213 genomic diversity among V4 MAGs. Winter associated Opi-474 and Opi-242 contained
214 70 and 19 GH genes, respectively (**Table S4**). Compared to V1 winter MAGs, the
215 abundance of GH transcripts was elevated in the Opi-474 transcriptome. Among the
216 expressed GHs were GH78 genes involved in the use of rhamnose-containing
217 polysaccharides.

218 A common feature of Opi-474 and Opi-242 was the expression of a predicted
219 fatty acid α-oxidation II pathway (**Fig. 4D**). Fatty acid α-oxidation is implicated in the
220 metabolism of phytol, a long chain alcohol constituent of chlorophyll (Jansen and
221 Wanders, 2006). Phytol is first converted to phytanoyl-CoA, which then enters the α-
222 oxidation pathway where a methyl group at the C-3 position is removed before passage to
223 the more common β-oxidation pathway (Jansen and Wanders, 2006). Both genomes
224 encoded proteins of the phytanoyl-coA dioxygenase family, which hydroxylates the

methyl-branched fatty acid preparing it for downstream cleavage and passage to the β-oxidation pathway. A putative pathway containing alcohol and aldehyde dehydrogenases necessary for transforming phytol to phytanoyl-CoA was also identified. The presence of a putative α-oxidation pathway and upstream steps for introducing phytol may allow these organisms to scavenge carbon and energy from chlorophyll degradation products.

Evidence for chemolithotrophic energy conservation was observed in V4. A complete gene cluster associated with aerobic hydrogen oxidation was specifically identified in Opi-474, including genes encoding a Group 1d oxygen-tolerant hydrogenase (Greening *et al.*, 2016) and associated maturation and nickel incorporation proteins (**Fig. 4C**). Expression of a number of these genes was detected under the ice in Croche, suggesting the use of hydrogen as an electron donor in energy metabolism.

Complex seasonal dynamics in V2 and V3 MAGs

Compared to V1/V4 MAGs, V2 and V3 MAGs exhibited more complex distribution patterns across lakes and seasons. In V3, Pedo-303 and Pedo-1123 branched deeply within V3, and were present throughout the year in Montjoie, while Pedo-510 formed a clade with MAG recovered from Trout Bog and was present throughout the year in Montjoie and Simoncouche (**Fig. 3**). V3 Quebec MAGs exhibited less than 66 % AAI with MAGs from Trout Bog and the Tous reservoir.

The most complex seasonal patterns were observed for the four V2 MAGs. Two MAGs (Xiphi-315 and Xiphi-554) were members of the *Xiphinematobacter*. Xiphi-315 was common but variable in Croche summer samples, while Xiphi-554 was restricted to

the June epilimnion in Simoncouche. Quebec *Xiphinematobacter* MAGs (Xiphi-315 and Xiphi-554) were within the same clade as three MAGs from the Tous and Amadorio reservoirs, but shared at most 63% AAI with them, reflecting high genome variation between MAGs from Quebec lakes and other freshwater environments. Interestingly, Xiphi-554 was the most widely distributed in other freshwater systems based on metagenome fragment recruitment patterns (**Table S4**)

The two other V2 MAGs (Chth-244 and Chth-196) were members of the *Chthoniobacter*. Chth-196 was the most broadly distributed MAG of all in Quebec lakes, being common in Montjoie and Simoncouche but exhibiting high gene expression levels in all. However, Chth196 was the only MAG that did not recruit any other freshwater metagenome reads (**Table S5**), indicating very unique populations in Quebec lakes. Chth-244 was common throughout the year in Montjoie although transcripts were only detected during ice cover.

With respect to physiological adaptations to life under ice, Pedo-303 (V3) and Chth-196 (V2) are of interest because differences in the number of transcripts between summer and winter was detected for both MAGs, providing insight into metabolic responses to different seasons. Although Pedo-303 and Chth-196 were distantly related from a phylogenetic perspective, they exhibited some similarities at the metabolic level. For example, ammonia and urea appear to be important nitrogen sources for both groups. Urea transport proteins and urease subunits were identified in Chth-196 and generally had higher expression during winter than summer in Simoncouche. Urea transport and urease

genes were also present in Pedo-303 and more transcripts were detected during ice cover than the ice-free period in Montjoie (**Fig. 5C**). Furthermore, Pedo-303 had genes implicated in ammonia transport and assimilation, including three copies of the nitrogen regulatory proteins P-II1, two Amt family ammonium transporters, and two glutamine synthetases, which had higher levels of expression in the winter (**Fig. 5D**). In Chth-196, three Amt family transporters, one glutamine synthetase, and three nitrogen regulatory proteins P-II-1 were found, and these genes were expressed more in Simoncouche during ice cover, but more in Montjoie during the summer.

Pedo-303 and Chth-196 both encoded genes for the use of a range of organic carbon compounds. In Pedo-303, rhamnose and xylose utilization genes were expressed at both time points, but transcript abundance was higher in the winter. In Chth-196, genes for rhamnose and ribose transport and degradation were identified and expressed during winter. The use of glycolate, a photorespiration product of phytoplankton, was also suggested by the presence of a glycolate utilization operon (Fe-S subunit, FAD-binding subunit, oxidase) in Pedo-303. Transcripts for glycolate use were only detected during winter in Montjoie.

Another notable observation within V2 and V3 MAGs was the relatively high number of transcripts for genes involved in motility and chemotaxis. Genes for flagellar biosynthesis were identified in numerous MAGs (Chth-196, Chth-244, Opi-474, Pedo-510 and Xiphi-554), but evidence for active use of the flagellar machinery was found in Chth-196 and Pedo-510 only. In Chth-196, transcripts encoding flagellar motility proteins

288 were generally more abundant in summer in Montjoie, but in winter in Simoncouche
289 (**Fig. 5A**). In addition, Chth-196 encoded genes involved in chemotaxis behaviour were
290 more abundant in winter than summer in Simoncouche (**Fig. 5B**).

291 **Discussion**

292 In this study, we conducted the first meta-omics assessment of bacterial
293 metabolism and gene expression patterns in seasonally ice-covered lakes. In combination
294 with other recent metagenomic studies (Cabello-Yeves, Ghai, *et al.*, 2017; Cabello-
295 Yeves, Zemskaya, *et al.*, 2017; He *et al.*, 2017), our findings contribute to the emerging
296 view that Verrucomicrobia is an important, but previously overlooked component of lake
297 microbial ecosystems. The association of Verrucomicrobia MAGs with ice-covered
298 conditions suggests certain populations exhibit a winter preference. However, an
299 alternative explanation for greater under-ice abundance is that Verrucomicrobia prefer as
300 habitat the deeper hypolimnetic regions of northern lakes, and thus they could appear
301 winter-associated following autumn mixing of the water column. Although we cannot
302 rule out this possibility, it seems unlikely given that sampling occurred several months
303 following the onset of fall overturn. Furthermore, we did not observe evidence for
304 hypolimnetic preferences via anaerobic metabolism in the Verrucomicrobia MAGs
305 despite the fact that Croche and Simoncouche have anoxic hypolimnia during summer.
306 However, Verrucomicrobia have been observed in the oxic hypolimnion of the Tous and
307 Amadorio reservoirs (Cabello-Yeves, Ghai, *et al.*, 2017). Moreover, studies of the sub-ice

308 microbial communities of deep Lake Baikal have shown Verrucomicrobia are among the
309 most dominant groups at the surface (Cabello-Yeves, Zemskaya, *et al.*, 2017) and that
310 their abundances are linked to diatom blooms (Bashenkhaeva *et al.*, 2015). Hence, it
311 appears that Verrucomicrobia populations represented by our MAGs are resident
312 members of the microbial community in ice-covered periods in northern lakes, and
313 knowledge of their genomic and metabolic traits can contribute to our understanding of
314 lake metabolism and nutrient cycling in winter.

315 Ice-covered conditions may favour microbes that have low resource requirements
316 (*i.e.* oligotrophs) and small genome size is a common adaptation to oligotrophic
317 conditions (Carini *et al.*, 2012; Neuenschwander *et al.*, 2017). Previous analysis of
318 Verrucomicrobia MAGs have shown a wide range in genome sizes (Cabello-Yeves, Ghai,
319 *et al.*, 2017). We speculated that winter-associated MAGs may be specialized to low
320 carbon/energy conditions and that this would manifest in overall smaller genomes
321 compared to those associated with the ice-free period. However, we did not observe
322 generalizable differences in estimated genome size across seasons within Quebec lakes.
323 Rather, we observed that overall estimated genome size was on average smaller in
324 Verrucomicrobia MAGs from more oligotrophic lakes (Quebec, Lake Baikal, and
325 freshwater reservoirs) compared to eutrophic lakes (Mendota and Trout Bog). Wisconsin
326 MAGs had the largest average estimated genome size in Subdivision 1, 2 and 3 compared
327 to Quebec, Baikal, Tous and Amadorio MAGs. The largest average estimated genome
328 size in V4 was in Tous and Amadorio MAGs, while the smallest was among Quebec

MAGs. Based on our results, small genome size does not seem to be distinguishing characteristic of winter association, at least for Verrucomicrobia, but instead reflects adaptations to oligotrophic freshwater environments in general.

Verrucomicrobia-phytoplankton coupling

Phytoplankton can persist and form transient blooms during ice cover (Twiss *et al.*, 2012) and a number of studies have implicated Verrucomicrobia in the degradation of phytoplankton-derived carbohydrates(Paver and Kent, 2010; Parveen *et al.*, 2013), including in ice-covered lakes (Bižić-Ionescu *et al.*, 2014; Beall *et al.*, 2016). For example, Verrucomicrobia were among the main bacteria associated with diatom-dominated under ice blooms in Lake Baikal (Bashenkhaeva *et al.*, 2015). Furthermore, Verrucomicrobia OTUs, including *XipA1*/ *XipB1* (V2) and *Opitutacea* (V4) were strongly correlated with algal carbon in Finnish subarctic lakes (Roiha *et al.*, 2016). Finally, Verrucomicrobia (*Verrucomicrobiae* (V1) and *Opitutacea* (V4)) abundance significantly increased following the release of extracellular polymeric substances released by diatoms in intertidal zones (Bohórquez *et al.*, 2017)

Numerous traits of the Quebec lake Verrucomicrobia suggest a capacity to couple growth to phytoplankton, including winter phytoplankton blooms. Based on 16S rRNA gene analysis, Verrucomicrobia subdivisions were commonly abundant in both the free-living and particle-attached fractions of the community, indicating an ability to switch between these lifestyles. During winter, cells may persist in a free-living state but also

associate with phytoplankton during bloom onset and progression. If so, then the MAGs analysed here originated from the free-living persisters. Interestingly, GH genes were identified in all MAGs, with no strong correlation between estimated genome size and number of GH found. However, we found a wider diversity of GHs was expressed in winter compared to summer, perhaps in readiness for persister cells to quickly respond to carbohydrate that becomes available, both from phytoplankton or terrestrial sources. Expression of genes for the use of phytoplankton-derived saccharides including fucose and rhamnose were identified, suggesting that Verrucomicrobia interact with phytoplankton during winter. Future studies comparing gene expression patterns between particle-attached and free-living cells, would be informative in understanding the metabolic shifts that occur during these transient bloom events.

Intriguingly, expression of a glycolate degradation operon was detected in ice cover associated MAGs from Lake Montjoie. Glycolate is a photorespiration product (Fogg, 1983) that is typically produced by phytoplankton growing under high light stress (Parker and Armbrust, 2005; Davis *et al.*, 2013). Glycolate-utilizing bacteria have been shown to be transcriptionally responsive to phytoplankton blooms (Lau *et al.*, 2007). Glycolate production during winter may be unexpected owing to low light penetration through snow and the absence of light stress (Maxwell *et al.*, 1994). However, low-light adapted phytoplankton may have a lower tolerance to light overall. Rapid increases in light intensity due to blowing/melting snow could lead to light stress, and a corresponding pulse of glycolate into the environment. Another possibility is that

371 glycolate produced during the ice-free period is relatively long-lived in the water column.
372 Glycolate is present in the ocean year-round, and can account for as much as 33% of the
373 dissolved organic carbon pool (Carlson and Ducklow, 1996; Leboulanger *et al.*, 1997). In
374 any case, it is believed that the main role of glycolate in heterotrophic metabolism is as an
375 energy source (Wright and Shah, 1977; Edenborn and Litchfield, 1987). Recently, a study
376 of carbon utilization in temperate lakes showed that a relatively larger portion of
377 phytoplankton-derived carbon was allocated to respiration (and hence energy
378 conservation), while terrestrial carbon was allocated to biosynthesis (Guillemette *et al.*,
379 2016). Therefore, glycolate may be an important contributor to the maintenance energy of
380 a persisting cell or could facilitate growth by serving as an energy source, fuelling the
381 subsequent assimilation of terrestrial carbon in winter.

382 Phytoplankton blooms and resource availability under ice are often patchy.
383 Motility and chemotaxis may be useful adaptations to efficiently exploit hotspots of
384 organic and inorganic nutrients (Stocker and Seymour, 2012). This idea is supported by a
385 number of Verrucomicrobia MAGs that expressed genes for flagella and chemotactic
386 abilities. Of particular interest were two closely related winter MAGs from Montjoie
387 (Chth-244) and Simoncouche (Chth-196). Flagella and chemotaxis gene expression was
388 evident in the Montjoie MAG. Despite the presence of flagella and chemotaxis genes in
389 the Simoncouche MAG, only the flagella genes were detectably expressed. The
390 difference in chemotactic activity may reflect the difference in lake physical conditions;
391 while Montjoie is unstratified in winter, Simoncouche is inversely stratified. Hence, the

ability to detect and respond to chemical gradients may be more advantageous in stratified systems. Operating at the micro-scale rather than meters, this notion of motility towards hotspots of nutrients is analogous to findings in phytoplankton motility. For example, in a 20-year study of phytoplankton functional traits in an ice-covered lake in Germany, authors found that during mild winters a mixed water column favoured non-motile phytoplankton (Özkundakci *et al.*, 2016), as in Montjoie. Therefore, the difference in water column mixing between lakes is potentially reflected in bacterial functional traits as well.

Scavenging of the phytol moiety of chlorophyll

In addition to responding to algal growth, bacterial scavenging of dead and degraded phytoplankton and other detritus may be an important metabolic strategy for life under ice. Labile algal carbohydrates and proteins would likely be the first biomolecules scavenged from the environment, but carbon-rich lipids may also sustain growth during winter. In freshwater reservoirs, Verrucomicrobia metagenomes contained several enzymes (aryl sulfatases, beta-galactosidases and sialidases) for degradation of glycosphingolipids and are suggested to be involved in the degradation of plant or algae derived sulfur-containing lipids (Cabello-Yeves, Ghai, *et al.*, 2017). In our study, winter MAGs encoded for a predicted α-oxidation pathway. The α-oxidation pathway is implicated in the degradation of phytol, the long-chained alcohol moiety of chlorophyll. Early studies showed growth of bacteria on phytol as a sole source of carbon and energy (Hoag *et al.*, 1969), however the ecological significance of this metabolism is unknown.

413 The concentration of free phytol in the water column may be low, but we can envision
414 microenvironments where phytol concentrations are elevated. One micro-environment
415 may be formed by the fecal pellets of zooplankton. Early studies have found that the fecal
416 pellets of zooplankton, specifically coccoliths feeding on an algal-rich diet, were rich in
417 phytanic acid, the end product of phytol oxidation that feeds into the α-oxidation pathway
418 (Moussa, 1988). The phytanic acid-rich particles would be available in the water column
419 for consumption by bacteria possessing the α-oxidation pathway

420 Additionally, it is possible that phytol may accumulate in particles that sink into
421 the anoxic hypolimnion of lakes. Since the α-oxidation pathway requires oxygen (through
422 the essential phytanoyl CoA hydroxylase), phytol may accumulate under anoxic
423 conditions, although anaerobic degradation of phytol in sediments has been reported
424 (Rontani *et al.*, 1999). Nevertheless, during breakdown of stratification in autumn, phytol
425 accumulated in the hypolimnion could serve as a carbon reserve supporting microbial
426 metabolism throughout the winter. A recent study in Lake Simoncouche has shown that
427 zooplankton are able to store fatty acids from phytoplankton to survive over the winter
428 (Grosbois *et al.*, 2017). Similarly, bacteria might be able to use zooplankton and
429 phytoplankton-derived lipids (e.g. phytanic acid) to maintain baseline metabolism under
430 ice. These observations point to the complex relationships among bacteria,
431 phytoplankton, and zooplankton occurring under the ice in freshwater lakes.

Chemolithotrophy under ice

Chemolithotrophic growth is traditionally thought to be restricted to sediments or the chemocline of stratified lakes where there is a sufficient supply of reduced inorganic compounds for metabolism. Here we detected the expression of sulfur and hydrogen gene clusters, suggesting that chemolithotrophic energy metabolism may be employed by bacteria under the ice. Similar sulfur oxidation genes were also recently reported in *Beta-proteobacteria* MAGs from Lake Baikal (Cabello-Yeves, Zemskaya, *et al.*, 2017). Heterotrophic organisms capable of supplementing their energy demand using inorganic compounds would have an advantage over those that cannot if the availability of organic carbon is limited in winter. A wide diversity of so-called "heterotrophic sulfur-oxidizing" bacteria are found in marine systems (Teske *et al.*, 2000). In these cases, sulfur oxidation allows bacteria to allocate organic carbon for biosynthesis instead of respiration, allowing them to thrive in a wider range of habitats (Teske *et al.*, 2000; Podgorsek *et al.*, 2004). Similarly, the winter water column availability of H_2 (previously restricted to the anaerobic hypolimnion), combined with the availability of oxygen under ice, may allow aerobic hydrogen-oxidizing to use this energy rich fermentation end-product for growth or maintenance energy.

Urea as an important N source in lake

The observation of urea utilization genes in our MAGs further suggests the potential for trophic interactions. Although urea is usually present in lakes at ambient

concentrations below 1 µM-N, it can contribute 50% or more of the total N used by planktonic communities (Solomon et al., 2010). Interestingly, urea availability is predicted to be regulated mainly by the decomposition of algae under anoxic conditions (e.g. hypolimnion/sediments), followed by redistribution in the water column (Bogard *et al.*, 2012) and thus potentially serving as a valuable nitrogen source in winter. All urease subunits were found and expressed in the winter in Chth-196 and Pedo-303. A recognized source of N for polar phytoplankton, the importance of urea for bacteria has received less attention. In the arctic, urea utilization was detectable in the prokaryotic size fraction, but not that corresponding to the phytoplankton fraction, showing the potential importance of urea for cold-adapted metabolism, and adaptation to low energy environments (Alonso-Sáez *et al.*, 2012). In Lake Baikal, urea utilization genes were found in *Cyanobacteria, Acidobacteria, Nitrospearea, Beta-proteobacteria and Thaumarchaeota* MAG (Cabello-Yeves, Zemskaya, *et al.*, 2017). The expression of the urea utilization genes by Verrucomicrobia that we observed suggests that urea can be beneficial for bacterial species living under the ice in temperate and boreal lakes as well, given the similar constraints of temperature, light and organic nutrient availability.

Implications

There is a growing acknowledgment among aquatic ecologists as to the need to study the full annual cycle of lakes in order to understand lake dynamics, particularly as lake temperature increases worldwide (O'Reilly *et al.*, 2015). This is exemplified by a

473 recent quantitative synthesis on under ice ecology in which winter-summer patterns of
474 nutrients, phytoplankton, and zooplankton were investigated (Hampton *et al.*, 2017)
475 Though full-year time-series of these variables are rare, studies that also include
476 microbial community structure and function variables are even more so although
477 impressive long-term microbial time-series of the ice-free period do exist (Bendall *et al.*,
478 2016; Linz *et al.*, 2017). In addition to a few other studies, our findings substantiate the
479 idea that understanding the under-ice microbiome is important for predicting the
480 dynamics of seasonally ice-covered lakes in the future (Bertilsson *et al.*, 2013). Although
481 this study was illuminating regarding under ice microbial metabolism, we focused solely
482 on winter-associated Verrucomicrobia. Community wide winter-summer patterns remain
483 to be elucidated and will likely vary tremendously between lakes. Hence, a much wider
484 range of studies must be executed before generalizable patterns can be reported.
485 Moreover, meta-omics studies such as this are, in essence, hypothesis-generating and
486 future work that includes targeted enrichment/cultivation and *in situ* rate measurement-
487 based approaches are required to validate and quantify microbial metabolic contributions
488 to nutrient cycling in lake environments during winter.

Methods

Sampling

Microbial samples were collected from three freshwater lakes: Lake Croche (45°59'N, 74°01'W), Lake Montjoie (45°24'N, 72°14'W) and Lake Simoncouche (48°14'N, 71°15'W) in conjunction with the Groupe de Recherche Interuniversitaire en Limnologie et en Environnement Aquatique (GRIL) Monitoring Program. During 3 years (2013-2015), epilimnion and metalimnion samples were collected biweekly during ice-free periods and monthly during ice-covered periods of the year (see **Table S1** for a full description of samples collected for microbial analyses). In winter, samples were collected from just below the ice, as well as deeper in the water column for lakes Simoncouche and Croche, while a single integrated sample of water collected from multiple depths was collected from Montjoie. Lake water, collected in acid-washed bottles, was pre-filtered through 53 μm mesh, followed by sequential filtration on to a 3 μm polycarbonate filter to collect particle-attached cells followed by a 0.22 μm Sterivex filter to collect free-living cells. 1.8 ml of sucrose-based lysis buffer was added to samples collected for DNA extraction, 1.8 ml of RNA later was added to samples collected for RNA extraction, and filters were stored at -80° C until processing.

Water column profiles of environmental variables (**Table S2**) were measured including temperature, dissolved oxygen, pH, specific conductivity and oxidation potential, directly in the field using a multiparameter sonde (YSI, OH, USA). Water samples were collected at the same depths as the microbial samples. Analyses of total

phosphorus (TP), total nitrogen (TN), total dissolved phosphorus (TDP), total dissolved nitrogen (TDN), dissolved organic carbon (DOC), nitrate (NO_3^-), nitrite (NO_2^-), ammonium (NH_4^+), ions (Cl^-, PO_4^{3-}, SO_4^{2-}, Ca^{2+}, Mg^{2+}, Na^+), and Chlorophyll *a* were performed in the GRIL laboratory at Université de Montréal (Montreal, Canada).

Nucleic acid extraction

DNA was extracted from 0.22 µm Sterivex filters using a phenol/chloroform-method modified from Zhou (1996). Sterivex filters were thawed on ice and the storage buffer was removed. The storage buffer was concentrated into Amicon 30 kD filter (500 µl at a time) followed by centrifugation for 20 minutes at 10,000 g. 500 µl storage buffer was repeatedly added until concentrated to a final volume of 100 µl. Buffer exchange was conducted twice by washing storage buffer with 500 µl of TENP buffer (600 mg Tris, 740 mg EDTA, 580 mg NaCl, 2 g Polyvinylpyrrolidon and 100 ml milliQ, pH 8). We then broke open the Sterivex filter and removed the filter. The filter was split into halves and placed inside a 2 ml Eppendorf tube. To conduct the cell lysis and digestion, 0.37 grams of 0.7 mm pre-sterilized Zirconium beads, 60 µl of 20 % SDS, 100 µl concentrated buffer exchanged filtrate, 500 µl TENP buffer and 500 µl phenol-chloroform-isoamylalcohol (PCI) 25:24:1 were added to the 2 ml Eppendorf tube containing the shredded filter. Then, the sample was vortexed for 10 minutes. Samples were incubated for 10 min in a 60 °C water bath followed by incubation on ice for 1 min. The samples were centrifuged for 6 min at 10,000 rpm and 4°C. Supernatant was transferred to a clean 1.5 ml Eppendorf tube, 500 µl phenol-chloroform-isoamylalcohol (PCI) 25:24:1 was

added and samples were vortexed briefly. Then, samples were centrifuged for 6 min at 10,000 rpm and 4 °C. The supernatant was transferred to a new 1.5 ml Eppendorf tube. The PCR step was repeated until there was no longer any white precipitate at the interface (usually 2 times). DNA was precipitated by adding 120 µl of 3 M sodium acetate followed by 1 ml of 96% ethanol. DNA was precipitated at -20 °C for at least 1.5 hours, followed by centrifugation for 60 minutes at 13,000 rpm and 4 °C. The supernatant was decanted, and the pellet was washed with 850 µl of 80 % ethanol. We incubated samples for 10 minutes on ice followed by short vortexing and then centrifuging samples for 15 minutes at 13,000 rpm and 4 °C. The supernatant was removed and the DNA was resuspended in 50 µl TE or Tris-HCl (pH 7.5-8).

RNA was extracted from 0.22 µm Sterivex filters with a modified protocol (Shi *et al.*, 2009; Stewart *et al.*, 2010) which employs both the mirVana miRNA isolation kit (Invitrogen) and the RNeasy RNA cleanup kit (Qiagen). Samples were thawed and had the RNAlater (Invitrogen) surrounding the Sterivex filter removed (approximately 1700 µl) and discarded. 1700 µl of mirVana lysis buffer was added to the Sterivex filter and vortexed to lyse bacterial cells attached to the filter. Total RNA was then extracted from the lysate according to the mirVana protocol. Purified sample (100 ul) was treated with 2 ul DNase (New England Biotech) incubated at 65ºC for 1-2 hours to remove genomic DNA, and concentrated using the RNeasy RNA cleanup kit (Qiagen).

550 **16S rRNA gene sequencing and analysis**

551 16S rRNA gene sequence data was generated from 143 samples (**Table S1**). The
552 V3 region of the 16S rRNA gene was amplified using the universal primers (341F: 5'-
553 CCTACGGGRSGCAGCAG-3' and 515R: 5'-TTACCGCGGCKGCTGVCAC-3')
554 (Klindworth *et al.*, 2013). Two-step PCR reactions (modified from Berry et al., 2011)
555 were conducted in 25 µl volume contained 0.5 µM $MgCl_2$, 0.2 mM deoxynucleotide, 0.2
556 µM each primer and 1U of Phire Hot Start II DNA Polymerase (Finnzymes Thermo
557 Fisher Scientific). The template was amplified using non-barcoded PCR primers for 20
558 cycles, followed by 1:50 dilution of the PCR product and 10 additional cycles of
559 amplifications with barcoded PCR primers. The thermal program consisted of an initial
560 95 °C denaturation step for 4 min, a cycling program of 95 °C for 30 s, 52 °C for 30 s,
561 and 72 °C for 60 s, and a final elongation step at 72 °C for 7 min. Reverse primers were
562 barcoded with specific IonXpress sequence to identify samples. PCR products were
563 purified using QIAquick Gel Extraction Kit (Qiagen), quantified using Quantifluor
564 dsDNA System (Promega), pooled at equimolar concentration and sequenced using an
565 Ion Torrent PGM system on a 316 chip with the ION Sequencing 200 kit as described in
566 Sanschagrin and Yergeau (2014).

567 V3 region 16S rRNA sequences were analyzed using MOTHUR (Schloss *et al.*,
568 2009). Sequences with an average quality of <17, length <100 bp or that did not match
569 the IonXpress barcode and both the PCR forward and reverse primer sequences were
570 discarded. Potential chimeric sequences were identified using UCHIME (Edgar *et al.*,

2011) and discarded. Verrucomicrobia sequences were identified by taxonomic analysis in MOTHUR.

Metagenome sequencing, assembly, annotation and binning

DNA sequencing of 24 samples (**Table S1**) was performed at the Department of Energy Joint Genome Institute (JGI) (Walnut Creek, CA, USA) on the HiSeq 2500-1TB (Illumina) platform. Paired-end sequences of 2 × 150 bp were generated for all libraries. A combined assembly of all raw reads was generated using MEGAHIT (https://github.com/JGI-Bioinformatics/megahit) with kmer sizes of 23,43,63,83,103,123. Gene prediction and annotation was performed using the DoE JGI IMG/M functional annotation pipeline (Markowitz *et al.*, 2014). Metagenomic binning based on tetranucleotide frequency and differential coverage was performed with MetaWatt version 3.5.2 (Strous *et al.*, 2012) using the following settings: base pair cutoff of >2000 on the co-assembly, and a relative weight of binning coverage of 0.75. The identity of bins was assessed using the phylogenetic analysis of a concatenation of single copy core genes implemented in MetaWatt, MAFFT aligner and FastTreeMP for tree inference (Price *et al.*, 2010). Genome completeness and contamination was assessed with CheckM (Parks *et al.*, 2015), which relies on pplacer (Matsen *et al.*, 2010), prodigal (Hyatt *et al.*, 2012) and HMM (Eddy, 2011). Fourteen MAGS with low contamination (<5%) and substantially or near-complete (>70%) were obtained, as described by the bin quality terminology proposed in Parks *et al.* 2015. Despite a completeness value of 58%, we also

591 included Opi-242, given its abundance during the ice-covered period of Lake
592 Simoncouche.

593 **Ecological association of Verrucomicrobia bins**

594 We performed a canonical constrained analysis (CCA) using the *vegan* package
595 (Oksanen *et al.*, 2017) in R (R Core Development Team, 2016). The 20 environmental
596 variables measured were tested for normality using a Shapiro test, and transformed using
597 *powerTransform*, followed by a Box-Cox transformation *(bcPower)*, using the *car*
598 package (Fox *et al.*, 2016) for all variables for which $p<0.05$ in the Shapiro test. A linear
599 model was fitted to each pair of variable, and we removed variables for which the
600 correlation coefficient was greater than 0.70 from the CCA. In brief, we performed the
601 CCA using 14 of the environmental variables, and the complete species matrix containing
602 24 samples and 54 MAG.

603 **Phylogeny of Verrucomicrobia**

604 A concatenated gene phylogeny tree was created based on 4 of the 5 genes as in
605 He *et al.* (2017), in which the 15 Quebec MAGs were put in the contact of the 19 MAGs
606 from Lake Mendota and Trout Bog (Wisconsin), and 7 MAGs from ice-covered Lake
607 Baikal (Cabello-Yeves, Zemskaya, *et al.*, 2017) and 17 MAGs from the Tous and
608 Amadorio reservoirs (Cabello-Yeves, Ghai, *et al.*, 2017). Because none of the 15 MAGs
609 in our study contained the DNA polymerase III beta subunit (but had other subunits such
610 as alpha), the concatenated gene phylogeny contained instead the following four genes:

611 TIGR01391 (DNA primase), TIGR01011 (Small subunit ribosomal protein S2),
612 TIGR00460 (Methionyl-tRNA-formyltransferase), and TIGR00362 (Chromosomal
613 replication initiation factor). 17 out of the 18 MAGs from the Tous and Amadorio
614 Reservoirs had these genes. To create the phylogenetic tree, sequences for each gene
615 were aligned using MUSCLE (Edgar, 2004) and a maximum-likelihood tree was created
616 for each individual gene to ensure that these proteins were conserved and represented the
617 phylogenetic relationships between groups, before concatenation of the sequences using
618 Mesquite (Maddison & Maddison, 2017). The confidence score of each amino acid
619 position in the multiple sequence alignment was calculated using ZORRO (Wu *et al.*,
620 2012), and amino acid position that had a score below 0.5 was manually deleted using
621 Mesquite (Maddison & Maddison, 2017) to obtain a more accurate phylogenetic
622 inference. MEGA6.06 (Tamura *et al.*, 2013) was used to generate a maximum likelihood
623 phylogeny of 65 taxa, using a bootstrap of 100 iterations. The substitution model was the
624 Jones-Taylor-Thornton (JTT) model for amino acids. The rates among site were Gamma
625 distribution, with 4 gamma categories. The maximum-likelihood (ML) heuristic method
626 was Nearest-Neighbour Interchange (NNI), and the initial tree was NJ. Finally, the
627 branch swap filter was Very Strong. Despite that four instead of five genes were used, the
628 phylogenetic structure of the tree (and relationships between taxa) is the same as when
629 using 5 genes, as in He *et al.* (2017). In order to include all 18 MAGs from (Cabello-
630 Yeves, Ghai, *et al.*, 2017), an automated concatenated gene phylogeny was created using

631 PhyloPhlan (Segata *et al.*, 2013) (**Figure S2**). The phylogenetic structure of the
632 PhyloPhlan tree was consistent with the manually curated tree (**Figure 3)**.

633 **Comparative genomics and functional annotation**

634 The distribution of protein-encoding gene content between genomes was
635 determined using proteinortho (Lechner *et al.*, 2011). Inference of protein function and
636 metabolic reconstruction was based on the IMG annotations provided by the JGI,
637 including KEGG, Pfam, EC numbers, and Metacyc annotations. Metabolic reconstruction
638 was also facilitated by generated pathway genome databases for each MAG using the
639 pathologic software available through Pathway Tools (Karp *et al.*, 2009). In addition, we
640 annotated carbohydrate-active enzymes using dbCan (Yin *et al.*, 2012) and hydrogenase
641 classes using HydDB (Sondergaard *et al.*, 2016).

642 **Metatranscriptomic analysis of Verrucomicrobia gene expression patterns**

643 cDNA library preparation and sequencing of 24 samples (**Table** S**1)** was
644 performed at the Department of Energy Joint Genome Institute (JGI) (Walnut Creek, CA,
645 USA) on the HiSeq 2500-1TB (Illumina) platform. Paired-end sequences of 2 × 150 bp
646 were generated for all libraries. The metatranscriptome dataset comprises of 24 samples
647 (6 Croche, 7 Montjoie and 11 Simoncouche). BBMAP (https://jgi.doe.gov/data-and-
648 tools/bbtools/bb-tools-user-guide/bbmap-guide/) was used to map the raw quality filtered
649 reads to the CDS in each 15 Verrucomicrobia MAG, using the option *min_id*=0.97
650 instead of the default one of 0.7. The RPKB, recruitments per kilo bases of billion reads

per sample, was calculated instead of the RPKM (recruitments per kilo bases of million mapped reads) to control for differences in raw reads between samples.

ANI and AAI calculations and fragment recruitment

ANI values were calculated with ANICalculator (https://ani.jgi-psf.org/html/anicalculator.php) using the default settings, but the alignment fraction was too low (median 0.02-0.03 %) to produce meaningful results. Therefore, we assessed genome similarity using AAI values calculated using CompareM (https://github.com/dparks1134/CompareM) and the default settings. Fragment recruitment was performed using BBMAP (https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbmap-guide/) and the following settings: *minid*=94, *maxlen*=500, *idtag*=t. To create the heatmap of relative coverage values, we divided the number of reads mapped by the size of the metagenome for each MAG.

Data accessibility

The 24 metagenomes and the associated co-assembly can be downloaded at the JGI (IMG Genome ID: 3300010885). The 24 metatranscriptome and the associated co-assembly can be downloaded at IMG Genome ID: 3300013295. MAG data is available at NCBI under the accession numbers XXX-XXX

Acknowledgments

We would like to thank Morgan Botrel and Stéphanie Massé for coordinating the Lake Sentinels Project, a program funded by the Groupe de recherche interuniversitaire en limnologie et en environnement aquatique (GRIL), and for sampling the Lakes Croche and Montjoie. We thank Guillaume Grosbois, Tobias Schneider and Maxime Wauthy for sampling the Lake Simoncouche. Funding from the Canadian Natural Sciences and Engineering Research Council (NSERC) Discovery Grants (D.W., M.R., B.B., and Y.H.) and the Canada Research Chair Program (D.W., M.R., Y.H.) are also acknowledged. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231. P.T. was funded by NSERC-CREATE ÉcoLac, and Fonds de Recherche du Québec Nature et Technologie (FRQNT).

Conflict of Interest Statement

The authors have no conflict of interest to declare.

References

Alonso-Sáez, L., Waller, A.S., Mende, D.R., Bakker, K., Farnelid, H., Yager, P.L., et al. (2012) Role for urea in nitrification by polar marine Archaea. *Proc. Natl. Acad. Sci.* **109**: 17989–17994.

Auguet, J.C., Nomokonova, N., Camarero, L., and Casamayor, E.O. (2011) Seasonal changes of freshwater ammonia-oxidizing archaeal assemblages and nitrogen species in oligotrophic alpine lakes. *Appl. Environ. Microbiol.* **77**: 1937–1945.

Bashenkhaeva, M. V., Zakharova, Y.R., Petrova, D.P., Khanaev, I. V., Galachyants, Y.P., and Likhoshway, Y. V. (2015) Sub-Ice Microalgal and Bacterial Communities in Freshwater Lake Baikal, Russia. *Microb. Ecol.* **70**: 751–765.

Beall, B.F.N., Twiss, M.R., Smith, D.E., Oyserman, B.O., Rozmarynowycz, M.J., Binding, C.E., et al. (2016) Ice cover extent drives phytoplankton and bacterial community structure in a large north-temperate lake: implications for a warming climate. *Environ. Microbiol.* **18**: 1704–1719.

Bendall, M.L., Stevens, S.L., Chan, L.-K., Malfatti, S., Schwientek, P., Tremblay, J., et al. (2016) Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. *ISME J.* **10**: 1589–1601.

Bertilsson, S., Burgin, A., Carey, C.C., Fey, S.B., Grossart, H.P., Grubisic, L.M., et al. (2013) The under-ice microbiome of seasonally frozen lakes. *Limnol. Oceanogr.* **58**: 1998–2012.

Bižić-Ionescu, M., Amann, R., and Grossart, H.P. (2014) Massive regime shifts and high activity of heterotrophic bacteria in an ice-covered lake. *PLoS One* **9**: 1–17.

Bogard, M.J., Donald, D.B., Finlay, K., and Leavitt, P.R. (2012) Distribution and regulation of urea in lakes of central North America. *Freshw. Biol.* **57**: 1277–1292.

Bohórquez, J., McGenity, T.J., Papaspyrou, S., García-Robledo, E., Corzo, A., and Underwood, G.J.C. (2017) Different Types of Diatom-Derived Extracellular Polymeric Substances Drive Changes in Heterotrophic Bacterial Communities from Intertidal Sediments. *Front. Microbiol.* **8**:.

Cabello-Yeves, P.J., Ghai, R., Mehrshad, M., Picazo, A., Camacho, A., and Rodriguez-valera, F. (2017) Reconstruction of diverse verrucomicrobial genomes from metagenome datasets of freshwater reservoirs. *Front. Microbiol.* **8**:.

Cabello-Yeves, P.J., Zemskaya, T.I., Rosselli, R., Coutinho, F.H., Zakharenko, A.S., Blinov, V. V., and Rodriguez-Valera, F. (2017) Genomes of novel microbial lineages assembled from the sub-ice waters of Lake Baikal. *Appl. Environ. Microbiol.* AEM.02132-17.

Carini, P., Steindler, L., Beszteri, S., and Giovannoni, S.J. (2012) Nutrient requirements for growth of the extreme oligotroph “Candidatus Pelagibacter ubique” HTCC1062 on a defined medium. *ISME J.* **7**: 592–602.

Carlson, C.A. and Ducklow, H.W. (1996) Growth of bacterioplankton and consumption of dissolved organic carbon in the Sargasso Sea. *Aquat. Microb. Ecol.* **10**: 69–85.

Choo, Y.J., Lee, K., Song, J., and Cho, J.C. (2007) Puniceicoccus vermicola gen. nov.,

727 sp. nov., a novel marine bacterium, and description of Puniceicoccaceae fam. nov.,
728 Puniceicoccales ord. nov., Opitutaceae fam. nov., Opitutales ord. nov. and Opitutae
729 classis nov. in the phylum "Verrucomicrobia." *Int. J. Syst. Evol. Microbiol.* **57**: 532–
730 537.
731 Davis, M.C., Fiehn, O., and Durnford, D.G. (2013) Metabolic acclimation to excess light
732 intensity in Chlamydomonas reinhardtii. *Plant, Cell Environ.* **36**: 1391–1405.
733 Eddy, S.R. (2011) Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**:.
734 Edenborn, H.M. and Litchfield, C.D. (1987) Glycolate turnover in the water column of
735 the New York Bight apex. *Mar. Biol.* **95**: 459–467.
736 Edgar, R.C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high
737 throughput. *Nucleic Acids Res.* **32**: 1792–1797.
738 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011) UCHIME
739 improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200.
740 Fogg, G.E. (1983) The Ecological significane of extracellular products of phytoplankton.
741 *Bot. Mar.* **26**: 3–14.
742 Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-bovy, G., Ellison, S., et al. (2016)
743 Package "car." *CRAN Repos.* 171.
744 Ghosh, W. and Dam, B. (2009) Biochemistry and molecular biology of lithotrophic sulfur
745 oxidation by taxonomically and ecologically diverse bacteria and archaea. *FEMS*
746 *Microbiol. Rev.* **33**: 999–1043.
747 Greening, C., Biswas, A., Carere, C.R., Jackson, C.J., Taylor, M.C., Stott, M.B., et al.
748 (2016) Genomic and metagenomic surveys of hydrogenase distribution indicate H2
749 is a widely utilised energy source for microbial growth and survival. *ISME J.* **10**:
750 761–777.
751 Grosbois, G., Mariash, H.L., Schneider, T., and Rautio, M. (2017) Under-ice availability
752 of phytoplankton lipids is key to zooplankton winter survival. *Sci. Rep.* 1–37.
753 Guillemette, F., McCallister, S.L., and del Giorgio, P.A. (2016) Selective consumption
754 and metabolic allocation of terrestrial and algal carbon determine allochthony in
755 lakes. *ISME J.* 1–10.
756 Hampton, S.E., E Galloway, A.W., Powers, S.M., Ozersky, T., Woo, K.H., Batt, R.D., et
757 al. (2017) Ecology under lake ice. *Ecol. Lett.* **20**: 98–111.
758 He, S., Stevens, S.L., Chan, L., Bertilsson, S., del Rio, T.G., Tringe, S.G., et al. (2017)
759 Ecophysiology of Freshwater Verrucomicrobia Inferred from Metagenome-
760 Assembled Genomes. *mSphere* **2**: 1–17.
761 Hedlund, B.P., Gosink, J.J., and Staley, J.T. (1996) Phylogeny of Prosthecobacter, the
762 fusiform caulobacters: members of a recently discovered division of the bacteria. *Int.*
763 *J. Syst. Bacteriol.* **46**: 960–6.
764 Herlemann, D.P.R., Lundin, D., Labrenz, M., Jurgen, K., Zheng, Z., Aspeborg, H., and
765 Andersson (2013) Metagenomic De Novo Assembly of an Aquatic Representative
766 of the Verrucomicrobial Class Spartobacteria. *MBio* **4**: 1–9.
767 Hoag, K.B., Bradley, A.J., Tousimis, A.J., and Price, D.L. (1969) A Bacterium Capable
768 of using Phytol as Its Sole Carbon Source , Isolated from Algal Sediment of Mud

769 Lake , Florida. *Proc. Natl. Acad. Sci. U. S. A.* **63**: 748–752.
770 Hou, S., Makarova, K.S., Saw, J.H., Senin, P., Ly, B. V, Zhou, Z., et al. (2008) Complete
771 genome sequence of the extremely acidophilic methanotroph isolate V4,
772 Methylacidiphilum infernorum, a representative of the bacterial phylum
773 Verrucomicrobia. *Biol. Direct* **3**: 26.
774 Hyatt, D., Locascio, P.F., Hauser, L.J., and Uberbacher, E.C. (2012) Gene and translation
775 initiation site prediction in metagenomic sequences. *Bioinformatics* **28**: 2223–2230.
776 Jansen, G.A. and Wanders, R.J.A. (2006) Alpha-Oxidation. *Biochim. Biophys. Acta -*
777 *Mol. Cell Res.* **1763**: 1403–1412.
778 Kankaala, P., Huotari, J., Peltomaa, E., Saloranta, T., and Ojala, A. (2006)
779 Methanotrophic activity in relation to methane efflux and total heterotrophic
780 bacterial production in a stratified, humic, boreal lake. *Limnol. Oceanogr.* **51**: 1195–
781 1204.
782 Kant, R., van Passel, M.W.J., Sangwan, P., Palva, A., Lucas, S., Copeland, A., et al.
783 (2011) Genome sequence of "Pedosphaera parvula" Ellin514, an aerobic
784 verrucomicrobial isolate from pasture soil. *J. Bacteriol.* **193**: 2900–2901.
785 Karp, P.D., Paley, S.M., Krummenacker, M., Latendresse, M., Dale, J.M., Lee, T.J., et al.
786 (2009) Pathway Tools version 13.0: Integrated software for pathway/genome
787 informatics and systems biology. *Brief. Bioinform.* **11**: 40–79.
788 Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner,
789 F.O. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for
790 classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.*
791 **41**: 1–11.
792 Lau, W.W.Y., Keil, R.G., and Armbrust, E.V. (2007) Succession and diel transcriptional
793 response of the glycolate-utilizing component of the bacterial community during a
794 spring phytoplankton bloom. *Appl. Environ. Microbiol.* **73**: 2440–2450.
795 Leboulanger, C., Oriol, L., Jupin, H., and Desolas-gros, C. (1997) Diel variability of
796 glycolate in the eastern tropical Atlantic Ocean. *Deep Sea Res. Part I Oceanogr.*
797 *Res. Pap.* **44**: 2131–2139.
798 Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P.F., and Prohaska, S.J. (2011)
799 Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC*
800 *Bioinformatics* **12**: 124.
801 Linz, A.M., Crary, B.C., Shade, A., Owens, S., Gilbert, J.A., Knight, R., and McMahon,
802 K.D. (2017) Bacterial Community Composition and Freshwater Bog Lakes. **2**: 1–14.
803 Markowitz, V.M., Chen, I.M.A., Palaniappan, K., Chu, K., Szeto, E., Pillay, M., et al.
804 (2014) IMG 4 version of the integrated microbial genomes comparative analysis
805 system. *Nucleic Acids Res.* **42**: 560–567.
806 Martinez-Garcia, M., Brazel, D.M., Swan, B.K., Arnosti, C., Chain, P.S.G., Reitenga,
807 K.G., et al. (2012) Capturing single cell genomes of active polysaccharide
808 degraders: An unexpected contribution of verrucomicrobia. *PLoS One* **7**: 1–11.
809 Matsen, F.A., Kodner, R.B., and Armbrust, E.V. (2010) pplacer: linear time maximum-
810 likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference

811 tree. *BMC Bioinformatics* **11**: 538.
812 Maxwell, D.P., Falk, S., Trick, C.G., and Huner, N.P.A. (1994) Growth at Low
813 Temperature Mimics High-Light Acclimation in Chlorella vulgaris. *Plant Physiol.*
814 **105**: 535–543.
815 Maddison, W. P. and D.R. Maddison. 2018. Mesquite: a modular system for
816 evolutionary analysis. Version 3.40 http://mesquiteproject.org
817 Moussa, M.T. (1988) Zooplankton Fecal Pellets As a Source of Hydrocarbons in Chalk.
818 *J. Pet. Geol.* **11**: 347–354.
819 Neuenschwander, S.M., Ghai, R., Pernthaler, J., and Salcher, M.M. (2017)
820 Microdiversification in genome-streamlined ubiquitous freshwater Actinobacteria.
821 *ISME J.* 1–14.
822 Newton, R.J., Jones, S.E., Eiler, A., McMahon, K.D., and Bertilsson, S. (2011) A guide
823 to the natural history of freshwater lake bacteria.
824 O'Reilly, C.M., Rowley, R.J., Schneider, P., Lenters, J.D., Mcintyre, P.B., and Kraemer,
825 B.M. (2015) Rapid and highly variable warming of lake surface waters around the
826 globe. *Geophys. Res. Lett.* 1–9.
827 Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., et al.
828 (2017) vegan: Community Ecology Package. *R Packag. version 2.4-4*
829 https://CRAN.R-project.org/package=vegan.
830 Özkundakci, D., Gsell, A.S., Hintze, T., Täuscher, H., and Adrian, R. (2016) Winter
831 severity determines functional trait composition of phytoplankton in seasonally ice-
832 covered lakes. *Glob. Chang. Biol.* **22**: 284–298.
833 Parker, M.S. and Armbrust, E.V. (2005) Synergistic effects of light, temperature, and
834 nitrogen source on transcription of genes for carbon and nitrogen metabolism in the
835 centric diatom Thalassiosira pseudonana (Bacillariophyceae). *J. Phycol.* **41**: 1142–
836 1153.
837 Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015)
838 CheckM: assessing the quality of microbial genomes recovered from isolates, single
839 cells, and metagenomes. *Genome Res.* **25**: 1043–55.
840 Parveen, B., Mary, I., Vellet, A., Ravet, V., and Debroas, D. (2013) Temporal dynamics
841 and phylogenetic diversity of free-living and particle-associated Verrucomicrobia
842 communities in relation to environmental variables in a mesotrophic lake. *FEMS*
843 *Microbiol. Ecol.* **83**: 189–201.
844 Paver, S.F. and Kent, A.D. (2010) Temporal Patterns in Glycolate-Utilizing Bacterial
845 Community Composition Correlate with Phytoplankton Population Dynamics in
846 Humic Lakes. *Microb. Ecol.* **60**: 406–418.
847 Podgorsek, L., Petri, R., and Imhoff, J.F. (2004) Cultured and genetic diversity, and
848 activities of sulfur-oxidizing bacteria in low-temperature hydrothermal fluids of the
849 North Fiji Basin. *Mar. Ecol. Prog. Ser.* **266**: 65–76.
850 Powers, S.M. and Hampton, S.E. (2016) Winter Limnology as a New Frontier. *Limnol.*
851 *Oceanogr. Bull.* **25**: 103–108.
852 Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2 - Approximately maximum-

likelihood trees for large alignments. *PLoS One* **5**:.

Rontani, J., Bonin, P.C., and John, K. (1999) Biodegradation of Free Phytol by Bacterial Communities Isolated from Marine Sediments under Aerobic and Denitrifying Conditions Biodegradation of Free Phytol by Bacterial Communities Isolated from Marine Sediments under Aerobic and Denitrifying Conditions. **65**: 5484–5492.

Sangwan, P., Chen, X., Hugenholtz, P., and Janssen, P.H. (2004) Chthoniobacter flavus gen. nov., sp. nov., the first pure-culture representative of subdivision two, Spartobacteria classis nov., of the phylum Verrucomicrobia. *Appl. Environ. Microbiol.* **70**: 5875–5881.

Sanschagrin, S. and Yergeau, E. (2014) Next-generation Sequencing of 16S Ribosomal RNA Gene Amplicons. *J. Vis. Exp.* 3–8.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**: 7537–7541.

Segata, N., Börnigen, D., Morgan, X.C., and Huttenhower, C. (2013) PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat. Commun.* **4**:.

Shi, Y., Tyson, G.W., and DeLong, E.F. (2009) Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* **459**: 266–269.

Solomon, C.M., Collier, J.L., Berg, G.M., and Glibert, P.M. (2010) Role of urea in microbial metabolism in aquatic systems: A biochemical and molecular review. *Aquat. Microb. Ecol.* **59**: 67–88.

Sondergaard, D., Pedersen, C.N.S., and Greening, C. (2016) HydDB: A web tool for hydrogenase classification and analysis. *bioRxiv* 61994.

Stewart, F.J., Ottesen, E.A., and DeLong, E.F. (2010) Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J.* **4**: 896–907.

Stocker, R. and Seymour, J.R. (2012) Ecology and Physics of Bacterial Chemotaxis in the Ocean. *Microbiol. Mol. Biol. Rev.* **76**: 792–812.

Strous, M., Kraft, B., Bisdorf, R., and Tegetmeyer, H.E. (2012) The binning of metagenomic contigs for microbial physiology of mixed cultures. *Front. Microbiol.* **3**: 1–11.

Sundh, I., Bastviken, D., and Tranvik, L.J. (2005) Abundance , Activity , and Community Structure of Pelagic Methane-Oxidizing Bacteria in Temperate Lakes Abundance , Activity , and Community Structure of Pelagic Methane-Oxidizing Bacteria in Temperate Lakes. *Appl. Environ. Microbiol.* **71**: 6746–6752.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**: 2725–2729.

Teske, a, Brinkhoff, T., Muyzer, G., Moser, D.P., Rethmeier, J., and Jannasch, H.W. (2000) Diversity of Thiosulfate-Oxidizing Bacteria from Marine Sediments and

895 Hydrothermal Vents Diversity of Thiosulfate-Oxidizing Bacteria from Marine
896 Sediments and Hydrothermal Vents †. **66**: 3125–3133.
897 Twiss, M.R., McKay, R.M.L., Bourbonniere, R.A., Bullerjahn, G.S., Carrick, H.J., Smith,
898 R.E.H., et al. (2012) Diatoms abound in ice-covered Lake Erie: An investigation of
899 offshore winter limnology in Lake Erie over the period 2007 to 2010. *J. Great Lakes*
900 *Res.* **38**: 18–30.
901 Üveges, V., Tapolczai, K., Krienitz, L., and Padisák, J. (2012) Photosynthetic
902 characteristics and physiological plasticity of an Aphanizomenon flos-aquae
903 (Cyanobacteria, Nostocaceae) winter bloom in a deep oligo-mesotrophic lake (Lake
904 Stechlin, Germany). *Hydrobiologia* **698**: 263–272.
905 Weyhenmeyer, G.A., Livingstone, D.M., Meili, M., Jensen, O., Benson, B., and
906 Magnuson, J.J. (2011) Large geographical differences in the sensitivity of ice-
907 covered lakes and rivers in the Northern Hemisphere to temperature changes. *Glob.*
908 *Chang. Biol.* **17**: 268–275.
909 Wright, R.T. and Shah, N.M. (1977) The trophic role of glycolic acid in coastal seawater.
910 II. Seasonal changes in concentration and heterotrophic use in Ipswich Bay,
911 Massachusetts, USA. *Mar. Biol.* **43**: 257–263.
912 Wu, M., Chatterji, S., and Eisen, J.A. (2012) Accounting for alignment uncertainty in
913 phylogenomics. *PLoS One* **7**: 1–10.
914 Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012) DbCAN: A web
915 resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.*
916 **40**: 445–451.
917 Zhou, J., Bruns, M.A., and Tiedje, J.M. (1996) DNA recovery from soils of diverse
918 composition. *Appl. Environ. Microbiol.* **62**: 316–322.
919
920

Figure titles

Fig. 1. A. Locations of Lake Croche, Lake Montjoie and Lake Simoncouche in temperate and boreal regions of Quebec, Canada. **B.** Environmental variability of lake samples based on principal component analysis of 21 physiochemical variables. **C**. Relative abundance of Verrucomicrobia in the 16S rRNA amplicon datasets organized by lake and season. I-C and I-F refer to ice-covered and ice-free periods, respectively. Mid-lines represent median, upper and lower boundaries are 25% quartiles, and whiskers represents scores outside the middle 50%.

Fig. 2. A. Concatenated protein phylogeny of sampled lake MAGs and Verrucomicrobia reference genomes. **B. and C.** CCA of the 54 MAGs constrained to 14 environmental variables (arrows). The 24 samples are shown in the ordination space in panel b, whereas the 54 MAGs are shown in panel c. Abbreviations for environmental variables are described in Table S2.

Fig. 3 . A. Concatenated maximum likelihood phylogeny, using 100 boostraps, of selected Verrucomicrobia MAGs and reference genomes. **B.** Heatmap of the relative MAG abundance based on coverage (blue) and relative gene expression based on the number of transcripts recruited (red) to each MAG. Rectangular boxes represent ice-

941 covered samples, square boxes represent epilimnetic (top squares) and the metalimnetic
942 (bottom squares) samples.
943
944 **Fig. 4**. **A.** Venn diagrams of ortholog distributions between MAGs in V1 (left panel) and
945 V4 (right panel). **B.** Enzymatic reaction for limonene (top) and cymene (bottom)
946 degradation identified in V1-690. **C.** Gene clusters involved in sulfur oxidation in V1-690
947 and hydrogen oxidation in Opi-474. **D.** The α-oxidation pathway for phytol degradation
948 in Opi-474 and Opi-242. For panels **b** and **e**, E.C. numbers are shown, and shaded boxes
949 represent RPKB during winter. Numbers inside the boxes represent
950 "ScaffoldNumber_Loci".
951
952 **Fig. 5 A.** Expression of Chth-196 genes involved in flagellar motility. **B.** Expression of
953 Chth-196 genes involved in chemotaxis **C.** Expression of genes in the urease gene cluster
954 in Chth-196 and Pedo-303 **D.** Genes involved in ammonia utilization and degradation in
955 Chth-196 and Pedo-303. Since Chth-196 is present in Montjoie and Simoncouche, but
956 Pedos-303 is only present in Montjoie, only those lakes are shown. Genes for which
957 expression was positive (Reads per kilo bases per billions reads – RPKB) are highlighted
958 with colour.

Tables

Table 1. Genome characteristics of the 15 Verrucomicrobia MAGs, recovered from Quebec lakes metagenomes co-assembly.

Subdiv.	MAG	Sequence length (Mbp)[a]	N50[a]	Coverage[b]	GC Content (%)[a]	Gene Count[a]	Complete-ness (%)[a]	Contami-nation (%)[a]	Estimated bin size (Mbp)[c]
1	V1-33	4.08	49467	11.9	63.31	3225	78.24	0.68	5.18
1	V1-115	4.14	40207	9.9	61.77	3420	82.07	4.55	4.81
1	V1-159	2.63	31999	16.3	62.55	2373	95.92	0	2.74
1	V1-690	2.24	17084	12.3	60.46	2015	70.61	2.86	3.08
1	V1-1361	2.25	11724	23.8	60.29	2055	72.93	0.68	3.06
2	Chth-244	3.24	18974	13	65.1	3122	83.28	2.2	3.80
2	Chth-196	2.79	35078	15.1	58.27	2263	70.16	0	3.98
2	Xiphi-554	1.66	24984	10.4	42.77	1576	95.95	0.71	1.72
2	Xiphi-315	1.42	22457	38.9	54.51	1400	87.73	2.03	1.59
3	Pedos-303	4.47	17418	9	67.77	3693	89.7	2.7	4.85
3	Pedos-1123	2.73	11423	12.6	52.1	2471	95.05	2.59	2.80
3	Pedos-510	2.51	16489	25.8	54.52	2380	82.43	0.34	3.03
4	Opi-242	1.19	34025	63.8	54.32	1116	57.77	0	2.06
4	Opi-128	1.85	18982	107.1	42.47	1684	77.74	0.68	2.36
4	Opi-474	2.81	11480	51.6	65.6	2690	92.17	4.11	2.92

[a] Calculated using CheckM

[b] Calculated using MetaWatt

[c] Calculated as (sequence length/completeness)*(100-contamination)

966 **Supplementary Figures and Tables Titles**

967

968 **Figure S1.** Scatter plots of chlorophyll *a*, total phosphorus (TP) and total nitrogen (TN)
969 coloured by strata over the period of the study in the three Quebec lakes.

970

971 **Figure S2.** Concatenated gene phylogeny using Phylophlan showing the 15 Quebec
972 MAGs with all currently existing Verrucomicrobia MAGs from Lake Baikal, Trout Bog,
973 Lake Mendota, Tous and Amadorio reservoirs.

974

975 **Table S1.** Description of the 16S rRNA gene, metagenome and metatranscriptome
976 datasets employed in this study.

977

978 **Table S2.** Environmental variables measured in the field and in the laboratory.
979 Abbreviations are the same as those used in Figures 1B and 2B.

980

981 **Table S3.** Pair-wise amino acid identities (AAI) comparisons between all available
982 freshwater Verrucomicrobia MAGs, divided into subdivisions.

983

984 **Table S4.** Fragment recruitment results of 15 MAGs against 22 freshwater metagenomes.

985

986 **Table S5.** Expression values (RPKB) for GHs in each of the 15 MAGs.

Figure 1

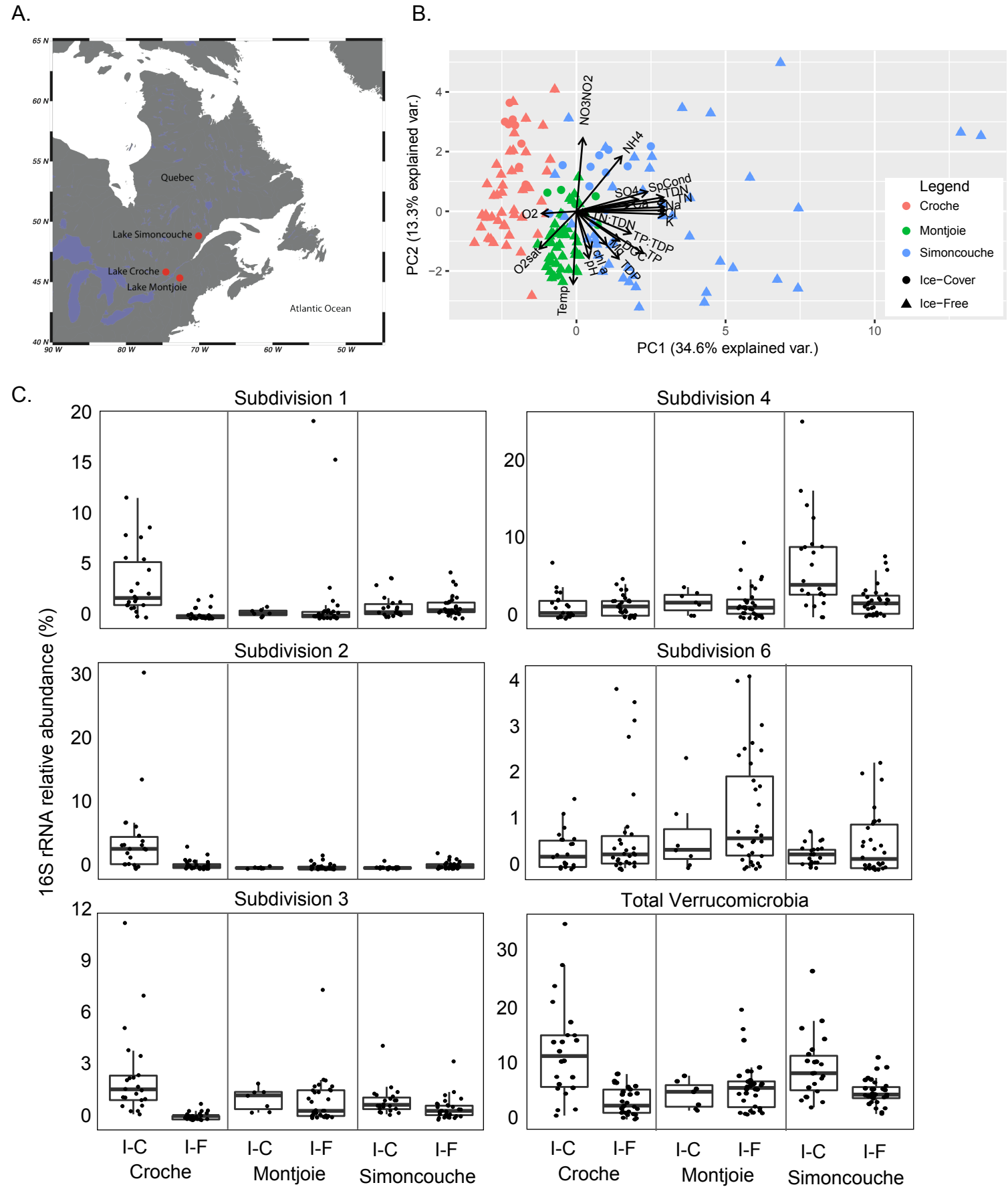
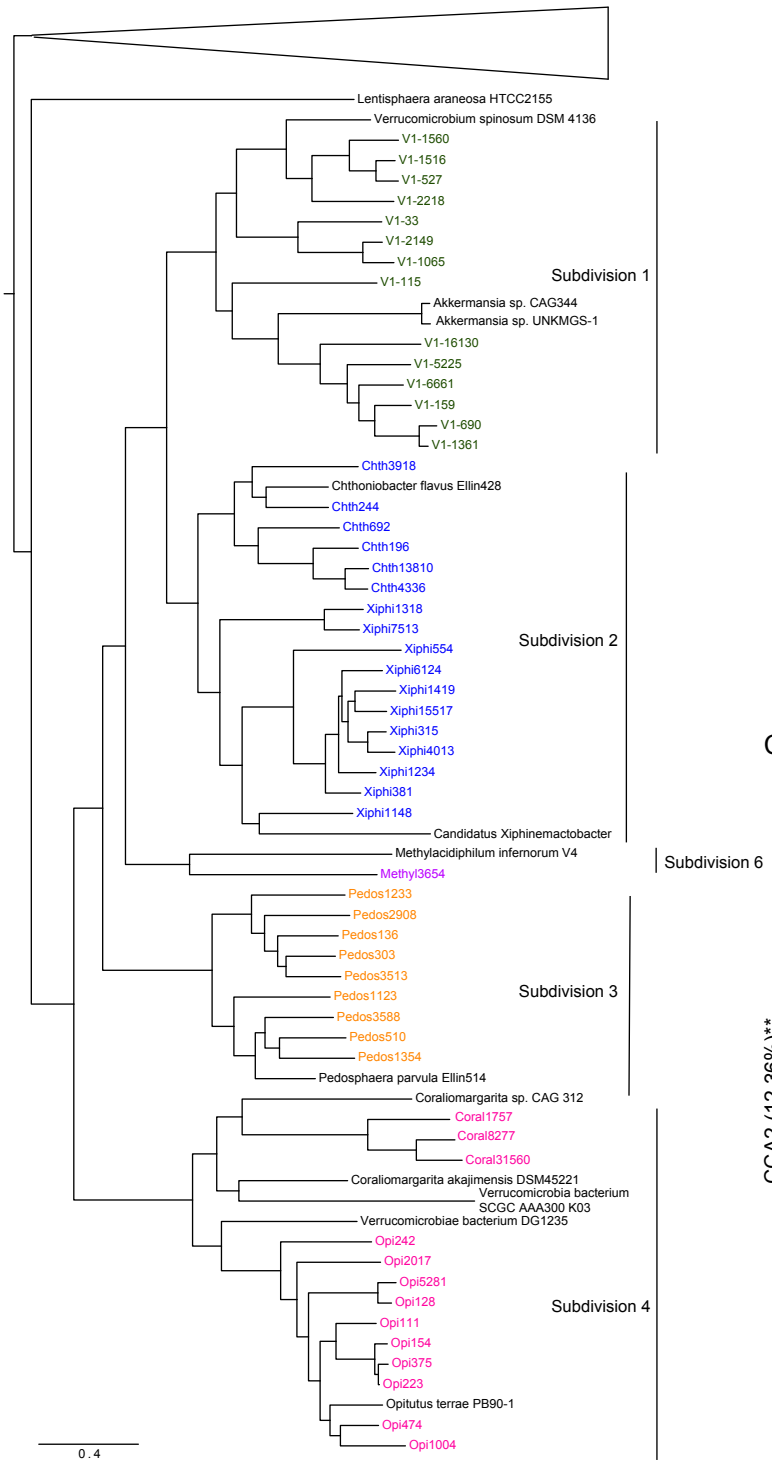
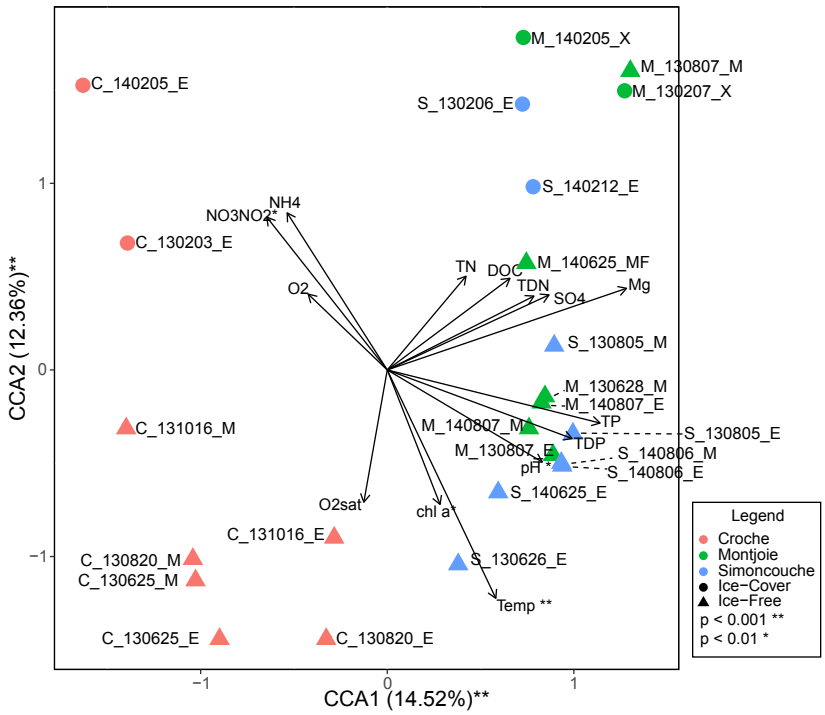


Figure 2

A. Phylogeny of Verrucomicrobia MAGs



B. CCA: Sites and environmental variables



C. CCA: MAGs ecological associations

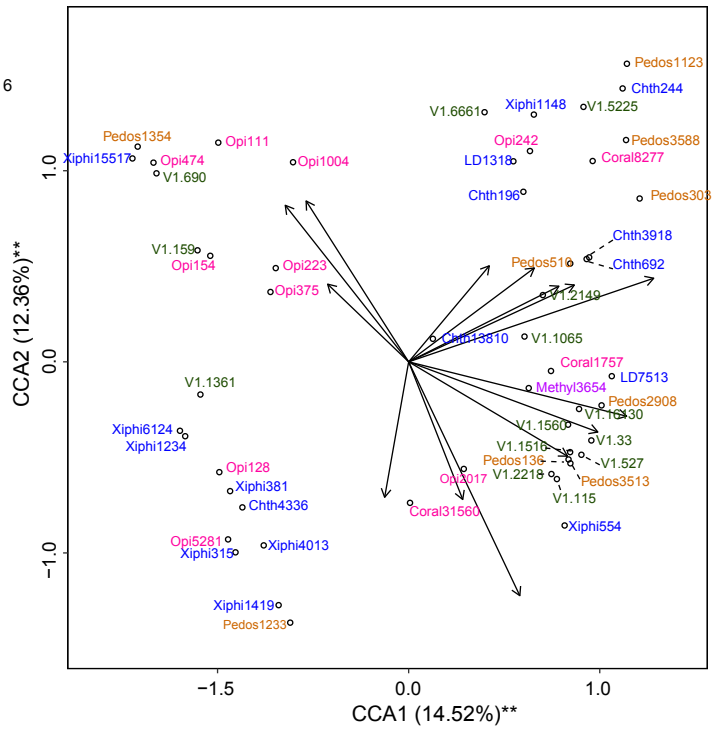


Figure 3

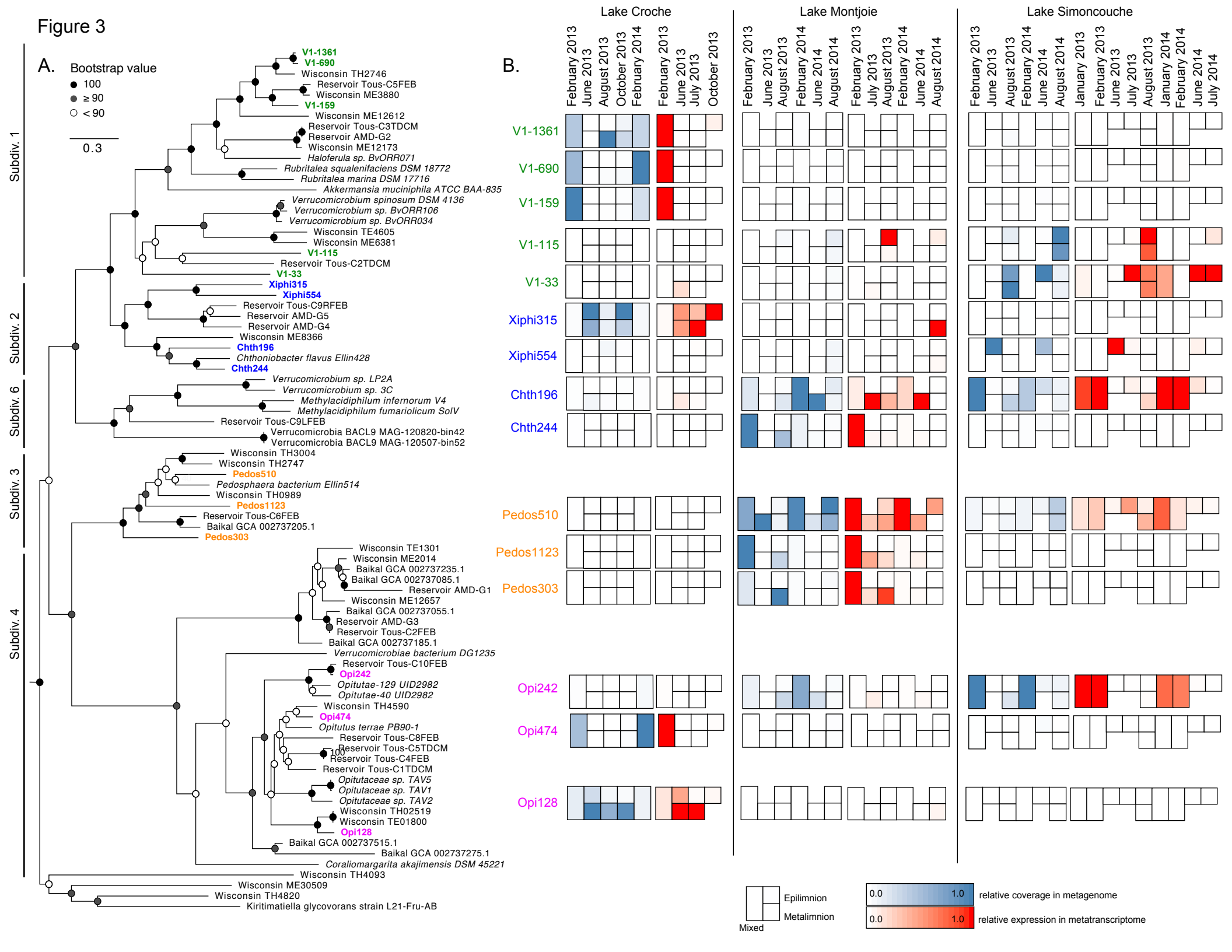

A.
Bootstrap value
100
≥ 90
< 90
0.3
Subdiv. 1
Subdiv. 2
Subdiv. 6
Subdiv. 3
Subdiv. 4
V1-1361
V1-690
Wisconsin TH2746
Reservoir Tous-C5FEB
Wisconsin ME3880
V1-159
Wisconsin ME12612
Reservoir Tous-C3TDCM
Reservoir AMD-G2
Wisconsin ME12173
Haloferula sp. BvORR071
Rubritalea squalenifaciens DSM 18772
Rubritalea marina DSM 17716
Akkermansia muciniphila ATCC BAA-835
Verrucomicrobium spinosum DSM 4136
Verrucomicrobium sp. BvORR106
Verrucomicrobium sp. BvORR034
Wisconsin TE4605
Wisconsin ME6381
V1-115
Reservoir Tous-C2TDCM
V1-33
Xiphi315
Xiphi554
Reservoir Tous-C9RFEB
Reservoir AMD-G5
Reservoir AMD-G4
Wisconsin ME8366
Chth196
Chthoniobacter flavus Ellin428
Chth244
Verrucomicrobium sp. LP2A
Verrucomicrobium sp. 3C
Methylacidiphilum infernorum V4
Methylacidiphilum fumariolicum SolV
Reservoir Tous-C9LFEB
Verrucomicrobia BACL9 MAG-120820-bin42
Verrucomicrobia BACL9 MAG-120507-bin52
Wisconsin TH3004
Wisconsin TH2747
Pedos510
Pedosphaera bacterium Ellin514
Wisconsin TH0989
Pedos1123
Reservoir Tous-C6FEB
Baikal GCA 002737205.1
Pedos303
Wisconsin TE1301
Wisconsin ME2014
Baikal GCA 002737235.1
Baikal GCA 002737085.1
Reservoir AMD-G1
Wisconsin ME12657
Baikal GCA 002737055.1
Reservoir AMD-G3
Reservoir Tous-C2FEB
Baikal GCA 002737185.1
Verrucomicrobiae bacterium DG1235
Reservoir Tous-C10FEB
Opi242
Opitutae-129 UID2982
Opitutae-40 UID2982
Wisconsin TH4590
Opi474
Opitutus terrae PB90-1
Reservoir Tous-C8FEB
Reservoir Tous-C5TDCM
100
Reservoir Tous-C4FEB
Reservoir Tous-C1TDCM
Opitutaceae sp. TAV5
Opitutaceae sp. TAV1
Opitutaceae sp. TAV2
Wisconsin TH02519
Wisconsin TE01800
Opi128
Baikal GCA 002737515.1
Baikal GCA 002737275.1
Coraliomargarita akajimensis DSM 45221
Wisconsin TH4093
Wisconsin ME30509
Wisconsin TH4820
Kiritimatiella glycovorans strain L21-Fru-AB
B.
Lake Croche
Lake Montjoie
Lake Simoncouche
February 2013
June 2013
August 2013
October 2013
February 2014
January 2013
July 2013
June 2014
August 2014
January 2014
July 2014
V1-1361
V1-690
V1-159
V1-115
V1-33
Xiphi315
Xiphi554
Chth196
Chth244
Pedos510
Pedos1123
Pedos303
Opi242
Opi474
Opi128
Epilimnion
Metalimnion
Mixed
0.0
1.0
relative coverage in metagenome
relative expression in metatranscriptome

Figure 4

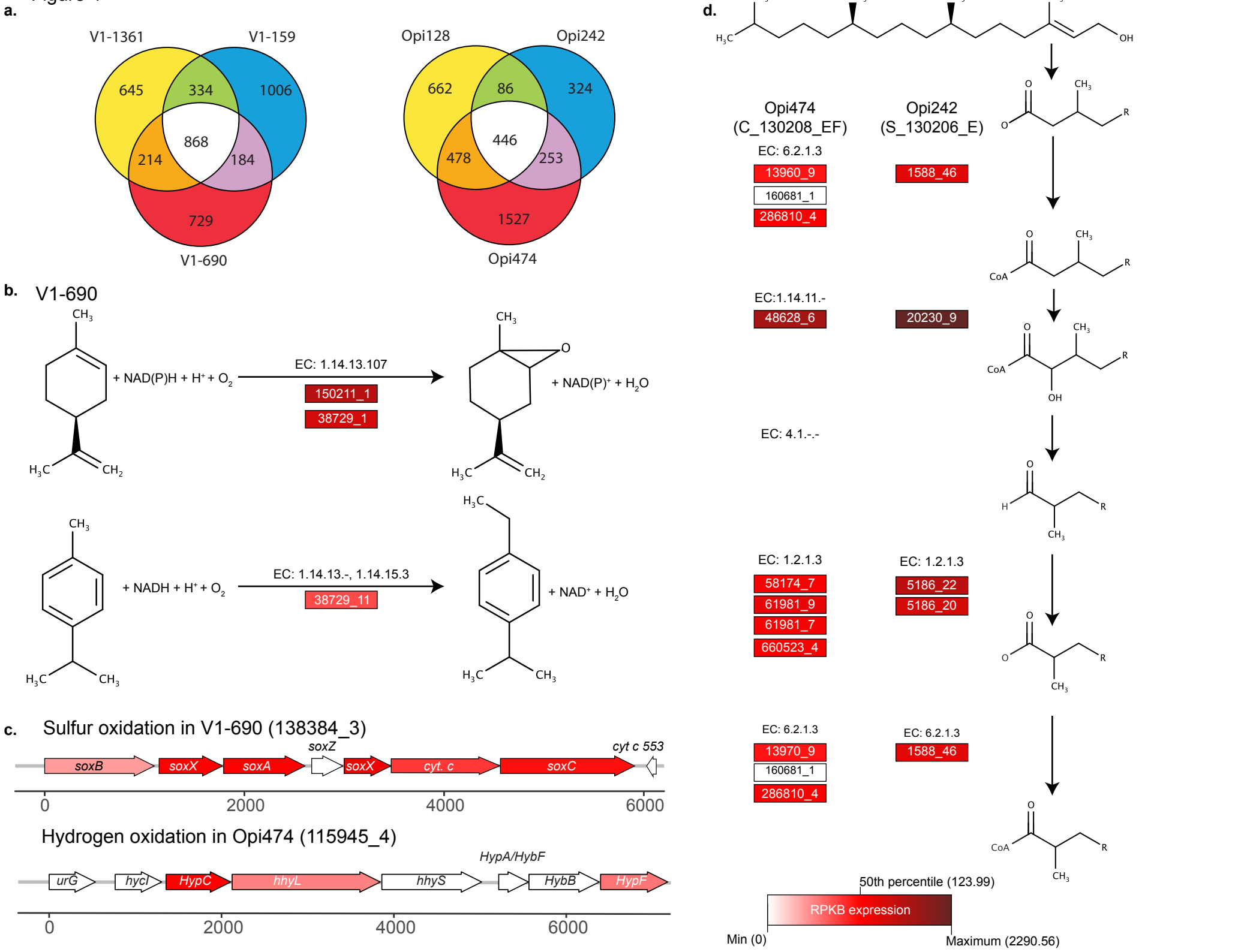
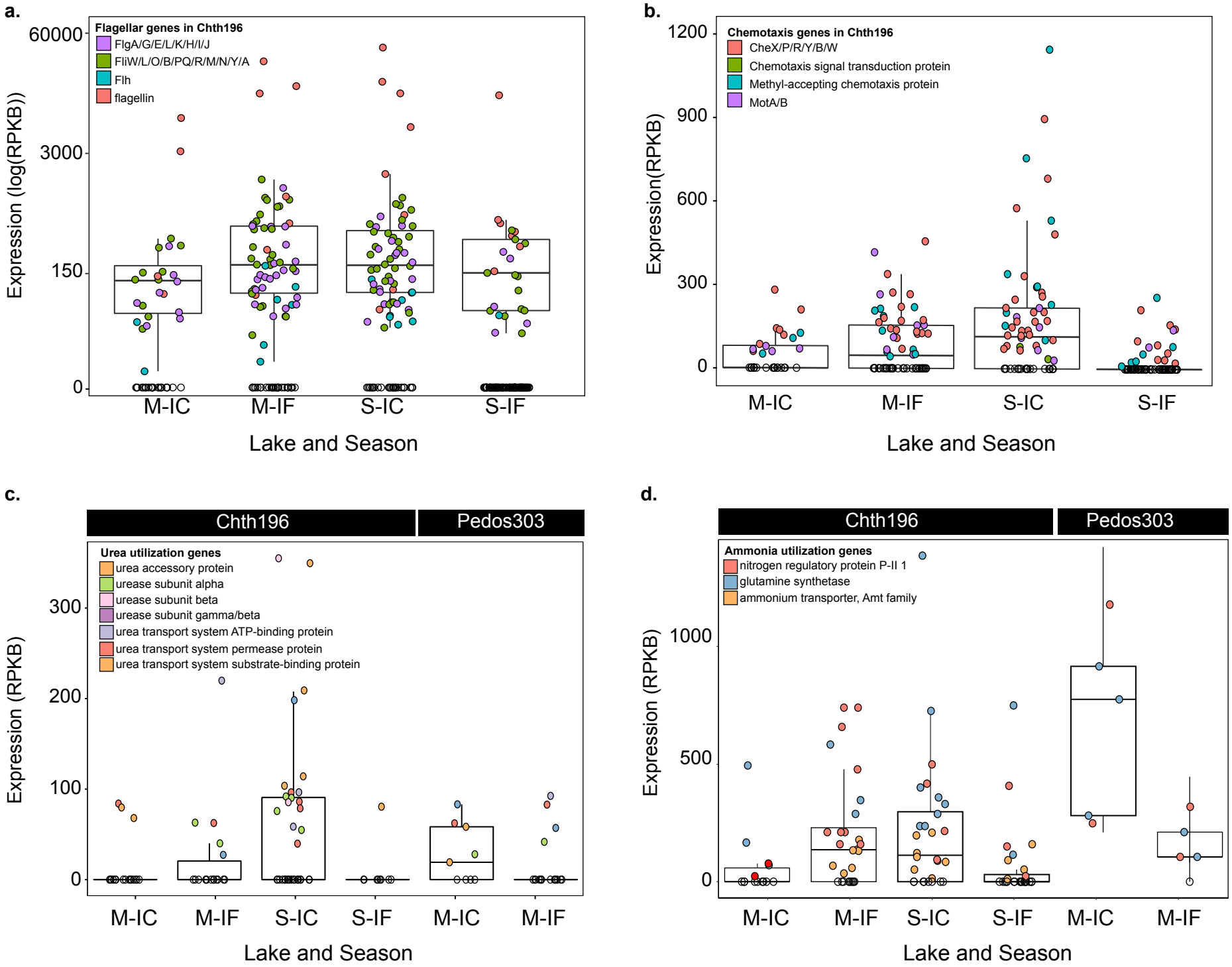

a.
V1-1361
V1-159
645
334
1006
868
214
184
729
V1-690
Opi128
Opi242
662
86
324
446
478
253
1527
Opi474
b.
V1-690
+ NAD(P)H + H+ + O2
EC: 1.14.13.107
150211_1
38729_1
+ NAD(P)+ + H2O
+ NADH + H+ + O2
EC: 1.14.13.-, 1.14.15.3
38729_11
+ NAD+ + H2O
c.
Sulfur oxidation in V1-690 (138384_3)
soxB
soxX
soxA
soxZ
soxX
cyt. c
soxC
cyt c 553
0
2000
4000
6000
Hydrogen oxidation in Opi474 (115945_4)
HypA/HybF
urG
hycI
HypC
hhyL
hhyS
HybB
HypF
0
2000
4000
6000
d.
Opi474
(C_130208_EF)
Opi242
(S_130206_E)
EC: 6.2.1.3
13960_9
160681_1
286810_4
1588_46
EC:1.14.11.-
48628_6
20230_9
EC: 4.1.-.-
EC: 1.2.1.3
58174_7
61981_9
61981_7
660523_4
EC: 1.2.1.3
5186_22
5186_20
EC: 6.2.1.3
13970_9
160681_1
286810_4
EC: 6.2.1.3
1588_46
50th percentile (123.99)
RPKB expression
Min (0)
Maximum (2290.56)

Figure 5


a.
Flagellar genes in Chth196
FlgA/G/E/L/K/H/I/J
FliW/L/O/B/PQ/R/M/N/Y/A
Flh
flagellin
Expression (log(RPKB))
60000
3000
150
0
M-IC
M-IF
S-IC
S-IF
Lake and Season
b.
Chemotaxis genes in Chth196
CheX/P/R/Y/B/W
Chemotaxis signal transduction protein
Methyl-accepting chemotaxis protein
MotA/B
Expression(RPKB)
1200
900
600
300
0
M-IC
M-IF
S-IC
S-IF
Lake and Season
c.
Chth196
Pedos303
Urea utilization genes
urea accessory protein
urease subunit alpha
urease subunit beta
urease subunit gamma/beta
urea transport system ATP-binding protein
urea transport system permease protein
urea transport system substrate-binding protein
Expression (RPKB)
300
200
100
0
M-IC
M-IF
S-IC
S-IF
M-IC
M-IF
Lake and Season
d.
Chth196
Pedos303
Ammonia utilization genes
nitrogen regulatory protein P-II 1
glutamine synthetase
ammonium transporter, Amt family
Expression (RPKB)
1000
500
0
M-IC
M-IF
S-IC
S-IF
M-IC
M-IF
Lake and Season

Supplementary Figure 1