



# Reconstruction of Diverse Verrucomicrobial Genomes from Metagenome Datasets of Freshwater Reservoirs

Pedro J. Cabello-Yeves<sup>1</sup>, Rohit Ghai<sup>2</sup>, Maliheh Mehrshad<sup>2</sup>, Antonio Picazo<sup>3</sup>, Antonio Camacho<sup>3</sup> and Francisco Rodriguez-Valera<sup>1\*</sup>

<sup>1</sup> Evolutionary Genomics Group, Departamento de Producción Vegetal y Microbiología, Universidad Miguel Hernández, San Juan de Alicante, Spain, <sup>2</sup> Department of Aquatic Microbial Ecology, Institute of Hydrobiology, Biology Centre of the Academy of Sciences of the Czech Republic, České Budějovice, Czechia, <sup>3</sup> Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, Valencia, Spain

## OPEN ACCESS

### Edited by:

Hongyue Dang,  
Xiamen University, China

### Reviewed by:

Juan Imperial,  
Universidad Politécnica de Madrid  
(UPM), Spain

Pelin Yilmaz,  
Max Planck Institute for Marine  
Microbiology (MPG), Germany  
Huib J. M. Op den Camp,  
Radboud University Nijmegen,  
Netherlands

### \*Correspondence:

Francisco Rodriguez-Valera  
frvalera@umh.es

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 17 August 2017

**Accepted:** 18 October 2017

**Published:** 02 November 2017

### Citation:

Cabello-Yeves PJ, 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.* 8:2131. doi: 10.3389/fmicb.2017.02131

The phylum *Verrucomicrobia* contains freshwater representatives which remain poorly studied at the genomic, taxonomic, and ecological levels. In this work we present eighteen new reconstructed verrucomicrobial genomes from two freshwater reservoirs located close to each other (Tous and Amadorio, Spain). These metagenome-assembled genomes (MAGs) display a remarkable taxonomic diversity inside the phylum and comprise wide ranges of estimated genome sizes (from 1.8 to 6 Mb). Among all *Verrucomicrobia* studied we found some of the smallest genomes of the Spartobacteria and Opitutae classes described so far. Some of the Opitutae family MAGs were small, cosmopolitan, with a general heterotrophic metabolism with preference for carbohydrates, and capable of xylan, chitin, or cellulose degradation. Besides, we assembled large copiotroph genomes, which contain a higher number of transporters, polysaccharide degrading pathways and in general more strategies for the uptake of nutrients and carbohydrate-based metabolic pathways in comparison with the representatives with the smaller genomes. The diverse genomes revealed interesting features like green-light absorbing rhodopsins and a complete set of genes involved in nitrogen fixation. The large diversity in genome sizes and physiological properties emphasize the diversity of this clade in freshwaters enlarging even further the already broad eco-physiological range of these microbes.

**Keywords:** freshwater *Verrucomicrobia*, metagenomics, rhodopsin, nitrogen fixation, genome streamlining

## INTRODUCTION

Freshwater *Verrucomicrobia* were originally classified by a largely morphology-based taxonomic approach into a broad taxon called the “Caulobacterales,” a diverse group of microbes bearing “prosthecae” (Henrici and Johnson, 1935). These prosthecate microbes were later understood to belong to multiple bacterial phyla, e.g., *Proteobacteria* and *Verrucomicrobia*. Several phylogenetic analyses, first based on 16S rRNA sequences from cultured isolates and environmental clone libraries, and expanded later using phylogenomic methods grouped them together with two other bacterial phyla, the *Planctomycetes* and *Chlamydiae*, together comprising what is now referred

to as the PVC superphylum (Wagner and Horn, 2006; Gupta et al., 2012). More recently, the phyla *Lentisphaerae* and OP3 have been added to this superphylum. While the presence of “wart-like structures,” i.e., *verrucae*, was identified as a distinguishing feature of the model organism, *Verrucomicrobium spinosum*, several members of *Verrucomicrobia* share a much more intriguing compartmentalized cellular organization with the *Planctomycetes* and also some *Chlamydiae*, suggesting a common ancestry for the group at large (Lee et al., 2009). The nature and function for these compartmentalized structures in the PVC superphylum remains as yet unclear, but hypotheses proposing homologous relationships to eukaryotic intracellular structures (Devos and Reynaud, 2010; Forterre and Gribaldo, 2010; Lonhienne et al., 2010; Santarella-Mellwig et al., 2010) that were advanced have been refuted (McInerney et al., 2011). On the other hand, in anaerobic ammonium oxidizing *Planctomycetes* the function of the organelle anammoxosome is well studied (Van Teeseling et al., 2013).

*Verrucomicrobia* have been detected in diverse habitats. Several species have been isolated from soil (Schlesner, 1987; Sangwan et al., 2004; Kant et al., 2011a), and culture-independent studies have estimated that *Verrucomicrobia* may account for up to 20% of total bacterial communities in soil (Bergmann et al., 2011). *Verrucomicrobial* isolates display a wide variety of features, e.g., the methanotrophs *Methylacidiphilum* and *Methylacidimicrobium* from volcanic environments (Dunfield et al., 2007; Pol et al., 2007; Islam et al., 2008; van Teeseling et al., 2014) able to perform nitrogen fixation (Hou et al., 2008; Khadem et al., 2010, 2012b; Erikstad and Birkeland, 2015), aerobes or anaerobes from rice paddy soils (Chin et al., 2001; Kant et al., 2011b). A considerable number have been isolated in association to animal guts, including termites (Isanapong et al., 2012; Wertz et al., 2012) and cephalotes (Lin et al., 2016), while mucin degraders have been found in the human microbiome (Derrien et al., 2004). Some isolates have been identified associated to fruits (grape parasites) and plants (Okazaki et al., 2014; Brown et al., 2015). Representatives from aquatic environments have been described in recent years, for instance: marine isolates associated to sponges, corals, and algae (Scheuermayer et al., 2006; Yoon et al., 2007, 2008).

Several ecological studies have detected *Verrucomicrobia* in suboxic ponds (Bri e et al., 2007), tropical lakes (De Wever et al., 2008), wetlands (Zhang et al., 2014), and associated to cyanobacterial blooms (Eiler and Bertilsson, 2004). Fluorescent *in situ* hybridization (FISH) experiments targeting specific *Verrucomicrobia* clades in humic lakes have shown that freshwater *Verrucomicrobia* are diverse and may have large populations up to ca. 19% of the total community (Arnds et al., 2010). Highly variable abundances (6–17%) have also been reported in shotgun metagenomic studies in freshwater lakes/reservoirs (Lonhienne et al., 2010; Oh et al., 2011; Eiler et al., 2014; Ghai et al., 2014).

However, very few genomes of *Verrucomicrobia* have been described from the aquatic habitat, i.e., water column, but some single cell-amplified genomes (SAGs) inferred to be polysaccharide-degraders have been sequenced from the

marine environment (Martinez-Garcia et al., 2012a) and two metagenome-assembled genomes (MAGs) (Tamagnini et al., 2002), one from the brackish environment of the Baltic Sea (Herlemann et al., 2013) and the other from a water treatment plant (*Verrucomicrobia* bacterium Ga0077533) are available. Only a single SAG has been described from the freshwater Lake Mendota (Martinez-Garcia et al., 2012a). Even though there is a growing number of *Verrucomicrobia* isolates, MAGs and SAGs in the databases, the genomic information from pelagic freshwater habitats is scarce and limited largely to 16S rRNA surveys. Therefore, the ecological role of these unusual microbes in freshwaters remains obscure.

Here we have used metagenomic tools to reconstruct novel freshwater *Verrucomicrobia* from two reservoirs (Amadorio and Tous) located in Southeast Spain. The large number of *verrucomicrobial* genomes (18) reconstructed from these freshwater reservoirs show broad differences in genomic make up within the same *verrucomicrobial* class *Opitutae*, likely reflecting specialization in different ecological strategies (oligotrophic K-strategists versus copiotrophic r-strategists). The data presented contribute a large number of novel *verrucomicrobial* groups detected for the first time through these MAGs.

## MATERIALS AND METHODS

### Sampling, Physicochemical Profiles, Sequencing, Assembling, and Annotation

A total of five datasets from two different Spanish freshwater reservoirs were used in this work. Four of them have already appeared in publications and came from winter season, when the water column was mixed. Two from Amadorio reservoir (Ghai et al., 2014) and another two from Tous reservoir (Cabello-Yeves et al., 2017). In this work we describe for the first time another metagenomic sample that was taken from Tous reservoir on July 21, 2016 at 13 m depth, when the water column was stratified. Following the same assembly and annotation methods previously described in Amadorio and Tous winter metagenomics studies (Ghai et al., 2014; Cabello-Yeves et al., 2017), sequencing of the 0.22  $\mu\text{m}$  filter was performed using Illumina HiSeq 4000 (Oklahoma Medical Research Foundation, United States). A total of 250 million sequence reads (paired end  $2 \times 150$  bp) were produced for the 13 m summer sample from Tous reservoir. The metagenome was assembled using the IDBA-UD assembler (Peng et al., 2012) with *mink* 70, *maxk* 100, *step* 10, and pre-correction parameters, obtaining a total of 4892 contigs larger than 10 Kb. Gene predictions on the assembled contigs were assessed using Prodigal in metagenomic mode (Hyatt et al., 2010), tRNAs were predicted using tRNAscan-SE (Lowe and Eddy, 1997), and ribosomal rRNA genes were identified using *ssu-align* (Nawrocki and Eddy, 2010) and *meta-rna* (Huang et al., 2009). Comparisons of predicted protein sequences against NCBI NR, COG (Tatusov et al., 2001), and TIGFRAM (Haft et al., 2001) databases were performed for taxonomic binning and functional annotation.

## Unassembled 16S rRNA Verrucomicrobial Classification in Different Freshwater Datasets

In order to estimate the percent of verrucomicrobial 16S rRNA reads in different metagenomics datasets we firstly prepared a non-redundant version of the RDP database by clustering all available 16S rRNA coding sequences (approximately 2.3 million) into approximately 800,000 sequences at 90% identity level using UCLUST (Edgar, 2010). This database was used to identify candidate 16S rRNA reads in the Illumina datasets (unassembled). If a read matched this database at an  $e$ -value  $< 1e-5$  it was considered a potential 16S rRNA gene fragment. These candidate reads were aligned to archaeal, bacterial, and eukaryal 16S/18S rRNA HMM models using ssu-align to identify true 16S/18S sequences (Nawrocki and Eddy, 2010). The final sequences were compared to the entire RDP database and classified into a high level taxon if the sequence identity was  $\geq 80\%$  and the alignment length was  $\geq 90$  bp. Sequences failing these thresholds were discarded.

## Identification of Verrucomicrobial Contigs and *de Novo* Genomic Reconstruction

*Verrucomicrobia* MAGs from Amadorio were reconstructed using the small and large fractions of Amadorio reservoir datasets (Ghai et al., 2014), while genomes from Tous winter (February) metagenomes were reconstructed using the assembly from 12 and 25 m samples (Cabello-Yeves et al., 2017). *Verrucomicrobia* genomes from Tous summer sample were reconstructed with a single sample from 13 m, as described above. Contigs longer than 10 kb were considered verrucomicrobial when hits (BLASTP) to this phylum were obtained for  $>60\%$  of the total genes presented in each contig. More specifically, contigs falling within these parameters were binned using taxonomy, principal component analysis of tetranucleotide frequencies, GC content, and coverage values in Tous and Amadorio metagenomes (Ghai et al., 2014; Cabello-Yeves et al., 2017). Tetranucleotide frequencies were computed using compseq program in the EMBOSS package (Rice et al., 2000). Principal component analysis was performed using the FactoMineR package in R (Lê et al., 2008). To estimate the genome size and completeness of the assembled genomes we used the CheckM package with the parameters taxonomy\_wf and specific *Verrucomicrobia* marker set of genes (Parks et al., 2015). Average Nucleotide Identity (ANI) among our MAGs was also calculated (Goris et al., 2007).

## Phylogenomic Trees

The phylogenomic placement of the newly described verrucomicrobial MAGs was assessed with PhyloPhlAn tool (Segata et al., 2013). The complete genomes of representatives from all known *Verrucomicrobia* classes and reconstructed MAGs of this study were added to the built-in tree of life in PhyloPhlAn. PhyloPhlAn uses USEARCH (Edgar, 2010) to identify the conserved proteins and subsequent alignments

against the built-in database are performed using MUSCLE (Edgar, 2004). Finally, an approximate maximum-likelihood tree is generated using FastTree (Price et al., 2010) with local support values using Shimodaira-Hasegawa test (Shimodaira and Hasegawa, 1999). We used 58 reference *Verrucomicrobia* genomes (MAGs, SAGs, and isolates) with  $>50\%$  of genome completeness assessed by CheckM (Parks et al., 2015) from the different existing subdivisions and classes inside the phylum.

## Single Gene Trees

A 16S rRNA tree was made with the *Verrucomicrobia* and *Planctomycetes* 16S rRNA references and those sequences retrieved from Amadorio and Tous metagenomes, either from binned or unbinned *Verrucomicrobia* MAGs. Sequences were aligned using MAFFT (Katoh et al., 2002), which uses the four-way consistency objective function for incorporating structural information. Trees were constructed with FastTree2 (Price et al., 2010), using a GTR model, a gamma approximation, and 100 bootstraps. We did not observe major differences between MAFFT and MUSCLE alignments for 16S phylogenies. The same method was applied for the NifHDK protein concatenate tree, which was constructed with NifHDK concatenate from a wide range of bacteria (Boyd, 2000; Boyd et al., 2011, 2015; Boyd and Peters, 2013; McGlynn et al., 2013). A batch of 200 rhodopsins from multiple origins were aligned using MUSCLE (Edgar, 2004) as previously described in other freshwater rhodopsin phylogenies (Ghai et al., 2014). A maximum-likelihood tree was constructed with FastTree2 (Price et al., 2010), JTT+CAT model, a gamma approximation, and 100 bootstraps.

## Metagenomic Read Recruitment

Metagenome reads mapping to our MAGs were performed against more than 150 publicly available freshwater and brackish metagenomes using BLASTN with at least 50 bp of length and  $>95\%$  identity and an  $e$ -value of  $\leq 1e-5$ . These hits were used to compute the abundance expressed as RPKG (reads recruited per kilobyte of genome per gigabyte of metagenome). RPKG is a normalized number comparable across metagenomes from different locations. All metagenomic datasets used in this work are publicly available: Amadorio (Ghai et al., 2014) and Tous reservoirs (Cabello-Yeves et al., 2017), Lake Lanier (Oh et al., 2011), Dexter reservoir (PRJNA312985), Lake Baikal (PRJNA396997, SRR5896114, and SRR5896115), Klamath Iron Gate Dam (PRJNA312830), Lake Houston (PRJNA312986), Amazon Lakes (Toyama et al., 2016), Lake Mendota (PRJNA330170, PRJNA330171, and PRJNA330042), Swedish Lakes, and Trout Bog (Eiler et al., 2014).

## Metabolic Analysis

The metabolic features of each genome were analyzed with RAST annotation pipeline database (Overbeek et al., 2013), BlastKoala with phylum-specific genes (Kanehisa et al., 2016), Batch Web CD-Search Tool with default parameters (Marchler-Bauer et al., 2016), and BLASTP, BLASTN, and TBLASTX (Altschul et al., 1997). We annotated each MAG protein set with the dbCAN

(Yin et al., 2012) and CAZY enzyme databases (Lombard et al., 2013).

## Microscopy and FISH

For microscopic counts of heterotrophic bacterioplankton, water samples were fixed *in situ* with a paraformaldehyde: glutaraldehyde solution to a final concentration in the sample of 1%: 0.05% (w/v) (Marie et al., 1997). Once in the laboratory, subsamples (5–10 mL) were filtered through 0.2- $\mu$ m-pore-size black filters (Nuclepore, Whatman). For quantification, a quarter of a filter was stained with 4,6-diamidino-2-phenylindole (DAPI) (Porter and Feig, 1980) (SIGMA) and counted with an inverted Zeiss III RS epifluorescence microscope (1250 $\times$ , resolution 0.02857 lm/pixel) using a G365 exciting filter, LP420 suppression filter for blue light, and G546 exciting filter, LP590 suppression filter for green light (MacIsaac and Stockner, 1993). Water samples fixed with paraformaldehyde to 2% final concentration then filtered within the next 2 h through white polycarbonate filters (0.2  $\mu$ m pore size). Sections were stained with both the different oligonucleotide probes and DAPI and subsequently mounted for microscopic evaluation. The previously described probe EUB338-III (GCTGCCACCCGTAGGTGT) was used to detect the phylum *Verrucomicrobia* (Daims et al., 1999). FISH-specific probes for the *Verrucomicrobia* taxa from Amadorio and Tous Reservoirs were also designed using the PRIMER3 tool (Rozen and Skaletsky, 1999) from specific sequences obtained from the metagenomes. Specifically, the 16s rRNA probe VrmC9LTo (ATAGGCCGCGGGCTCCTCAA) was designed to detect the specific *Verrucomicrobia*-Tous-C9LFEB from Tous Reservoir and the rRNA probe VrmG3Ama (AGGCCGCAAGCTCCTCCTGA) to detect *Opiritutae*-AMD-G3. All probes were labeled with the indocarbocyanine dye Cy3 in 5' (Thermo Fisher Scientific, Waltham, MA, United States). Probes matching sequences were checked for specificity using the Probe Match RDP (Cole et al., 2009), RDP Release 11, Update 5: September 30, 2016 (3,356,809 16S rRNAs database<sup>1</sup>). The FISH protocol was performed as previously described (Sekar et al., 2003). Hybridization conditions for VrmC9LTo and VrmG3Ama were adjusted by formamide (VWR BDH Prolabo) series applied to different subsamples, with a final optimal formamide concentration of 35% in both cases. Optimal concentration of formamide used for EUB338-III was 40%. Relative densities of hybridized bacteria were calculated as the product of their relative abundances on filter sections (percentage of DAPI-stained objects) and the DAPI-stained direct cell counts. The absolute abundance values were revised based on the values obtained in independent DAPI counts without the hybridization process. A minimum of 500 probe-stained cells were counted per sample. Total percent of heterotrophs was also determined independently. DAPI-stained objects were counted in order to confirm that the total percent did not vary with the hybridized cells. Images from FISH determinations were analyzed using the NIH ImageJ Software to determine cell dimensions<sup>2</sup>.

<sup>1</sup><https://rdp.cme.msu.edu/index.jsp>

<sup>2</sup><http://rsb.info.nih.gov/ij/index.html>

## RESULTS AND DISCUSSION

### Verrucomicrobial 16S rRNA Abundance in Freshwater Habitats

We have estimated the relative abundance of verrucomicrobial 16S rRNA in several available freshwater metagenomic datasets, ranging from oligotrophic to eutrophic and tropical to cold, distributed in different latitudes and altitudes. We have found a relative abundance between 2 and 19% (Supplementary Figure 1), much higher than some values estimated for marine water column and sediments (2 and 1.4%, respectively) (Freitas et al., 2012). In Lake Baikal, Lake Lanier, and Dexter reservoir the abundance of *Verrucomicrobia* was particularly high, reaching 20% of the total 16S rRNA reads of these datasets. No correlation of verrucomicrobial abundances was apparent either with latitude or with trophic status, although it is evident that they are found in all kinds of freshwater lakes or reservoirs.

### Reconstruction of Novel Verrucomicrobial Genomes

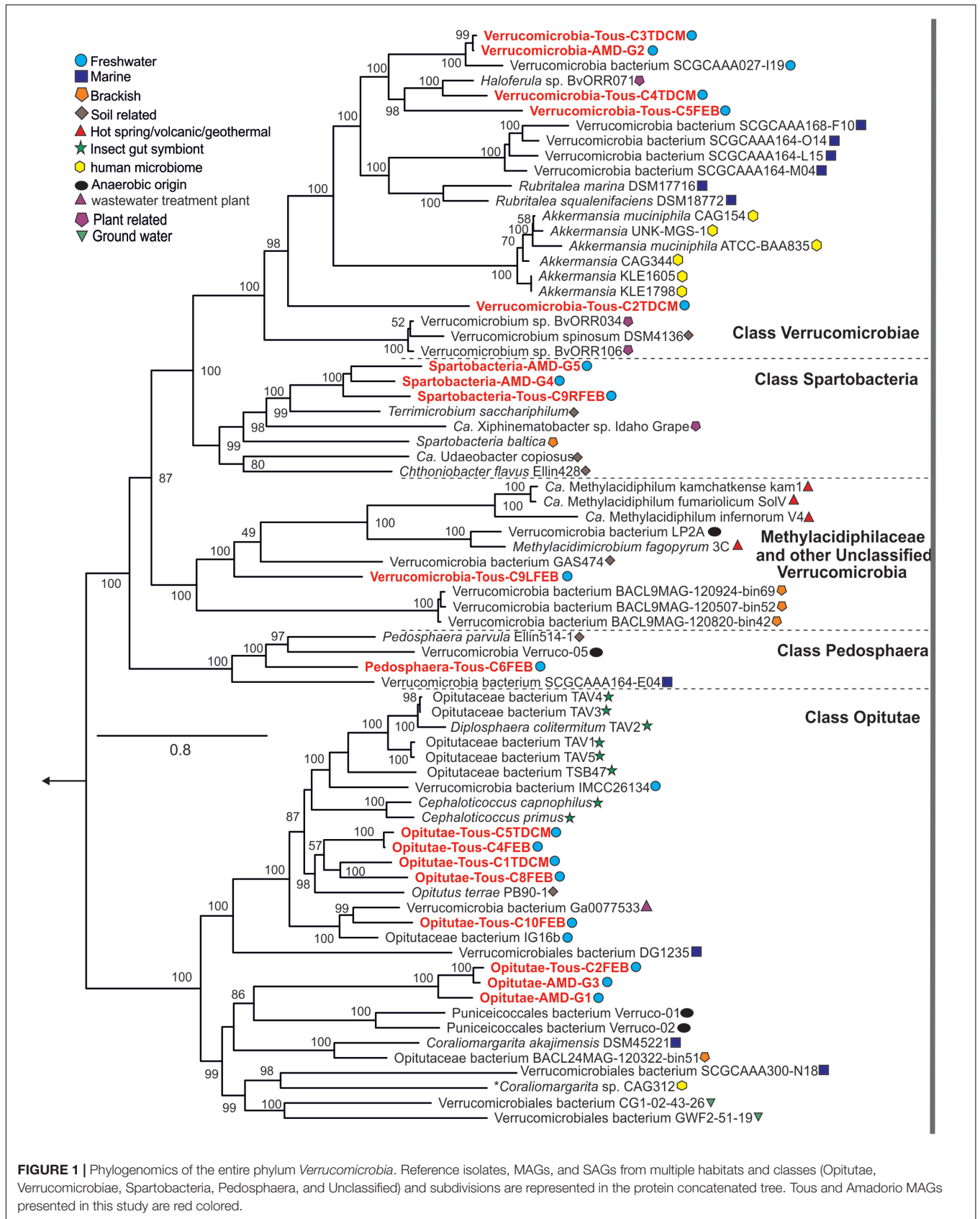
In order to simplify the nomenclature of the work presented here, all reconstructed *Verrucomicrobia* will be referred to as MAGs which come from an assorted clonal population from a single species. Based on the metagenomics sequencing, assembly, annotation, and binning procedures described before (see section “Materials and Methods”) we were able to reconstruct a total of 18 *Verrucomicrobia* genomes (Table 1) of which 15 were different genomes as suggested by ANI calculations. Only three pairs of genomes displayed an ANI higher than 98% suggesting they came from the same microbe (Supplementary Figure 2). A 100% ANI between *Opiritutae*-Tous-C4FEB and *Opiritutae*-Tous-C5TDCM was found, which suggested that the same microbe was obtained from two different samples (summer and winter) from the same reservoir. Along the same lines, another genome (*Verrucomicrobiae*-Tous-C3TDCM and *Verrucomicrobiae*-AMD-G2 with 100% ANI between them) was assembled twice from the two different reservoirs, Tous and Amadorio that are about 100 km apart. Finally, *Opiritutae*-AMD-G3 and *Opiritutae*-Tous-C2FEB shared 98% ANI values (same species). Estimations of genome sizes ranged from 1.83 to 6.18 Mb and eight genomes appeared to be more than 80% complete (Table 1). Some of these genomes appear to be the smallest verrucomicrobial representatives described yet (Supplementary Figure 3A), with a wide variation in GC content (51–70%), with the smaller genomes generally having a lower GC content (Supplementary Figure 3B).

Currently, three classes are defined within the phylum *Verrucomicrobia*: *Opiritutae* (subdivision 4), *Spartobacteria* (subdivision 2), and *Verrucomicrobiae* (subdivision 1). There are also two subdivisions (6 and 7) which comprise unclassified representatives and another one comprising *Pedospaera parvula* like microbes (subdivision 3). Recently, one divergent member from the verrucomicrobial subdivision 5 has been described as a novel phylum-level lineage (Spring et al., 2016). A complete phylogenomic tree with the reference verrucomicrobial genomes and the 18 MAGs assembled in this work is shown in Figure 1.

**TABLE 1** | Summary statistics for the 18 *Verrucomicrobia* metagenome-assembled genomes (MAGs) from Tous and Amadorio reservoirs.

<i>Verrucomicrobia</i> MAG	Number of CDS	Sequence length (Mbp)	GC	Median intergenic spacer (bp)	Coding density	Completeness	Contamination	Estimated genome size (Mb)	Origin	Class
<b>Opitutae-AMD-G1</b>	1479	1.54	64.07	38	90	67.84	0.14	2.27	Amadorio	Opitutae
<b>Opitutae-AMD-G3</b>	1720	1.90	64.52	37	94	85.58	3.74	2.14	Amadorio	Opitutae
<b>Opitutae-Tous-C10FEB</b>	1333	1.42	56.80	58	93	57.89	0.41	2.44	Tous	Opitutae
Opitutae-Tous-C1TDCM	3363	4.15	67.69	74	92	74.9	1.02	5.48	Tous	Opitutae
<b>Opitutae-Tous-C2FEB</b>	1236	1.33	64.70	35	94	68.16	2.86	1.89	Tous	Opitutae
Opitutae-Tous-C4FEB	3144	3.76	62.76	66	91	94.95	0.82	3.93	Tous	Opitutae
Opitutae-Tous-C5TDCM	3223	3.99	62.94	68	91	82.65	5.16	4.58	Tous	Opitutae
<b>Opitutae-Tous-C8FEB</b>	1810	2.17	69.98	34	95	67.76	0	3.21	Tous	Opitutae
Pedospaera-Tous-C6FEB	4618	5.69	64.58	76	91	87.85	35.78	4.16	Tous	Pedospaera
<b>Spartobacteria-AMD-G4</b>	1782	1.82	56.62	35	92	80.27	0.5	2.25	Amadorio	Spartobacteria
<b>Spartobacteria-AMD-G5</b>	617	0.56	58.59	18	94	29.59	0	1.89	Amadorio	Spartobacteria
<b>Spartobacteria-Tous-C9RFEB</b>	900	0.92	54.14	33	94	50.16	0	1.83	Tous	Spartobacteria
Verrucomicrobiae-AMD-G2	2798	3.26	52.06	63	91	96.08	0.82	3.36	Amadorio	Verrucomicrobiae
Verrucomicrobiae-Tous-C2TDCM	1612	1.85	61.72	63	93	55.07	0.41	3.34	Tous	Verrucomicrobiae
Verrucomicrobiae-Tous-C3TDCM	2628	2.96	52.00	62	91	87.62	0.41	3.37	Tous	Verrucomicrobiae
Verrucomicrobiae-Tous-C4TDCM	3690	4.28	63.94	69	91	66.98	3.27	6.18	Tous	Verrucomicrobiae
<b>Verrucomicrobiae-Tous-C5FEB</b>	1751	1.77	57.92	43	94	55.8	9.07	2.89	Tous	Verrucomicrobiae
Verrucomicrobia-Tous-C9LFEB	3585	4.28	57.29	68	90	90.61	2.04	4.63	Tous	Unclassified

Completeness, contamination, and estimated genome size (Mb) parameters were assessed by CheckM. Genomes considered small (<3.2 Mb and <60 bp of median intergenic spacer) are highlighted in bold.



**FIGURE 1 |** Phylogenomics of the entire phylum Verrucomicrobia. Reference isolates, MAGs, and SAGs from multiple habitats and classes (Opitutae, Verrucomicrobiae, Spartobacteria, Pedosphaera, and Unclassified) and subdivisions are represented in the protein concatenated tree. Tous and Amadorio MAGs presented in this study are red colored.

A 16S rRNA phylogenetic tree with reference verrucomicrobial sequences and those sequences retrieved from MAGs is also provided (Supplementary Figure 4). Among the 18 novel *Verrucomicrobia*, we reconstructed eight genomes belonging to the class Opiritutae, five to the class Verrucomicrobiae, three to class Spartobacteria, one to class Pedosphaera, and one related to the unclassified *Verrucomicrobia*, which contain representatives of the Methylococcaceae family (Op den Camp et al., 2009). It is quite remarkable that these recovered genomes appear in widely divergent taxonomic units attesting to the broad diversity of verrucomicrobial representatives inhabiting freshwater environments.

The genome sizes estimated from these MAGs suggest we have obtained two types of microbes (Figure 2), firstly small verrucomicrobial genomes, with relatively smaller intergenic spacers compared to those seen in reference marine (*Coralimargarita akajimensis* or *Rubritalea marina*), freshwater (*Opiritutaceae bacterium* IG16b, *Verrucomicrobia bacterium* SCGCAA027-I19), or brackish genomes (*Spartobacteria baltica*). These small genomes belong to the class Opiritutae within the family Puniceococcales (Opiritutae-AMD-G1, Opiritutae-AMD-G3, and Opiritutae-Tous-C2FEB) and the class Spartobacteria (*Spartobacteria-AMD-G4*, *Spartobacteria-AMD-G5*, and *Spartobacteria-Tous-C9RFEB*). They have estimated genome sizes of 1.83–2.27 Mb, with median intergenic spacers smaller than 38 bp (Supplementary Figure 3A). These *Spartobacteria* and Opiritutae–Puniceococcae genomes described here are among the smallest *Verrucomicrobia* genomes discovered so far. Secondly, we have reconstructed large *Verrucomicrobia* genomes, larger than 3.5 Mb (*Verrucomicrobia-Tous-C9LFEB* or *Opiritutae-Tous-C4FEB*), reaching in some cases 5 Mb (*Pedosphaera-Tous-C6FEB*), which exhibit larger genome sizes, although still smaller than the 6–8 Mb seen in soil representatives (e.g., *V. spinosum*), symbionts of termites (*Diplosphaera colitermitum* or *O. bacterium* TAV1, TAV3, TAV4, or TAV5) or associated to sugar beet plants (*Haloferula BVOr071*). Only a single genome was estimated to have a size >6 Mb (*Verrucomicrobiae-Tous-C4TDCM*, 6.18 Mb).

The verrucomicrobial genomes have been plotted by estimated genome sizes and median intergenic spacers (as indicators of how streamlined the genomes are) together with other typical freshwater microbes that go from the very streamlined K-strategists (acI *Actinobacteria* or LD12 *Alphaproteobacteria*) to relatively large genomes of r-strategists like *Limnohabitans* (Figure 2). It is evident how the genomes presented in this work cover the whole range of streamlining from one end to the other. Overall, we have obtained a great diversity of genomes from all known *Verrucomicrobia* classes, although we were able to reconstruct more genomes belonging to the class Opiritutae. We have focused our efforts in a comparative study between the two kinds of Opiritutae genomes, that appear to correspond also to a separation between endemic and cosmopolitan microbes (see below).

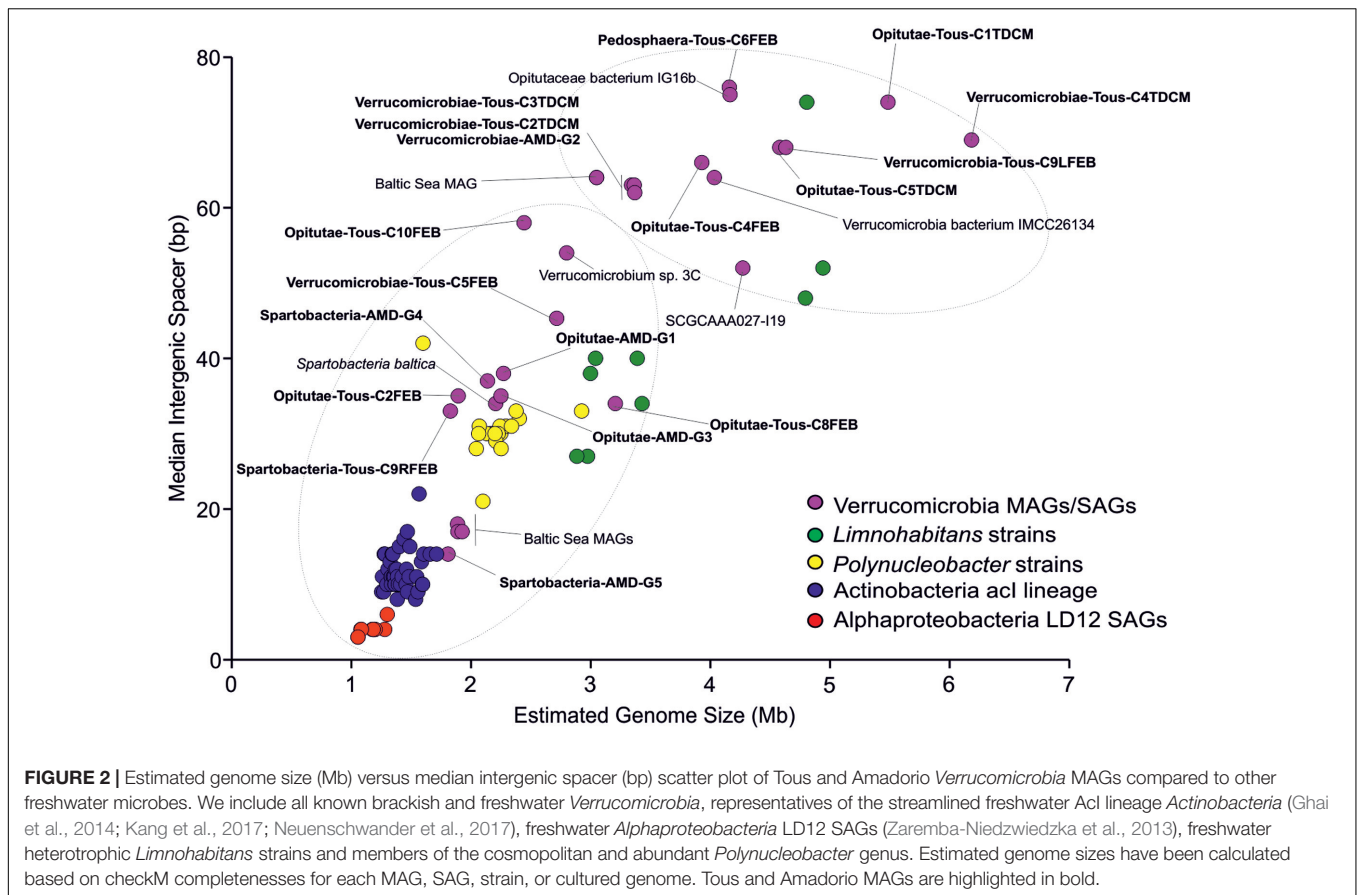
## Abundant versus Rare *Verrucomicrobia*

In this study we have retrieved MAGs coming from summer stratification and winter mixing periods from Tous reservoir.

Specifically, four of the Tous DCM (deep chlorophyll maximum) reconstructed *Verrucomicrobia* were found exclusively associated to the DCM of the stratified summer sample (Supplementary Figure 5). This seems to indicate that the DCM has a specific verrucomicrobial population associated to it. However, most MAGs were present at much higher abundances at the Tous 25 m metagenome. One likely explanation, being largely heterotrophic microbes specialized in the degradation of polymers (see below), is that they accumulate at the deepest part of the water body where particulate material sinking in the water column would be more abundant. A different distribution pattern was found for the photoheterotrophic Opiritutae-Tous-C4FEB (and the nearly identical Opiritutae-Tous-C5TDCM) and *Verrucomicrobiae-Tous-C5FEB*, which showed very similar recruitment values in the three depths tested for Tous, either 12 and 25 (winter sample) or 13 m (summer stratified sample). One of these, Opiritutae-Tous-C4FEB, was also found at similar abundances in Amadorio 10 m metagenome (winter sample). Both Opiritutae-Tous-C4FEB and *Verrucomicrobiae-Tous-C5FEB* contain proteorhodopsin genes (see below) and photoheterotrophic potential which could help extend their range to more superficial waters.

Abundances of all MAGs were tested by mapping the reads of different freshwater datasets to our MAGs (Supplementary Figure 6). In general, we observed that small Opiritutae genomes were more abundant in all tested metagenomes compared to the rest of the MAGs retrieved from different classes. The small Opiritutae-Puniceococcae (Opiritutae-AMD-G1, Opiritutae-AMD-G3, and Opiritutae-Tous-C2FEB) representatives mainly recruit in Tous 25 m metagenome, in contrast with the low abundances found in Tous 12 m metagenome, which suggests that these microbes are likely associated to the deeper parts of Tous reservoir. Interestingly, we also observed high abundances for Opiritutae-AMD-G1 and Opiritutae-AMD-G3/Opiritutae-Tous-C2FEB in Lake Baikal, the ultraoligotrophic largest freshwater body on Earth (Cabello-Yeves et al., submitted for publication) (suffering from 5-month frozen cycles) compared to the more temperate Spanish reservoirs. They were also found at lower abundances in Klamath Iron Gate Dam and lakes Mendota, Erken, and Lanier suggesting a good adaption to a wide range of lakes, ranging from oligotrophic (Tous reservoir) to eutrophic (Lake Mendota) and temperate (Tous or Amadorio) to cold (Lake Baikal or Swedish lakes). These data suggest that these microbes could be a very abundant subclass of Opiritutae members in many different freshwater bodies, as confirmed by the recruitment plots (Supplementary Figures 7, 8). We also detected values of 10 RPKG (1× coverage) for Opiritutae-AMD-G1/G3 in samples associated to cyanoblooms from Dexter reservoir, showing evidence of an association between (pico) cyanobacteria and *Verrucomicrobia* (Eiler and Bertilsson, 2004). Abundance was also assessed in the same datasets for some reference freshwater *Verrucomicrobia*, finding low RPKG values for all of them (Supplementary Figure 5).

The distribution of total planktonic *Verrucomicrobia* (EUBB-III) and the specific groups (*Verrucomicrobia-Tous-C9LFEB* for Tous and Opiritutae-AMD-G3 for Amadorio) in the vertical profile (samples were taken every 2 m depth and fixed



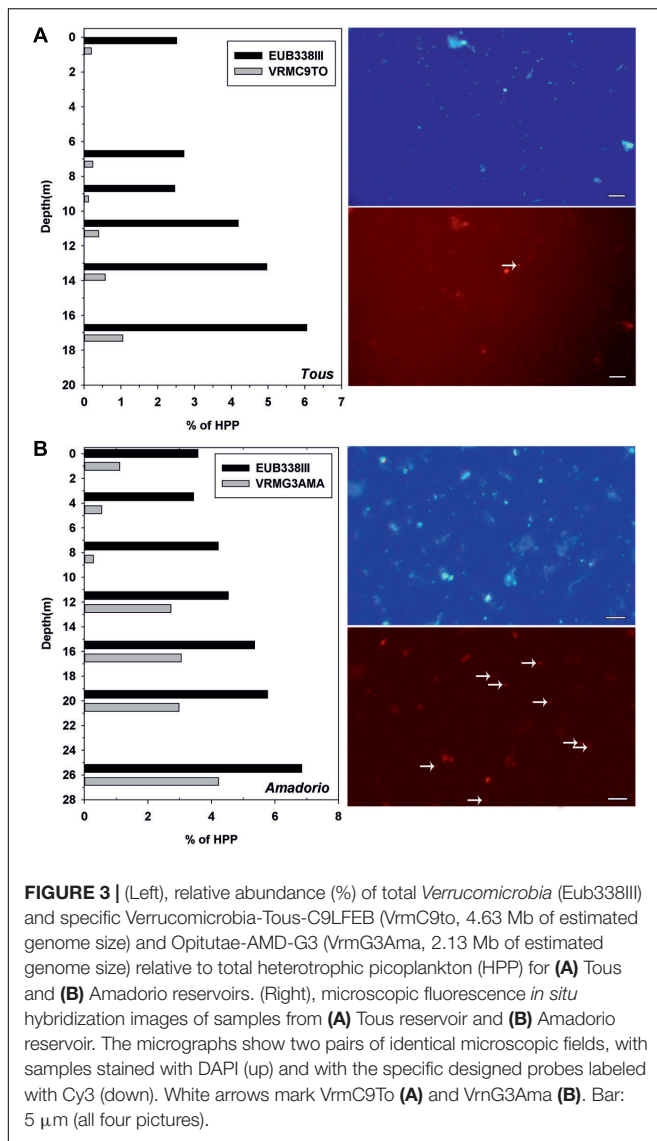
for microscopy) under summer stratification conditions was addressed by FISH in both reservoirs (Figure 3). The distribution pattern in the vertical profile of the percentage of total *Verrucomicrobia* with respect to the total heterotrophs was similar in both reservoirs, with their relative contribution increasing with depth. In both cases the lowest values were observed in the upper waters (epilimnion) with an average of 2.5% in Tous and a 2.8% in Amadorio, gradually increasing toward the bottom of the reservoirs with deep maxima of *Verrucomicrobia* contribution to total bacterioplankton of ca. 6% in Tous and 7% in Amadorio. FISH counts using the VrmC9LT-specific probe of *Verrucomicrobia*-Tous-C9LFEB show that this group represents only a small part of the total heterotrophs of Tous Reservoir during summer stratification, always below 1%, with a relatively constant contribution of ca. 15% of total *Verrucomicrobia* within the water column. The VrmG3Ama probe, which is specific of *Opitutae*-AMD-G3 from Amadorio reservoir, allowed us determining that this group was present at a relatively low contribution in relation to total heterotrophs (<1%), however, from 12 m depth to the bottom its abundance increased to 3–4% with respect to the total of heterotrophs and contributed over 50% of the total *Verrucomicrobia*. These data confirm that *Opitutae*-AMD-G3 (2.13 Mb of estimated genome size) is a very abundant microbe in Amadorio (and also in Tous as was previously assessed by recruitment), which fits with the proposal that these microbes are K-strategists, while

*Verrucomicrobia*-Tous-C9LFEB (large genome of 4.63 Mb of estimated genome size) contributed at very low numbers in Tous (as was also tested by recruitment), which fits with the idea that it possesses an opportunistic strategy.

## Carbohydrate-Based Metabolism in Freshwater *Verrucomicrobia*

In general, we found an aerobic heterotrophic metabolism with degradation of polysaccharides in all of our genomes, containing metabolic pathways involved in glycolate, glucarate/galactarate, glucuronate/galacturonate, gluconate, polysaccharides, monosaccharides, and aminosugars utilization and degradation. We also noticed a large amount of Ton/Tol/Exb receptors for transport of polymers, vitamins, and iron uptake, and carbon storage preference for glycogen (which is both synthesized and degraded) instead of poly-[(beta)]-hydroxybutyrate (PHB). Interestingly, *Verrucomicrobia* appear to be involved in the metabolism of two glycosphingolipids, galactosylceramides, and sulfatides, a pathway that is typically found in eukaryotes, presenting several enzymes like aryl-sulfatases, sialidases, and beta-galactosidases. Among the diverse microbes reconstructed in this work, these features were widespread and common to the majority of them, although certain pathways for mono and polysaccharide utilization were specific to some (see below). Despite the general aerobic and heterotrophic metabolism





inferred, we were able to identify some microbes with anaerobic reductase complexes (*Pedospaera*-Tous-C6FEB), some of them could be able to perform photoheterotrophic lifestyles through rhodopsin pumps and one of them could be capable of nitrogen fixation.

In addition to more simple sugars, *Verrucomicrobia* are known to be able to tackle more complex and recalcitrant polysaccharides (Martinez-Garcia et al., 2012a). We performed a protein annotation for each MAG in order to explore the presence of glycosyl hydrolases (GHs) and other polysaccharide degrading enzymes based on the CAZY (Lombard et al., 2013) and dbCAN (Yin et al., 2012) databases. An overview of the GH families presented in each MAG is shown in the Supplementary Data Sheet 1. In general, we observed that the reconstructed freshwater *Verrucomicrobia* are fundamentally chitin degraders as they chiefly contain GH109 ( $\alpha$ -*N*-acetylgalactosaminidase). We also detected some other abundant GHs families as GH33 (sialidase or neuraminidase), GH74 (endoglucanase and

xyloglucanase) and GH2 ( $\beta$ -galactosidase,  $\beta$ -mannosidase, or  $\beta$ -glucuronidase), GH13\_9 ( $\alpha$ -amylase), and GH77 (amylomaltase or 4- $\alpha$ -glucanotransferase). In general, we detected a growing proportion of GHs with the increase in MAG genome size and the total number of CDS. It also correlates with a higher number of transporters and metabolic pathways found in the larger genomes compared to the small genomes (see below). It is noteworthy that *Verrucomicrobia*-Tous-C9LFEB contained the highest number of GHs (218), a huge number compared to the other MAGs, which contained between 2 and 80 GHs. This microbe, as explained below, contains the highest and most remarkable variety of metabolic pathways of all the genomes compared here, including nitrogen fixation and efficient degradation of recalcitrant compounds such as chitin, cellulose, xylan, starch, or mannan, among others.

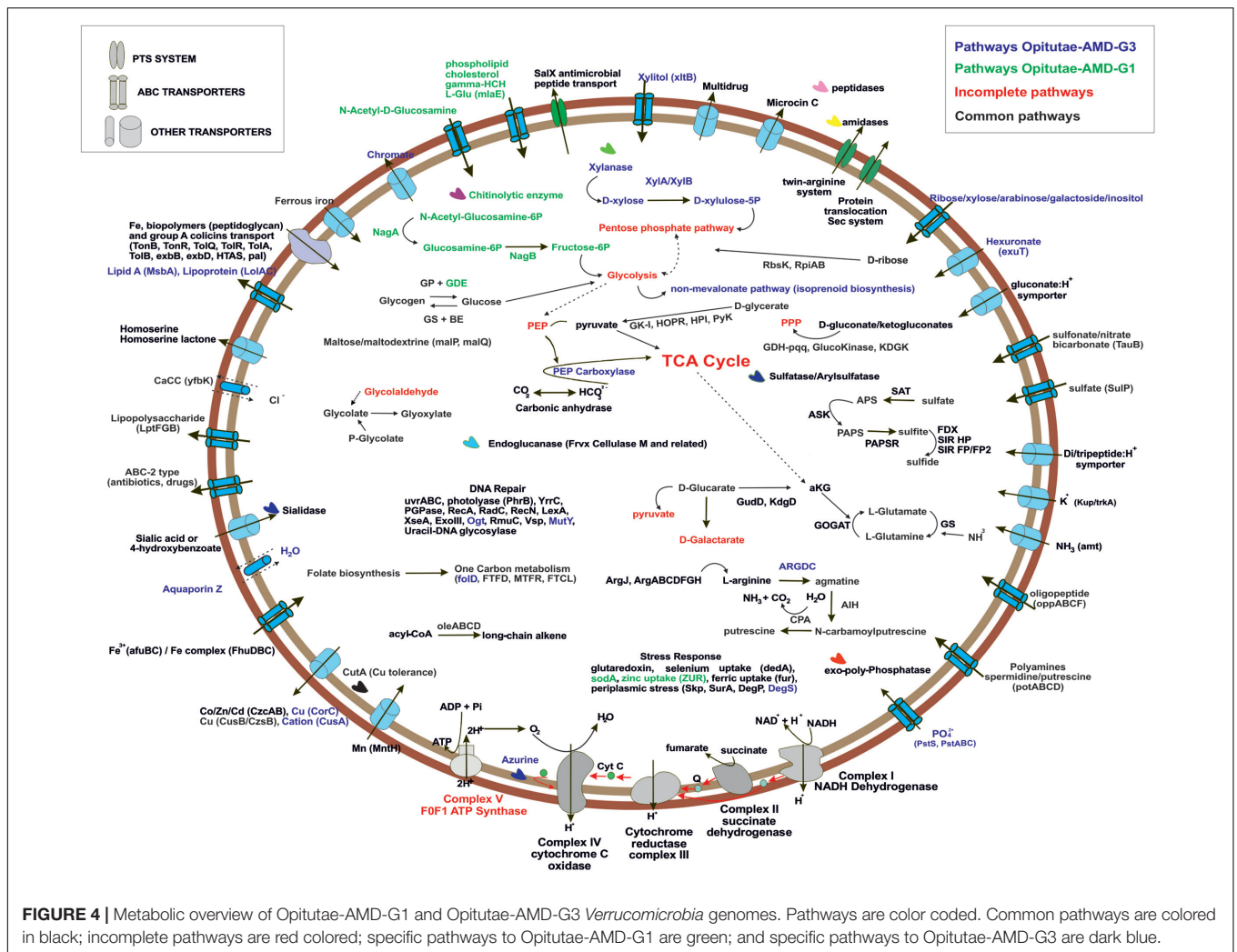
## Metabolic Overview of Small versus Large Verrucomicrobial Genomes

To simplify the work presented here, we have analyzed in depth the metabolic pathways found in four of our MAGs. We focused on the two small and cosmopolitan *Opiritutae*-AMD-G1 and *Opiritutae*-AMD-G3 (Figure 4), which have genome size estimations of around 2.2 Mb, being among the smallest *Verrucomicrobia* representatives ever found, at least in freshwater ecosystems. We also studied more in depth two of the largest genomes that appear sporadically in metagenomes (see above), *Opiritutae*-Tous-C4FEB (Figure 5) and *Verrucomicrobia*-Tous-C9LFEB (Supplementary Figure 9), which have estimated sizes of 3.9 and 4.6 Mb, respectively, and contain a rhodopsin and nitrogen fixation, respectively.

## C-metabolism

The ability of these microbes to use glycogen instead of PBH as carbon storage source is interesting, and has also been reported in acidophilic nitrogen fixing methanotrophs (Khadem et al., 2010, 2012a). Genes for glycogen biosynthesis (*glgA*, *glgB*, *glgC*, *glgP*) were common to both small and large genomes, although glycogen debranching enzyme (*glgX*) was only detected in the small *Opiritutae*-AMD-G1. A specific operon (*oleABCD*) that has been previously described for long chain alkene biosynthesis from acyl-CoA in some members of *Planctomycetes*, *Verrucomicrobia*, *Actinobacteria*, *Chloroflexi*, and *Betaproteobacteria* (Sukovich et al., 2010), was only found in these *Opiritutae* genomes, either large or small, and appears to be a specific carbon storage of this class.

One of the most remarkable features of these freshwater *Verrucomicrobia* (independently of their genome size) was their ability to effectively degrade different types of mono and polysaccharides (as described above). Sialidases (including one sialic acid transporter), which release *N*-acetylneuraminate residues from glycans, glycoproteins, and polysaccharides (Moncla et al., 1990; Kim et al., 2011) were found in the four genomes. Endoglucanases (Frvx), coding for cellulase M, were inferred from both large and small genomes. Degradation of xylan, a major group of hemicellulose present in plant cell walls and some algae, was detected in the big genomes and the small *Opiritutae*-AMD-G3 but not in the small *Opiritutae*-AMD-G1.



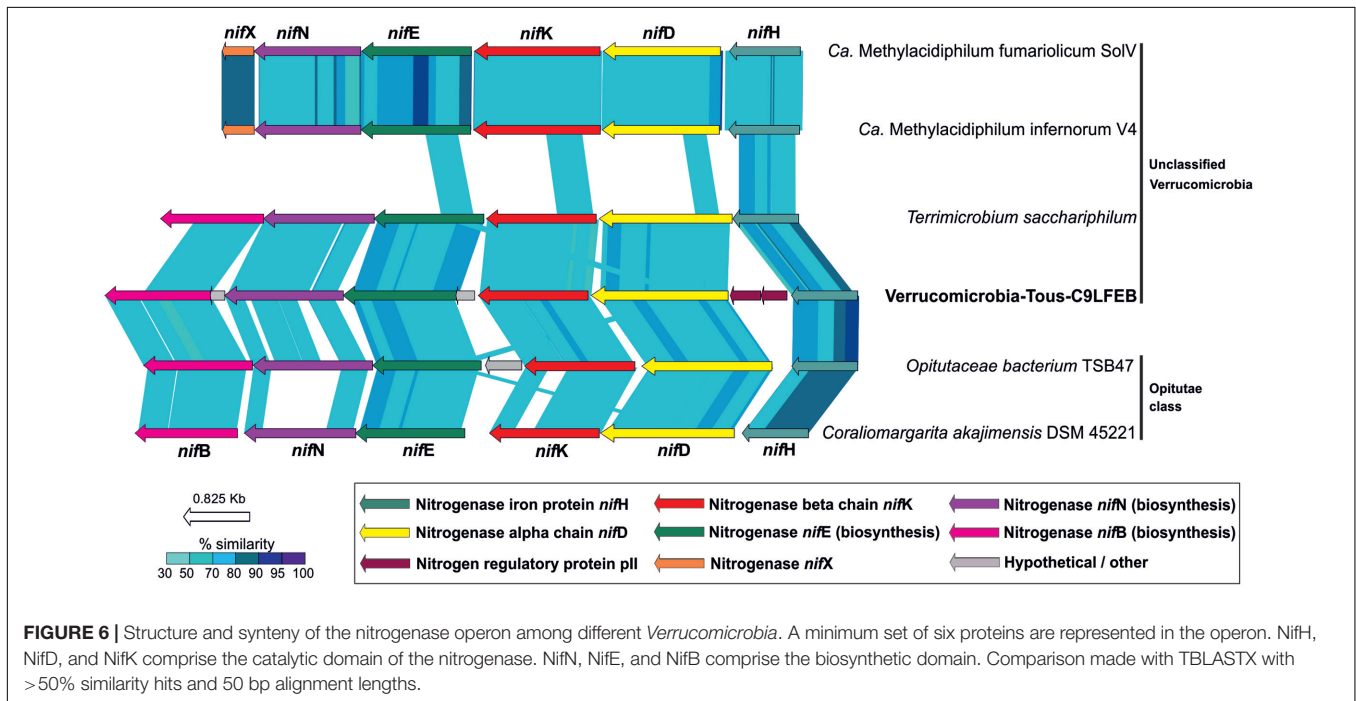
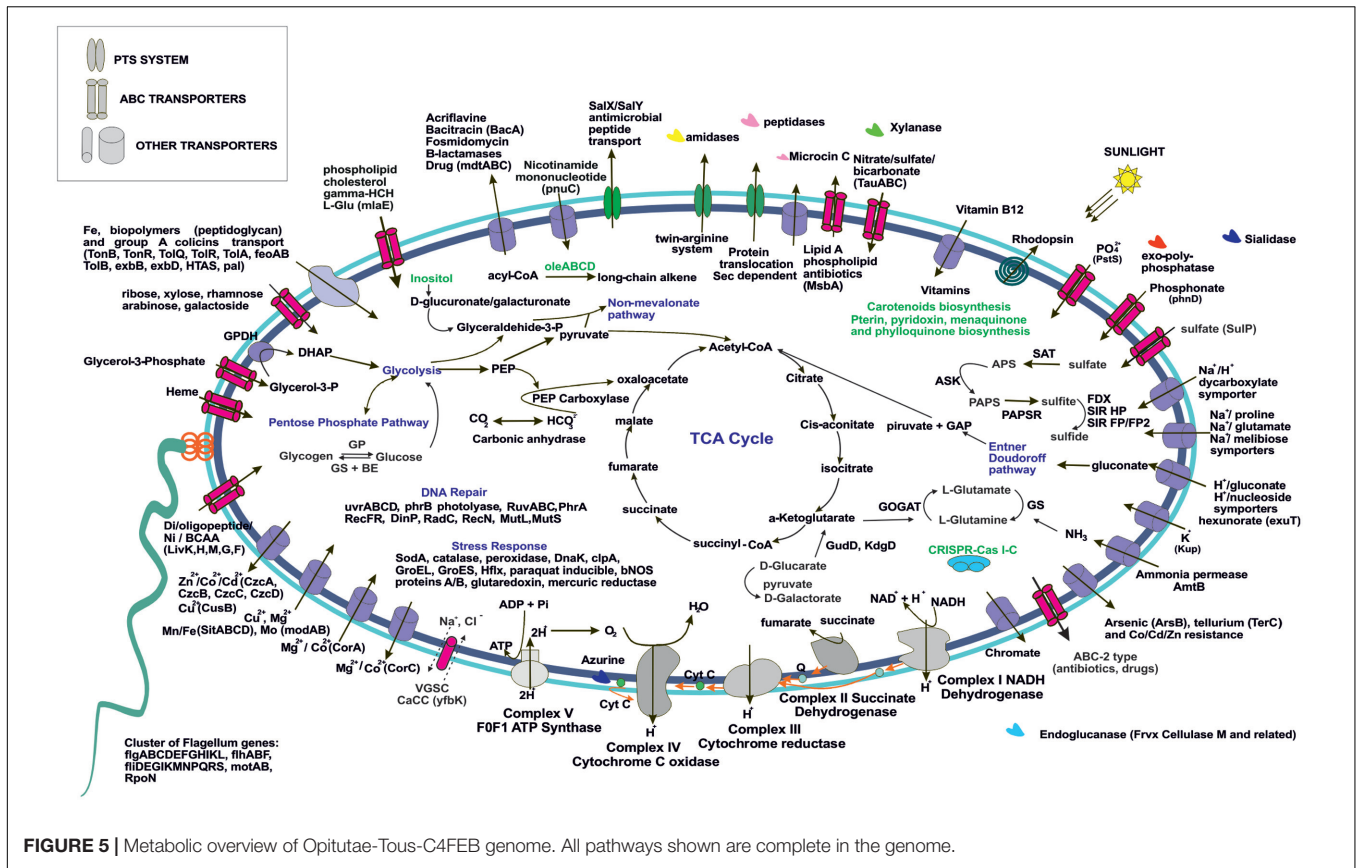
Conversely, we found a specific pathway for *N*-acetyl-D-glucosamine (chitin) degradation in small Opiritutae-AMD-G1 and big Opiritutae-Tous-C4FEB (although only one chitinolytic enzyme was found in the small Opiritutae-AMD-G1). Complete pathways with GHs for mannose and rhamnose utilization were only detected in the unclassified *Verrucomicrobia*-Tous-C9LFEB. The two large genomes were able to utilize sugar alcohol carbon sources, since they have a glycerol-3-phosphate transporter and a glycerol-3-phosphate dehydrogenase. Pathways for inositol catabolism to form glucuronate and transformation of the latter to glyceraldehyde-3-phosphate, as well as fructose utilization via PTS transport system were also common in these large genomes.

## Transporters

Generally, we detected a high number of Exb, Ton, and Tol biopolymer and iron transporters in all the reconstructed genomes, although we also noticed an increase in number of these transporter with increasing genome size. The presence of these transporters suggests an alternative route for incorporation and recycling of dissolved organic matter in freshwater bodies,

which was already described in marine ecosystems (Tang et al., 2012). In general we could identify a higher number of transporters for heavy metals (some of them conferring resistance to arsenite and tellurium), branched chain amino acids, di-oligo-peptides, heme, or multidrug resistance in the larger genomes.

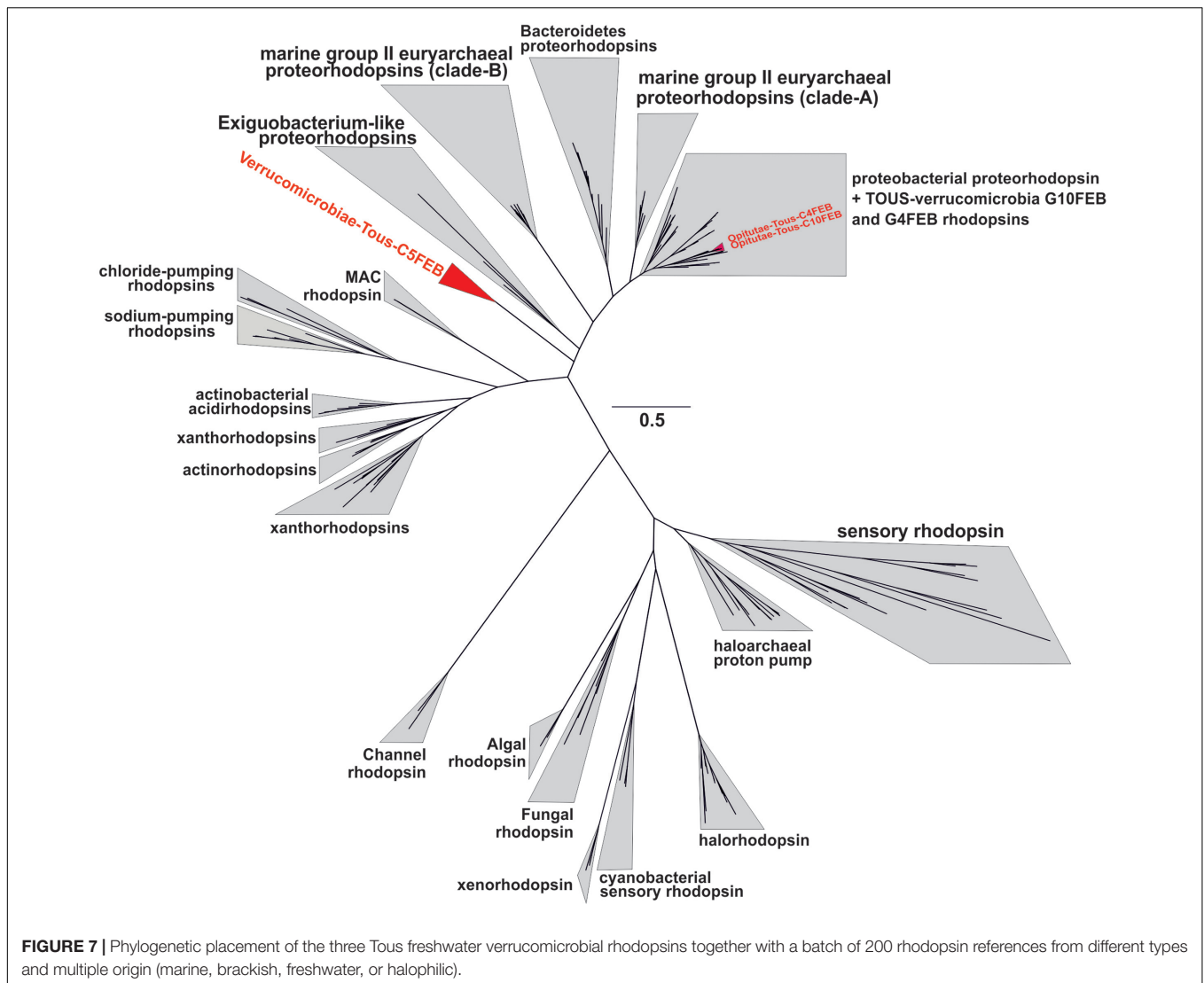
Both large and small genomes contain exo-polyphosphatases, inorganic pyrophosphatases, phosphate starvation proteins (PhoH), alkaline phosphatase, or inorganic phosphate transporters, which suggest that *Verrucomicrobia*, like freshwater *Actinobacteria*, are important in inorganic phosphorous sequestering which could contribute to prevent eutrophication (Ghai et al., 2014). The storage of polyphosphate was also assessed in the methanotroph *Ca. Methylacidiphilum fumarolicum* SolV (Khadem et al., 2012a). Polyamines have been described in some members of the PVC superphylum in freshwater systems, especially in *Planctomycetes* (Gripenburg et al., 1999); only small genomes presented a spermidine/putrescine transporter (potABCD) together with a cluster of putrescine biosynthesis from L-arginine and agmatine.



**Other Ecologically Relevant Genes**

Bacterial nitric oxide synthases have been described in the past years as enzymes involved in toxin biosynthesis, protection

against oxidative stress, and regulation of recovery from radiation damage (Correa-Aragunde et al., 2013). In *Opitutae-Tous-C4FEB* genome one gene encoding a nitric oxide synthase



was detected; the presence of this gene was also common to other reconstructed MAGs. However, we did not detect any potential toxin-related genes adjacent to this gene as has been described previously in *Streptomyces* (Kers et al., 2004).

We found a cluster of flagellum genes for motility functions exclusively in the large genomes (*Opitutae-Tous-C4FEB* and *Verrucomicrobia-Tous-C9LFEB*), which was totally absent in the rest of the reconstructed genomes of this work. A CRISPR-Cas type I-C was also found in *Opitutae-Tous-C4FEB*. Some of the main differences between small and large genomes were based on their ability to perform other different types of metabolism which couple with the general heterotrophic metabolism in *Verrucomicrobia*. For instance, *Opitutae-Tous-C4* (Figure 5), a 3.9 Mb microbe (estimated genome size) contains almost all of the pathways described in the small *Opitutae* (see above) plus other pathways suggesting a photoheterotrophic metabolism, including a complete pathway for carotenoids biosynthesis together with a proteorhodopsin proton-pump.

Carbonic anhydrases, phosphoenolpyruvate carboxylase, or PhrB photolyases for stress response to light were also detected. Compared to the small genomes, we found more vitamin (pterines, pyridoxine, menaquinone, and phylloquinone) and cofactor biosynthetic pathways, including one vitamin B12 transporter in *Opitutae-Tous-C4*.

## Nitrogen Fixation in Freshwater *Verrucomicrobia*

Although *Verrucomicrobia*, including methanotrophs and termite symbionts (Khadem et al., 2010; Wertz et al., 2012), have been previously reported to be capable of fixing nitrogen, we are describing for the first time the possible ability of a truly freshwater *Verrucomicrobia* (*Verrucomicrobia-Tous-C9LFEB*) for nitrogen fixation. Typically, a minimum set of six *nif* genes is required in bacteria for nitrogen fixation, three of them code for structural and catalytic components (*NifHDK*) and the other three for biosynthetic proteins (*NifENB*) (Dos Santos et al.,

2012). We found a complete *nif* operon containing the six *nifHDKENB* subunits.

A phylogeny of the NifHDK concatenate among different molybdenum, vanadium, and iron nitrogenases from different bacterial phyla representatives (Boyd et al., 2011, 2015; Boyd and Peters, 2013; McGlynn et al., 2013) (Supplementary Figure 10) suggests that the nitrogenase described here affiliates with a cluster of molybdenum *Verrucomicrobia* nitrogenases which have representatives in termite guts (*O. bacterium* TAV5 and TSB47), soil (*Terrimicrobium sacchariphilum*), and sea (*C. akajimensiis*). Similar to what can be seen for the entire operon (Figure 6) these nitrogenases are more divergent compared to the verrucomicrobial *Methylacidiphilum* representatives (Khadem et al., 2010).

A comparison of the *nif* operon found in *Verrucomicrobia*-Tous-C9LFEB with some representatives of the *Verrucomicrobia* phylum is shown in Figure 6. The *nif* operon is largely syntenic across these microbes, with the exception of an additional *nifX* gene in the *Methylacidiphilum* representatives. However, *Verrucomicrobia*-Tous-C9LFEB *nif* cluster is more similar to the one found before in *Opiritutae*, even though the genome itself is phylogenetically closer to unclassified *Verrucomicrobia* (Figure 1). Another peculiar feature of the operon is the apparent insertion/translocation of the gene encoding the pII nitrogen regulatory protein, between *nifD* and *nifH* genes. Moreover, the pII gene also appears to be duplicated.

Nitrogenase activity is commonly limited to anaerobic environments, where nitrogenase is not inhibited by oxygen, although some aerobic microorganisms like *Cyanobacteria* (Esteves-Ferreira et al., 2017) or *Azotobacter vinelandii* (Bishop et al., 1980) present strategies to perform nitrogen fixation under aerobic conditions. The discovery of the nitrogen fixing *Verrucomicrobia*-Tous-C9LFEB is remarkable since the study site (Tous reservoir) is totally oxic throughout the year. However, the anaerobic nitrogen fixing activity may be facilitated by anoxic microenvironments in cell aggregates or particles regardless of the oxygen content of the water column (Dang and Lovell, 2016). This microbe also contains nitrate and nitrite ammonification pathways, which could be used as an alternative pathway for ammonia production, especially when nitrogen fixation leads to high energetic expenses for the cell.

## Novel Verrucomicrobial Rhodopsins

It has been suggested that the *Verrucomicrobia* phylum could have representatives with photoheterotrophic metabolism via rhodopsin pumps (Martinez-Garcia et al., 2012b). In this manuscript we report three novel rhodopsins, the ones from the *Opiritutae*-Tous-C4FEB and *Opiritutae*-Tous-C10FEB genomes, clearly belong to the marine proteorhodopsin clade (Figure 7). Additionally, *Verrucomicrobia*-Tous-C5FEB contains one novel rhodopsin belonging to a new clade of rhodopsin proton pumps, its closest sequences being the rhodopsins from *Exiguobacterium* (Gushchin et al., 2013). This rhodopsin could be part of a novel branch characteristic of *Planctomycetes*-*Verrucomicrobia* PVC superphylum representatives. The key residues of these three rhodopsins, compared to other representatives (Supplementary Figure 11) and based on the

presence of the L-105 or M-105 residues in the retinal pocket, clearly show that the ones described here are green light absorbing proton pumps (Man et al., 2003). Actually, it is remarkable that all rhodopsins from Tous metagenomes were green-light variants.

## CONCLUSION

*Verrucomicrobia* is a relatively novel phylum which has acquired interest in the scientific community, mainly because of their global distribution in different types of environments and hosts. Analyzing metagenomes from freshwater reservoirs in Spain we could reconstruct a remarkable diversity of relatively complete verrucomicrobial-assembled genomes. They covered a broad range of genome sizes and streamlining levels (judging from the mean intergenic spacers). This great size variability of reconstructed genomes, ranging from about 2 to 6 Mb, would shed light onto the classification inside the phylum. We found a correlation between both ends of the genome size spectrum and the distribution in diverse freshwater and brackish metagenomes around the world with some small genomes being extremely widespread and abundant and others, particularly the large genomes, with much more restricted distribution. Although all the genomes seem to correspond to aerobic heterotrophs degrading polysaccharides, the larger genomes display much more transport potential and have likely more diversified metabolism. There is a group that appears to be adapted to the stratified period at DCM depths while the majority of genomes appear to be more abundant in deeper waters (Supplementary Figure 5), which are richer in sinking particles that could be the ecological niche of many of these microbes. The uneven distribution of larger genomes could just reflect a blooming strategy that makes their presence sporadic while the smaller genomes are probably more constant in their presence being less demanding and likely also slower growing (K-strategists).

We found a nitrogen fixing cluster, a characteristic never described before in free living freshwater *Verrucomicrobia* and (excluding cyanobacteria) not typical from freshwater well-oxygenated systems like Tous reservoir, which never undergoes anoxic conditions. The photoheterotrophic metabolism in some of the microbes presented here was confirmed with the discovery of freshwater *Planctomycetes*-*Verrucomicrobia* rhodopsins, likely a novel branch within this already hyperdiversified family of proteins. It is noteworthy that the small *Opiritutae*-AMD-G3 and *Opiritutae*-AMD-G1 described here are cosmopolitan in lakes from different environmental conditions, ranging from cold oligotrophic lakes (Lake Baikal or Swedish Lakes), temperate oligotrophic (Lake Lanier or Tous reservoir), and temperate mesotrophic (Amadorio) or eutrophic (Lake Mendota). We have visualized by FISH the microbes represented by *Opiritutae*-AMD-G3 and *Verrucomicrobia*-Tous-C9LFEB and would like to propose *Candidatus* names for both microbes: *Candidatus* *Opiritutus lacustris* AMD-G3 and *Candidatus* *Verrucomicrobium* *diatrophicus* Tous-C9LFEB.

## DATA ACCESSIBILITY

Tous reservoir metagenomic datasets have been deposited in the NCBI SRA database with BioProject number PRJNA342151 (Tous February 12 m: SRR4198666; Tous February 25 m: SRR4198832; Tous July 13 m: SRR5338504). Amadorio reservoir metagenomic datasets are available under BioProject number PRJNA238866. The 18 assembled *Verrucomicrobia* genomes have been deposited in the NCBI under Biosample identifiers SAMN06820742-SAMN06820759.

## AUTHOR CONTRIBUTIONS

RG, FR-V, and PC-Y conceived this work. PC-Y and AP performed the sample collection and filtration. PC-Y performed the DNA extraction. Analysis was carried out by PC-Y, RG, MM, and AP. Manuscript was written by PC-Y, FR-V, and RG. All authors read and approved the final manuscript.

## REFERENCES

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Arnds, J., Knittel, K., Buck, U., Winkel, M., and Amann, R. (2010). Development of a 16S rRNA-targeted probe set for *Verrucomicrobia* and its application for fluorescence *in situ* hybridization in a humic lake. *Syst. Appl. Microbiol.* 33, 139–148. doi: 10.1016/j.syapm.2009.12.005
- Bergmann, G. T., Bates, S. T., Eilers, K. G., Lauber, C. L., Caporaso, J. G., Walters, W. A., et al. (2011). The under-recognized dominance of *Verrucomicrobia* in soil bacterial communities. *Soil Biol. Biochem.* 43, 1450–1455. doi: 10.1016/j.soilbio.2011.03.012
- Bishop, P. E., Jarlenski, D., and Hetherington, D. R. (1980). Evidence for an alternative nitrogen fixation system in *Azotobacter vinelandii*. *Proc. Natl. Acad. Sci. U.S.A.* 77, 7342–7346. doi: 10.1073/pnas.77.12.7342
- Boyd, E., and Peters, J. W. (2013). New insights into the evolutionary history of biological nitrogen fixation. *Front. Microbiol.* 4:201. doi: 10.3389/fmicb.2013.00201
- Boyd, E. S., Costas, A. M. G., Hamilton, T. L., Mus, F., and Peters, J. W. (2015). Evolution of molybdenum nitrogenase during the transition from anaerobic to aerobic metabolism. *J. Bacteriol.* 197, 1690–1699. doi: 10.1128/JB.02611-14
- Boyd, E. S., Hamilton, T. L., and Peters, J. W. (2011). An alternative path for the evolution of biological nitrogen fixation. *Front. Microbiol.* 2:205. doi: 10.3389/fmicb.2011.00205
- Boyd, R. (2000). Herbicides and herbicide degradates in shallow groundwater and the Cedar River near a municipal well field, Cedar Rapids, Iowa. *Sci. Total Environ.* 248, 241–253. doi: 10.1016/S0048-9697(99)00546-X
- Bri e, C., Moreira, D., and L opez-Garc a, P. (2007). Archaeal and bacterial community composition of sediment and plankton from a suboxic freshwater pond. *Res. Microbiol.* 158, 213–227. doi: 10.1016/j.resmic.2006.12.012
- Brown, A. M., Howe, D. K., Wasala, S. K., Peetz, A. B., Zasada, I. A., and Denver, D. R. (2015). Comparative genomics of a plant-parasitic nematode endosymbiont suggest a role in nutritional symbiosis. *Genome Biol. Evol.* 7, 2727–2746. doi: 10.1093/gbe/evv176
- Cabello-Yeves, P. J., Haro-Moreno, J. M., Martin-Cuadrado, A.-B., Ghai, R., Picazo, A., Camacho, A., et al. (2017). Novel *Synechococcus* genomes reconstructed from freshwater reservoirs. *Front. Microbiol.* 8:1151. doi: 10.3389/fmicb.2017.01151
- Chin, K.-J., Liesack, W., and Janssen, P. H. (2001). *Opiritutus terrae* gen. nov., sp. nov., to accommodate novel strains of the division ‘*Verrucomicrobia*’ isolated

## ACKNOWLEDGMENTS

FR-V was supported by grant “VIREVO” CGL2016-76273-P [AEI/FEDER, EU], (cofunded with FEDER funds); Acciones de dinamizaci n “REDES DE EXCELENCIA” CONSOLIDER-CGL2015-71523-REDC from the Spanish Ministerio de Econom a, Industria y Competitividad and PROMETEO II/2014/012 “AQUAMET” from Generalitat Valenciana. MM was supported by the Czech Academy of Sciences (Postdoc program PPPLZ application number L200961651). RG was supported by the Grant Agency of the Czech Republic by the research grant 17-04828S. The authors would like to thank Ana-Bel en Martin-Cuadrado, Riccardo Rosselli, and Rafael Gonz alez-Serrano for assistance with sampling and filtration.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2017.02131/full#supplementary-material>

- from rice paddy soil. *Int. J. Syst. Evol. Microbiol.* 51, 1965–1968. doi: 10.1099/00207713-51-6-1965
- Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., et al. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* 37(suppl 1), D141–D145. doi: 10.1093/nar/gkn879
- Correa-Aragunde, N., Foresi, N., and Lamattina, L. (2013). Structure diversity of nitric oxide synthases (NOS): the emergence of new forms in photosynthetic organisms. *Front. Plant Sci.* 4:232. doi: 10.3389/fpls.2013.00232
- Daims, H., Br uhl, A., Amann, R., Schleifer, K.-H., and Wagner, M. (1999). The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. *Syst. Appl. Microbiol.* 22, 434–444. doi: 10.1016/S0723-2020(99)80053-8
- Dang, H., and Lovell, C. R. (2016). Microbial surface colonization and biofilm development in marine environments. *Microbiol. Mol. Biol. Rev.* 80, 91–138. doi: 10.1128/MMBR.00037-15
- De Wever, A., Van der Gucht, K., Muylaert, K., Cousin, S., and Vyverman, W. (2008). Clone library analysis reveals an unusual composition and strong habitat partitioning of pelagic bacterial communities in Lake Tanganyika. *Aquat. Microb. Ecol.* 50, 113–122. doi: 10.3354/ame01157
- Derrien, M., Vaughan, E. E., Plugge, C. M., and de Vos, W. M. (2004). *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* 54, 1469–1476. doi: 10.1099/ijs.0.02873-0
- Devos, D. P., and Reynaud, E. G. (2010). Intermediate steps. *Science* 330, 1187–1188. doi: 10.1126/science.1196720
- Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C., and Dixon, R. (2012). Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* 13:162. doi: 10.1186/1471-2164-13-162
- Dunfield, P. F., Yuryev, A., Senin, P., Smirnova, A. V., Stott, M. B., Hou, S., et al. (2007). Methane oxidation by an extremely acidophilic bacterium of the phylum Verrucomicrobia. *Nature* 450, 879–882. doi: 10.1038/nature06411
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461
- Eiler, A., and Bertilsson, S. (2004). Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. *Environ. Microbiol.* 6, 1228–1243. doi: 10.1111/j.1462-2920.2004.00657.x
- Eiler, A., Zaremba-Niedzwiedzka, K., Mart nez-Garc a, M., McMahon, K. D., Stepanauskas, R., Andersson, S. G., et al. (2014). Productivity and salinity structuring of the microplankton revealed by comparative freshwater

- metagenomics. *Environ. Microbiol.* 16, 2682–2698. doi: 10.1111/1462-2920.12301
- Erikstad, H.-A., and Birkeland, N.-K. (2015). Draft genome sequence of “*Candidatus Methylocidiphilum kamchatkense*” strain Kam1, a thermoacidophilic methanotrophic Verrucomicrobium. *Genome Announc.* 3:e00065-15. doi: 10.1128/genomeA.00065-15
- Esteves-Ferreira, A. A., Cavalcanti, J. H. F., Vaz, M. G. M. V., Alvarenga, L. V., Nunes-Nesi, A., and Araújo, W. L. (2017). Cyanobacterial nitrogenases: phylogenetic diversity, regulation and functional predictions. *Genet. Mol. Biol.* 40(1 Suppl. 1), 261–275. doi: 10.1590/1678-4685-gmb-2016-0050
- Forster, P., and Gribaldo, S. (2010). Bacteria with a eukaryotic touch: a glimpse of ancient evolution? *Proc. Natl. Acad. Sci. U.S.A.* 107, 12739–12740. doi: 10.1073/pnas.1007720107
- Freitas, S., Hatosy, S., Fuhrman, J. A., Huse, S. M., Welch, D. B. M., Sogin, M. L., et al. (2012). Global distribution and diversity of marine Verrucomicrobia. *ISME J.* 6, 1499–1505. doi: 10.1038/ismej.2012.3
- Ghai, R., Mizuno, C. M., Picazo, A., Camacho, A., and Rodriguez-Valera, F. (2014). Key roles for freshwater Actinobacteria revealed by deep metagenomic sequencing. *Mol. Ecol.* 23, 6073–6090. doi: 10.1111/mec.12985
- Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijs.0.64483-0
- Gripenburg, U., Ward-Rainey, N., Mohamed, S., Schlesner, H., Marxsen, H., Rainey, F. A., et al. (1999). Phylogenetic diversity, polyamine pattern and DNA base composition of members of the order Planctomycetales. *Int. J. Syst. Evol. Microbiol.* 49, 689–696. doi: 10.1099/00207713-49-2-689
- Gupta, R. S., Bhandari, V., and Naushad, H. S. (2012). Molecular signatures for the PVC clade (Planctomycetes, Verrucomicrobia, Chlamydiae, and Lentisphaerae) of bacteria provide insights into their evolutionary relationships. *Front. Microbiol.* 3:327. doi: 10.3389/fmicb.2012.00327
- Gushchin, I., Chervakov, P., Kuzmichev, P., Popov, A. N., Round, E., Borshchevskiy, V., et al. (2013). Structural insights into the proton pumping by unusual proteorhodopsin from nonmarine bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12631–12636. doi: 10.1073/pnas.1221629110
- Haft, D. H., Loftus, B. J., Richardson, D. L., Yang, F., Eisen, J. A., Paulsen, I. T., et al. (2001). TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res.* 29, 41–43. doi: 10.1093/nar/29.1.41
- Henrici, A. T., and Johnson, D. E. (1935). Studies of freshwater bacteria II. Stalked bacteria, a new order of schizomycetes I. *J. Bacteriol.* 30, 61–93.
- Herlemann, D. P., Lundin, D., Labrenz, M., Jürgens, K., Zheng, Z., Aspeborg, H., et al. (2013). Metagenomic *de novo* assembly of an aquatic representative of the verrucomicrobial class *Spartobacteria*. *mBio* 4:e00569-12. doi: 10.1128/mBio.00569-12
- Hou, S., Makarova, K. S., Saw, J. H., Senin, P., Ly, B. V., Zhou, Z., et al. (2008). Complete genome sequence of the extremely acidophilic methanotroph isolate V4, *Methylocidiphilum inferorum*, a representative of the bacterial phylum Verrucomicrobia. *Biol. Dir.* 3:26. doi: 10.1186/1745-6150-3-26
- Huang, Y., Gilna, P., and Li, W. (2009). Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* 25, 1338–1340. doi: 10.1093/bioinformatics/btp161
- Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119
- Isanapong, J., Goodwin, L., Bruce, D., Chen, A., Detter, C., Han, J., et al. (2012). High-quality draft genome sequence of the Opitutaceae bacterium strain TAV1, a symbiont of the wood-feeding termite *Reticulitermes flavipes*. *J. Bacteriol.* 194, 2744–2745. doi: 10.1128/JB.00264-12
- Islam, T., Jensen, S., Reigstad, L. J., Larsen, Ø., and Birkeland, N.-K. (2008). Methane oxidation at 55°C and pH 2 by a thermoacidophilic bacterium belonging to the Verrucomicrobia phylum. *Proc. Natl. Acad. Sci. U.S.A.* 105, 300–304. doi: 10.1073/pnas.0704162105
- Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006
- Kang, I., Kim, S., Islam, M. R., and Cho, J.-C. (2017). The first complete genome sequences of the acI lineage, the most abundant freshwater Actinobacteria, obtained by whole-genome-amplification of dilution-to-extinction cultures. *Sci. Rep.* 7:42252. doi: 10.1038/srep42252
- Kant, R., Van Passel, M. W., Palva, A., Lucas, S., Lapidus, A., del Rio, T. G., et al. (2011a). Genome sequence of *Chthoniobacter flavus* Ellin428, an aerobic heterotrophic soil bacterium. *J. Bacteriol.* 193, 2902–2903. doi: 10.1128/JB.00295-11
- Kant, R., Van Passel, M. W., Sangwan, P., Palva, A., Lucas, S., Copeland, A., et al. (2011b). Genome sequence of “*Pedospaera parvula*” Ellin514, an aerobic verrucomicrobial isolate from pasture soil. *J. Bacteriol.* 193, 2900–2901. doi: 10.1128/JB.00299-11
- Katoh, K., Misawa, K., Kuma, K. I., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi: 10.1093/nar/gkf436
- Kers, J. A., Wach, M. J., Krasnoff, S. B., and Widom, J. (2004). Nitration of a peptide phytotoxin by bacterial nitric oxide synthase. *Nature* 429, 79–82. doi: 10.1038/nature02504
- Khadem, A. F., Pol, A., Jetten, M. S., and den Camp, H. J. O. (2010). Nitrogen fixation by the verrucomicrobial methanotroph ‘*Methylocidiphilum fumariolicum*’ SolV. *Microbiology* 156, 1052–1059. doi: 10.1099/mic.0.036061-0
- Khadem, A. F., van Teeseling, M. C., van Niftrik, L., Jetten, M. S., den Camp, H. J. O., and Pol, A. (2012a). Genomic and physiological analysis of carbon storage in the verrucomicrobial methanotroph “*Ca. Methylocidiphilum fumariolicum*” SolV. *Front. Microbiol.* 3:345. doi: 10.3389/fmicb.2012.00345
- Khadem, A. F., Wiczorek, A. S., Pol, A., Vuilleumier, S., Harhangi, H. R., Dunfield, P. F., et al. (2012b). Draft genome sequence of the volcano-inhabiting thermoacidophilic methanotroph *Methylocidiphilum fumariolicum* strain SolV. *J. Bacteriol.* 194, 3729–3730. doi: 10.1128/JB.00501-12
- Kim, S., Oh, D.-B., Kang, H. A., and Kwon, O. (2011). Features and applications of bacterial sialidases. *Appl. Microbiol. Biotechnol.* 91, 1–15. doi: 10.1007/s00253-011-3307-2
- Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* 25, 1–18. doi: 10.18637/jss.v025.i01
- Lee, K.-C., Webb, R. I., Janssen, P. H., Sangwan, P., Romeo, T., Staley, J. T., et al. (2009). Phylum Verrucomicrobia representatives share a compartmentalized cell plan with members of bacterial phylum Planctomycetes. *BMC Microbiol.* 9:5. doi: 10.1186/1471-2180-9-5
- Lin, J. Y., Russell, J. A., Sanders, J. G., and Wertz, J. T. (2016). *Cephalotococcus* gen. nov., a new genus of ‘Verrucomicrobia’ containing two novel species isolated from *Cephalotes* ant guts. *Int. J. Syst. Evol. Microbiol.* 66, 3034–3040. doi: 10.1099/ijsem.0.001141
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178
- Lonhienne, T. G., Sagulenko, E., Webb, R. I., Lee, K.-C., Franke, J., Devos, D. P., et al. (2010). Endocytosis-like protein uptake in the bacterium Gemmata obscuriglobus. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12883–12888. doi: 10.1073/pnas.1001085107
- Lowe, T. M., and Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955–964. doi: 10.1093/nar/25.5.0955
- MacIsaac, E., and Stockner, J. G. (1993). “Enumeration of phototrophic picoplankton by autofluorescence microscopy,” in *Handbook of Methods in Aquatic Microbial Ecology*, eds P. F. Kemp, B. F. Sherr, E. B. Sherr, and J. J. Cole (Boca Raton, FL: Lewis Publishers) 187–197.
- Man, D., Wang, W., Sabehi, G., Aravind, L., Post, A. F., Massana, R., et al. (2003). Diversification and spectral tuning in marine proteorhodopsins. *EMBO J.* 22, 1725–1731. doi: 10.1093/emboj/cdgl83
- Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., et al. (2016). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* 45, D200–D203. doi: 10.1093/nar/gkw1129
- Marie, D., Partensky, F., Jacquet, S., and Vaulot, D. (1997). Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. *Appl. Environ. Microbiol.* 63, 186–193.
- Martinez-Garcia, M., Brazel, D. M., Swan, B. K., Arnosti, C., Chain, P. S., Reitenga, K. G., et al. (2012a). Capturing single cell genomes of active

- polysaccharide degraders: an unexpected contribution of *Verrucomicrobia*. *PLOS ONE* 7:e35314. doi: 10.1371/journal.pone.0035314
- Martinez-Garcia, M., Swan, B. K., Poulton, N. J., Gomez, M. L., Masland, D., Sieracki, M. E., et al. (2012b). High-throughput single-cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton. *ISME J.* 6, 113–123. doi: 10.1038/ismej.2011.84
- McGlynn, S. E., Boyd, E. S., Peters, J. W., and Orphan, V. J. (2013). Classifying the metal dependence of uncharacterized nitrogenases. *Front. Microbiol.* 3:419. doi: 10.3389/fmicb.2012.00419
- McInerney, J. O., Martin, W. F., Koonin, E. V., Allen, J. F., Galperin, M. Y., Lane, N., et al. (2011). Planctomycetes and eukaryotes: a case of analogy not homology. *Bioessays* 33, 810–817. doi: 10.1002/bies.201100045
- Moncla, B., Braham, P., and Hillier, S. (1990). Sialidase (neuraminidase) activity among gram-negative anaerobic and capnophilic bacteria. *J. Clin. Microbiol.* 28, 422–425.
- Nawrocki, E. P., and Eddy, S. R. (2010). *ssu-align: A Tool for Structural Alignment of SSU rRNA Sequences*. Available at: <http://eddylib.org/software/ssu-align/>
- Neuenschwander, S. M., Ghai, R., Pernthaler, J., and Salcher, M. M. (2017). Microdiversification in genome-streamlined ubiquitous freshwater Actinobacteria. *ISME J.* doi: 10.1038/ismej.2017.156 [Epub ahead of print].
- Oh, S., Caro-Quintero, A., Tsementzi, D., DeLeon-Rodriguez, N., Luo, C., Poretsky, R., et al. (2011). Metagenomic insights into the evolution, function, and complexity of the planktonic microbial community of Lake Lanier, a temperate freshwater ecosystem. *Appl. Environ. Microbiol.* 77, 6000–6011. doi: 10.1128/AEM.00107-11
- Okazaki, K., Iino, T., Kuroda, Y., Taguchi, K., Takahashi, H., Ohwada, T., et al. (2014). An assessment of the diversity of culturable bacteria from main root of sugar beet. *Microbes Environ.* 29, 220–223. doi: 10.1264/jsme2.ME13182
- Op den Camp, H. J., Islam, T., Stott, M. B., Harhangi, H. R., Hynes, A., Schouten, S., et al. (2009). Environmental, genomic and taxonomic perspectives on methanotrophic *Verrucomicrobia*. *Environ. Microbiol. Rep.* 1, 293–306. doi: 10.1111/j.1758-2229.2009.00022.x
- Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2013). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Peng, Y., Leung, H. C., Yiu, S.-M., and Chin, F. Y. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi: 10.1093/bioinformatics/bts174
- Pol, A., Heijmans, K., Harhangi, H. R., Tedesco, D., Jetten, M. S., and Den Camp, H. J. O. (2007). Methanotrophy below pH 1 by a new *Verrucomicrobia* species. *Nature* 450, 874–878. doi: 10.1038/nature06222
- Porter, K., and Feig, Y. (1980). The use of DAPI for identification and enumeration of bacteria and blue-green algae. *Limnol. Oceanogr.* 25, 943–948. doi: 10.4319/lo.1980.25.5.0943
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2—approximately maximum-likelihood trees for large alignments. *PLOS ONE* 5:e9490. doi: 10.1371/journal.pone.0009490
- Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. *Trends Genet.* 16, 276–277. doi: 10.1016/S0168-9525(00)02024-2
- Rozen, S., and Skaletsky, H. (1999). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132, 365–386. doi: 10.1385/1-59259-192-2:365
- 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. doi: 10.1128/AEM.70.10.5875-5881.2004
- Santarella-Mellwig, R., Franke, J., Jaedicke, A., Gorjanacz, M., Bauer, U., Budd, A., et al. (2010). The compartmentalized bacteria of the planctomycetes-verrucomicrobia-chlamydiae superphylum have membrane coat-like proteins. *PLOS Biol.* 8:e1000281. doi: 10.1371/journal.pbio.1000281
- Scheuermayer, M., Gulder, T. A., Bringmann, G., and Hentschel, U. (2006). *Rubritalea marina* gen. nov., sp. nov., a marine representative of the phylum ‘*Verrucomicrobia*’, isolated from a sponge (*Porifera*). *Int. J. Syst. Evol. Microbiol.* 56, 2119–2124. doi: 10.1099/ijs.0.64360-0
- Schlesner, H. (1987). *Verrucomicrobium spinosum* gen. nov., sp. nov.: a fimbriated prosthecate bacterium. *Syst. Appl. Microbiol.* 10, 54–56. doi: 10.1016/S0723-2020(87)80010-3
- 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.* 4:2304. doi: 10.1038/ncomms3304
- Sekar, R., Pernthaler, A., Pernthaler, J., Warnecke, F., Posch, T., and Amann, R. (2003). An improved protocol for quantification of freshwater *Actinobacteria* by fluorescence in situ hybridization. *Appl. Environ. Microbiol.* 69, 2928–2935. doi: 10.1128/AEM.69.5.2928-2935.2003
- Shimodaira, H., and Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1114. doi: 10.1093/oxfordjournals.molbev.a026201
- Spring, S., Bunk, B., Spröer, C., Schumann, P., Rohde, M., Tindall, B. J., et al. (2016). Characterization of the first cultured representative of *Verrucomicrobia* subdivision 5 indicates the proposal of a novel phylum. *ISME J.* 10, 2801–2816. doi: 10.1038/ismej.2016.84
- Sukovich, D. J., Seffernick, J. L., Richman, J. E., Gralnick, J. A., and Wackett, L. P. (2010). Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. *Appl. Environ. Microbiol.* 76, 3850–3862. doi: 10.1128/AEM.00436-10
- Tamagnini, P., Axelsson, R., Lindberg, P., Oxelfelt, F., Wünschiers, R., and Lindblad, P. (2002). Hydrogenases and hydrogen metabolism of cyanobacteria. *Microbiol. Mol. Biol. Rev.* 66, 1–20. doi: 10.1128/MMBR.66.1.1-20.2002
- Tang, K., Jiao, N., Liu, K., Zhang, Y., and Li, S. (2012). Distribution and functions of TonB-dependent transporters in marine bacteria and environments: implications for dissolved organic matter utilization. *PLOS ONE* 7:e41204. doi: 10.1371/journal.pone.0041204
- Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., et al. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 29, 22–28. doi: 10.1093/nar/29.1.22
- Toyama, D., Kishi, L. T., Santos-Júnior, C. D., Soares-Costa, A., de Oliveira, T. C. S., de Miranda, F. P., et al. (2016). Metagenomics analysis of microorganisms in freshwater lakes of the Amazon Basin. *Genome Announc.* 4:e01440-16. doi: 10.1128/genomeA.01440-16
- Van Teeseling, M. C., Neumann, S., and Van Niftrik, L. (2013). The anammoxosome organelle is crucial for the energy metabolism of anaerobic ammonium oxidizing bacteria. *J. Mol. Microbiol. Biotechnol.* 23, 104–117. doi: 10.1159/000346547
- van Teeseling, M. C., Pol, A., Harhangi, H. R., van der Zwart, S., Jetten, M. S., den Camp, H. J. O., et al. (2014). Expanding the verrucomicrobial methanotrophic world: description of three novel species of *Methylacidimicrobium* gen. nov. *Appl. Environ. Microbiol.* 80, 6782–6791. doi: 10.1128/AEM.01838-14
- Wagner, M., and Horn, M. (2006). The *Planctomycetes*, *Verrucomicrobia*, *Chlamydiae* and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr. Opin. Biotechnol.* 17, 241–249. doi: 10.1016/j.copbio.2006.05.005
- Wertz, J. T., Kim, E., Breznak, J. A., Schmidt, T. M., and Rodrigues, J. L. (2012). Genomic and physiological characterization of the *Verrucomicrobia* isolate *Diplosphaera colitermitum* gen. nov., sp. nov., reveals microaerophily and nitrogen fixation genes. *Appl. Environ. Microbiol.* 78, 1544–1555. doi: 10.1128/AEM.06466-11
- Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40, W445–W451. doi: 10.1093/nar/gks479
- Yoon, J., Matsuo, Y., Katsuta, A., Jang, J.-H., Matsuda, S., Adachi, K., et al. (2008). *Haloferula rosea* gen. nov., sp. nov., *Haloferula harenae* sp. nov., *Haloferula phyci* sp. nov., *Haloferula helveola* sp. nov. and *Haloferula sargassicola* sp. nov., five marine representatives of the family *Verrucomicrobiaceae* within the phylum ‘*Verrucomicrobia*’. *Int. J. Syst. Evol. Microbiol.* 58, 2491–2500. doi: 10.1099/ijs.0.2008/000711-0



- Yoon, J., Yasumoto-Hirose, M., Katsuta, A., Sekiguchi, H., Matsuda, S., Kasai, H., et al. (2007). *Coralimargarita akajimensis* gen. nov., sp. nov., a novel member of the phylum 'Verrucomicrobia' isolated from seawater in Japan. *Int. J. Syst. Evol. Microbiol.* 57, 959–963. doi: 10.1099/ijs.0.64755-0
- Zaremba-Niedzwiedzka, K., Viklund, J., Zhao, W., Ast, J., Sczyrba, A., Woyke, T., et al. (2013). Single-cell genomics reveal low recombination frequencies in freshwater bacteria of the SAR11 clade. *Genome Biol.* 14:R130. doi: 10.1186/gb-2013-14-11-r130
- Zhang, J., Zhang, X., Liu, Y., Xie, S., and Liu, Y. (2014). Bacterioplankton communities in a high-altitude freshwater wetland. *Ann. Microbiol.* 64, 1405–1411. doi: 10.1007/s13213-013-0785-8

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Cabello-Yeves, Ghai, Mehrshad, Picazo, Camacho and Rodriguez-Valera. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.