RESEARCH ARTICLE

In Silico Analysis of the Metabolic Potential and Niche Specialization of Candidate Phylum "Latescibacteria" (WS3)

Noha H. Youssef1,*, Ibrahim F. Farag1, Christian Rinke2, Steven J. Hallam3,4, Tanja Woyke2, Mostafa S. Elshahed1

1 Department of Microbiology and Molecular Genetics, Oklahoma State University Stillwater, Oklahoma, United States of America, 2 DOE Joint Genome Institute, Walnut Creek, California, United States of America, 3 University of British Columbia, Vancouver, BC, Canada, 4 Graduate Program in Bioinformatics, University of British Columbia, Vancouver, BC, Canada

* Noha@Okstate.edu

Abstract

The “Latescibacteria” (formerly WS3), member of the Fibrobacteres–Chlorobi–Bacteroidetes (FCB) superphylum, represents a ubiquitous candidate phylum found in terrestrial, aquatic, and marine ecosystems. Recently, single-cell amplified genomes (SAGs) representing the “Latescibacteria” were obtained from the anoxic monimolimnion layers of Saki-naw Lake (British Columbia, Canada), and anoxic sediments of a coastal lagoon (Etoliko lagoon, Western Greece). Here, we present a detailed in-silico analysis of the four SAGs to gain some insights on their metabolic potential and apparent ecological roles. Metabolic reconstruction suggests an anaerobic fermentative mode of metabolism, as well as the capability to degrade multiple polysaccharides and glycoproteins that represent integral components of green (Charophyta and Chlorophyta) and brown (Phaeophyceae) algae cell walls (pectin, alginate, ulvan, fucan, hydroxyproline-rich glycoproteins), storage molecules (starch and trehalose), and extracellular polymeric substances (EPSs). The analyzed SAGs also encode dedicated transporters for the uptake of produced sugars and amino acids/oligopeptides, as well as an extensive machinery for the catabolism of all transported sugars, including the production of a bacterial microcompartment (BMC) to sequester propionaldehyde, a toxic intermediate produced during fucose and rhamnose metabolism. Finally, genes for the formation of gas vesicles, flagella, type IV pili, and oxidative stress response were found, features that could aid in cellular association with algal detritus. Collectively, these results indicate that the analyzed “Latescibacteria” mediate the turnover of multiple complex organic polymers of algal origin that reach deeper anoxic/microoxic habitats in lakes and lagoons. The implications of such process on our understanding of niche specialization in microbial communities mediating organic carbon turnover in stratified water bodies are discussed.
Introduction

Over the past few decades, small subunit ribosomal RNA (SSU or 16S rRNA) gene-based surveys have prompted a drastic reevaluation of the scope of phylum level diversity within the domain Bacteria. Current taxonomic outlines indicate that the majority of recognized bacterial phyla (54.1% using SILVA database [1], 65.48% using Greengenes database [2]) have no pure culture representatives (candidate phyla). Many of these candidate phyla, so-called microbial dark matter (MDM) are globally distributed and display significant levels of intra-phylum level diversity [3–7]. Recent advances in cell sorting and whole genome amplification and assembly have facilitated the acquisition of single-cell amplified genomes (SAGs) derived from numerous candidate phyla [8–16]. Metabolic reconstruction with these SAGs provides a unique opportunity to uncover the ecological and biogeochemical roles played by these enigmatic microbial groups.

One such candidate phylum is WS3 (Wurtsmith aquifer Sequences-3), whose members were first identified in a 16S rRNA gene-based survey of anoxic sediments obtained from a hydrocarbon- and chlorinated-solvents-contaminated aquifer in northern Michigan, USA in 1998 [17]. Since then, their presence has been documented across a wide range of habitats including marine hydrothermal vents, gas hydrate-bearing habitats, deep sea hypersaline anoxic lakes, and oil-exposed microbial mats [18–28]. Recently, as part of an extensive single cell genomic study of 9 different habitats, Rinke et al. [29] reported on the recovery of four SAGs from WS3 single cells. Phylogenomic-based analysis using conserved marker genes indicated the monophyletic nature of WS3 as part of the Fibrobacteres–Chlorobi–Bacteroidetes (FCB) superphylum together with “Marinimicrobia” (SAR406), “Cloacimonetes” (WWE1), Gemmatimonadetes, and Caldithrix. The name “Latescibacteria” (hiding small rods) was suggested for the candidate phylum.

However, little is known about the biological capabilities of this phylum, and no systematic attempts have been made to reconstruct its metabolic potential. Thus, we here present a detailed analysis of the metabolic and physiological capabilities, and putative ecological roles of four “Latescibacteria” SAGs obtained from two different aquatic environments. Our analysis suggests that the “Latescibacteria” recovered from Sakinaw Lake and Etoliko lagoon transform algal detritus sinking from sunlit surface waters into fermentation products with the potential to contribute to microbial food webs in anaerobic waters below.

Materials and Methods

Origin of “Latescibacteria” SAGs

“Latescibacteria” SAGs analyzed in this study were obtained from two different locations [29]: Three SAGs originated from a single sample obtained from the anaerobic monimolimnion of Sakinaw lake (British Columbia, Canada) at 49°40’30”N, 124°2’2.4”W coordinates, and a depth of 120m (Gies et al 2014). A fourth SAG was obtained by sampling anaerobic sediments in Etoliko Lagoon, a coastal lagoon in the south of Aetolia-Acarnania, Greece, at the deepest point (~27.5 m) at 38°28’59.54”N, 21°19’17.44”E. Single cell sorting and lysis, whole genome amplification, identification via 16S rRNA gene sequencing of amplified genomes, as well as SAG sequencing, assemblies and estimates of genome completion were previously described [29]. The four “Latescibacteria” SAGs were deposited under Genbank assembly IDs: NZ_ASMB0000000.1, NZ_AQSL0000000.1, ASWY0000000.1, and AQRO0000000.1, and in Integrated Microbial Genomics (IMG) under SAG IDs: SCGC AAA252-D10, SCGC AAA252-B13 and SCGC AAA252-E07 for Sakinaw lake SAGs, and SCGC AAA257-K07 for
the Etoliko lagoon SAG. These SAGs will henceforth be referred to as S-D10, S-B13, and S-E07 for Sakinaw Lake SAGs, and E-K07 for Etoliko Lagoon SAG. The type species for "Latescibacteria" is S-E07, for which the name Candidatus "Latescibacter anaerobius" has been proposed [29].

Detailed analysis was conducted on S-E07, which has the highest estimated genome completion (73.02%) among the "Latescibacteria" SAGs. The closely related S-B13 (94% 16S rRNA gene sequence similarity to SAG S-E07) with 57.1% estimated genome completion was used to confirm shared gene content and fill pathway holes when needed. Only general metabolic features for SAG S-D10 (94% 16S rRNA gene sequence similarity to S-E07, and 96% to S-B13) are discussed, given its low percentage of estimated genome completeness (38.2%). Due to the observed differences between the 3 Sakinaw Lake SAGs, and the Etoliko lagoon SAG E-K07 (85–86% 16S rRNA gene sequence similarity to Sakinaw Lake SAGs), as well as its low estimated genome completion (23.02%), analysis of SAG E-K07 was restricted to identifying variation in conserved genes or pathways between "Latescibacteria" SAGs from two distinct locations.

Genome annotation, general genomic features, and metabolic reconstruction

The IMG platform (http://img.jgi.doe.gov) was used for genome functional annotation. Detailed metabolic reconstruction of relevant pathways was performed with both KEGG [30] and Metacyc [31] databases. As part of the IMG annotation pipeline, CRISPR elements are detected with CRT [32] and PILERCR [33]. Predictions from both methods are concatenated and in case of overlapping elements, the shorter one is removed. Overall annotation followed procedures outlined in [13]: In brief, proteases, peptidases, and protease inhibitors were identified with Blastp against the Merops database [34]. Transporters were identified with the transporter classification database (TCDB) [35]. dbCAN HMMs [36] were used to identify carbohydrate active enzymes (CAZymes) including glycoside hydrolases (GH), polysaccharide lyases (PL), and carboxyl esterases (CE) following the classification scheme of the Carbohydrate active enzyme (CAZy) database [37].

Results

Phylogenetic affiliation and general genomic features of "Latescibacteria" SAGs

All four SAGs were affiliated with the candidate order PBS_III_9 based on phylogenetic analysis of the candidate phylum "Latescibacteria" using 1198 near-full length 16S rRNA gene sequences (Fig 1, Table A in S1 File, and Supplementary Text in S1 File). Sakinaw lake SAGs belonged to family I, while Etoliko lagoon SAG E-K07 belonged to family VI within this order (Fig 1B). General genomic features for each SAG are shown in Table 1.

Metabolic potential of Sakinaw Lake SAGs

Anabolic pathways identified in S-E07 and S-B13 include machinery for the production of amino acids, cofactors, fatty acids, purines and pyrimidines, terpenoid unit backbone, and glycerophospholipids. In addition, the SAGs encode near-complete replication, transcriptional, and translational machineries. The presence of genes for lipopolysaccharide (LPS) biosynthesis and pathway for LPS insertion in the outer membrane suggests a Gram-negative cell wall (Supplementary Text).

Catabolical pathways identified in S-E07 and S-B13 indicate a heterotrophic lifestyle. Moreover, the apparent absence of a respiratory chain suggests sole dependence on fermentative
pathways and substrate level phosphorylation for coupled energy release and conservation. Both S-E07 and S-B13 encode a diverse array of carbohydrate active enzymes (CAZymes), with a conspicuous enrichment (Genes/Mbp), and diversity (number of different families) of polysaccharide lyases (PLs) (Fig 2, Figure A in S1 File, Table B in S1 File). In contrast, the SAGs are relatively depauperate in genes encoding glycoside hydrolases (GHs) including enzymes involved in the degradation of cellulose (1 putative endoglucanase (GH5), 3 putative β-glucosidases (GH3, GH116, and GH9), and no putative cellobiohydrolase), and enzymes involved in the degradation of xylans (xylanases, and β-xylosidases).

Table 1. General genomic features of “Latescibacteria” SAGs.

|                      | SCGC AAA252-E07 | SCGC AAA252-B13 | SCGC AAA252-D10 | SCGC AAA257-K07 |
|----------------------|-----------------|-----------------|-----------------|-----------------|
| Genome size, Mb      | 2.3             | 1.49            | 0.5             | 1.77            |
| Estimated size, %    | 73.02           | 57.09           | 38.17           | 23.02           |
| Estimated size, Mb    | 3.15            | 2.61            | 1.31            | 7.69            |
| GC %                 | 42.07           | 40.86           | 40.86           | 42.12           |
| RNA genes            |                 |                 |                 |                 |
| 5S rRNA Count        | 1               | 1               | 1               | 1               |
| 16S rRNA Count       | 1               | 0*              | 1               | 1               |
| 23S rRNA Count       | 1               | 1               | 1               | 1               |
| tRNA Count           | 27              | 18              | 10              | 19              |
| Number of CDS        | 1951            | 1558            | 647             | 1534            |
| with function prediction | 1451          | 1158            | 433             | 1073            |
| without function prediction    | 500            | 400             | 214             | 461             |

* The S-B13 16S rRNA couldn’t be retrieved via the whole genome shotgun approach, however the affiliation of S-B13 to CP-“Latescibacteria” was confirmed through analyzing the amplified and Sanger-sequenced full-length 16S rRNA gene
Interestingly, many of the polymers that S-E07 and S-B13 are predicted to degrade are integral components of cell walls of the green algal phyla Charophyta (most commonly encountered in freshwater habitats), and Chlorophyta (widely distributed in freshwater, marine, and terrestrial habitats), as well as the brown algal Class Phaeophyceae. Green and brown algal cell walls are complex, with a diverse array of structural fibrillar polymers enmeshed in complex matrices with crystalline polymer components (Fig 3). Both S-E07 and S-B13 encode genes necessary for the conversion of these cell wall components, including pectin, alginate, ulvans, fucans, hydroxyproline-rich glycoproteins (HRGP), e.g. arabinogalactan proteins (AGP) and extensins, and xyloglucan (Table 2, Figure B in S1 File, Supplementary Text in S1 File). Moreover, the SAGs also encode pathways mediating the conversion of soluble organic compounds commonly utilized for storage in algae (e.g. starch and trehalose). A more in depth description of these capabilities follows.

Algal cell wall degradation potential

1. Pectins. Pectins are components of the amorphous matrix and outer lattice of Charophyta cell wall (Fig 3A) [38]. Both S-E07 and S-B13 encode machinery for depolymerizing the pectic polysaccharide homogalacturonan (HG) (Table 2). They encode carboxyl esterases (CE8...
and CE12) for the removal of accessory acetyl and methyl groups attached to the backbone, pectin lyase and pectate lyase (PL1, and PL10) to breakdown the backbone to oligosaccharides with 4-deoxy-α-D-galact-4-enuronosyl groups at their non-reducing ends, exopolygalacturonate lyase (PL9) to cleave digalacturonate unit, and oligogalacturonide lyase (PL22) to degrade the digalacturonate units to 5-dehydro-4-deoxy-D-glucuronate and galacturonic acid as the final end products of HG degradation [39, 40]. In addition to HG, S-E07 and S-B13 encode all
Table 2. Polymers potentially targeted by Latescibacteria, their distribution and occurrence in algae, structure, degradation enzymes encoded in the Latescibacterial SAGs, potential degradation products, their transport systems encoded in the SAGs, and ultimate central catabolic pathway.

| Polymer                    | Distribution                      | Degradation                  | Products                                                                 | Transport system | Central pathway |
|----------------------------|-----------------------------------|------------------------------|--------------------------------------------------------------------------|------------------|-----------------|
| **1. Pectin**              |                                   |                              |                                                                          |                  |                 |
| Homogalacturonan           | Land plants and Charophyta green algae | Pectin methyl esterase | Pectinate/ Pectate (demethylated)                                        | EMP              |                 |
|                            |                                   | Pectin acetyl esterase | Deacetylated polymer                                                      |                  |                 |
|                            |                                   | Pectin lyase, pectate lyase | Oligosaccharides with 4-deoxy-α-D-galact-4-enuronosyl groups at their non-reducing ends |                  |                 |
|                            |                                   | Exopolygalacturonate lyase | digalacturonate                                                            |                  |                 |
|                            |                                   | Oligogalacturonide lyase | 5-dehydro-4-deoxy-D-glucuronate                                           | ExuT symporter  |                 |
| Rhamnogalacturonan I       |                                   | Pectin methyl esterase | Demethylated RGI                                                           |                  |                 |
|                            |                                   | Pectin acetyl esterase | Deacetylated RGI                                                           |                  |                 |
|                            |                                   | Rhamnogalacturonan endolyase | Oligosaccharides with L-rhamnopyranose at the reducing end and 4-deoxy-4,5-unsaturated D-galactopyranosyl uronic acid at the non-reducing end |                  |                 |
|                            |                                   | Rhamnogalacturonan exolyase | disaccharide 2-O-(4-deoxy-beta-L-threo-hex-4-enopyranuronosyl)-alpha-L-rhamnopyranose |                  |                 |
|                            |                                   | d-4,5-unsaturated β-glucuronyl hydrolase | rhamnose                                                                 | Rhamnose:proton symporter | BMC            |
|                            |                                   | β-galactosidase | 5-dehydro-4-deoxy-D-glucuronate                                           | ExuT symporter  | EMP             |
|                            |                                   | Galactose          | Galactose: sodium symporter                                               |                  |                 |
| Alginates                  | Brown Algae                       | Poly β-D-mannuronate lyase | Oligosaccharides with 4-deoxy-α-L-erythro-hex-4-enopyranuronosyl groups at their non-reducing ends |                  |                 |
|                            |                                   | Oligoalginate lyase | 4-deoxy-α-L-erythro-hex-4-enopyranuronosyl spontaneously converted to 5-dehydro-4-deoxy-D-glucuronate | ExuT symporter  | EMP             |
| **2. Fucans**              |                                   |                             |                                                                          |                  |                 |
| Homofucan                  | Brown Algae                       | Sulfatase                  | Unsulfated homofucans                                                     |                  |                 |
|                            |                                   | Fucoidan lyase-like protein (hypothetical protein) | Unsaturated, non-sulfated di- and tetrasaccharides                         |                  |                 |
|                            |                                   | α-L-fucosidase | Fucose                                                                     | Fucose:proton symporter | BMC            |
| Xylofucogalactan/ xylofucomannan |                                   | Sulfatase                  | Unsulfated fucans                                                         |                  |                 |
|                            |                                   | α-L-fucosidase | Fucose                                                                     | Fucose:proton symporter | BMC            |
|                            |                                   | β-glucuronidase | Glucuronic acid                                                           | ExuT symporter  | EMP             |
| Xylofucoglucuronan         |                                   | Sulfatase                  | Unsulfated fucans                                                         |                  |                 |
|                            |                                   | α-L-fucosidase | Fucose                                                                     | Fucose:proton symporter | BMC            |

(Continued)
### Table 2. (Continued)

| Polymer                     | Distribution                          | Degradation                  | Products                                      | Transport system               | Central pathway |
|-----------------------------|---------------------------------------|------------------------------|-----------------------------------------------|--------------------------------|-----------------|
| **3. Ulvans**               | Chlorophyta                           | Sulfatase                    | Unsulfated ulvans                             |                                |                 |
|                             |                                       | Heparin lyase                | Unsaturated, non-sulfated di- and             |                                |                 |
|                             |                                       |                              | tetrasaccharides                              |                                |                 |
|                             |                                       | d-4,5-unsaturated β-         | 5-dehydro-4-deoxy-D-glucuronate              | ExuT symporter                 | EMP             |
|                             |                                       | glucuronyl hydrolase         |                                               |                                |                 |
|                             |                                       |                              | Rhamnose                                      | Rhamnose:proton symporter      | BMC             |
|                             |                                       |                              | Xylose                                        | Xylose:proton symporter        | PPP             |
| **4. Xyloglucan**           | Land plants, some green algae         | endo-β-1,4-glucanase         | A mixture of oligosaccharides                 |                                |                 |
|                             |                                       | α-1,2-fucosidase             | Fucose                                        | Fucose:proton symporter        | BMC             |
|                             |                                       | β-galactosidase              | Galactose                                     | Galactose:sodium symporter     | EMP             |
|                             |                                       | β-glucosidase                | Glucose                                       | Glucose:sodium symporter      | EMP             |
| **5. Hydroxyproline-rich glycoprotein (HRGP)** |                       |                              |                                               |                                |                 |
| **Extensin**                | Land plants, some green algae         | β-L-arabinofuranosidase      | Arabinose                                      | ABC transporter                 | PPP             |
| **Arabinoagelactan protein (AGP)** | Land plants, some green algae         | endo-β-1,6-galactanase (T)   | galactan oligosaccharides                     |                                |                 |
|                             |                                       | β-glucuronidase              | Glucuronic acid                               |                                |                 |
|                             |                                       | α-Fucosidase                 | Fucose                                        | Fucose:proton symporter        | BMC             |
|                             |                                       | α-rhamnosidase               | Rhamnose                                      | Rhamnose:proton symporter      | BMC             |
|                             |                                       | β-galactosidase              | Galactose                                     | Galactose:sodium symporter     | EMP             |
| **6. Others**               | All organisms                         | Non-specific endopeptidases  | Oligopeptides                                 | ABC transporter                 |                 |
|                             |                                       | Dipeptidases                 | Dipeptides                                    | ABC transporter                 |                 |
|                             |                                       | Dipeptide peptidase, aminopeptidases, or Carboxypeptidases | Free amino acids | Symporters for Pro, Ala, Asp, Glu, Gly, cationic aaABC transporter for Pro | EMP (Asp and Glu) |
| **Starch**                  | Land plants, and green algae storage compounds | α-amylase                    | Oligosaccharides                              |                                |                 |
| **Trehalose**               | Brown algae storage compound           | α-glucosidase                | Glucose                                       | Glucose:sodium symporter       | EMP             |
| **Poly-D-galactosamine**    | Some fungi such as Aspergillus, and Neurocrassa | Endo1,4-poly-D- galactosaminidase | Galactosamine                                  | PTS                            | EMP             |

doi:10.1371/journal.pone.0127499.t002
the necessary machinery to degrade rhamnogalacturonan I (RGI) (Table 2). These include carboxyl esterases (CE8 and CE12), rhamnogalacturonan endolyase (PL11) that attack the backbone to produce oligosaccharides with L-rhamnopyranose at the reducing end and 4-deoxy-4,5-unsaturated D-galactopyranosyl uronic acid at the non-reducing end, rhamnogalacturonan exolyase (PL11) that attacks those oligosaccharides to release the disaccharide 2-O-(4-deoxy-beta-L-threo-hex-4-enopyranuronosyl)-alpha-L-rhamnopyranose from the reducing end, and d-4,5-unsaturated beta-glucuronyl hydrolase (GH88) that degrades those disaccharides to rhamnose and 5-dehydro-4-deoxy-D-glucuronate. The SAGs also encode beta-galactosidase (GH42) for removal of galactosyl sugar substitutions [39, 40].

2. Alginate. Algicates are present in the brown algal cell walls enmeshing fibrillar cellulose and also in the interfibrillar layers with fucans (Fig 3C) [41]. Both S-E07 and S-B13 encode PLs for the complete degradation of alginate (Table 2). These PLs include alginate lyases (PL6, PL15, PL17) that break down the alginate backbone producing oligosaccharides with 4-deoxy-\(\alpha\)-L-erythro-hex-4-enopyranuronosyl groups at their non-reducing ends, as well as oligoalginate lyase (PL15, and PL17) that exolytically cleave these oligosaccharides into monosaccharides and releases 4-deoxy-\(\alpha\)-L-erythro-hex-4-enopyranuronose from the non-reducing end. The produced 4-deoxy-\(\alpha\)-L-erythro-hex-4-enopyranuronose is spontaneously converted into 5-dehydro-4-deoxy-D-glucuronate as the final end product of alginate degradation [42].

3. Fucans. In addition to pectin and alginate, S-E07, and S-B13 also encode machinery for fucan degradation. Fucans are present, together with alginates, in brown algal cell walls interfibrillar matrix (Fig 3C) [41]. Fucans exhibit wide variations in chemical structures, ranging from the highly sulfated homofucan polymers to the highly branched high-uronic-acid, low-sulfate-containing polymers (xylofucoglucan, xylofucogalactan, xylofucomannan, xylofucoglucuronan) [41]. However, mechanistic details on the degradation of fucans are still in their infancy. Genomic analysis of “Latescibacteria” SAGs that S-E07, and S-B13 have the capacity to transform several fucans including homofucans, sulfated-xylofucoglucan, and sulfated-xylofucoglucuronan. Indeed, a potential homofucan-degrading enzyme with sequence similarity to Mariniflexible fucanivorans fucoidan lyase could attack the backbone releasing unsaturated, non-sulfated fucan di- and tetrasaccharides. The SAGs also encode many \(\alpha\)-fucosidases (GH29, and GH95), that could attack those oligosaccharides and release fucosyl residues from the reducing end. Genomic evidence for the degradation of the highly branched high-uronic-acid, low-sulfate-containing polymers include many \(\alpha\)-fucosidases (GH29, and GH95), and one \(\alpha\)-glucuronidase (GH67).

4. Ulvans. Ulvans are present in the amorphous interfibrillar matrix of Chlorophyta cell walls (Fig 3B) [43–45]. Ulvan backbones are made of a few repeating disaccharides (Supplementary Text). However, the exact composition of ulvans is largely unknown. One important characteristic of ulvans is the presence of unusual sugars, e.g. iduronic acid, in its backbone [44]. Iduronic acid is also an important constituent of mammalian glycosaminoglycans (GAGs), e.g. heparan sulfate, dermatan sulfate, heparin [46]. “Latescibacteria” SAGs harbor several PLs annotated as heparin and heparan lyase (PL12 and PL21). Structural similarity in sugar composition between ulvans and mammalian GAGs such as heparin suggest that those polysaccharide lyases (annotated as PL12 and PL21 with heparinase activity) might be potential ulvan lyases responsible for ulvan backbone cleavage to produce di- and tetrasaccharides with an unsaturated beta-glucuronyl residue located at the non-reducing end [47]. SAGs also harbor several copies of unsaturated glucuronyl hydrolases (GH88) that could potentially act on the oligosaccharides produced and release 5-dehydro-4-deoxy-D-glucuronate and other sugar residues, e.g. rhamnose, and xylose, as end products.

5. Xyloglucan. Xyloglucan is a component of Charophyta and Chlorophyta cell wall usually present in association with cellulose microfibrils (Fig 3A and 3B) [48–50]. Both S-E07 and
S-B13 encode machinery to degrade xyloglucan, a component of Charophyta and Chlorophyta cell walls usually present in association with cellulose microfibrils (Fig 3A and 3B) [48–50], including endo-β-1,4-glucanases (GH74), that cleave the xyloglucan backbone at locations of unsubstituted glycosyl moieties and give rise to a mixture of oligosaccharides, α-1,2-fucosidase (GH95), and β-galactosidases (GH2, GH42) that attack those oligosaccharides to give rise to XXXG xyloglucans. The latter oligosaccharide can be attacked by oligoxyloglucan β-glycosidase (GH3) generating isoprimeverose (Xyl-α(1,6)-Glu), and glucose. However, no homologs of oligoxyloglucan β-glycosidase were identified.

6. Hydroxyproline-rich, other O-linked, and N-linked glycoproteins. Hydroxyproline-rich glycoproteins (HRGP) are minor components in green algal cell walls (Fig 3) [51, 52]. Both S-E07, and S-B13 SAGs encode β-L-arabinofuranosidase (GH127) that specifically targets arabinose residues attached to hydroxyproline in extensins [53] and release the sugar monomer arabinose. The SAGs also encode machinery for arabinogalactan protein (AGP) degradation including endo-β-1,6-galactanases (GH30) that hydrolyses the β-1,6-galactan side chains and gives rise to galactan oligosaccharides, β-galactosidases (GH2, GH42), β-glucuronidase (GH79), α-fucosidase (GH29, GH95), and α-rhamnosidase (GH28, GH78, GH106) that attack the produced oligosaccharides and release substituting sugar monomers, e.g. galactose, glucuronic acid, fucose, and rhamnose [54]. In addition to HRGP degradation potential, the SAGs encode several α-N-acetylgalactosaminidases (GH109) that specifically release N-acetylgalactosaminyl residues from O-linked glycoproteins [55], as well as several α-mannosidases (GH38) that could potentially release mannosyl residues from N-linked glycoproteins [56]. Recently, sialic acid (neuraminic acid), a 9-carbon sugar acid was identified in green algal N-linked glycoproteins [57]. While a sialidase (GH33) homologue was not identified in the SAGs, they do encode for all the enzymes required for sialic acid degradation, including sialate O-acetylersterase, N-acetylneuraminic lyase, and N-acyl-D-glucosamine 2-epimerase that will collectively degrade sialic acid into pyruvate and N-acetyl-glucosamine (NAG).

7. Degradation of cell wall proteins. Both S-E07 and S-B13 SAGs encode multiple peptidases that can attack the peptide moiety of glycoproteins in algal cell walls (Table C in S1 File). The majority of these peptidases (∼66% in S-E07, and 63.4% in S-B13) are thought to be nutritional, where they non-specifically break down proteins into oligopeptides (protease families C25, M06, M10, M20, M41, M48, M50, S01, S08, S09, S41, S54, and U62), dipeptides (protease family M19), and free amino acids (protease families M24, M28, S49, T03).

8. Sulfatase activity on sulfated polysaccharide. Both S-E07 and S-B13 encode multiple sulfatases (n = 14 in S-E07 and n = 3 in S-B13) belonging to the family of arylsulfatases (pfam 00884). Many of the polymers in marine algal cell walls are sulfated, e.g. ulvans, homofucans, sulfated-xylofucoglucan, and sulfated-xylofucogluconan [58]. Removal of the sulfate groups from such polysaccharides prior to their degradation facilitates access of GHs and PLs to side chains and backbones [59]. The SAGs also harbor the essential anaerobic sulfatase maturation enzyme-coding gene [60] for post-translational modification of a critical Cys or Ser in the active site to a C-α-formylglycine [61].

Degradation of algal storage compounds and additional polymers of non-algal origin

In addition to algal cell wall components, both S-E07 and S-B13 encode GHs that could potentially target algal intracellular carbon storage compounds, or secreted polysaccharides sourced from other organisms. The SAGs encode GHs specific for starch (α-amylase belonging to GH119, GH57, GH13, and α-glucosidase belonging to GH97), as well as for trehalose (trehalase/maltase belonging to GH65) degradation. Starch is recognized as an important
intracellular storage compound in green algae and green plants [62], while trehalose is an intracellular storage compound in brown algae [41]. In addition, S-E07 and S-B13 encode β-fructofuranosidase (GH32) specific for sucrose, and endo1,4-poly-D-galactosaminidase (GH114) specific for poly-D-galactosamine (Table 2).

**Extracellular polymeric substance (EPS) as additional potential source of energy for the “Latescibacteria”**

EPS forms extensive mucilaginous sheath outside the algal cell wall and function in adhesion, gliding motility, biofilm formation, and protection. Although the exact chemistry of EPS is not entirely known, EPS was shown to be composed mainly of polysaccharides (up to 75%), with minor protein content (2–10%). The polysaccharide fraction is rich in uronic acids, as well as monosaccharides, mainly glucose, galactose, mannose, xylose, arabinose, fucose, and rhamnose [63, 64]. As mentioned above, “Latescibacteria” SAGs harbor genes involved in the uptake and catabolism of all such components.

**“Latescibacteria” SAGs harbor extensive transport systems for sugars, and amino acids/oligopeptides uptake**

Both S-E07 and S-B13 encode several non-specific porins for transport of substrates across the outer membrane, coupled to specialized transporters in the inner membrane, including multiple secondary (symport), ABC (ATP-binding cassette), and phosphotransferase system (PTS) transporters for the uptake of a wide array of monomers, e.g. those putatively produced from the degradation of all polymers described above (Fig 4, Table 2). Uronic acids and uronic acid derivatives are potentially imported using a single common transporter (a sugar phosphate permease transporter of the major facilitator superfamily similar to ExuT transporter of *Ralstonia solanacearum* [65]). Fucose, rhamnose, as well as xylose are potentially imported via dedicated proton symporters, while glucose and galactose are potentially imported via dedicated sodium symporters. Moreover, the SAGs encode components of dedicated ABC transporters for arabinose, ribose, and oligopeptides and dipeptides as well as components of the PTS specific for N-acetylgalactosamine, fructose, and mannose import. The SAGs also encode a complete two-component signal transduction system for sensing di/tricarboxylates, e.g. malate, citrate, (DctBD), as well as a tripartite ATP-independent di/tricarboxylate transport system (TRAP) (DctPQM) [66].

**Catabolism of imported sugars**

Both S-E07 and S-B13 encode extensive pathways for the catabolism of a wide array of sugars, sugar acids, amino sugars, amino acids, as well as citrate and malate. Monomer degradation pathways in the SAGs are predicted to converge on one of three central metabolic routes, (i) feeding into the EMP pathway (for glucose, galactose, mannose, fructose, sugar acids, amino sugars, aspartate, and citrate and malate), (ii) feeding into PPP (for xylose, ribose, and arabinose), or (iii) the special fucose and rhamnose degradation pathways to propionate and propanol.

Monomer catabolism is detailed in the supplementary text and Fig 5. Briefly, the genomes encode a complete glycolytic pathway for metabolism of various C6 sugars to pyruvate, including glucose, galactose, mannose, and fructose and the amino sugars N-acetylgalactosamine, N-acetylglucosamine, and D-galactosamine. The genomes also encode the necessary enzymes for channeling the C6 sugar acids galacturonic acid, glucuronic acid, and 5-dehydro-4-deoxy-D-glucuronate to the central metabolite 2-dehydro-3-deoxy-D-gluconate (KDG), which is...
subsequently converted to pyruvate and glyceraldehyde-3-phosphate (GAP), that feed into the EMP. In addition, the amino acid aspartate, as well as dicarboxylates (malate) and tricarboxylates (citrate) that could potentially serve as C and energy source are catabolized via conversion to oxaloacetate and subsequently to phosphoenolpyruvate (PEP). On the other hand, the C5 sugars xylose, ribose, and arabinose are metabolized via the non-oxidative branch of the pentose phosphate pathway by first conversion to xylulose-5-P. Collectively, the metabolism of these compounds via the EMP or the PPP results in the production of pyruvate. Pyruvate could potentially be converted to acetyl-CoA via the action of pyruvate:ferredoxin oxidoreductase. Indeed, as indicated previously, the SAGs encode the machinery necessary for substrate-level phosphorylation including acetyl CoA synthase, as well as propanediol transacetylase and acetate kinase, both of which convert acetyl-CoA to acetate with concomitant ATP production (Fig 5).

Fucose and rhamnose metabolism requires a different catabolic pathway and partially occurs in an intracellular bacterial microcompartment (BMC) to protect against cellular damage.
by containing the reactive intermediate propionaldehyde [67, 68]. Both S-E07 and S-B13 encode a dedicated pathway for the degradation of fucose and rhamnose to lactaldehyde and dihydroxyacetone-phosphate. Several genes encoding for BMC structural shell proteins with BMC domains (pfam 00936, as well as pfam 03319) were identified in the SAGs consistent with a recent observation by Axen and colleagues exploring the taxonomic distribution of BMCs across bacterial phyla [69]. Inside the BMC, lactaldehyde is converted to 1,2-propanediol (1, 2-PD). Although homologues for 1,2-PD dehydratase, the enzyme responsible for conversion of 1,2-PD to propionaldehyde, were not identified in S-E07 and S-B13, both SAGs harbor NAD-dependent aldehyde dehydrogenase, and NADH-dependent alcohol dehydrogenase for conversion of propionaldehyde to propionyl-CoA, and propanol, respectively. Propionyl-CoA can then be converted to propionate with the concomitant production of 1 mole of ATP per propionate produced.
Additional genomic features

Both S-E07 and S-B13 encode machinery for pili and flagella production, enabling potential attachment to surfaces [70], as well as gas vesicles production for maintaining a position in the water column with the most favorable growth conditions [71] (Supplementary Text). In addition, the SAGs encode multiple oxidative stress enzymes that counter harmful effects of changing oxygen tension caused by vertical migration in the stratified water column while in pursuit of decaying algal cells or other food particles. These include ruberythrin, rubredoxin, rubredoxin oxidoreductase, superoxide reductase (desulfoferredoxin), ferritin-like protein, NADPH-dependent alkyl hydperoxide reductase, and glutathione peroxidase [72], as well as machinery for bacillithiol biosynthesis, a thiol implicated in peroxide sensing [72–75].

General features of Etoliko lagoon SAG E-K07

While the Etoliko lagoon SAG E-K07 shared similar metabolic potential with respect to algal cell wall polymer degradation to the Sakinaw Lake SAGs several unique features were apparent. In addition to harboring a large genome (estimated size 7.7 Mbp, Table 1) E-K07 encodes machinery for the following: (1) Degradation of the amino acids Thr, D-Cys, Glu, and Met, (2) Neuraminidase (GH33) gene for cleavage of sialic acid residues from N-linked glycoproteins, and endo-β-1,4-glucuronan lyase (PL20) [76], that targets β-(1→4)-glucuronan, a minor polysaccharide present in green algal cell walls [77], and (3) A papain (peptidase family C01), and a hyicolyisin-like peptidase (family M30), possibly involved in matrix degradation. Also, E-K07 SAG encodes several stress response pathways, signal transduction, and defense mechanisms that were not identified in Sakinaw Lake SAGs. These include (1) oxidative stress enzymes catalase and ferroxidase, (2) CRISPR-associated genes including the 6 core cas genes (cas1–cas6), as well as the CRISPR-associated csn1 gene [78], and (3) type VI secretion system including ten of the thirteen core tss genes [79].

Discussion

Our analysis of four “Latescibacteria” SAGs obtained from the anaerobic monimolimnion water column of Sakinaw Lake, and the anaerobic sediments of Etoliko lagoon revealed extensive saccharolytic and proteolytic capabilities, with preference for specific polysaccharides and glycoproteins such as pectins, alginates, fucans, ulvans, xyloglucans, starch, extensins, and arabinogalactan protein originating from algal cell walls and EPS. While the degradation of some of these polymers (e.g. pectins and alginates) have been fairly well characterized at the genomic, enzymatic, and organismal levels [39, 40, 42], limited information is available regarding the pathways, genes, and microorganisms mediating the degradation of others (e.g. fucans, ulvans, extensins and arabinogalactan proteins) [44, 46, 47, 53, 54, 80]. More importantly, our knowledge of the degradation of many of these compounds is based on the study of model aerobic organisms with little knowledge of such pathways in anaerobes.

We argue that the observed patterns of polymer degradation, and monomer/oligomer transport and catabolism reflect niche specialization within “Latescibacteria” for survival and substrate acquisition in aquatic ecosystems. Specifically, we hypothesize that “Latescibacteria” SAGs analyzed in Sakinaw lake and Etoliko lagoon are involved in the degradation of a considerable fraction of algal cell wall polysaccharides and glycoprotein, algal EPS, and algal storage molecules within the detritus of green and brown algae originating at the oxic and photic zones and sinking to the anoxic and aphotic zones through sedimentation. Primary productivity is an important source for organic matter deposited in lakes [81, 82]. Algal cells represent up to 90% of such sinking organic matters, especially in stratified lakes like Sakinaw. Prior studies have demonstrated that CO₂ fixation by algae represents the major source of organic carbon input.
in Sakinaw Lake, with the water column being the main site for the degradation of fixed organic carbon [83]. The stratified nature and lack of upwelling within meromictic lakes results in greater accumulation of organic matter into the lake’s deeper anoxic layers [84]. The overall contribution of algal detritus to lacustrine sediments is often enhanced by the frequent occurrence of algal blooms, an ecological phenomenon predicted to increase due to global warming trends, and the progressive increase in fertilizers usage [85]. This has been reported in the lagoon systems of Western Greece, where the occurrence of algal blooms and subsequent sedimentation of organic matter represent one of the driving forces for the observed progressive eutrophication and anoxia within this ecosystem [86, 87].

It should also be noted that, in addition to polymers putatively degraded “Latescibacteria”, algal cells are known to produce considerable quantities of oils (up to 60% of their weight), especially under unfavorable conditions (e.g. N and P starvation, temperature, salinity, or pH shifts, or heavy metal accumulation) [88, 89]. Interestingly, the analyzed SAGs lack all enzymes of the fatty acid degradation pathway to acetyl CoA. Similarly, cellulose represents an important constituent of green and brown algal cell wall [43], but the analyzed “Latescibacteria” SAGs display an extremely sparse cellulose degradation capacity (Figure A in S1 File). We reason that readily degradable components within algal detritus, e.g. cellular lipids and fatty acids, free proteins, and cellulose, are promptly utilized by microorganisms in the algal phycosphere [90–92], as well as by aerobic and anaerobic copiotrophs in the surrounding water column during the sedimentation process. Thus “Latescibacteria” residing in the deeper anaerobic layers of Sakinaw lake and Etoliko lagoon sediments have evolved to specialize in the degradation of the more recalcitrant substrates that accumulate as algal detritus descends to deeper anoxic layers in stratified aquatic ecosystems. Indeed, studies in meromictic lakes have demonstrated that degradation of algal blooms occurs during sedimentation leading to biomass loss and chemical structure alteration of the algal blooms with depth [81, 82].

The proposed ecological role for members of the “Latescibacteria” strongly suggests cellular attachment to sinking algal detritus. “Latescibacteria” SAGs encode genes for flagella and pili production, and formation of gas vesicles; traits that could enhance cellular capacity for tracking and attachment to particulate organic matter. A recent survey of microbial communities in the oxygen starved Black Sea with considerable primary productivity within the upper oxic zone, shows higher relative abundance of “Latescibacteria” in particulate-associated samples derived from the deep anoxic zone when compared to water samples from the same location [24].

In addition to the major contribution to sinking organic matter in water bodies, algal biomass degradation under anaerobic conditions has recently received additional attention due to its potential use for biogas production [93–99]. Surprisingly, little is currently known regarding the microbial community involved in algal biomass degradation under anaerobic conditions [93]. Thus analysis of “Latescibacteria” SAGs directly contributes to our understanding of potential bacterial lineages involved in the anaerobic turnover of algal cell components.

Finally, the “Latescibacteria” SAGs encode numerous biosynthetic capabilities and a rich repertoire of catabolic enzymes and transporters with the potential to promote growth on a large number of substrates. Such capabilities are in contrast to multiple recently obtained genomes of several uncultured bacterial and archaeal CP, where sparse anabolic capabilities, small genome size, and apparent dependence on syntrophic interactions for growth were observed [9, 13]. As such, the reported physiological properties (anaerobic nature and predicted slow growth rate due to possession of a relatively large genome size and a single rRNA operon), metabolic capabilities (distinct preference to specific polymers and sugars/sugar acids, auxotrophy to specific amino acids), and ecological distribution (preference to anaerobic and eutrophic
habitats) should be considered when designing strategies for the isolation of members of the “Latescibacteria”.

Supporting Information

S1 File. Supporting Information document containing supplementary text, Tables A-C, and Figures A-B, accompany this manuscript. Table A. Genbank accession numbers, candidate order, and study site of all near-full-length 16S rRNA gene sequences affiliated with “Latescibacteria” that were used to construct phylogenetic trees shown in Fig 1. Table B. Total number of glycosyl hydrolases (GHs), polysaccharide lyases (PLs), and carbohydrate esterases (CEs) in the two most complete “Latescibacteria” SAGs compared to other lignocellulolytic and alginolytic organisms. Table C. Number of peptidases belonging to various Merops peptidase families identified in “Latescibacteria” genomes and their possible physiological roles. Figure A. Total number of “Latescibacteria” genes belonging to the different families of glycosyl hydrolases (GHs) and polysaccharide lyases (PLs) shown on the X-axis for SAGs S-E07 and S-B13. Figure B. Schematic representation of polymers shown in Table 2.

Author Contributions

Conceived and designed the experiments: NY CR TW ME. Performed the experiments: NY IF CR. Analyzed the data: NY IF CR SJH TW ME. Contributed reagents/materials/analysis tools: SJH. Wrote the paper: NY IF ME.

References

1. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013; 41: D590–6. doi:10.1093/nar/gks1219 PMID: 23193283
2. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved Green ages taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 2012; 6:610–8. doi:10.1038/ismej.2011.139 PMID: 22134646
3. Winsley T, Snape I, McKinlay J, Stark J, van Dorst JM, Ji M, et al. The ecological controls on the prevalence of candidate division TM7 in polar regions. Front Microbiol. 2014; 5:345. doi: 10.3389/fmicb.2014.00345 PMID: 25076941
4. Farag IF, Davis JP, Youssef NH, Elshahed MS. Global patterns of abundance, diversity and community structure of the Aminicenantes (candidate phylum OP8). PloS one. 2014; 9:e92139. doi:10.1371/journal.pone.0092139 PMID: 24637619
5. Portillo MC, Sririn V, Kanoksilapatham W, Gonzalez JM. Differential microbial communities in hot spring mats from Western Thailand. Extremophiles. 2009; 13:321–31. doi:10.1007/s00792-008-0219-x PMID: 19109691
6. Ferrari B, Winsley T, Ji M, Neillan B. Insights into the distribution and abundance of the ubiquitous candidatus Saccharibacteria phyllum following tag pyrosequencing. Sci Rep. 2014; 4:3957. doi: 10.1038/srep03957 PMID: 24492458
7. Ohkuma M, Sato T, Noda S, Ui S, Kudo T, Hongoh Y. The candidate phyylum ‘Termite Group 1’ of bacteria: phylogenetic diversity, distribution, and endosymbiont members of various gut flagellated protists. FEMS Microbiol Ecol. 2007; 60:467–76. PMID: 17391329
8. Kamke J, Rinke C, Schwientek P, Mavromatis K, Ivanova N, Szymbra A, et al. The candidate phyllum Poribacteria by single-cell genomics: new insights into phylogeny, cell-compartmentation, eukaryote-like repeat proteins, and other genomic features. PloS one. 2014; 9:e87353. doi: 10.1371/journal.pone.0087353 PMID: 24498082
9. Kantor RS, Wrighton KC, Handley KM, Sharon I, Hug LA, Castelle CJ, et al. Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. mBio. 2013; 4: e00708–13. doi:10.1128/mBio.00708-13 PMID: 24149512
10. McLean JS, Lombardo MJ, Badger JH, Edlund A, Novotny M, Yee-Greenbaum J, et al. Candidate phyllum TM6 genome recovered from a hospital sink biofilm provides genomic insights into this...
uncultivated phylum. Proc Natl Acad Sci USA. 2013; 110:E2390–9. doi: 10.1073/pnas.1219809110
PMID: 23754396

11. Wrighton KC, Castelle CJ, Wilkins MJ, Hug LA, Sharon I, Thomas BC, et al. Metabolic interdependencies between phylogenetically novel fermenters and respiratory organisms in an unconfined aquifer. ISME J. 2014; 8:1452–63. doi: 10.1038/ismej.2013.249 PMID: 24621521

12. Takami H, Noguchi H, Takaki Y, Uchiyama I, Toyoda A, Nishi S, et al. A deeply branching thermophilic bacterium with an ancient acetyl-CoA pathway dominates a subsurface ecosystem. PloS one. 2012; 7:e30559. doi: 10.1371/journal.pone.0030559 PMID: 22303444

13. Youssef NH, Rinke C, Stepanauskas R, Farag I, Woyke T, Elshahed MS. Insights into the metabolism, lifestyle and putative evolutionary history of the novel archaeal phylum 'Diapherotrites'. ISME J. 2015; 9:447–60. doi: 10.1038/ismej.2014.141 PMID: 25083931

14. Kamke J, Sczyrba A, Ivanova N, Schwientek P, Rinke C, Mavromatis K, et al. Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges. ISME J. 2013; 7:2287–300. Epub 2013/07/12. doi: 10.1038/ismej.2013.111 PMID: 23842652

15. Campbell JH, O’Donoghue P, Campbell AG, Schwientek P, Sczyrba A, Woyke T, et al. UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. Proc Natl Acad Sci. 2013; 110:5541–5. doi: 10.1073/pnas.1303090110 PMID: 23509275

16. Wilson MC, Mori T, Ruckert C, Uria AR, Helf MJ, Takada K, et al. An environmental bacterial taxon with a large and distinct metabolic repertoire. Nature. 2014; 506:558–62. doi: 10.1038/nature12959 PMID: 24476823

17. Dojka MA, Hugenholtz P, Haack SK, Pace NR. Microbial diversity in a hydrocarbon- and chlorinated solvent-contaminated aquifer undergoing intrinsic bioremediation. Appl Environ Microbiol. 1998; 64:3869–77. PMID: 9758812

18. Pereira AD, Leal CD, Dias MF, Etchebehere C, Chemicharo CA, da Araujo JC. Effect of phenol on the nitrogen removal performance and microbial community structure and composition of an anammox reactor. Bioreour Technol. 2014; 166:103–11. doi: 10.1016/j.biortech.2014.05.043 PMID: 24907569

19. Schabereiter-Gurtner C, Saiz-Jimenez C, Pinar G, Lubitz W, Rolleke S. Phylogenetic diversity of bacteria associated with Paleolithic paintings and surrounding rock walls in two Spanish caves (Lion in and La Garma). FEMS Microbiol Ecol. 2004; 47:235–47. doi: 10.1016/S0168-6496(03)00280-0 PMID: 19712338

20. Ikenaga M, Guevara R, Dean AL, Pisani C, Boyer JN. Changes in community structure of sediment bacteria along the Florida coastal everglades marsh-mangrove-seagrass salinity gradient. Microb Ecol. 2010; 59:284–95. doi: 10.1007/s00248-009-9572-2 PMID: 19705193

21. Reed AJ, Lutz RA, Vetriani C. Vertical distribution and diversity of bacteria and archaea in sulfide and methane-rich cold seep sediments located at the base of the Florida Escarpment. Extremophiles. 2006; 10:199–211. PMID: 16465452

22. Hernandez-Raquet G, Budzinski H, Caumette P, Dabert P, Le Menach K, Muyzer G, et al. Molecular diversity studies of bacterial communities of oil polluted microbial mats from the Etang de Berre (France). FEMS Microbiol Ecol. 2006; 58:550–62. PMID: 17117996

23. Briggs BR, Pohlman JW, Torres M, Riedel M, Brodie EL, Colwell FS. Macroscopic biofilms in fracture-dominated sediment that anaerobically oxidize methane. Appl Environ Microbiol. 2011; 77:6780–7. doi: 10.1128/AEM.00288-11 PMID: 21821755

24. Carbonetto B, Rascovan N, Alvarez R, Mentaberry A, Vazquez MP. Structure, composition and metagenomic profile of soil microbomes associated to agricultural land use and tillage systems in Argentine Pampas. PloS one. 2014; 9:e99949. doi: 10.1371/journal.pone.0099949 PMID: 24923965

25. Yakimov MM, La Cono V, Stepak VZ, La Spada G, Arcadi E, Messina E, et al. Microbial life in the Lake Medee, the largest deep-sea salt-saturated formation. Sci Rep. 2013; 3:3554. doi: 10.1038/srep03554 PMID: 24392146
29. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, et al. Insights into the phylogeny and coding potential of microbial dark matter. Nature. 2013; 499:431–7. doi: 10.1038/nature12352 PMID: 23851394

30. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2014; 42:D199–205. doi: 10.1093/nar/gkt1076 PMID: 24214961

31. Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Res. 2014; 42:D459–D71. doi: 10.1093/nar/gkt1103 PMID: 24225315

32. Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpides NC, et al. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics. 2007; 8:209. PMID: 17577412

33. Anonymous. PILER Genomic repeat analysis software. 2009.

34. Rawlings ND, Waller M, Barrett AJ, Bateman A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 2014; 42:D503–D71. doi: 10.1093/nar/gkt953 PMID: 24157837

35. Saier MH Jr, Reddy VS, Tamang DG, Vastermark A. The transporter classification database. Nucleic Acids Res. 2014; 42:D251–D8. doi: 10.1093/nar/gkt1097 PMID: 24225317

36. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012; 40:W445–51. doi: 10.1093/nar/gks479 PMID: 22645317

37. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014; 42:D490–D5. doi: 10.1093/nar/gkt1178 PMID: 24270786

38. Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpides NC, et al. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics. 2007; 8:209. PMID: 17577412

39. Anonymous. PILER Genomic repeat analysis software. 2009.

40. Rawlings ND, Waller M, Barrett AJ, Bateman A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 2014; 42:D503–D71. doi: 10.1093/nar/gkt953 PMID: 24157837

41. Saier MH Jr, Reddy VS, Tamang DG, Vastermark A. The transporter classification database. Nucleic Acids Res. 2014; 42:D251–D8. doi: 10.1093/nar/gkt1097 PMID: 24225317

42. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012; 40:W445–51. doi: 10.1093/nar/gks479 PMID: 22645317

43. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014; 42:D490–D5. doi: 10.1093/nar/gkt1178 PMID: 24270786

44. Domozych DS, Sorensen I, Popper ZA, Ochs J, Andreas A, Fangel JU, et al. Pectin metabolism and assembly in the cell wall of the charophyte green alga *Penium margaritaceum*. Plant Physiol. 2014; 165:105–18. doi: 10.1104/pp.114.236257 PMID: 24652345

45. Abbott DW, Boraston AB. Structural biology of pectin degradation by Enterobacteriaceae. Microbiol Mol Biol Rev. 2008; 72:301–16. doi: 10.1128/MMBR.00038-07 PMID: 18535148

46. Benoit I, Coutinho P, Schols H, Gerlach J, Henrissat B, de Vries R. Degradation of different pectins by fungi: correlations and contrasts between the pectinolytic enzyme sets identified in genomes and the growth on pectins of different origin. BMC Genomics. 2012; 13:321. PMID: 22812459

47. Michel G, Tonon T, Scovento D, Cck JM, Klaareg B. The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular matrix polysaccharides in Eukaryotes. New Phytol. 2010; 188:82–97. doi: 10.1111/j.1469-8177.2010.03374.x PMID: 20618907

48. Kabisch A, Otto A, Konig S, Becher D, Albrecht D, Schuler M, et al. Functional characterization of polysaccharide utilization loci in the marine Bacteroidetes *'Gramella forsetii' KT0803*. ISME J. 2014; 8:1492–502. doi: 10.1038/ismej.2014.4 PMID: 24522261

49. Domozych DS. Algal cell walls. eLS. Chichester: John Wiley & Sons; 2011.

50. Chiellini F, Morelli A. Ulvan: A Versatile Platform of Biomaterials from Renewable Resources: INTECH Open Access Publisher; 2011.

51. Jiao G, Yu G, Zhang J, Ewart HS. Chemical structures and bioactivities of sulfated polysaccharides from marine algae. Mar Drugs. 2011; 9:196–223. doi: 10.3390/md9020196 PMID: 21586795

52. Ferro DR, Provasoli A, Ragazzi M, Casu B, Torri G, Bossennec V, et al. Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences. Carbohydr Res. 1990; 195:157–67. PMID: 2331699

53. Nyvall Collen P, Saasi JF, Rogniaux H, Marfaing H, Helbert W. Ulvan lyases isolated from the Flavobacteria *Persicrivaga ulvanivorans* are the first members of a new polysaccharide lyase family. J Bacteriol. 2011; 193:2063–71. doi: 10.1128/jb.01728-10 PMID: 21566795

54. Domozych DS, Sorensen I, Willats WG. The distribution of cell wall polymers during antheridium development and spermatogenesis in the Charophycean green alga, *Chara corallina*. Ann Bot. 2009; 104:1045–56. doi: 10.1093/aob/mcp198 PMID: 19896537

55. Ikegaya H, Hayashi T, Kaku T, Iwata K, Sonobe S, Shimmen T. Presence of xyloglucan-like polysaccharide in *Spirogyra* and possible involvement in cell-cell attachment. Physcol Res. 2009; 56:216–22.

56. Lahaye M, Jegou D, Buleon A. Chemical characteristics of insoluble glucans from the cell wall of the marine green alga *Ulva lactuca* (L.) Thuret. Carbohydr Res. 1994; 262:115–25.
1. Estevez JM, Fernandez PV, Kasulin L, Dupree P, Ciancia M. Chemical and in situ characterization of macromolecular components of the cell walls from the green seaweed Codium fragile. Glycobiology. 2009; 19:212–28. doi: 10.1093/glycob/cwn101 PMID: 18832454

2. Domozych DS, Ciancia M, Fangel JU, Mikkelsen MD, Ulvskov P, Willats WG. The cell walls of green algae: a journey through evolution and diversity. Front Plant Sci. 2012; 3:82. doi: 10.3389/fpls.2012.00082 PMID: 22639667

3. Fujita K, Sakamoto S, Ono Y, Wakao M, Suda Y, Kitahara K, et al. Molecular cloning and characterization of a beta-L-arabinobiosidase in Bifidobacterium longum that belongs to a novel glycoside hydrolase family. J Biol Chem. 2011; 286:5143–50. doi: 10.1074/jbc.M110.190512 PMID: 21149454

4. Knoch E, Dilkopinol A, Geshi N. Arabinogalactan proteins: focus on carbohydrate active enzymes. Front Plant Sci. 2014; 5:198. doi: 10.3389/fpls.2014.00198 PMID: 24966680

5. Bakunina I, Nedashkovskaya O, Balabanova L, Zvyagintseva T, Rasskasov V, Mikhailov V. Comparative analysis of glycoside hydrolases activities from phylogenetically diverse marine bacteria of the genus Arenibacter. Mar Drugs. 2013; 11:1977–98. doi: 10.3390/md11061977 PMID: 23752354

6. Cobucci-Ponzano B, Conte F, Strazzulli A, Capasso C, Fiume I, Pocsfalvi G, et al. The molecular characterization of a novel GH38 alpha-mannosidase from the crenarchaeon Sulfolobus solfataricus revealed its ability in de-mannosylating glycoproteins. Biochimie. 2010; 92:1895–907. doi: 10.1016/j.biochi.2010.07.016 PMID: 20696204

7. Mamedov T, Yusibov V. Green algae Chlamydomonas reinhardtii possess endogenous sialylated Nglycans. FEBS Open Bio. 2011; 1:15–22. doi: 10.1016/j.fob.2011.10.003 PMID: 23650571

8. Popper ZA, Michel G, Herve C, Domozych DS, Willats WG, Tuohy MG, et al. Evolution and diversity of plant cell walls: from algae to flowering plants. Ann Rev Plant Biol. 2011; 62:567–90. doi: 10.1146/annurev-arplant-042110-103809 PMID: 21351878

9. Gerken HG, Donohoe B, Knoshaug EP. Enzymatic cell wall degradation of Chlorella vulgaris and other microalgae for biofuels production. Planta. 2013; 237:239–53. doi: 10.1007/s00425-012-1765-0 PMID: 23011569

10. Benjdia A, Leprince J, Guillot A, Vaudry H, Rabot S, Berteaux O. Anaerobic sulfatase-maturating enzymes: radical SAM enzymes able to catalyze in vitro sulfatase post-translational modification. J Am Chem Soc. 2007; 129:3462–3. PMID: 17335281

11. Dierks T, Miech C, Hummerjohann J, Schmidt B, Kertesz MA, von Figura K. Posttranslational formation of formylglycine in prokaryotic sulfatases by modification of either cysteine or serine. J Biol Chem. 1998; 273:25560–4. PMID: 9748219

12. Deschamps P, Haferkamp I, d’Hulst C, Neuhaus HE, Ball SG. The relocation of starch metabolism to chloroplasts: when, why and how. Trends Plant Sci. 2008; 13:574–82. doi: 10.1016/j.tplants.2008.08.009 PMID: 18824400

13. Bellinger BJ, Gretz MR, Domozych DS, Kiemle SN, Hagertheye SE. Composition of extracellular polymeric substances from periphyton assemblages in the Florida everglades. J Phycol. 2010; 46:484–96.

14. Mishra A, Kavita K, Jha B. Characterization of extracellular polymeric substances produced by microalgae Dunaliella salina. Carbohydr Pol. 2011; 83:852–7.

15. Gonzalez ET, Allen C. Characterization of aRalstonia solanacearum operon required for polygalacturonate degradation and uptake of galacturonic acid. Mol Plant-Micro Int. 2003; 16:536–44.

16. Valentini M, Storelli N, Lapouge K. Identification of C(4)-dicarboxylate transport systems in Pseudomonas aeruginosa PA01. J Bacteriol. 2011; 193:4307–16. doi: 10.1128/JB.05074-11 PMID: 21725012

17. Havemann GD, Bobik TA. Protein content of polyhedral organelles involved in coenzyme B12-dependent degradation of 1,2-propanediol in Salmonella enterica serovar Typhimurium LT2. J Bacteriol. 2003; 185:5086–95. PMID: 12923081

18. Petit E, LaTouf WG, Coppi MV, Warnick TA, Currie D, Romashko I, et al. Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by Clostridium phytofermentans. PLoS One. 2013; 8:e54337. doi: 10.1371/journal.pone.0054337 PMID: 23382892

19. Axen SD, Erbilgin O, Kerfeld CA. A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. PLoS Comput Biol. 2014; 10:e1003898. doi: 10.1371/journal.pcbi.1003898 PMID: 25340524

20. Dunne WM. Bacterial adhesion: Seen any good biofilms lately? Clin Microbiol Rev. 2002; 15:155–66. PMID: 11932228

21. Walsby AE. Gas vesicles. Microbiol Rev. 1994; 58:94–144. PMID: 8177173
72. Briukhanov AL, Netrusov AI. Aerotolerance of strictly anaerobic microorganisms and factors of defense against oxidative stress: a review. Prikladnaia Biokhimia i Mikrobiologiia. 2007; 43:635–52. PMID: 18173105
73. Liu R, Ochman H. Stepwise formation of the bacterial flagellar system. Proc Natl Acad Sci USA. 2007; 104:7116–21. PMID: 17438286
74. Pfeifer F. Distribution, formation and regulation of gas vesicles. Nat Rev Microbiol. 2012; 10:705–15. doi:10.1038/nrmicro2834 PMID: 22941504
75. Takhar HK, Kemp K, Kim M, Howell PL, Burrows LL. The platform protein is essential for type IV pilus biogenesis. Journal Biol Chem. 2013; 288:9721–8. doi: 10.1074/jbc.M113.453506 PMID: 23413032
76. Konno N, Ishida T, Igarashi K, Fushinobu S, Habu N, Samejima M, et al. Crystal structure of polysaccharide lyase family 20 endo-β-1,4-glucuronan lyase from the filamentous fungus Trichoderma reesei. FEBS Lett. 2009; 583:1323–6. doi: 10.1016/j.febslet.2009.03.034 PMID: 19306878
77. Redouan E, Cedric D, Emmanuel P, Mohamed EG, Bernard C, Philippe M, et al. Improved isolation of glucuronan from algae and the production of glucuronic acid oligosaccharides using a glucuronan lyase. Carbohydr Res. 2009; 344:1670–5. doi:10.1016/j.carres.2009.05.031 PMID: 19616199
78. Meyers PA. Preservation of elemental and isotopic source identification of sedimentary organic matter. Chem Geol. 1994; 114:289–302.
79. Meyers PA, Ishiwatari R. Lacustrine organic geochemistry—an overview of indicators of organic matter sources and diagenesis in lake sediments. Org Geochem. 1993; 20:867–900.
80. Perry KA. The chemical limnology of two meromictic lakes with emphasis on pyrite formation. Vancouver: University of British Columbia; 1990.
81. Bresciani M, Bolpagni R, Laini A, Matta E, Bartoli M, Giardino C. Multitemporal analysis of algal blooms with MERIS images in a deep meromictic lake. Eur J Remote Sens. 2013; 46:445–58.
82. Paerl HW, Scott JT. Throwing fuel on the fire: synergistic effects of excessive nitrogen inputs and global warming on harmful algal blooms. Environ Sci Technol. 2010; 44:7756–8. doi:10.1021/es102665e PMID: 20804137
83. Kormas KA, Nicolaidou A, Reizopoulou S. Temporal variations of nutrients, chlorophyll A and particulate matter in three coastal lagoons of Amvrakikos Gulf (Ionian Sea, Greece). Mar Ecol. 2001; 22:201–13.
84. Chisti Y. Biodiesel from microalgae. Biotechnol Adv. 2007; 25:294–306. PMID: 17350212
85. Sharma KK, Schuhrmann H, Schenk PM. High lipid induction in microalgae for biodiesel production. Energies. 2012; 5:1532–53.
86. Buchan A, LeCleir GR, Gulvik CA, Gonzalez JM. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. Nat Rev Microbiol. 2014; 12:686–98. doi: 10.1038/nrnmicro3326 PMID: 25134618
87. Sapp M, Schwaderer AS, Wiltshire KH, Hoppe HG, Gerdts G, Wichels A. Species-specific bacterial communities in the phycosphere of microalgae? Microb Ecol. 2007; 53:683–99. PMID: 17264999
88. Mann AJ, Hahnke RL, Huang S, Werner J, Xing P, Babeyron T, et al. The genome of the alga-associated marine flavobacterium Formosa agariphila KMM 3901T reveals a broad potential for degradation of algal polysaccharides. Appl Environ Microbiol. 2013; 79:6813–22. doi: 10.1128/AEM.01937-13 PMID: 23995932
89. Ward AJ, Lewis DM, Green FB. Anaerobic digestion of algae biomass: A review. Alg Res. 2014; 5:204–14.
90. Chisti Y. Biodiesel from microalgae beats bioethanol. Trends Biotechnol. 2008; 26:126–31. doi: 10.1016/j.tibtech.2007.12.002 PMID: 18221809
96. Sutherland DL, Turnbull MH, Broady PA, Craggs RJ. Effects of two different nutrient loads on microalgal production, nutrient removal and photosynthetic efficiency in pilot-scale wastewater high rate algal ponds. Water Res. 2014; 66c:53–62.

97. Prajapati SK, Kaushik P, Malik A, Vijay VK. Phycoremediation and biogas potential of native algal isolates from soil and wastewater. Bioresource Technol. 2013; 135:232–8. doi:10.1016/j.biortech.2012.08.069 PMID: 22985826

98. Vergara-Fernández A, Vargas G, Alarcón N, Velasco A. Evaluation of marine algae as a source of biogas in a two-stage anaerobic reactor system. Biomass Bioenergy. 2008; 32:338–44.

99. Mahadevaswamy M, Venkataraman LV. Bioconversion of poultry droppings for biogas and algal production. Agric Wastes. 1986; 18:93–101.

100. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013; 30:2725–9. doi:10.1093/molbev/mst197 PMID: 24132122

101. Eder M, Lutz-Meindl U. Analyses and localization of pectin-like carbohydrates in cell wall and mucilage of the green alga Netrium digitus. Protoplasma. 2010; 243:25–38. doi:10.1007/s00709-009-0040-0 PMID: 19340523