Systematic and Applied Microbiology 43 (2020) 126088

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: http://www.elsevier.com/locate/syapm

Candidatus Abditibacter, a novel genus within the *Cryomorphaceae*, thriving in the North Sea

Anissa Grieb, T. Ben Francis, Karen Krüger, Luis H. Orellana, Rudolf Amann, Bernhard M. Fuchs*

Max Planck Institute for Marine Microbiology, Celsiusstr.1, 28359 Bremen, Germany

ARTICLE INFO

Article history:
Received 21 January 2020
Received in revised form 22 April 2020
Accepted 22 April 2020

Keywords:
North Sea
Helgoland
metagenome assembled genome
Vis6
bacterioplankton

ABSTRACT

Coastal phytoplankton blooms are frequently followed by successive blooms of heterotrophic bacterial clades. The class *Flavobacteriia* within the *Bacteroidetes* has been shown to play an important role in the degradation of high molecular weight substrates that become available in the later stages of such blooms. One of the flavobacterial clades repeatedly observed over the course of several years during phytoplankton blooms off the coast of Helgoland, North Sea, is Vis6. This genus-level clade belongs to the family *Cryomorphaceae* and has been resistant to cultivation to date. Based on metagenome assembled genomes, comparative 16S rRNA gene sequence analyses and fluorescence *in situ* hybridization, we here propose a novel candidate genus Abditibacter, comprising three novel species *Candidatus* Abditibacter vernus, *Candidatus* Abditibacter forsetii and *Candidatus* Abditibacter autumni. While the small genomes of the three novel photoheterotrophic species encode highly similar gene repertoires, including genes for degradation of proteins and algal storage polysaccharides such as laminarin, two of them – *Ca.* A. vernus and *Ca.* A. forsetii – seem to have a preference for spring blooms, while *Ca.* A. autumni almost exclusively occurs in late summer and autumn.

© 2020 The Author(s). Published by Elsevier GmbH. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

Marine phytoplankton fixes the same amount of CO_2 as land plants despite representing only 0.2% of the global biomass of primary producers [22]. In coastal and upwelling regions, phytoplankton blooms can be initiated by increases in solar irradiation, nutrient availability and reduced grazing pressure particularly in spring [11]. These blooms are characterized by rapid increases in phytoplankton growth followed by a decline in population density after several days to weeks. A large part of the organic matter released by living and decaying phytoplankton during and after these bloom events is remineralized to inorganic nutrients and CO_2 by heterotrophic bacteria [11,57], mostly *Alphaproteobacteria*, *Gammaproteobacteria* and *Bacteroidetes* [11]. In particular, members of the class *Flavobacteriia* within the phylum *Bacteroidetes* play an important role in the degradation of high molecular weight (HMW) compounds, such as proteins and polysaccharides (e.g. [21,37,42,59]). There are correlations between flavobacterial abundances and specific phytoplankton species such as diatoms or flagellates [42,49], yet the interaction of phytoplankton and bacterioplankton can also be interpreted as a recurrent substrate-based succession [59,60].

Helgoland Roads is a long term ecological research station, located off the island of Helgoland in the German Bight, where spring phytoplankton blooms have been studied in detail for one decade [12,23,29,32,59,60]. Among the flavobacterial clades recurrently responding to the diatom-dominated blooms were novel clades related to the known genera *Ulvibacter*, *Polaribacter*, *Formosa* and, within the *Cryomorphaceae*, a clade referred to as Vis6 [12,59,60]. The *Ulvibacter* related genus was recently described as *Candidatus* (*Ca.*) Prosiliicoccus based on metagenome assembled genomes (MAGs) and other data [23]. *Polaribacter* and *Formosa* strains have been isolated from samples from 2010 [29] and have been studied in detail by Xing *et al.* [66] and Unfried *et al.* [62]. A taxonomic and ecological study of the clade Vis6 is lacking, despite the fact that this group is recurrently abundant in nearly all spring blooms investigated at Helgoland Roads with relative abundances up to 20% [60], and 16S rRNA gene sequences related to Vis6 have also been found in other phytoplankton blooms (e.g. [34,60,64]).

Vis6 was first found during a research cruise crossing different oceanic provinces from the East Greenland current to the North Atlantic subtropical gyre in September 2006. A fluorescence *in situ* hybridization (FISH) probe was designed based on 16S rRNA gene

* Corresponding author.

https://doi.org/10.1016/j.syapm.2020.126088
0723-2020/© 2020 The Author(s). Published by Elsevier GmbH. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

libraries, named Vis6-814 [25]. FISH counts showed highest abundances (up to 3.4 ± 1.3 x10^3 cells per ml) in the northern stations with decreasing numbers southwards. Despite substantial cultivation effort (e.g. [14,17,29]), no cultured representative is available to date and the global distribution and relevance of this clade remains unclear. Vis6 related sequences have been detected as abundant phytoplankton bloom responders (e.g. [34,60,64]), but were sometimes referred to as "*Owenweeksia*", due to *Owenweeksia hongkongensis* being at that time the closest cultured relative based on 16S rRNA gene sequence comparison [12,29]. The family *Cryomorphaceae*, to which both *Owenweeksia* and Vis6 belong, is polyphyletic and consists of nine genera [10]. It was named after the first and at that time sole member *Cryomorpha ignava* which was isolated from a quartz stone sublithic cyanobacterial biofilm in Eastern Antarctica [9]. The family seems to be widely distributed as indicated by different sequence databases such as SILVA, NCBI and GTDB, where highly related sequences are listed as "uncultured" or "unclassified".

In this study, we wanted to close a gap in our knowledge about this abundant bloom-responding flavobacterial clade. Here, the Vis6 clade is characterized based on (i) metagenomic data including a functional description based on gene annotation, (ii) 16S rRNA gene sequences for phylogenetic classification and (iii) abundance data based on 16S rRNA amplicon sequencing and cell counts from fluorescence *in situ* hybridization (FISH) from several years. Based on these data, we propose a novel candidate genus, *Candidatus* Abditibacter, within the *Cryomorphaceae* family that includes three species: *Candidatus* Abditibacter forsetii, *Candidatus* Abditibacter vernus and *Candidatus* Abditibacter autumni.

Materials and methods

Sampling

In this study, we combined previously published sequencing and cell count data with new analyses. All samples were taken off the coast of Helgoland, North Sea, at the long-term ecological research station Helgoland Roads (54° 11.3' N, 7° 54.0' E) in varying sampling intervals in the years 2009-2013 and 2016. A strong emphasis was put on the spring time of each year to analyze the bacterial responses to spring phytoplankton blooms. In 2017, one water sample was taken on September 20.

The samples were taken as described in Teeling *et al.* [59]. In brief, the community DNA, collected on 0.2 μm filters after prefiltration with 10 and 3 μm cut-offs, was extracted and sequenced for metagenomic analyses. Unfractionated seawater, fixed with 1% formaldehyde, was collected on 0.2 μm filters for catalyzed reporter deposition (CARD)-FISH and cell counting. From the sample from September 20, 2017, cells hybridized with a Vis6 specific FISH probe have been enriched by flow cytometry prior to sequencing [28] in addition to a shotgun metagenome. Phytoplankton data, including chlorophyll *a* concentrations, total cell counts and CARD-FISH counts for the years 2009-2012 have been published in Teeling *et al.* [60]. Chlorophyll *a* concentrations from the years 2013 and 2016 were assessed from fluorescence data using an algal group analyzer (bbe moldaenke, Kiel-Kronshagen, Germany) [51,65].

16S rRNA phylogenetic reconstruction

A 16S rRNA reference tree was created based on the SILVA database release 132 SSU Ref (www.arb-silva.de). Sequence curation and phylogenetic tree reconstruction was done with the ARB software [38]. The alignment of Vis6 sequences (selected based on a match with probe Vis6-814 [25]) was manually improved for 381 high quality sequences of a length of >1400 bp. All type strains within the *Flavobacteriia* were additionally selected, resulting in 708 sequences that served as outgroup for the tree construction. With the total of 1102 sequences, two trees with the neighbor joining (NJ) method were calculated, one with an additional 30% *Bacteroidetes* variability filter and one without. Two trees were calculated with RAxML8 tree construction method [56], one with 30% *Bacteroidetes* and one without any variability filter. Of those four trees a consensus tree using the consensus tree option implemented in ARB was generated and manually refined as described in Peplies *et al.* [47].

Amplicon data

The acquisition and processing of 16S rRNA amplicon data from 2010-2012 were previously published in Chafee *et al.* [12], and done in a similar manner for the years 2013 and 2016. In brief, the V4 region of the 16S rRNA gene was amplified from the 0.2-3 μm size fractions and sequenced. The sequences were clustered based on oligotyping using Minimum Entropy Decomposition [19,20]. The 3-10 μm size fraction was analyzed in the same way. Differentially abundant oligotypes were determined with the DESeq2 package [36].

Oligotypes with > 1% relative read abundance in the amplicon sequences that have been classified as *Cryomorphaceae*, as well as 16S rRNA gene sequences from the genome taxonomy database (GTDB; [46]) were selected. These sequences were added to the Vis6 16S rRNA tree using parsimony criteria with the tool ARB_Parsimony in ARB.

Metagenome sequencing, assembly and binning

Thirty-eight metagenomes from 2010-2012 were sequenced at the DOE Joint Genome Institute as described in Teeling *et al.* [60]. Assembly of the reads and binning of contigs was done according to Krüger *et al.* [32]. Metagenome raw reads, assemblies and MAGs have been submitted to the European Nucleotide Archive (ENA) under accession number PRJEB28156. The MAGs analyzed in our study have the same identifiers as in ENA. MAGs from these analyzed metagenomes were clustered into Mash-clusters using Mash [44], which represent approximate species clusters of highly similar MAGs close to 95% average nucleotide identity (ANI). These Mash-clusters have been published in Krüger *et al.* [32] and two of them (mc_3 and mc_12) were classified as Vis6 clade, based on 16S rRNA gene sequences, and analyzed in this study. Briefly, fragments longer than 250 bp carrying 16S rRNA gene sequences were identified with CheckM ssu_finder [45] in the MAGs from 2010-2012 and classified by adding them to the Vis6 consensus 16S rRNA tree with parsimony criteria using ARB [38]. From a total of 38 sequences falling into the Vis6 cluster, 37 derived either from mash-cluster mc_3 or mc_12. The flow-sorting based acquisition of MAGs from a sample taken on September 20, 2017 was described in Grieb *et al.* [28]. In brief, samples were sorted based on the fluorescence FISH signal of the Vis6-814 probe [25] prior to sequencing, assembly and binning. From the same sampling day, a Vis6 MAG was obtained from a shotgun bulk metagenome. This metagenome as well as the described mini-metagenomes were sequenced at the DOE Joint Genome Institute under the IMG GOLD [41] study ID Gs0130320.

Completeness, contamination and heterogeneity of the MAGs were estimated by CheckM [45]. Adopting quality thresholds from Bowers *et al.* [8], MAGs were classified as high, medium and low. Only the medium (<10% contamination, > 50% completeness) and high quality (<5% contamination, > 90% completeness, ≥18 tRNA) genomes were used for analyses. This included 13 high quality and 13 medium quality MAGs from mc_3, 29 high quality and 9 medium quality from mc_12 and 8 medium quality MAGs from September

2017. All MAG identifiers, their bin sizes and quality are listed in Table S1.

Phylogenomic reconstruction

One representative MAG (selected bases on highest quality values) of mc_3, mc_12 and September 2017 was placed in a GTDB reference tree [46]. The genome based phylogeny was calculated using GTDBtk v0.3.1, and GTDB R89 as the reference data package [46]. The tree itself was created with the de_novo_wf pipeline, using the bacterial marker set (–bac120_ms), and the phylum p_Deinococcota as the –outgroup_taxon. The ANI and amino acid identity (AAI) between MAGs and references were calculated using ani.rb and aai.rb from the enveomics collection [53].

Temporal and spatial distribution

The relative abundance of cells targeted with the Vis6-814 probe [25] was determined by CARD-FISH in relation to the total cell counts as determined by 4',6-diamidino-2-phenylindole (DAPI) staining. CARD-FISH was done according to Pernthaler *et al.* [48]. Counts from 2009-2012 were taken from Teeling *et al.* [60], the counts from 2013 and 2016 were done as part of this study. We selected a representative MAG from both mc_ 3 and mc_12 and recruited the reads from the metagenomes from 2010-2012 and 2016 to these MAGs as described in Francis *et al.* [23]. Relative abundance was calculated as the percentage of recruited reads from the total number of reads per sampling date. For calculating the abundance of oligotypes, the relative abundance of reads of the amplicon data-set was calculated as described in Chafee *et al.* [12].

Data for global distribution was collected using IMNGS [33]. As query sequences, representatives of each 16S rRNA gene cluster were used. These were the sequence FQ032803 [26], the 16S rRNA gene sequences from the MAGs 20120524_Bin_102_1, 20110509_Bin_54_1 and 3300031407_1. Minimum target size was 200 and an identity threshold of 99% was used. Percent of reads in each sequencing run was calculated from the IMNGS output, and the corresponding geographic positions for each sequencing run were collected from NCBI. A cutoff of at least 10 reads matching the query was used for plotting.

Gene prediction and annotation

Gene prediction and annotation was done according to Francis *et al.* [23]. Briefly, an initial annotation was done with prokka [54] and refined by running HMMer [16] searches against profiles of SusC-like (TIGR04056) and SusD-like (PF07980.10, PF12741.6, PF12771.6, PF14322.5) domains, sulfatases (PF00884) and proteases (MEROPS). Sulfatases were further classified using BLAST search against the SulfAtlas database [1,4]. TonB dependent receptors were searched using HMMer with the profiles TIGR01352, TIGR01776, TIGR01778, TIGR01779, TIGR01782, TIGR01783, TIGR01785, TIGR01786, TIGR02796, TIGR02797, TIGR02803, TIGR02804, TIGR02805, TIGR04056 and TIGR04057. In the resulting list of annotations we performed word searches for carbohydrate active enzymes (CAZymes) that included glycoside hydrolases (GHs), glycosyltransferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CE) and carbohydrate-binding modules (CBMs). For the peptidase:CAZyme ratio only GHs, PLs and CEs were taken into account. To detect polysaccharide utilization loci (PULs), we searched for susC and susD like genes in proximity to these CAZymes. For all annotations the average of all MAGs with medium or high quality belonging to one species was calculated.

Additionally, three MAGs each of mc_3 (20100303_Bin_80_1, 20110324_Bin_51_2, 20120524_Bin_102_1) and mc_12 (20100420_Bin_31_1, 20110509_Bin_54_1, 20120510_Bin_94_1) were selected for annotation by the IMG pipeline [13]. These are available by the ER comparative analysis system IMG/MER [13] under the submission IDs 208456, 208439, 208438, 208426, 208427 and 208362. The annotations of the MAGs from September 2017 have been published in Grieb *et al.* [28] and are available under the GOLD study ID Gs130320. These annotations were used to reconstruct metabolic pathways.

Ortholog groups

Gene coding sequences were predicted using Prodigal v2.6.3 (default parameters) [30] for the assembled genomes 3300031407_1, 20120524_Bin_102_1, and 20110509_Bin_54_1. Protein orthology was determined using the predicted protein coding sequences in OrthoFinder v2.3.3 (default options) [18]. Orthologous groups were plotted using the UpSetR [15] package implemented for R 3.5.1 [58].

Morphological characterization

Filtered seawater (0.2-3 μm fraction) was hybridized with the Vis6-814 CARD-FISH probe according to Pernthaler *et al.* [48]. Structured illumination microscopy (SIM) images were acquired with a Zeiss ELYRA PS.1 (Carl Zeiss, Germany) system coupled with an iXON897 EM-CCD camera and a 63x plan apochromatic oil immersion objective. Cell dimensions were determined using ACMEtool3 software (M. Zeder; http://www.technobiology.ch/) on images taken with AxioVision v 4.8.2.0 software (Carl Zeiss, Germany) and a AxioImager D2 (Carl Zeiss, Germany) epifluorescence microscope using a Plan-Apochromat 100x/1.40 oil objective.

Results

Taxonomic classification

Based on the 16S rRNA gene sequence analyses, Vis6 formed three clusters - referred to as cluster A, B and C - independent of the phylogenetic tree construction method. A phylogenetic consensus tree is shown in Fig. 1. Sequence identity was above 98.7% within the sequences of clusters A and B, whereas Cluster C was more diverse with sequence identities > 97.4% among sequences of this group. Sequence identities between the clusters A, B and C were between 94.5% and 98.6% (Table S2), suggesting that they all originate from one genus, but comprise different species [67]. A few sequences, targeted by the Vis6-814 probe, were deep branching and with 16S rRNA sequence identities < 94% to clusters A, B and C, and therefore clearly outside the genus.

From the study of Chafee and coworkers [12], three oligotypes could be classified within the genus-level Vis6 group: oligotype_6900, oligotype_6896 and oligotype_6921. The oligotype identifiers from the study of Chafee and coworkers were altered when we analyzed the data together with the sequences from 2016. Oligotype_6900 is identical to oligotype_2940 in Chafee *et al.* [12], which was referred to "*Owenweeksia* related". In our study, this oligotype_6900 was affiliated to Vis6 cluster B, whereas oligotype_6896 and oligotype_6921 were affiliated to Vis6 cluster C. Another oligotype, oligotype_16147, was affiliated to a sequence targeted by the Vis6-814 probe, but was not affiliated to the genus-level Vis6 group. This could have been due to the short sequence length of the representative oligotype sequences. Other oligotypes that have been classified as *Cryomorphaceae* in Chafee *et al.* [12] were either not affiliated to the Vis6 group or could not be stably placed in the 16S rRNA gene tree due to short sequence lengths. MAGs of the Mash-clusters mc_3 and mc_12 included 16S rRNA gene sequences, which were affiliated to cluster B. The 16S rRNA gene sequences from the MAGs from September 2017 were affiliated to cluster C.

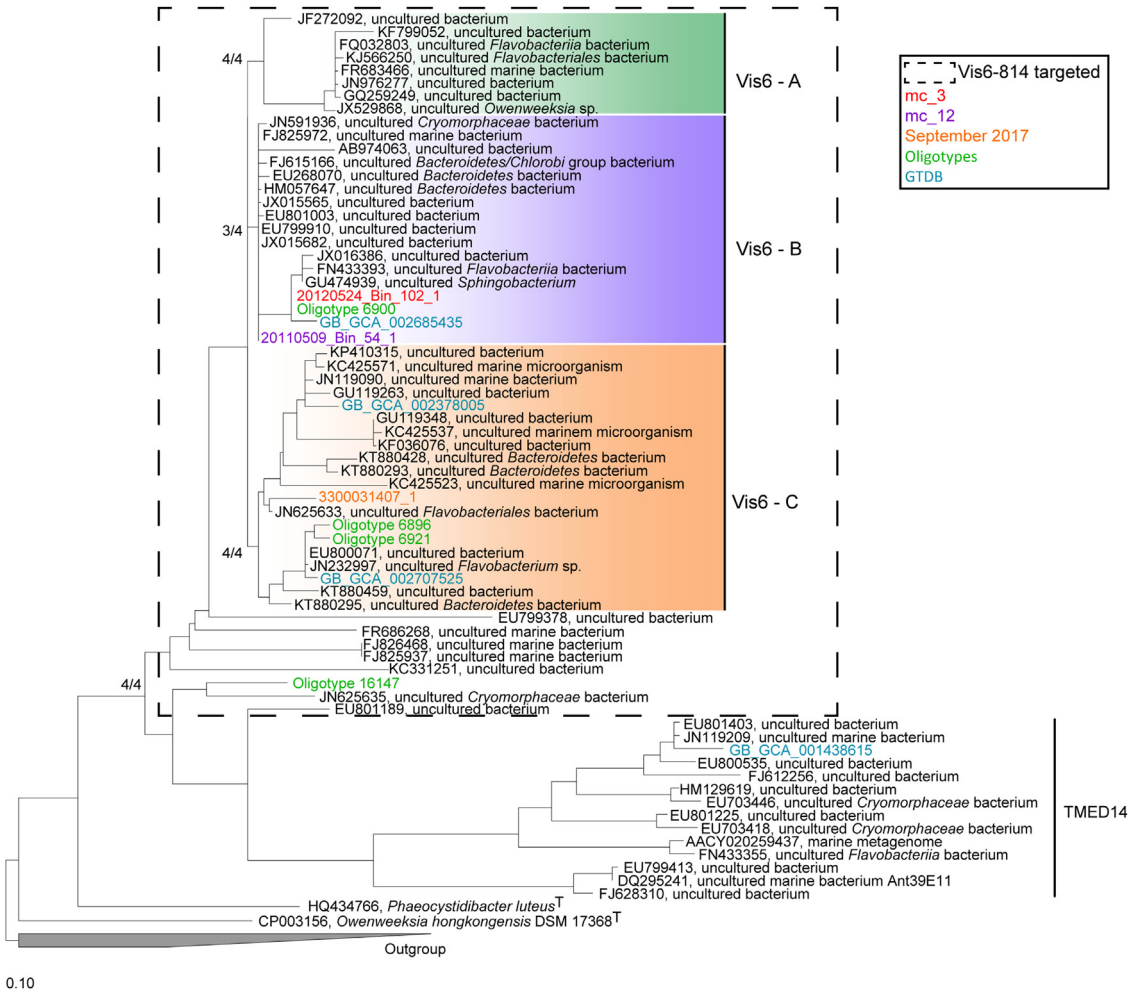

Fig. 1. Consensus tree of the Vis6 group within the *Cryomorphaceae* family, based on 16S rRNA gene sequences from the SILVA database. The numbers indicate how many of the four generated trees, used for the consensus tree, branched at that position. The dashed line indicates sequences targeted by the Vis6-814 probe, including few sequences that are deep-branching between Vis6 and "TMED14". Added to the tree were 16S rRNA gene sequences from MAGs of all three Vis6 candidate species, four oligotypes that were closely related to Vis6 and four closely related 16S rRNA sequences from the GTDB. The classification of GB_GCA_001438615 as TMED14 was adopted from GTDB.

Analyzing the 16S rRNA gene sequences contained in the genomes from GTDB, we found that sequences from genus UBA10364 were affiliated to the Vis6 clusters B and C. The closest neighboring cluster contained the 16S rRNA gene sequences of MAG GCA_001438205.1 from candidate genus TM14 (former "Coccinistipes").

The GTDB genus UBA10364 contained 11 species, of which we calculated the ANI values of each representative sequence to the Vis6 genome sequences from this study (Fig. 2). Sequences from mc_3 and mc_12 have been published in Teeling *et al.* [60] and were therefore considered when the GTDB database was created [46]. Consequently, whole genome comparison showed that a representative sequence of mc_3 (20100303_Bin_80_1) shared 99.9% ANI with UBA_10364 sp002387615 (GCA_002387615.1) and a sequence of mc_12 (20100420_Bin_31_1) shared 98.7% ANI with Sum29DL08_bin30 (GCA_003045825.1). All other species from GTDB share < 95% ANI with our analyzed MAGs. Sequences from September 2017 formed a distinct cluster with no representative from GTDB.

The representative MAGs (from mc_ 3, mc_12 and September 2017) represent three different species based on 77-81% ANI between each other. Within the three species clusters (mc_3, mc_12 and September 2017), ANI values were > 99%. We will refer to MAGs from mc_12 as candidate species 1, to MAGs of mc_3 as candidate species 2 and to MAGs from September 2017 as candidate species 3.

The closest cultured representative of the Vis6 sequences, based on 16S rRNA gene sequence analysis, was *Phaeocystidibacter luteus* with 90% sequence identity (Fig. 1). The closest cultured representative, based on whole genome analysis, was *Owenweeksia hongkongensis* with 50% AAI (Fig. 3).

Temporal distribution

Abundance estimations based on FISH with the genus-specific probe Vis6-814 ([60] and Table S3) indicated a growth pattern of Vis6 which peaked shortly after phytoplankton blooms. This was in particular seen in 2013, but also in 2009 and 2012 (Fig. 4). One exception was 2010, where Vis6-814 FISH counts reached 20% relative abundance already at the beginning of the spring phytoplankton bloom and had a second, smaller peak after the summer bloom. The lowest cell counts within all years (2009-2013, 2016) were on February 2, 2012 with 1.34 $x10^3$ cells per ml seawater. Highest cell counts were on May 22, 2009 with 2.2 $x10^5$ cells per ml seawater.

Based on frequency of partial 16S rRNA reads, oligotype_6900 was the most abundant of the analyzed oligotypes. Its relative abundances showed similar patterns as the results of FISH analyses. The relative abundances of oligotype_6900 increased together with the FISH counts in 2010, 2012 and 2016. Of the analyzed oligotypes, only oligotype_6900 was detected in the first six months of all

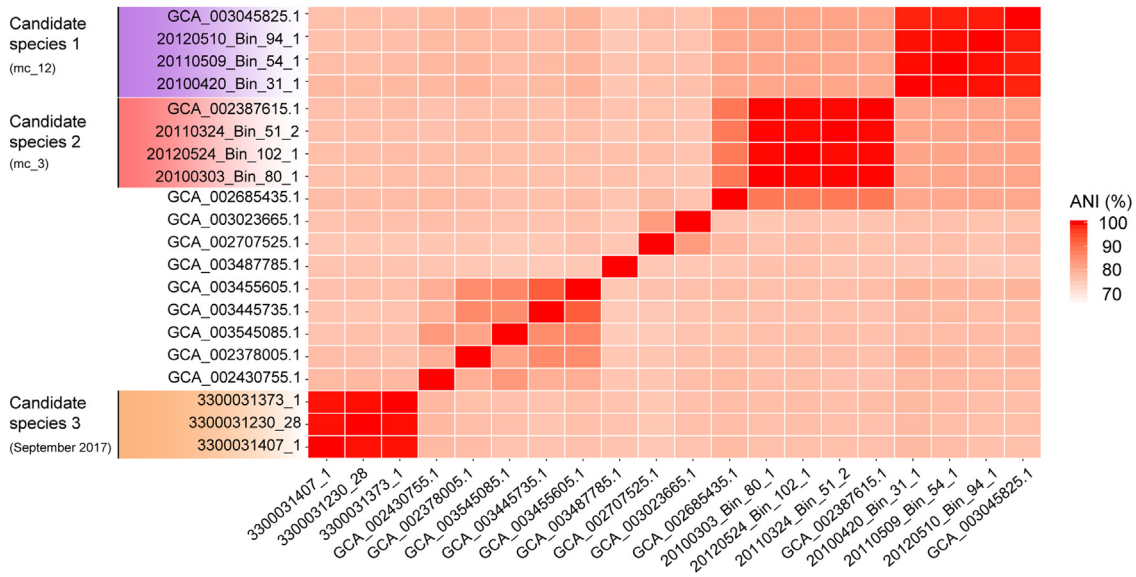

Fig. 2. Average nucleotide identities (ANI) between three representative MAGs of each candidate species and one representative sequence of each species from genus "UBA10364" from GTDB. The ANI between the three candidate species clusters was 77-81%. The ANI within each species cluster was > 99%.

examined years. The other two oligotypes, 6896 and 6921, were detected only later in the years and were therefore only present in the datasets of 2010-2012 where sampling occured throughout the year. Oligotype 6896 was always more abundant than oligotype 6921.

In the analyzed years, metagenomes have only been sequenced during spring. Resulting MAGs were affiliated to candidate species 1 (mc_12) and 2 (mc_3). Within the analyzed metagenomes, candidate species 1 was more abundant in 2010 and 2011, candidate species 2 peaked in 2012 and both were present in nearly equal numbers in 2016. It seems that the oligotype_6900 has a similar abundance pattern as candidate species 1 and 2 combined.

Within the analyzed datasets based on metagenomics read recruitment, we observed that candidate species 1 and 2 only occurred in spring, but not in autumn, whereas candidate species 3 occurred in autumn, but not in spring (Table S4).

Spatial distribution

The global distribution of 16S rRNA gene sequences closely related to Vis6 shows that the majority of samples derived from coastal surface waters (Fig. 5). We observed patterns for a latitudinal separation in different Vis6 clusters, based on 16S rRNA gene amplicons. Sequences affiliated to cluster A were found primarily in polar regions, sequences affiliated to cluster B were found in polar and temperate regions and C showed tendencies towards warmer, temperate zones.

Functional annotation

The estimated genome sizes of the three candidate species of Vis6 ranged between 1.7 and 2.4 Mbp and the GC content ranged between 44% and 47% (Table 1). The number of CAZymes per Mbp was between 17 and 19 and the number of peptidases per Mbp was between 35 and 43.

Based on KEGG annotation, genes for TCA cycle, glycolysis, the non-oxidative part of the pentose phosphate pathway and fatty acid metabolism were present in all three candidate species (Table S5). No differences in presence of analyzed KEGG modules were found between the three candidate species. The annotation of transporters also yielded similar transporters for the three candidate species. Annotated genes encoded lipopolysaccharide export, multidrug efflux pumps, polysaccharide transport, peptide transport, MFS transporters, the Tol biomer transport system and transporter for manganese, iron, zinc, nickel, magnesium and ammonium as well as mechanosensitive channels (Table S6). Candidate species 3 had a higher number of O-antigen ligases and membrane proteins annotated than the two spring species. All three candidate species had a gene encoding for proteorhodopsin. They can possibly replenish oxaloacetate for the TCA cycle by anaplerotic CO_2 fixation. Genes for carbonic anhydrase, converting CO_2 to bicarbonate, using a phosphoenolpyruvate carboxylase, were also detected.

Table 1
Data on genome size, GC content, number of CAZymes and peptidases of the three candidate species.

	Candidate species 1 (mc_12)	Candidate species 2 (mc_3)	Candidate species 3 (September 2017)
Genome size (Mbp)	1.7	1.9	2.4
GC content (%)	45	44	48
CAZymes (GH, PL, CE) / Mbp	17	19	18
Peptidases / Mbp	42	43	35

The total number of CAZymes (PL, CE and GH) per Mbp was almost the same in all three candidate species (17-19 CAZymes per Mbp) (Table S7) as well as the number of GHs (Fig. 6). GH13, encoding mostly alpha-glucan degradation, GH16 and GH17, both encoding ß-glucan degradation, were found in all candidate species. Abundant was also the family GH74 endoglucanases that could indicate a degradation of xyloglucans [2].

Candidate species 3 stood out by the increased presence of GH73, likely cleaving ß-1,4-glycosidic linkage between N-acetylglucosaminyl and N-acetylmuramyl moieties in the carbohydrate backbone of bacterial peptidoglycans [35]. N-acytylglucosaminidase activity has been shown for the GH73 family, but also for the GH3 family which was present in all three candidate species [63]. GH99 (endo-mannosidase), GH100 (invertase) and GH109 (N-Acetyl-galactosaminidase) [35] were annotated in candidate species 3, but absent or only very rare in the other two species. The same applied to GH37, a glycoside hydrolase family coding for trehalases [35]. Trehalase activity is also predicted for GH65, which was annotated in all three species. The family GH46

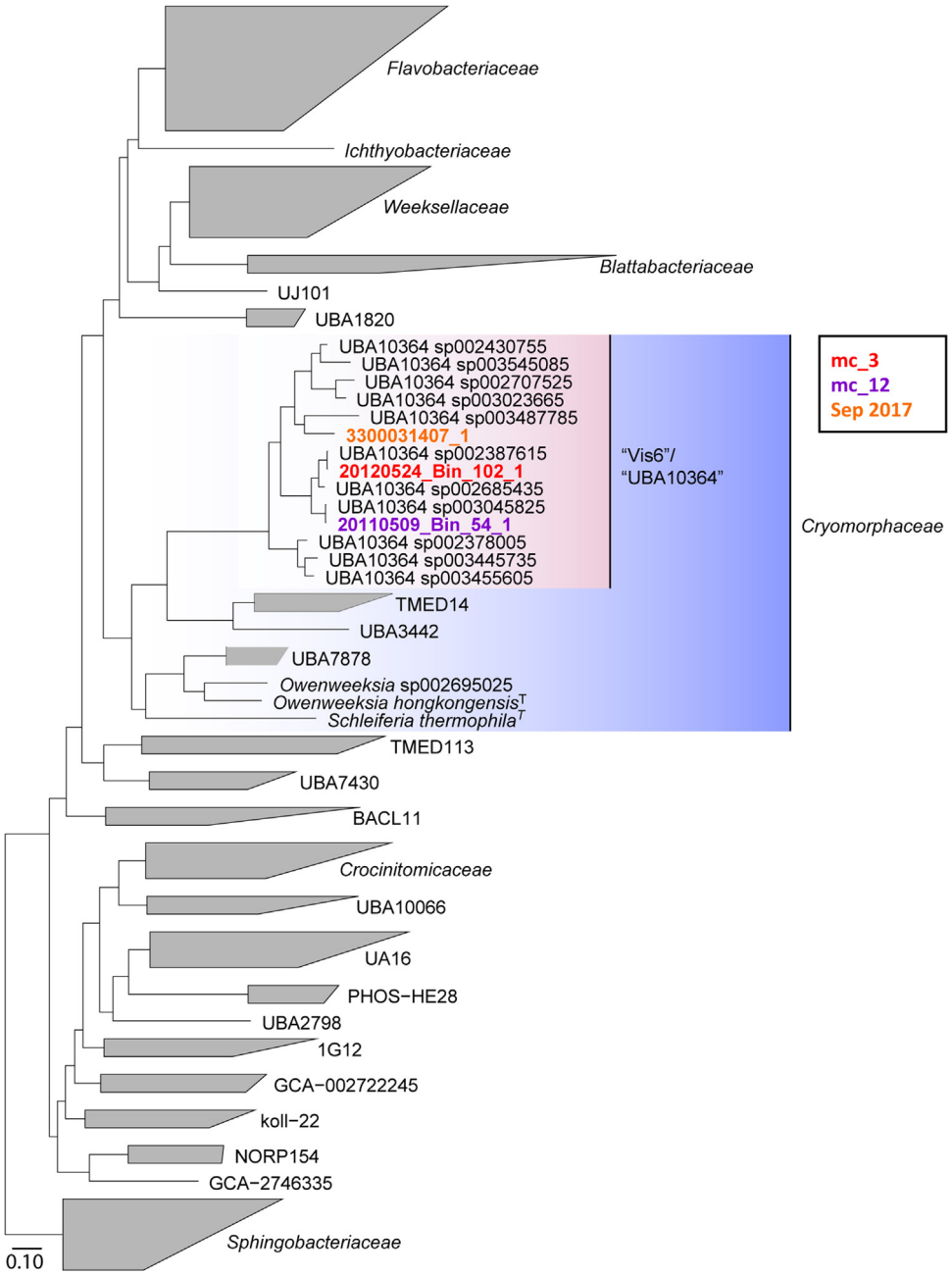

Fig. 3. Phylogenomic reconstruction of Vis6 within the family *Cryomorphaceae* and the order *Flavobacteriales*, based on GTDB. Representative MAGs for the three candidate species, 20120524_Bin_102_1 (mc_3), 20110509_Bin_54_1 (mc_12) and 3300031407_1 (September 2017) were used. The candidate genus described here was previously named UBA10364 in the GTDB.

(chitosan degrading, [35]) was only annotated for candidate species 2.

The number of CBMs per Mbp was lower for candidate species 3 (7 CBM/Mbp) compared to 12 CBMs per Mbp for candidate species 1 and 2 (Figure S1). CBM44 (cellulose and xyloglucan binding) was the most abundant CBM in all three species, followed by CBM50 (peptidoglycan binding). Of all CEs, CE1 (targeting peptidoglycan, xylan and chitin) was most often annotated in all three species (Figure S2). Candidate species 1 and 3 had a CE3, which was not annotated for candidate species 3. Candidate species 3 had only one polysaccharide lyase – PL12, a heparin-sulfate lyase (Figure S3). Candidate species 1 had only PL22 (oligogalacturonate lyase). Candidate species 2 had PL22, PL1_10 and PL9. The annotated GTs were similar between the three candidate species (Figure S4).

Some of the annotated CAZymes were co-localized with each other and with SusC and SusD genes, forming PULs. A putative alpha-glucan degrading PUL, containing GH13 and GH65, was predicted for all three candidate species (Table S8 and Figure S5), as well as a putative beta-glucan degrading PUL, containing GH30, GH17, GH16 and GT4. The latter PUL type is reminiscent of a laminarin degrading PUL with GH30 removing the side chains of laminarin and GH17 degrading the polysaccharides to oligosaccharides [62]. An additional beta-glucan PUL, containing CBM6/GH5_46 and GH16, was annotated in candidate species 1 and 2. Candidate species 3 had two putative PULs, not found in the other two candidate species, of which the first contained GH97 (hydrolyzing a-glycosidic linkages; [31]) and GH37 (hydrolyzing trehalose into glucose; [35]) and the second GH92 and GH20 (containing ß-N-acetylglucosaminidases and lacto-N-biosidases; [35]). No sul-

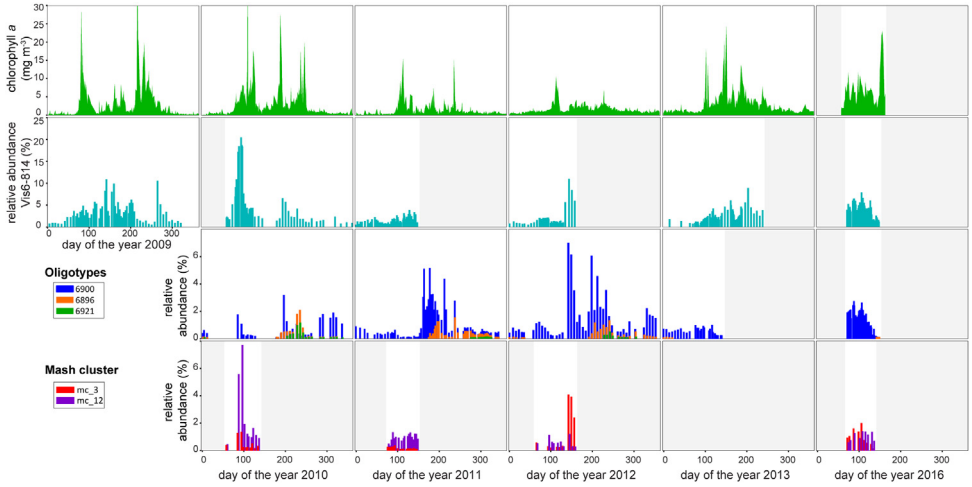

Fig. 4. Data from Helgoland roads of the years 2009-2013 and 2016, showing the temporal distribution of Vis6. From top to bottom: (1) Chlorophyll *a* concentrations, (2) FISH counts based on CARD-FISH with the Vis6-814 probe relative to total cell counts, (3) Relative abundance of oligotypes that were classified as Vis6 in relation to total number of oligotypes, (4) Relative abundance of Vis6 MAGs, calculated by read mapping of metagenome reads against the MAGs of mc_3 and mc_12. Greyed out areas indicate that no data was available.

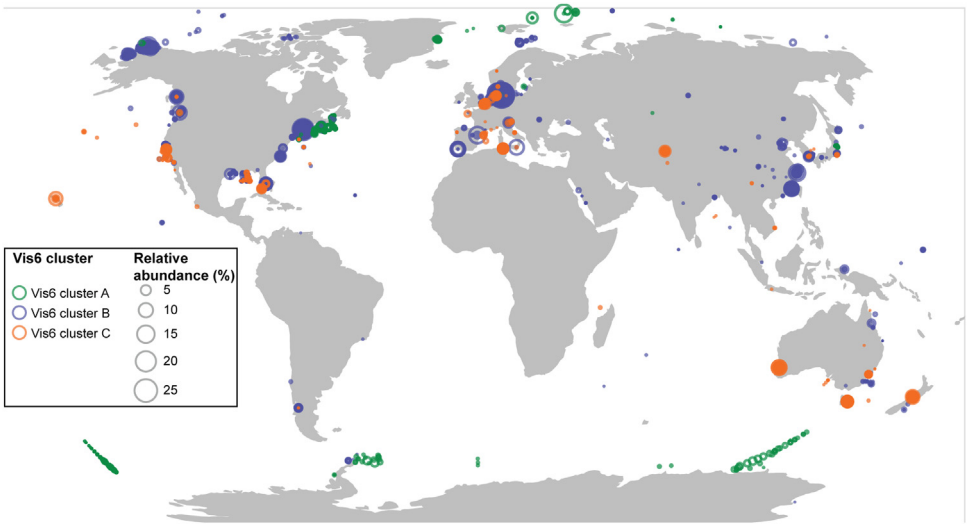

Fig. 5. Global distribution of Vis6, based on IMNGS analysis of 16S rRNA sequences affiliated to the candidate genus. Results are separated by the 16S rRNA gene based clusters A, B and C (compare Fig. 1). The bubble sizes indicate the percentage of matching reads per sequencing run.

fatases were detected in the MAGs of candidate species 3, one sulfatase per genome was annotated for candidate species 2 and about half of the MAGs of the candidate species 1 had a sulfatase encoded (16 out of 38 MAGs). All detected sulfatases were classified as sulfatase family S1 [4].

The peptidase families detected within the genomes of all three candidate species were very similar. Among the most abundant families were metallopeptidases (M1, M9, M16, M23, M38), cysteine peptidases (C26, C44) and serine peptidases (S8, S9, S12, S16, S33, S41, S46, S54) (Figure S6).

The essential proteins for gliding motility of *Bacteroidetes* are suggested to be GldB, GldD, GldH and GldJ in addition to the PorSS/type IX secretion system (GldK, GldL, GldM, GldN, SprA, SprE, SprT) [39]. Genes gldD, H, J, L, M and N were detected for candidate species 1 and 2 (Table S6). Genes for GldH, M, SprA, SprB and SprT were detected for candidate species 3. More than 8 unspecified gliding-motility annotations per MAG of candidate species 1 and 2 were found, which were not annotated in the MAGs of candidate species 3. Few genes involved in surface adhesion [27] were annotated like PKD and additionally FG-GAP for candidate species 3.

Ortholog groups

Half of the ortholog groups were shared by all three candidate species (1111 of 2215) (Fig. 7). The largest proportion of ortholog groups, that were only present in one of the species, was found in candidate species 3 (537). Candidate species 1 and 2 were more similar based on this analysis (sharing 207 groups) and more distant to candidate species 3. Most of the orthologs only present in candidate species 3 were of unknown function (Table S9). The kinases and integrases were notable by their higher presence in candidate species 3 compared to the other two species.

Cell morphology and habitat

Cells targeted by the Vis6-814 probe were rod-shaped (Figure S7). The cell dimensions, measured on four sampling dates in three different years, were all in the same range between 0.9 ± 0.2 μm and 1.25 ± 0.2 μm in length and between 0.48 ± 0.04 μm and 0.62 ± 0.08 μm in width.

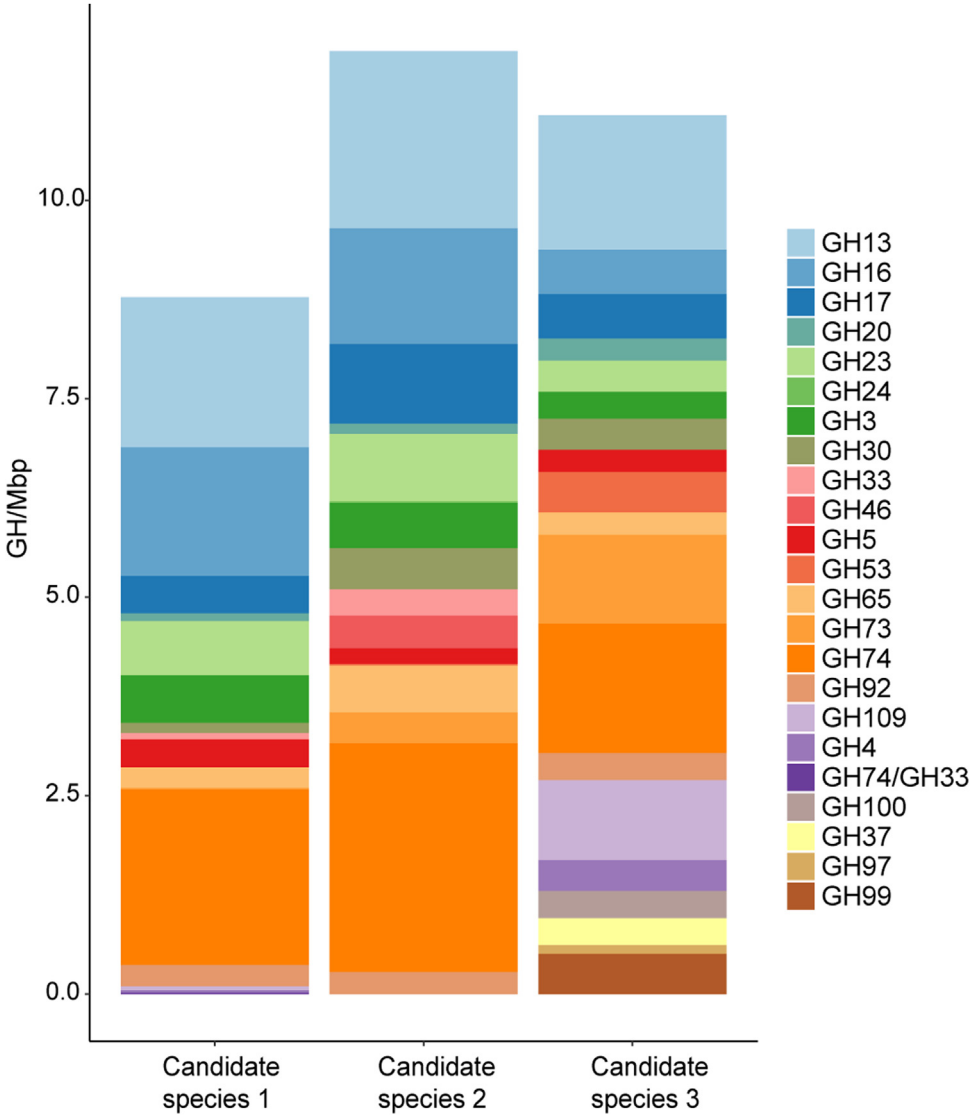

Fig. 6. Average number of glycoside hydrolases (GHs) per Mbp of MAGs from the three candidate species.

Oligotypes affiliated to the candidate genus were found in both free-living (0.2-3 μm) and the particle associated (3-10 μm) fraction, but enriched in the free-living fraction, particularly in spring (Figure S8). oligotype_6921 and oligotype_6896 were enriched in the particle associated fractions in autumn.

Discussion

Based on the polyphasic data, we describe here three candidate species within a new candidate genus of the family *Cryomorphaceae*. The description of the novel candidate genus is based on a high coherence within the MAGs of >99% ANI, and an AAI of 50-51% between the candidate genus and the genome of the closest described relative *Owenweeksia hongkongensis* [52]. As a name for the new candidate genus we suggest *Candidatus* Abditibacter, which alludes to a genus of hidden, rod-shaped bacteria, being nearly always present in the coastal marine surface waters examined, but most of the time not dominant. We suggest the renaming of the genus UBA10364 from GTDB to *Ca.* Abditibacter. The species sp003045825 thereby represents candidate species 1 and sp002387615 represents candidate species 2. Candidate species 3 represents a species not yet present in the GTDB.

Three species of Ca. Abditibacter with seasonal variation

We hypothesize a seasonal alternation of the analyzed *Ca.* Abditibacter species based on the observation that MAGs of candidate species 3 could not be retrieved from the spring metagenomes and MAGs of candidate species 1 and 2 could not be retrieved from the September 2017 metagenome. We named candidate species 1 *Ca.* Abditibacter vernus due to its main detection in spring. Candidate species 2 was named *Ca.* Abditibacter forsetii after Forseti, the god that was worshipped on Helgoland, for the sampling location off the coast of Helgoland. Candidate species 3 was named *Ca.* Abditibacter autumni due to its occurrence in autumn. The connection between the various data used for candidate species description is shown in Table 2.

The two spring species *Ca.* A. forsetii and *Ca.* A. vernus were not distinguishable on 16S rRNA gene sequence level. Together with the oligotype_6900 they were affiliated to the *Ca.* Abditibacter 16S rRNA gene cluster B with > 98.8% sequence identity, suggesting that oligotype_6900 represents both species. This assumption is supported by the observation that the combined abundance patterns of *Ca.* A. forsetii and *Ca.* A. vernus match the abundance patterns of oligotype_6900 and also the microscopic FISH counts from the *Ca.* Abditibacter specific probe Vis6-814.

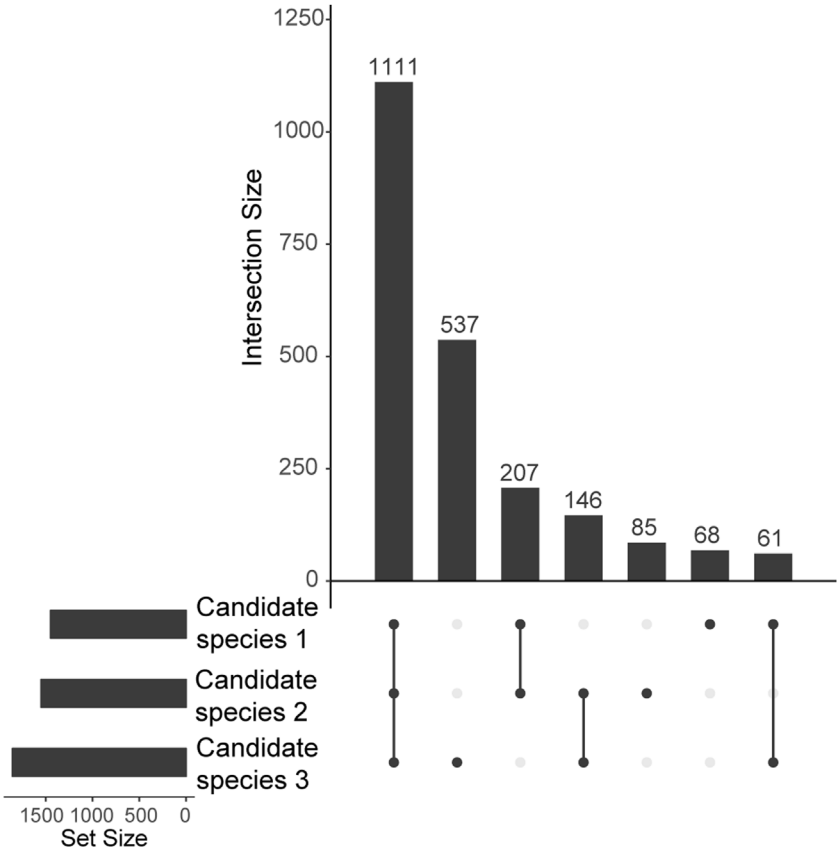

Fig. 7. Number of ortholog groups (COG cluster) of the three candidate species. Shown are the total number of orthologs (Set Size) and the number of orthologs shared (Intersection Size).

Table 2
Sets of data analyzed for the description of the three *Candidatus* Abditibacter species.

Candidate species	*Ca.* Abditibacter	MAGs	GTDB species	16S rRNA cluster
Species 1	*Ca.* A. vernus	mc_12	sp003045825	Vis6 cluster B
Species 2	*Ca.* A. forsetii	mc_3	sp002387615	Vis6 cluster B
Species 3	*Ca.* A. autumni	September 2017	-	Vis6 cluster C

The two oligotypes 6896 and 6921, which are 99.6% identical, were affiliated to the *Ca.* Abditibacter 16S rRNA gene sequence cluster C. In the same cluster C, the 16S rRNA gene sequences from the *Ca.* A. autumni MAGs were grouped. We therefore assume that these two oligotypes represent *Ca.* A. autumni. This corroborates the assumption that *Ca.* A. autumni flourishes in late summer and autumn.

Taxonomic classification within the family Cryomorphaceae

All described members of the *Cryomorphaceae* are either rods or filamentous rods with a GC content ranging from 34-45% [10]. Most described genera showed gliding motility and require elevated salt concentrations for growth. *Ca.* Abditibacter matches these traits by being rod-shaped and dwelling in seawater. Gliding motility may be indicated for *Ca.* Abditibacter based on the detection of yet incomplete set of gliding motility genes, but could also be an evolutionary relict. The GC content is with 44-48% at the higher end of the range described for the *Cryomorphaceae* family.

Prevalence in coastal areas linked to phytoplankton blooms

As the name of the family *Cryomorphaceae* implies, first isolates were retrieved from cold waters, which could indicate an adaption to cold conditions in this family. Indeed, the map of Vis6 16S rRNA gene sequences indicates that in particular cluster A species have been found preferentially in polar regions. Cluster C species seem to have a tendency towards warmer waters and tropical regions, but have also been detected in temperate regions. However, the 16S rRNA dataset provides only a limited view into the global distribution of *Ca.* Abditibacter, since the resolution of 16S rRNA is not sufficient to discriminate between the different *Ca.* Abditibacter species. A distribution analysis based on whole genome sequences of *Ca.* Abditibacter species could probably provide a more detailed picture.

In the initial FISH experiments by Gomez-Pereira *et al.* [25], *Ca.* Abditibacter abundance correlated to chlorophyll *a* concentrations and to total flavobacterial abundance. Our data also shows that *Ca.* Abditibacter abundance peaks simultaneously with or shortly after chlorophyll *a*, which was also seen in other studies where *Ca.* Abditibacter related 16S rRNA gene sequences were retrieved from a diatom dominated phytoplankton bloom [34], from the deep chlorophyll maxima [55] and from a phytoplankton bloom in a mesocosm study [43]. Other isolation sources further indicate that *Ca.* Abditibacter spp. occur mainly in coastal surface waters, but also related to phytoplankton blooms in the North Atlantic [25] and the Southern Ocean [55].

Table 3
Protologue for *Candidatus* Abditibacter species.

Species name	Abditibacter vernus	Abditibacter forsetii	Abditibacter autumni
Genus name	Abditibacter	Abditibacter	Abditibacter
Specific epiphet	vernus	forsetii	autumni
Genus status	Gen. nov.	Gen. nov.	Gen. nov.
Genus etymology	Abditibacter (Ab.di.ti.bac'ter. L. past part. abditus hidden, kept secret, concealed; N.L. masc. n. bacter a rod; N.L. masc. n. Abditibacter a hidden rod	Abditibacter (Ab.di.ti.bac'ter. L. past part. abditus hidden, kept secret, concealed; N.L. masc. n. bacter a rod; N.L. masc. n. Abditibacter a hidden rod	Abditibacter (Ab.di.ti.bac'ter. L. past part. abditus hidden, kept secret, concealed; N.L. masc. n. bacter a rod; N.L. masc. n. Abditibacter a hidden rod
Species status	Sp. nov.	Sp. nov.	Sp. nov.
Species etymology	vernus (ver'nus. L. masc. adj. vernus pertaining to spring, vernal)	forsetii (for.se'ti.i. N.L. gen. n. forsetii of Forseti, a god in Scandinavian mythology that lived on Helgoland)	autumni (au.tum'ni. L. gen. n. autumni of the autumn)
Country of origin	DEU	DEU	DEU
Region of origin	Helgoland	Helgoland	Helgoland
Source of sample	Coastal seawater	Coastal seawater	Coastal seawater
Sampling date	2011-05-09	2012-05-24	2017-09-20
Latitude	54° 11.3' N	54° 11.3' N	54° 11.3' N
Longitude	7° 54.0' E	7° 54.0' E	7° 54.0' E
Designation of the type MAG	20110509_Bin_54_1	20120524_Bin_102_1	3300031407_1
Metagenome accession number	PRJEB28156	PRJEB28156	IMG: 3300031407
Genome size	1725	1829	2488
GC mol%	45.24	44.36	47.52
Relationship to O2	Aerobic	Aerobic	Aerobic
Energy metabolism	Chemoorganotroph	Chemoorganotroph	Chemoorganotroph
Assembly	1 sample	1 sample	1 sample
Sequencing technology	Illumina HiSeq 2500	Illumina HiSeq 2500	Illumina NextSeq-HO
Binning software used	CONCOCT, anvi'o v3	CONCOCT, anvi'o v3	Manual binning, anvi'o v3
Assembly software used	SPAdes v3.10	SPAdes v3.10	SPAdes v3.11
Habitat	Coastal seawater	Coastal seawater	Coastal seawater
Biotic relationship	Free-living	Free-living	Free-living

Free-living polysaccharide and peptide degraders

Based on the annotation of genes encoding transporters for polysaccharides and peptides, and the respective degradation enzymes, we propose for all three candidate species a largely heterotrophic lifestyle based on polysaccharide and peptide utilization. Relatively small genomes (1.7 - 2.4 Mbp) and high peptidase to CAZyme ratios (2-2.5) are typical for heterotrophic bacteria appearing during the degradation of phytoplankton blooms [7,32,66]. The proteorhodopsin gene is more commonly found in planktonic than in algae associated *Bacteroidetes* [66] and likely provides an advantage during phases of substrate limitation [24]. The proton gradient generated by proteorhodopsin energizes the inner membrane, for example for transport so that substrates can be used for anabolism instead of respiration and the TCA cycle can be replenished by anaplerotic bicarbonate fixation [27].

We assume a predominantly free-living lifestyle for *Ca.* Abditibacter based on microscopic observations and 16S rRNA gene analysis. CARD-FISH hybridization on unfractionated seawater did not indicate particle attachment of *Ca.* Abditibacter. Oligotype analysis also showed an enrichment of *Ca.* Abditibacter in the free-living fraction (0.2-3 μm), except for two out of three oligotypes in autumn. This could indicate that some *Ca.* Abditibacter species survive in the particle associated fraction over autumn and winter, though we did not see microscopic evidence in our September sample.

The genomes of *Ca.* Abditibacter species contained PULs for the degradation of the simple storage molecules of phytoplankton. These include alpha-glucans like glycogen in cyanobacteria [3] or beta-glucans like chrysolaminarin in diatoms [5]. The CAZyme repertoires of the two spring species *Ca.* A. forsetii and *Ca.* A. vernus were more alike than the repertoire of *Ca.* A. autumni (Fig. 6). The latter stood out by a larger number of GHs specific for cleaving the β-1,4-glycosidic linkage between N-acetylglucosaminyl (NAG) and N-acetylmuramyl (NAM) moieties in the carbohydrate backbone of bacterial peptidoglycans. Enzymes involved in cell wall hydrolysis are needed for cell division, cell wall rearrangement and also for the recycling of cell walls in sudden carbon depletion [63]. These enzymes could also indicate a utilization of peptidoglycans present in the dissolved and particulate organic matter fraction [6,40].

Sulfatases were not (*Ca.* A. autumni) or rarely (*Ca.* A. vernus and *Ca.* A. forsetii) detected for the described species, suggesting that sulfated polysaccharides are not utilized. Similar to *Formosa* spp. we did not detect mannitol dehydrogenases for *Ca.* Abditibacter which suggests a specialization on chrysolaminarin that does not have mannitol side chains and is preferentially produced by diatoms [62]. A slight correlation of clade "*Owenweeksia*", including a sequence closely related to *Ca.* Abditibacter, to the diatom species *Pseudo-Nitzschia* was found by Needham and Fuhrman [42]. The PUL spectrum of *Ca.* Abditibacter identified in this study suggests that the association with diatoms is caused by chrysolaminarin utilization.

Besides polysaccharides, proteins and peptides seem to be an important carbon source for *Ca.* Abditibacter species, which is not uncommon for *Bacteroidetes* [21]. The metallopeptidases M1, M23, M16 and the serine proteases S9 and S41 were among the most abundant peptidase families in our dataset, which corresponds with findings of Gomez-Pereira *et al.* [26], who analyzed bacteroidetal fosmids retrieved from the North Atlantic Ocean. Peptidases of family M23 lyse peptidoglycans of bacterial cell walls, either as a defensive or feeding mechanism [50,63]. This does not necessarily point to a pathogenic lifestyle as peptidoglycan can be taken up directly from the environment and serve as an energy source [6,61].

Ca. Abditibacter likely plays a significant role in the degradation of organic matter derived in particular from phytoplankton blooms. The 16S rRNA gene sequence analyses in IMNGS shows a rather cosmopolitan occurrence of *Ca.* Abditibacter in coastal and open ocean surface water. The data even hint towards polar occurrence of *Ca.* Abditibacter cluster A and a preference for higher temperatures of *Ca.* Abditibacter cluster C. This hypothesis of different growth temperature optima could also explain the seasonal distribution of the three described *Ca.* Abditibacter species in the temperate

climate of the North Sea and could be tested in the future with isolates.

Description of Ca. *Abditibacter gen. nov*

Candidatus Abditibacter (Ab.di.ti.bac'ter. L. past part. abditus hidden, kept secret, concealed; N.L. masc. n. bacter a rod; N.L. masc. n. Abditibacter a hidden rod)

Members of the genus *Ca.* Abditibacter are psychro- to mesophilic, photo-heterotrophic, marine bacteria, living primarily of peptides and polysaccharides. Three species have been defined within this genus, based on ANI values. *Ca.* Abditibacter cells can be visualized by the FISH probes Vis6-814 and Vis6-871. The genus was detected in densities of $1.34x10^3$ to $2.2x10^5$ cells ml^{-1} in all marine surface water of Helgoland during spring. The GC content of the three species is between 44% and 48%. Cells are rod shaped with an approximate size of 1.1 μm x 0.5 μm. The genus *Ca.* Abditibacter belongs to the family *Cryomorphaceae*, order *Flavobacteriales*, class and phylum *Bacteroidetes*.

Description of Ca. *Abditibacter spp.*

Candidatus Abditibacter vernus (ver'nus. L. masc. adj. vernus pertaining to spring, vernal).

This species was detected in seawater samples from the North Sea during spring phytoplankton blooms in multiple years. Its estimated genome size is 1.7 Mbp with a GC content of 45%.

Candidatus Abditibacter forsetii (for.se'ti.i. N.L. gen. n. forsetii of Forseti, a god in Scandinavian mythology that was worshipped on Helgoland).

This species was detected in seawater samples from the North Sea during spring phytoplankton blooms in multiple years. Its estimated genome size is 1.9 Mbp with a GC content of 44%.

Candidatus Abditibacter autumni (au.tum'ni. L. gen. n. autumni of the autumn). This species was detected in seawater samples from the North Sea in September 2017. Its estimated genome size is 2.4 Mbp with a GC content of 48%.

A tabular overview of the three novel candidate species is summarized in Table 3.

Declarations of interest

none.

Acknowledgements

We thank Aharon Oren, The Hebrew University of Jerusalem, for his advice on bacterial nomenclature. We acknowledge Laura Zeugner and Andreas Ellrott for expert technical assistance with microscopy. This work was supported by the German Research Foundation (DFG) through the project 'Proteogenomics Of Marine Polysaccharide Utilization' (POMPU; DFG FOR2406), the U.S. Department of Energy Joint Genome Institute (Contract No. DE-AC02-05CH11231), and the Max Planck Society.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.syapm.2020.126088.

References

[1] Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25 (17), 3389–3402.

[2] Arnal, G., Stogios, P.J., Asohan, J., Attia, M.A., Skarina, T., Viborg, A.H., Henrissat, B., Savchenko, A., Brumer, H. (2019) Substrate specificity, regiospecificity, and processivity in glycoside hydrolase family 74. J Biol Chem 294 (36), 13233–13247.

[3] Ball, S.G., Morell, M.K. (2003) From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. Annu. Rev. Plant Biol. 54, 207–233.

[4] Barbeyron, T., Brillet-Guéguen, L., Carré, W., Carrière, C., Caron, C., Czjzek, M., Hoebeke, M., Michel, G. (2016) Matching the diversity of sulfated biomolecules: creation of a classification database for sulfatases reflecting their substrate specificity. PLOS One 11 (10), e0164846.

[5] Beattie, A., Percival, E., Hirst, E.L. (1961) Studies on metabolism of Chrysophyceae - comparative structural investigations on leucosin (chrysolaminarin) separated from diatoms and laminarin from brown algae. Biochem. J. 79 (3), p. 531-&.

[6] Benner, R., Kaiser, K. (2003) Abundance of amino sugars and peptidoglycan in marine particulate and dissolved organic matter. Limnol. Oceanogr. 48 (1), 118–128.

[7] Bennke, C.M., Krüger, K., Kappelmann, L., Huang, S., Gobet, A., Schüler, M., Barbe, V., Fuchs, B.M., Michel, G., Teeling, H., Amann, R.I. (2016) Polysaccharide utilisation loci of Bacteroidetes from two contrasting open ocean sites in the North Atlantic. Environ. Microbiol. 18 (12), 4456–4470.

[8] Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., Schulz, F., Jarett, J., Rivers, A.R., Eloe-Fadrosh, E.A., Tringe, S.G., Ivanova, N.N., Copeland, A., Clum, A., Becraft, E.D., Malmstrom, R.R., Birren, B., Podar, M., Bork, P., Weinstock, G.M., Garrity, G.M., Dodsworth, J.A., Yooseph, S., Sutton, G., Glöckner, F.O., Gilbert, J.A., Nelson, W.C., Hallam, S.J., Jungbluth, S.P., Ettema, T.J.G., Tighe, S., Konstantinidis, K.T., Liu, W.-T., Baker, B.J., Rattei, T., Eisen, J.A., Hedlund, B., McMahon, K.D., Fierer, N., Knight, R., Finn, R., Cochrane, G., Karsch-Mizrachi, I., Tyson, G.W., Rinke, C., Genome Standards, C., Lapidus, A., Meyer, F., Yilmaz, P., Parks, D.H., Eren, A.M., Schriml, L., Banfield, J.F., Hugenholtz, P., Woyke, T. (2017) Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. 35 (8), 725–731.

[9] Bowman, J.P., Mancuso, C., Nichols, C.M., Gibson, J.A.E. (2003) Algoriphagus ratkowskyi gen. nov., sp nov., Brumimicrobium glaciale gen. nov., sp nov., Cryomorpha ignava gen. nov., sp nov and Crocinitomix catalasitica gen. nov., sp nov., novel flavobacteria isolated from various polar habitats. Int. J. Syst. Evol. Microbiol. 53, 1343–1355.

[10] Bowman, J.P. (2014) The family Cryomorphaceae, in The prokaryotes: other major lineages of Bacteria and the Archaea. In: Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F. (Eds.), Springer Berlin Heidelberg: Berlin, Heidelberg, pp. 539–550.

[11] Buchan, A., LeCleir, G.R., Gulvik, C.A., Gonzalez, J.M. (2014) Master recyclers: features and functions of bacteria associated with phytoplankton blooms. Nat. Rev. Microbiol. 12 (10), 686–698.

[12] Chafee, M., Fernandez-Guerra, A., Buttigieg, P.L., Gerdts, G., Eren, A.M., Teeling, H., Amann, R.I. (2018) Recurrent patterns of microdiversity in a temperate coastal marine environment. ISME Journal 12 (1), 237–252.

[13] Chen, I.A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., Huntemann, M., Varghese, N., White, J.R., Seshadri, R., Smirnova, T., Kirton, E., Jungbluth, S.P., Woyke, T., Eloe-Fadrosh, E.A., Ivanova, N.N., Kyrpides, N.C. (2019) IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. Nucleic Acids Res 47 (D1), D666–D677.

[14] Connon, S.A., Giovannoni, S.J. (2002) High-Throughput Methods for Culturing Microorganisms in Very-Low-Nutrient Media Yield Diverse New Marine Isolates. Appl. Environ. Microbiol. 68 (8), 3878–3885.

[15] Conway, J.R., Lex, A., Gehlenborg, N. (2017) UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics 33 (18), 2938–2940.

[16] Eddy, S.R. (2011) Accelerated Profile HMM Searches. PLOS Computational Biology 7 (10), e1002195.

[17] Eilers, H., Pernthaler, J., Glöckner, F.O., Amann, R. (2000) Culturability and In situ abundance of pelagic bacteria from the North Sea. Appl. Environ. Microbiol. 66 (7), 3044–3051.

[18] Emms, D.M., Kelly, S. (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 16 (1), 157.

[19] Eren, A.M., Maignien, L., Sul, W.J., Murphy, L.G., Grim, S.L., Morrison, H.G., Sogin, M.L. (2013) Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. Methods Ecol. Evol. 4 (12), 1111–1119.

[20] Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., Delmont, T.O. (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3, e1319.

[21] Fernández-Gómez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S.G., González, J.M., Pedrós-Alió, C. (2013) Ecology of marine Bacteroidetes: a comparative genomics approach. ISME Journal 7 (5), 1026–1037.

[22] Field, C.B., Behrenfeld, M.J., Randerson, J.T., Falkowski, P. (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. Science 281 (5374), 237–240.

[23] Francis, T.B., Krüger, K., Fuchs, B.M., Teeling, H., Amann, R.I. (2019) Candidatus Prosiliicoccus vernus, a spring phytoplankton bloom associated member of the Flavobacteriaceae. Syst. Appl. Microbiol. 42 (1), 41–53.

[24] Gómez-Consarnau, L., Akram, N., Lindell, K., Pedersen, A., Neutze, R., Milton, D.L., González, J.M., Pinhassi, J. (2010) Proteorhodopsin phototrophy promotes survival of marine bacteria during starvation. PLoS. Biol. 8 (4), e1000358.

[25] Gomez-Pereira, P.R., Fuchs, B.M., Alonso, C., Oliver, M.J., van Beusekom, J.E.E., Amann, R. (2010) Distinct flavobacterial communities in contrasting water masses of the North Atlantic Ocean. ISME Journal 4 (4), 472–487.
[26] Gomez-Pereira, P.R., Schuler, M., Fuchs, B.M., Bennke, C., Teeling, H., Waldmann, J., Richter, M., Barbe, V., Bataille, E., Glockner, F.O., Amann, R. (2012) Genomic content of uncultured Bacteroidetes from contrasting oceanic provinces in the North Atlantic Ocean. Environ. Microbiol. 14 (1), 52–66.
[27] González, J.M., Fernández-Gómez, B., Fernàndez-Guerra, A., Gómez-Consarnau, L., Sánchez, O., Coll-Lladó, M., del Campo, J., Escudero, L., Rodríguez-Martínez, R., Alonso-Sáez, L., Latasa, M., Paulsen, I., Nedashkovskaya, O., Lekunberri, I., Pinhassi, J., Pedrós-Alió, C. (2008) Genome analysis of the proteorhodopsin-containing marine bacterium Polaribacter sp. MED152 (Flavobacteria). PNAS 105 (25), 8724–8729.
[28] Grieb, A., Bowers, Robert M., Oggerin, M., Goudeau, D., Lee, J., Malmstrom, Rex R., Woyke, T., Fuchs, Bernhard M. (2020) A Pipeline for Targeted Metagenomics of Environmental Bacteria. Microbiome 8 (February (1)), 21, http://dx.doi.org/10.1186/s40168-020-0790-7.
[29] Hahnke, R.L., Bennke, C.M., Fuchs, B.M., Mann, A.J., Rhiel, E., Teeling, H., Amann, R., Harder, J. (2015) Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea. Environ. Microbiol. 17 (10), 3515–3526.
[30] Hyatt, D., Chen, G.-L., Locascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11, 119.
[31] Kitamura, M., Okuyama, M., Tanzawa, F., Mori, H., Kitago, Y., Watanabe, N., Kimura, A., Tanaka, I., Yao, M. (2008) Structural and functional analysis of a glycoside hydrolase family 97 enzyme from Bacteroides thetaiotaomicron. J. Biol. Chem. 283 (52), 36328–36337.
[32] Krüger, K., Chafee, M., Ben Francis, T., Glavina del Rio, T., Becher, D., Schweder, T., Amann, R.I., Teeling, H. (2019) In marine Bacteroidetes the bulk of glycan degradation during algae blooms is mediated by few clades using a restricted set of genes. ISME J.
[33] Lagkouvardos, I., Joseph, D., Kapfhammer, M., Giritli, S., Horn, M., Haller, D., Clavel, T. (2016) IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. Sci. Rep. 6, 33721.
[34] Liu, M., Xiao, T., Sun, J., Wei, H., Wu, Y., Zhao, Y., Zhang, W. (2013) Bacterial community structures associated with a natural spring phytoplankton bloom in the Yellow Sea, China. Deep Sea Research Part II: Topical Studies in Oceanography 97, 85–92.
[35] Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42(Database issue, D490–D495.
[36] Love, M.I., Huber, W., Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15 (12), 550.
[37] Lucas, J., Wichels, A., Teeling, H., Chafee, M., Scharfe, M., Gerdts, G. (2015) Annual dynamics of North Sea bacterioplankton: seasonal variability superimposes short-term variation. FEMS Microbiol. Ecol. 91 (9), fiv099.
[38] Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S., Hermann, S., Jost, R., Konig, A., Liss, T., Lussmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., Schleifer, K.H. (2004) ARB: a software environment for sequence data. Nucleic Acids Res. 32 (4), 1363–1371.
[39] McBride, M.J., Zhu, Y. (2013) Gliding motility and Por secretion system genes are widespread among members of the phylum Bacteroidetes. J. Bacteriol. 195 (2), 270–278.
[40] McCarthy, M.D., Hedges, J.I., Benner, R. (1998) Major bacterial contribution to marine dissolved organic nitrogen. Science 281 (5374), 231–234.
[41] Mukherjee, S., Stamatis, D., Bertsch, J., Ovchinnikova, G., Katta, H.Y., Mojica, A., Chen, I.A., Kyrpides, N.C., Reddy, T. (2019) Genomes OnLine database (GOLD) v.7: updates and new features. Nucleic Acids Res 47 (D1), D649–D659.
[42] Needham, D.M., Fuhrman, J.A. (2016) Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. Nat. Microbiol. 1, 16005.
[43] Newbold, L.K., Oliver, A.E., Booth, T., Tiwari, B., DeSantis, T., Maguire, M., Andersen, G., van der Gast, C.J., Whiteley, A.S. (2012) The response of marine picoplankton to ocean acidification. Environ. Microbiol. 14 (9), 2293–2307.
[44] Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., Phillippy, A.M. (2016) Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol. 17 (1), 132.
[45] Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 25 (7), 1043–1055.
[46] Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A., Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat. Biotechnol. 36 (10), p. 996-+.
[47] Peplies, J., Kottmann, R., Ludwig, W., Glockner, F.O. (2008) A standard operating procedure for phylogenetic inference (SOPPI) using (rRNA) marker genes. Syst. Appl. Microbiol. 31 (4), 251–257.
[48] Pernthaler, A., Pernthaler, J., Amann, R. (2002) Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. Appl. Environ. Microbiol. 68 (6), 3094–3101.
[49] Pinhassi, J., Sala, M.M., Havskum, H., Peters, F., Guadayol, Ò., Malits, A., Marrasé, C. (2004) Changes in bacterioplankton composition under different phytoplankton regimens. Appl. Environ. Microbiol. 70 (11), 6753–6766.
[50] Rawlings, N.D., Barrett, A.J., Bateman, A. (2010) MEROPS: the peptidase database. Nucleic Acids Res. 38, D227–D233.
[51] Reintjes, G., Fuchs, B.M., Scharfe, M., Wiltshire, K.H., Amann, R., Arnosti, C. (2020) Short-term changes in polysaccharide utilization mechanisms of marine bacterioplankton during a spring phytoplankton bloom. Environ. Microbiol. 22 (5), 1884–1900, http://dx.doi.org/10.1111/1462-2920.14971.
[52] Rodriguez-R, L., Konstantinidis, K. 2014 Bypassing cultivation to identify bacterial species, Microbe Magazine.
[53] Rodriguez-R, L.M., Konstantinidis, K.T. (2016) The enveomics collection: A toolbox for specialized analyses of microbial genomes and metagenomes. PeerJ Preprints 4, e1900v1.
[54] Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics 30 (14), 2068–2069.
[55] Singh, S.K., Kotakonda, A., Kapardar, R.K., Kankipati, H.K., Sreenivasa Rao, P., Sankaranarayanan, P.M., Vetaikorumagan, S.R., Gundlapally, S.R., Nagappa, R., Shivaji, S. (2015) Response of bacterioplankton to iron fertilization of the Southern Ocean, Antarctica. Front. Microbiol. 6, 863.
[56] Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30 (9), 1312–1313.
[57] Stubbins, A., Dittmar, T. 2014 Dissolved organic matter in aquatic systems, in Treatise on Geochemistry, Second Edition, Elsevier, pp. , 125–156.
[58] Team, R.C. 2018 R foundation for statistical computing, R: A Language and Environment for Statistical Computing, Vienna, Austria: 2014.
[59] Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., Kassabgy, M., Huang, S.X., Mann, A.J., Waldmann, J., Weber, M., Klindworth, A., Otto, A., Lange, J., Bernhardt, J., Reinsch, C., Hecker, M., Peplies, J., Bockelmann, F.D., Callies, U., Gerdts, G., Wichels, A., Wiltshire, K.H., Glockner, F.O., Schweder, T., Amann, R. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. Science 336 (6081), 608–611.
[60] Teeling, H., Fuchs, B.M., Bennke, C.M., Kruger, K., Chafee, M., Kappelmann, L., Reintjes, G., Waldmann, J., Quast, C., Glockner, F.O., Lucas, J., Wichels, A., Gerdts, G., Wiltshire, K.H., Amann, R.I. (2016) Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. eLife 5, e11888.
[61] Uehara, T., Suefuji, K., Jaeger, T., Mayer, C., Park, J.T. (2006) MurQ etherase is required by Escherichia coli in order to metabolize anhydro-N-acetylmuramic acid obtained either from the environment or from its own cell wall. J. Bacteriol. 188 (4), 1660–1662.
[62] Unfried, F., Becker, S., Robb, C.S., Hehemann, J.-H., Markert, S., Heiden, S.E., Hinzke, T., Becher, D., Reintjes, G., Krüger, K., Avcı, B., Kappelmann, L., Hahnke, R.L., Fischer, T., Harder, J., Teeling, H., Fuchs, B., Barbeyron, T., Amann, R.I., Schweder, T. (2018) Adaptive mechanisms that provide competitive advantages to marine bacteroidetes during microalgal blooms. ISME Journal 12 (12), 2894–2906.
[63] Vermassen, A., Leroy, S., Talon, R., Provot, C., Popowska, M., Desvaux, M. (2019) Cell wall hydrolases in bacteria: insight on the diversity of cell wall amidases, glycosidases and peptidases toward peptidoglycan. Front. Microbiol. 10 (331).
[64] West, N.J., Obernosterer, I., Zemb, O., Lebaron, P. (2008) Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. Environ. Microbiol. 10 (3), 738–756.
[65] Wiltshire, K.H. 2016 Hydrochemistry at time series station Helgoland Roads, North Sea, in 2013, PANGAEA.
[66] Xing, P., Hahnke, R.L., Unfried, F., Markert, S., Huang, S., Barbeyron, T., Harder, J., Becher, D., Schweder, T., Glöckner, F.O., Amann, R.I., Teeling, H. (2015) Niches of two polysaccharide-degrading Polaribacter isolates from the North Sea during a spring diatom bloom. ISME Journal 9 (6), 1410–1422.
[67] Yarza, P., Yilmaz, P., Pruesse, E., Glockner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., Euzeby, J., Amann, R., Rossello-Mora, R. (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12 (9), 635–645.