

A taxonomic framework for emerging groups of ecologically important marine gammaproteobacteria based on the reconstruction of evolutionary relationships using genome-scale data

Stefan Spring^{1*}, Carmen Scheuner¹, Markus Göker¹ and Hans-Peter Klenk^{1,2}

¹ Department Microorganisms, Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany, ² School of Biology, Newcastle University, Newcastle upon Tyne, UK

OPEN ACCESS

Edited by:

Marcelino T. Suzuki,
Sorbonne Universities (UPMC) and
Centre National de la Recherche
Scientifique, France

Reviewed by:

Fabiano Thompson,
Federal University of Rio de Janeiro,
Brazil
Jose M. Gonzalez,
University of La Laguna, Spain

*Correspondence:

Stefan Spring,
Department Microorganisms, Leibniz
Institute DSMZ – German Collection of
Microorganisms and Cell Cultures,
Inhoffenstr. 7 B, 38124 Braunschweig,
Germany
ssp@dsMZ.de

Specialty section:

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

Received: 17 January 2015

Accepted: 21 March 2015

Published: 09 April 2015

Citation:

Spring S, Scheuner C, Göker M and
Klenk H-P (2015) A taxonomic
framework for emerging groups of
ecologically important marine
gammaproteobacteria based on the
reconstruction of evolutionary
relationships using genome-scale
data. *Front. Microbiol.* 6:281.
doi: 10.3389/fmicb.2015.00281

In recent years a large number of isolates were obtained from saline environments that are phylogenetically related to distinct clades of oligotrophic marine gammaproteobacteria, which were originally identified in seawater samples using cultivation independent methods and are characterized by high seasonal abundances in coastal environments. To date a sound taxonomic framework for the classification of these ecologically important isolates and related species in accordance with their evolutionary relationships is missing. In this study we demonstrate that a reliable allocation of members of the oligotrophic marine gammaproteobacteria (OMG) group and related species to higher taxonomic ranks is possible by phylogenetic analyses of whole proteomes but also of the RNA polymerase beta subunit, whereas phylogenetic reconstructions based on 16S rRNA genes alone resulted in unstable tree topologies with only insignificant bootstrap support. The identified clades could be correlated with distinct phenotypic traits illustrating an adaptation to common environmental factors in their evolutionary history. Genome wide gene-content analyses revealed the existence of two distinct ecological guilds within the analyzed lineage of marine gammaproteobacteria which can be distinguished by their trophic strategies. Based on our results a novel order within the class *Gammaproteobacteria* is proposed, which is designated *Cellvibrionales* ord. nov. and comprises the five novel families *Cellvibrionaceae* fam. nov., *Haliaceae* fam. nov., *Microbulbiferaceae* fam. nov., *Porticoccaceae* fam. nov., and *Spongiibacteraceae* fam. nov.

Keywords: bacterioplankton, trophic guilds, oligotroph, K-selection, phylogenomics

Introduction

For several years an increasing effort in the characterization of bacteria from marine environments led to the isolation of a large number of gammaproteobacteria that belong to phylogenetic groups originally defined by 16S rRNA gene sequences retrieved from environmental samples. Many of

these novel isolates can be allocated to the oligotrophic marine gammaproteobacteria (OMG) group as defined by Cho and Giovannoni (2004). The five distinct clades of the OMG group (OM60/NOR5, OM182, SAR92, KI89A, and BD1-7) are phylogenetically related, but do not form a robust monophyletic lineage in 16S rRNA gene based trees. Consequently, it seems likely that at least some of these clades have evolved independently from each other as adaptation to oligotrophic conditions encountered in seawater. Within the OMG group members of the OM60/NOR5 and SAR92 clades are of special interest, because cultivation-independent surveys have shown that these bacteria can reach a high seasonal abundance especially in coastal environments (Stingl et al., 2007; Yan et al., 2009). In addition, genome analyses revealed the potential of harvesting light for mixotrophic growth in several members of the OMG group, which made these isolates interesting study objects for further research (e.g., Stingl et al., 2007; Spring and Riedel, 2013). Genes involved in the complex pathway of aerobic anoxygenic photosynthesis were so far only found among members of the OM60/NOR5 clade (Spring et al., 2013, 2014), whereas genes for the synthesis of proteorhodopsin, a light-driven proton pump, seems to be more widely spread and were identified in the genomes of representatives of the SAR92 (Courties et al., 2014), BD1-7 (Oh et al., 2010a), and OM60/NOR5 clades (Jang et al., 2011b).

Currently, most of the novel gammaproteobacterial strains that were obtained by high-throughput cultivation approaches and have been affiliated to the OMG group were not described as novel species and thus remain unclassified (Cho and Giovannoni, 2004). On the other hand, some isolates that were recently described as novel species were either not allocated to higher taxonomic ranks at the family and order level or their assignment to higher order taxa is inconsistent with the phylogenetic relationships among the *Gammaproteobacteria* as shown for example in recent trees (although without branch support) distributed by the all-species living tree project (Yarza et al., 2008; release LTPs115, <http://www.arb-silva.de/projects/living-tree/>). This leads to a confusing classification scheme of environmental important marine gammaproteobacteria that is not in line with the presumable evolutionary relationships. Part of the confusion may have been caused by a close relationship of recently isolated strains affiliated to the OMG group with species regarded as non-oligotrophic and belonging to established genera like *Cellvibrio* and *Microbulbifer*, which have been originally affiliated to the families *Pseudomonadaceae* and *Alteromonadaceae*, respectively. Consequently, unclassified members of the OMG group are intermingled with species that have been affiliated with several different orders of *Gammaproteobacteria* leading to a classification scheme comprising various paraphyletic taxa.

However, the definition of meaningful higher taxonomic units comprising members of environmental important clades of the OMG group and related species could be of great value for the cultivation-independent analyses of microbial populations in marine environments (e.g., Ruff et al., 2014). Possible applications include the automatic classification of ribosomal RNA gene amplicon reads obtained by high-throughput sequencing using curated databases (e.g., SILVAngs, <https://www.arb-silva.de/ngs/#about:>) or the correlation of group-specific oligonucleotide probes used for *in situ* hybridization experiments with the phenotypic characteristics associated with well-defined microbial taxa. For this reason the aim of this study was to define a taxonomic framework for major clades of the OMG group and related described species that is based on the unveiled evolutionary relationships among these bacteria and thus less arbitrary than the current classification.

It was shown previously, that the identification of clades within the *Gammaproteobacteria* is difficult based solely on comparative analyses of 16S rRNA gene sequences due to the limited number of informative characters and the slow evolutionary rate of this molecule. Especially, if a large number of closely related species is compared the obtained branching patterns are often unstable, as indicated by a low bootstrap support of distinct tree topologies (e.g., Yamamoto and Harayama, 1998; Cho and Giovannoni, 2004; Spring et al., 2013). In contrast, housekeeping genes encoding large proteins seem to be more promising for phylogenetic analyses because calculations are based on a larger dataset and analyses of protein sequences are less prone to the nucleotide compositional bias seen in 16S rRNA (Wu and Eisen, 2008; Mulet et al., 2010). Among the known housekeeping proteins the RNA polymerase beta subunit encoded by the *rpoB* gene seems to be the most promising candidate as universal phylogenetic marker, especially for the classification of gammaproteobacteria. Initially, the potential application of the RpoB protein as reliable phylogenetic marker was demonstrated for bacteria by Rowland et al. (1993) and for Archaea by Klenk and Zillig (1994). Later it was shown by Mollet et al. (1997) that phylogenetic trees of *Enterobacteriaceae* based on *rpoB* genes were more compatible with their accepted classification than those obtained with 16S rRNA. Several further studies have demonstrated that *rpoB* genes can be used as a good proxy of the respective genome sequences, which is indicated for example by a strong correlation between the G + C content of the *rpoB* gene and the corresponding genome (Adékambi et al., 2009). Above all, among several tested housekeeping proteins the phylogenetic signal of the RpoB protein shows the highest correlation with the corresponding amino acid identity values of the respective genomes (Konstantinidis and Tiedje, 2005).

With the increasing number of genome sequences that became available for members of the OMG group and related taxa, especially in the course of large-scale sequencing projects (Wu et al., 2009; Yooseph et al., 2010; Kyrpides et al., 2014), a more advanced reconstruction of evolutionary relationships based on genome-scale data became possible (Delsuc et al., 2005; Klenk and Göker, 2010), thereby avoiding stochastic errors related to the use of single genes (Wu and Eisen, 2008). Initially, this approach was questioned by an assumed high incidence of lateral gene transfer between distantly related taxa of prokaryotes, which would lead to a substantial loss of phylogenetic signal inherited by genome sequences (e.g., Gogarten et al., 2002). However, in the case of *Gammaproteobacteria* it was clearly shown that a high phylogenetic concordance exists among a large number of gene families, hence it is possible to infer reliable and robust tree topologies from genome-scale data (Lerat et al., 2003). In following studies the superiority of using proteome data (i.e., concatenated

protein alignments) over single genes for the phylogenetic resolution of *Gammaproteobacteria* was asserted (Wu and Eisen, 2008; Williams et al., 2010).

Based on the reasons given above we used in the present study datasets based on the complete RpoB protein and whole proteomes besides a thorough analysis of the available 16S rRNA gene sequences in order to unravel the evolutionary relationships among members of the OMG group and related gammaproteobacteria.

As a result of our study the novel order *Cellvibrionales* is proposed to accommodate recently isolated marine gammaproteobacteria belonging to the OM60/NOR5, SAR92, and BD1-7 clades of the OMG group as well as related described species with a so far absent or misleading classification into higher order taxa.

Materials and Methods

Phylogenetic Analyses Based on Single Molecules

Comprehensive datasets of 16S rRNA gene sequences of *Gammaproteobacteria* comprising cultured and uncultured members of the OMG group were derived from the alignment included in the non-redundant version of the SILVA Ref database (Quast et al., 2013; SSU Ref NR 99 release 119). Firstly, based on the manually curated trees included in the SILVA Ref NR 119 (<http://www.arb-silva.de/download/arb-files/>) and LTPs 115 (<http://www.arb-silva.de/projects/living-tree/>) databases a range of relevant type strains with a recognizable relationship to major clades of the OMG group were determined. Then, additional outgroup sequences were selected that represent type genera of all major families and orders within the class *Gammaproteobacteria*. Whenever possible only 16S rRNA gene sequences of type strains of type species with validly published names as given in the List of Prokaryotic Names with Standing in Nomenclature (LPSN; Parte, 2014) were used in order to avoid the inclusion of accidentally misclassified strains. The number of outgroup sequences representing a distinct family was limited to a maximum number of six to keep the calculation time required for bootstrap analyses practicable. The accession numbers of the used 16S rRNA gene sequences are given in **Supplementary Table S1**.

Based on the obtained 16S rRNA gene phylogeny a range of full-length RpoB protein sequences of genome-sequenced strains were selected that covers isolates and described species affiliated to clades of the OMG group as well as representative type species of *Pseudomonadales*, *Oceanospirillales*, *Alteromonadales*, and *Enterobacteriaceae* as outgroup sequences. RpoB protein sequences were obtained from GenBank or Uniprot and aligned using the ClustalW algorithm implemented in the ARB package. The accession numbers of the used RpoB proteins are given in **Table 1**.

Phylogenetic trees based on datasets of single sequences were reconstructed using programs implemented in the ARB software package (Ludwig et al., 2004), PAUP* (Swofford, 2002), RAXML (Stamatakis, 2006), or TNT (Goloboff et al., 2008). No filter or weighting masks were applied to constrain the used positions of the alignment. Phylogenetic distances were calculated with the correction of Felsenstein for nucleotide sequences

(Felsenstein, 1982) or with PAM for amino acid sequences, when the ARB neighbor-joining program was used. Maximum parsimony trees were inferred with PAUP* and for the 16S rRNA gene data set, which was larger in terms of organisms, with TNT. Maximum likelihood trees were reconstructed using the RAXML maximum likelihood program (version 7.0.3) with the GTRGAMMA model for DNA and PROTCATLGF for the RpoB data, which was determined as optimal empirical model using RAXML. In each case the robustness of the tree topologies was evaluated by performing 1000 rounds of bootstrap replicates.

Phylogenetic Analyses Based on Genome-Scale Data

The selection of species used for phylogenetic analyses using genome-scale data was identical to the data set used for the RpoB protein trees. Total protein sequences of 59 strains belonging to the class *Gammaproteobacteria* and one outgroup taxon (*Magnetococcus marinus* MC-1^T) were obtained from the IMG website (<https://img.jgi.doe.gov/cgi-bin/er/main.cgi>) with the exception of the proteomes of the strains MOLA455 and *Congregibacter litoralis* KT71^T which were obtained from the NCBI website (<http://www.ncbi.nlm.nih.gov/genome>). The accession numbers of all used genome sequences are given in **Table 1**. Genome sequences were phylogenetically investigated using the phylogenomics pipeline developed at the Leibniz Institute DSMZ including the software applications NCBI BLAST (Altschul et al., 1997), OrthoMCL (Li et al., 2003), MUSCLE (Edgar, 2004), RASCAL (Thompson et al., 2003), and GBLOCKS (Talavera and Castresana, 2007). This method was firstly described by Spring et al. (2010) and later successfully applied to the phylogenetic structuring of several unresolved taxonomic groups of prokaryotes (e.g., Anderson et al., 2011; Abt et al., 2012; Meier-Kolthoff et al., 2014a; Scheuner et al., 2014). In brief, clusters of orthologs were generated using OrthoMCL, inparalogs were removed, the remaining sequences were aligned with MUSCLE and filtered with RASCAL and GBLOCKS. Then, three different supermatrices were compiled: (i) Filtered alignments comprising at least four sequences were concatenated to form a supermatrix. (ii) The supermatrix was cleaned from relatively uninformative genes using MARE (Meusemann et al., 2010) under default values (except that deleting organisms was disallowed). (iii) Those alignments containing 60 sequences (one sequence per genome) were concatenated to form the core-genes matrix.

The OrthoMCL clusters were also converted to an ortholog-content matrix representing the presence or absence of a gene within a certain genome and clusters of orthologs. Further, clusters of homologous sequences were determined using a re-implementation of the TribeMCL algorithm (Enright et al., 2002), applying an *e*-value threshold of 10^{-5} and an MCL inflation parameter of 2.0. The clusters of homologs were converted to a gene-content matrix in analogy to the ortholog-content matrix.

Maximum likelihood and maximum-parsimony trees were inferred from each matrix with RAXML (Stamatakis, 2006) and PAUP* (Swofford, 2002), respectively, as previously described (e.g., Spring et al., 2010).

TABLE 1 | Dataset of RpoB protein sequences and genome sequences used for the reconstruction of RpoB-based trees or supermatrix trees based on whole proteomes.

Strain	RpoB protein acc. no.	Genome acc. no.	Classification ^a
CELLVIBRIONALES ord. nov.			
<i>Cellvibrio japonicus</i> Ueda107 ^T	B3PK30	CP000934	<i>Cellvibrionaceae</i>
<i>Cellvibrio</i> sp. BR	I3I669	AICM00000000	<i>Cellvibrionaceae</i>
<i>Gilvimirinus chinensis</i> [T] DSM 19667 ^T	WP_020211074	ARIX00000000	<i>Cellvibrionaceae</i>
<i>Marinimicrobium agarilyticum</i> DSM 16975 ^T	WP_027330681	AUHU00000000	<i>Cellvibrionaceae</i>
<i>Saccharophagus degradans</i> [T] 2-40 ^T	Q21M93	CP000282	<i>Cellvibrionaceae</i>
<i>Simiduia agarivorans</i> [T] DSM 21679 ^T	K4KJ11	CP003746	<i>Cellvibrionaceae</i>
<i>Teredinibacter turnerae</i> [T] T7902 ^T	WP_018274091	ARAH00000000	<i>Cellvibrionaceae</i>
<i>Congregibacter litoralis</i> [T] KT71 ^T	A4A6A0	AAOA00000000	<i>Haliaceae</i>
<i>Congregibacter</i> sp. NOR5-3	B8KNT7	ACCX00000000	<i>Haliaceae</i>
Gammaproteobacterium HTCC2080	A0Z129	AAV000000000	<i>Haliaceae</i>
Gammaproteobacterium HTCC2148	B7S356	ABXQ00000000	<i>Haliaceae</i>
Gammaproteobacterium HIMB55	H3NW15	AGIF00000000	<i>Haliaceae</i>
Gammaproteobacterium IMCC3088	F3L4H4	AEIG00000000	<i>Haliaceae</i>
<i>Haliae salexigens</i> [T] DSM 19537 ^T	J9Z3P3	AUHJ00000000	<i>Haliaceae</i>
<i>Luminiphilus sylvensis</i> [T] NOR5-1B ^T	B8KWF4	ACCY00000000	<i>Haliaceae</i>
<i>Pseudohaliae rubra</i> [T] DSM 19751 ^T	R4J8J3	AUVB00000000	<i>Haliaceae</i>
<i>Microbulbifer agarilyticus</i> S89	WP_010130329	AFPJ00000000	<i>Microbulbiferaceae</i>
<i>Microbulbifer variabilis</i> ATCC 700307 ^T	WP_020411122	AQYJ00000000	<i>Microbulbiferaceae</i>
Gammaproteobacterium HTCC2207	Q1YNY6	AAPI00000000	<i>Porticoccaceae</i>
Gammaproteobacterium MOLA455	W2UEW1	AZIN00000000	<i>Porticoccaceae</i>
" <i>Porticoccus hydrocarbonoclasticus</i> " MCTG13d	U740DRAFT_0413	JQMM00000000	<i>Porticoccaceae</i>
<i>Dasania marina</i> [T] DSM 21967 ^T	WP_026244726	ARDZ00000000	<i>Spongiibacteraceae</i>
Gammaproteobacterium BDW918	I2JM08	AJMK00000000	<i>Spongiibacteraceae</i>
Gammaproteobacterium HTCC2143	A0YHK1	AAVT00000000	<i>Spongiibacteraceae</i>
<i>Spongiibacter marinus</i> (=Melitea salexigens) [T] DSM 19753	WP_027875072	AULP00000000	<i>Spongiibacteraceae</i>
<i>Spongiibacter tropicus</i> DSM 19543 ^T	WP_022960717	ATUS00000000	<i>Spongiibacteraceae</i>
OUTGROUP REFERENCES			
<i>Alcanivorax borkumensis</i> [T] SK2 ^T	Q0VSM2	AM286690	<i>Alcanivoracaceae</i>
<i>Alcanivorax pacificus</i> W11-5 ^T	K2HC73	AJGP00000000	<i>Alcanivoracaceae</i>
<i>Alcanivorax dieselolei</i> B5 ^T	K0CEN5	CP003466	<i>Alcanivoracaceae</i>
<i>Alteromonas macleodii</i> [T] ATCC 27126 ^T	J9YDH2	CP003841	<i>Alteromonadaceae</i>
<i>Glaciecola punicea</i> [T] ACAM 611 ^T	H5T848	BAET00000000	<i>Alteromonadaceae</i>
<i>Glaciecola nitratireducens</i> FR1064 ^T	G4QDW1	CP003060	<i>Alteromonadaceae</i>
<i>Paraglaciecola polaris</i> LMG 21857 ^T	K6ZKT0	BAER00000000	<i>Alteromonadaceae</i>
<i>Salinimonas chungwhensis</i> [T] DSM 16280 ^T	WP_026294703	ARJT00000000	<i>Alteromonadaceae</i>
<i>Enterobacter cloacae</i> subsp. <i>cloacae</i> [T] ATCC 13047 ^T	D5CBL4	CP001918	<i>Enterobacteriaceae</i>
<i>Rahnella aquatilis</i> [T] CIP 78.65 ^T	H2IVU9	CP003244	<i>Enterobacteriaceae</i>
<i>Xenorhabdus nematophila</i> [T] ATCC 19061 ^T	D3VEF0	FN667742	<i>Enterobacteriaceae</i>
<i>Providencia alcalifaciens</i> [T] DSM 30120 ^T	B6X9R3	ABXW00000000	<i>Enterobacteriaceae</i>
<i>Hahella chejuensis</i> [T] KCTC 2396 ^T	Q2S905	CP000155	<i>Hahellaceae</i>
<i>Hahella ganghwensis</i> DSM 17046 ^T	WP_020405297	AQXX00000000	<i>Hahellaceae</i>
<i>Halomonas elongata</i> [T] DSM 2581 ^T	E1V5K0	FN869568	<i>Halomonadaceae</i>
<i>Halomonas zhanjiangensis</i> DSM 21076 ^T	WP_018917740	ARIT00000000	<i>Halomonadaceae</i>
<i>Chromohalobacter salexigens</i> DSM 3043 ^T	Q1R0I2	CP000285	<i>Halomonadaceae</i>
<i>Kushneria aurantia</i> [T] DSM 21353 ^T	WP_026351855	ARNK00000000	<i>Halomonadaceae</i>
<i>Perluclidibaca piscinae</i> [T] DSM 21586 ^T	WP_022957094	ATUE00000000	<i>Moraxellaceae</i>
<i>Moraxella bovoculi</i> 237 ^T	KDN25377	AOMT00000000	<i>Moraxellaceae</i>
<i>Psychrobacter arcticus</i> 273-4 ^T	Q4FQH3	CP000082	<i>Moraxellaceae</i>
<i>Acinetobacter calcoaceticus</i> [T] DSM 30006 ^T	N9DU07	AIEC00000000	<i>Moraxellaceae</i>

(Continued)

TABLE 1 | Continued

Strain	RpoB protein acc. no.	Genome acc. no.	Classification ^a
<i>Acinetobacter radioresistens</i> DSM 6976 ^T	K6VIF1	APQF00000000	<i>Moraxellaceae</i>
<i>Oceanospirillum maris</i> subsp. <i>maris</i> DSM 6286 ^T	WP_028304193	AUGW00000000	<i>Oceanospirillaceae</i>
<i>Oceanospirillum beijerinckii</i> subsp. <i>beijerinckii</i> DSM 7166 ^T	WP_028301410	AULT00000000	<i>Oceanospirillaceae</i>
<i>Marinospirillum minutulum</i> [T] DSM 6287 ^T	WP_027849030	AULO00000000	<i>Oceanospirillaceae</i>
<i>Pseudomonas aeruginosa</i> [T] PAO1	Q51561	AE004091	<i>Pseudomonadaceae</i>
<i>Pseudomonas sutzeri</i> ATCC 17588 ^T	F8H9S9	CP002881	<i>Pseudomonadaceae</i>
<i>Pseudomonas syringae</i> pv. <i>tabaci</i> ATCC 11528	F3K5M0	AEAP00000000	<i>Pseudomonadaceae</i>
<i>Pseudomonas veronii</i> 1YdBTEX2	WP_017848798	AOUH00000000	<i>Pseudomonadaceae</i>
<i>Azotobacter vinelandii</i> ATCC BAA-1303	C1DKK5	CP001157	<i>Pseudomonadaceae</i>
<i>Saccharosporillum impatiens</i> [T] DSM 12546 ^T	WP_028671302	AUIH00000000	" <i>Saccharosporillaceae</i> "
<i>Reinekea blandensis</i> MED297 ^T	A4BF53	AAOE00000000	" <i>Saccharosporillaceae</i> "
ROOT			
<i>Magnetococcus marinus</i> [T] MC-1 ^T	A0L5W6	CP000471	<i>Magnetococcaceae</i>

The [T] after a species epithet indicates a type species; the superscript T following a strain designation a type strain.

^aFor members of the proposed order *Cellvibrionales* the classification of strains at family level is based on the results of this study, whereas the classification of the outgroup references has been adopted from the NCBI Taxonomy Database (<http://www.ncbi.nlm.nih.gov/taxonomy>).

Results and Discussion

Evolutionary Relationships between Members of the OMG Group and Related Species Phylogeny Based on 16S rRNA Genes

Phylogenetic analyses were based on a comprehensive dataset of 200 nearly complete 16S rRNA gene sequences resulting in an alignment with a minimum sequence length of 1115 nucleotides and 1172 distinct alignment patterns in the maximum likelihood analysis. In reconstructed trees containing representatives of most recognized families within the class *Gammaproteobacteria* the OMG group separates at least into two major lineages. In the tree shown in **Supplementary Figure S1** members of the clades OM60/NOR5, BD1-7, and SAR92 group together, whereas members of the OM182 and KI89A clades form a separate lineage that is not closely related to the main group. Since the first phylogenetic assessment of the OMG group by Cho and Giovannoni (2004) a large number of isolates with a close phylogenetic relationship to members of the OMG group were cultured and described as novel species. So far all of these isolates could be either affiliated to the OM60/NOR5 clade or were closely related to members of the BD1-7 and SAR92 clades, whereas no type strains or genome sequenced isolates became available belonging to the OM182 or KI89A clade which prevents a more detailed phylogenetic assessment of the latter two groups. In addition, the SAR86 clade representing another emerging group of oligotrophic marine gammaproteobacteria (Dupont et al., 2012) was not considered in this study because of its large phylogenetic distance to the targeted OMG group and a lack of cultured representatives. Most of the already classified species that cluster together with the main lineage of the OMG group were allocated to various established higher taxa within the *Gammaproteobacteria*, including *Pseudomonadaceae*, *Alteromonadaceae*, and *Oceanospirillaceae* (**Supplementary Table S1**). However, as shown in **Supplementary Figure S1** the type genera of these

families are only distantly related to members of the main lineage of the OMG group. Based on phylogenetic analyses using 16S rRNA sequence data a recognizable relationship to the main lineage of the OMG group could be only assumed for members of the genus *Pseudomonas*, whereas the other type genera appear to be unrelated. In **Supplementary Figure S2** a subtree of the comprehensive 16S rRNA gene based tree of *Gammaproteobacteria* is presented that shows exclusively the position of sequences that are either closely related or allocated to clades of the main lineage of the OMG group. The lowest 16S rRNA gene identity values observed among members of this group were around 85% which corresponds to the minimum identity typically found in an order defined according to modern taxonomic standards (Yarza et al., 2014). Thus, we propose to define a novel taxon at order level for this deep-branching phylogenetic lineage within the class *Gammaproteobacteria* and suggest to designate it *Cellvibrionales*, because the first described type species within this group was *Cellvibrio mixtus* (Blackall et al., 1985). Note, however, that sequence identity should not be used to assign strains to taxa because more similar organisms are not necessarily more closely related (Meier-Kolthoff et al., 2014b), hence it is important to determine the minimum sequence identity for subtrees in the existing phylogenetic tree. Within this order-level lineage several monophyletic groups can be distinguished three of which contain representatives originally affiliated to the OMG group, whereas two comprise exclusively described species that were allocated to established families of *Gammaproteobacteria*. The sequence cluster represented by the OM60/NOR5 clade also comprises the genera *Congregibacter*, *Haliaea*, *Halioglobus*, *Luminiphilus*, and *Pseudohaliaea*; the BD1-7 clade forms a group together with the genera *Spongiibacter*, *Zhongshania*, *Sinobacterium*, and *Dasania*, whereas the SAR92 clade is closely related to the genus *Porticoccus* and several sequences affiliated to the C1-B045 clade that were first detected in deep-sea hydrothermal sediments (Teske et al., 2002). Clades that are mainly comprised

of formally described cultured strains are represented by members of the genus *Microbulbifer* on the one hand and *Cellvibrionales* species and related taxa on the other hand. It is important to note, however, that the monophyletic assemblage shown in **Supplementary Figure S2** was only obtained using the neighbor-joining reconstruction method and is not supported by significant bootstrap values. In addition, the bootstrap support for the depicted phylogenetic subgroups except the OM60/NOR5 and *Microbulbifer* cluster are insignificant (below 75%) with most applied reconstruction methods, so that a formal proposal of higher order taxonomic groups within this phylogenetic branch requires further investigation.

Phylogeny Based on RpoB Proteins

In this study a dataset of 60 complete RpoB proteins derived from the available genome sequences of members of the OMG group and related gammaproteobacteria was used (**Table 1**). The resulting amino acid alignment comprised sequences with a minimum length of 1337 positions and 924 distinct alignment patterns in the maximum likelihood analysis. In the reconstructed phylogenetic tree shown in **Figure 1** the average bootstrap support for the five identified subgroups of the proposed order *Cellvibrionales* is considerably higher than in the 16S rRNA gene based tree. To exclude the possibility that the improved stability of the RpoB tree topology is related to the reduced number of selected sequences compared to the 16S rRNA gene based tree shown in **Supplementary Figure S1**, trees with the same set of strains but based on 16S rRNA sequences were reconstructed. However, it turned out that the stability of the 16S rRNA gene based trees was not significantly improved using the reduced dataset compared to the more comprehensive set of sequences. In addition, the monophyly of the *Cellvibrionales* lineage was not retained using the reduced 16S rRNA gene dataset (data not shown). Hence, the improved reliability of the RpoB based phylogeny is independent of the number of used sequences and caused by the higher information content of this large conserved protein compared to the 16S rRNA. However, despite the improved stability of the phylogenetic reconstruction several subgroups of the *Cellvibrionales*, like the clade represented by *Cellvibrionales* species and closely related taxa was still not resolved with high confidence values in the RpoB based tree. We used, therefore, genome-scale data to further improve the reliability of the reconstructed phylogenetic relationships.

Phylogeny Based on Genome-Scale Data

Supermatrix trees

The basic supermatrix comprised 10,059 genes and 2,801,974 characters (i); a further matrix containing 928 genes and 223,650 characters was obtained after cleaning with MARE (ii). The core-genes matrix contained 341 genes and 111,505 characters (iii). For all three matrices the selected model for maximum likelihood reconstructions was PROTCATLGF; that is the LG model of amino acid evolution (Le and Gascuel, 2008) in conjunction with the CAT model approximation of rate heterogeneity (Stamatakis, 2006) and empirical amino acid frequencies. The resulting trees had a log likelihood of $-53,403,618.46$ (i), $-9,353,599.37$ (ii), and $-5,308,688.73$ (iii), respectively. The supermatrix maximum

likelihood (ML) tree is shown in **Figure 2**. The best maximum-parsimony (MP) trees found had a length of 9,726,609 (i), 1,786,102 (ii), and 1,031,863 (iii) steps (not counting uninformative characters). As expected the increased number of characters in the supermatrix analyses resulted in a more stable tree topology compared to the single gene analyses with nearly maximum bootstrap support for the proposed *Cellvibrionales* lineage including the five identified subgroups. Interestingly, the close relationship of the *Cellvibrionales* to members of the *Pseudomonadaceae* that became apparent in the 16S rRNA gene based tree was not evident in the RpoB or supermatrix trees. This finding supports our proposal to separate members of the main lineage of the phylogenetically defined OMG group and related taxa from the *Pseudomonadales* or any other established order within the *Gammaproteobacteria*.

Gene-content analyses

In addition to supermatrix trees genome wide gene-content analyses can be used to obtain useful information about the relatedness of environmental bacteria. The obtained gene-content matrix contained 32,576 characters, and the resulting best trees had a log likelihood of $-329,588.46$ and a length of 72,280 steps, respectively. The ortholog-content matrix comprised 44,726 characters and yielded best trees with a log likelihood of $-462,997.17$ and a parsimony score of 102,753, respectively. The gene-content ML tree is shown in **Figure 3**. Although, the overall topology of the supermatrix tree was confirmed in the gene-content analyses, bootstrap support for the five subgroups within the *Cellvibrionales* was markedly reduced indicating a lower phylogenetic resolution due to a reduced number of characters or a greater effect of lateral gene transfer events on the obtained phylogenetic signal (Philippe and Douady, 2003; Kloesges et al., 2011). However, the gene-content analyses revealed another interesting aspect about this branch of the *Gammaproteobacteria*: In contrast to the sequence-matrix based phylogenetic trees a divergence between members of the originally identified OMG clades OM60/NOR5, BD1-7, and SAR92 and established taxa belonging to genera like *Cellvibrionales* or *Microbulbifer* emerged. The unveiled separation of this phylogenetic branch into two major groups probably reflects two different trophic strategies, because members of the OMG clades are considered as oligotrophs, whereas representatives of the other group were mainly isolated on nutrient rich complex media and can be considered as copiotrophs. The lack of congruence in the overall branching pattern of the two different trophic groups with the phylogenetic relationships of their representatives indicate that the adaptation to oligotrophic growth conditions in the SAR92 clade on the one hand and the OM60/NOR5 and BD1-7 clades on the other hand has been probably developed from different copiotrophic ancestors.

Correlation of Phenotypic Traits with the Proposed Evolutionary Relationships Chemotaxonomy

Based on the available descriptions of the cultured members of the proposed order *Cellvibrionales* chemotaxonomic data were collected and correlated with the five identified subgroups of this

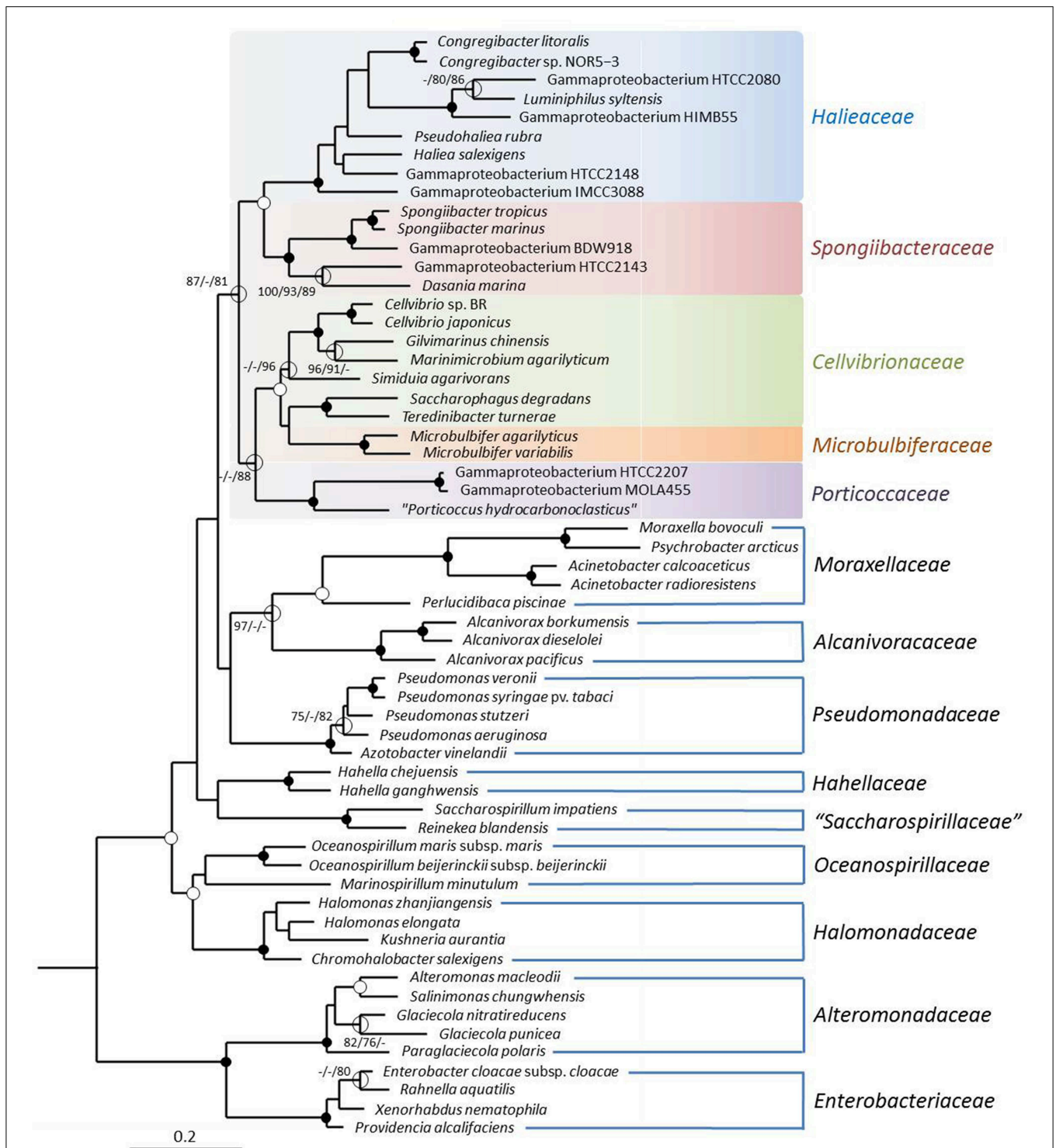


FIGURE 1 | Phylogenetic tree inferred from a dataset of complete RpoB protein sequences. The tree topology was reconstructed with the maximum-likelihood method using the PROTCATLGF model. The RpoB sequence of *Magnetococcus marinus* MC-1^T was used as an outgroup (not shown). Additional trees were reconstructed using maximum parsimony and neighbor-joining methods. Support of a distinct branching by bootstrap analyses is indicated by symbols. Black dots at a distinct node indicate that bootstrap values of 95% or above (percentages of 1000 resamplings) were

obtained with three different reconstruction methods, while white dots indicate that values of 95% or above were obtained with only two reconstruction methods. Hollow circles indicate that bootstrap values of 75% or above were obtained with at least one reconstruction method. In such cases the values above 75% are given from left to right for the maximum-likelihood, maximum parsimony and neighbor-joining method. The branches are scaled in terms of the expected number of substitutions per site.

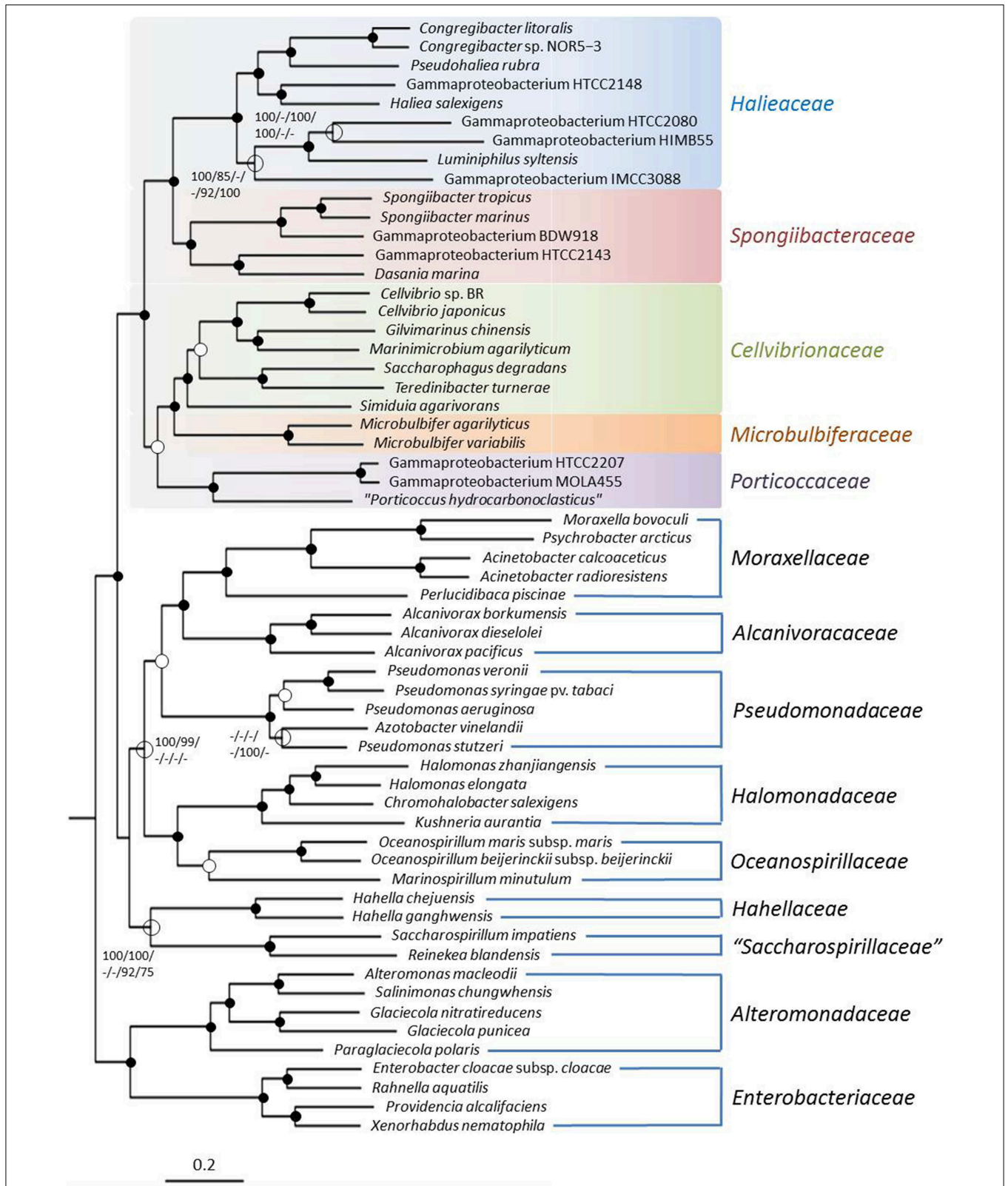


FIGURE 2 | Phylogenetic tree inferred from the 2,801,974 aligned amino acid character containing supermatrix. The tree topology was reconstructed under the maximum likelihood criterion and was rooted with *Magnetococcus marinus* MC-1^T (not shown). Black dots at a distinct node

indicate that bootstrap values of 95% or above (percentages of 1000 resamplings) were obtained with six different inference methods, while white dots indicate that values of 95% or above were obtained with four or five (Continued)

FIGURE 2 | Continued

reconstruction methods. Hollow circles indicate that bootstrap values of 75% or above were obtained with at least one inference method. In such cases values above 75% are given from left to right for the ML

supermatrix, MP supermatrix, ML MARE-filtered supermatrix, MP MARE-filtered supermatrix, ML core-genes matrix and MP core-genes matrix analyses. The branches are scaled in terms of the expected number of substitutions per site.

lineage. The obtained results along with the used literature are shown in **Table 2**. With only very few exceptions the major respiratory lipoquinone of representatives of the *Cellvibrionales* is ubiquinone 8 (Q8), which provides a strong argument to separate this group from the pseudomonads *in sensu stricto* which are characterized by ubiquinone 9 (Q9) as major respiratory lipoquinone (Oyaizu and Komagata, 1983; Moore et al., 2006).

It further turned out that several phylogenetic groups within this order are characterized by distinct cellular fatty acid patterns that allow their differentiation from related clades. The clade comprising *Microbulbifer* species can be identified by the prevalence of *iso*-branched cellular fatty acids, whereas the presence of *anteiso*-branched fatty acids in significant amounts is specific for cells of *Porticoccus* species. These data clearly demonstrate that the revealed phylogenetic structure of this branch of the class *Gammaproteobacteria* can be correlated with phenotypic traits of the respective strains, so that the establishment of higher order taxa for the identified phylogenetic groups is corroborated by phenotypic data.

In contrast to the lipid patterns, the guanine plus cytosine content of genomic DNA (DNA G + C values) showed only little variation between the five different phylogenetic groups, with the highest average value found in the OM60/NOR5 clade (52–66 mol%) and the lowest median represented by the *Porticoccus*/SAR92 cluster (48–53 mol%).

Ecophysiology

The dichotomy in the *Cellvibrionales* lineage that becomes apparent in the gene-content analyses shown in **Figure 3** could reflect an adaptation to different ecological niches that resulted in two basically different genetic inventories. Most members of the OMG clades OM60/NOR5, BD1-7, SAR92, and closely related strains that represent one monophyletic group in **Figure 3** are numerically abundant in marine environments and were mainly isolated using nutrient-poor media, so that they are considered as typical oligotrophs (Cho and Giovannoni, 2004). The metabolism of these bacteria is adapted to allow growth under conditions of extreme carbon source limitation and nutrient depletion, so that many of these organisms have the potential to use alternative sources of energy for mixotrophic growth. For instance several members of the OM60/NOR5 clade encode *sox* genes for the utilization of reduced inorganic sulfur compounds and photosynthesis genes for the harvesting of light in the euphotic zone of the oceans. However, it seems that only light, but not thiosulfate, is used by members of this phylogenetic group as additional energy source and stimulates growth (Spring and Riedel, 2013; Spring, 2014). In stark contrast members of the genera *Cellvibrio*, *Microbulbifer* and related taxa that comprise the second group in the gene-content tree (**Figure 3**) are typical copiotrophic bacteria that prefer nutrient-rich conditions for growth. A further typical characteristic of many of these copiotrophic strains is a

specialization on complex polysaccharides as substrates, which is not observed among the oligotrophic representatives of the OMG group (Spring et al., 2013). Especially, the capability to degrade complex carbohydrates like agarose, chitin, or cellulose that has been detected in many members of the genera *Cellvibrio*, *Microbulbifer*, *Marinimicrobium*, or *Simiduia* could provide these copiotrophic bacteria with a selective advantage that enables them to occupy distinct nutrient-rich ecological niches, like for instance marine snow. Consequently, the dichotomy depicted in the gene-content analyses reflects probably the existence of two basically different ecological guilds that emerged independent of the evolutionary relationships within this branch of the *Gammaproteobacteria*. However, it should be noted that several representatives belonging to the oligotrophic guild of the *Cellvibrionales* show an improved growth under nutrient-rich conditions (Spring and Riedel, 2013) so that we prefer to define the members of this group in terms of traditional ecology as typical *K*-strategists and the copiotrophic members as *r*-strategists. As outlined by Andrews and Harris (1986) the advantage of the latter definitions is that they are not restricted to the level of substrate concentrations allowing growth and, therefore, are more widely applicable. On the other hand, the theory of *r*- and *K*-selection was originally developed to describe the ecology of highly developed animals and thus has some shortcomings in respect to very small asexual proliferating organisms like bacteria. An alternative approach specific for prokaryotes was developed by Schmidt and colleagues who try to define both trophic strategies based on different response times to nutrient pulses. According to their hypothesis typical oligotrophic bacteria could be characterized as slow responders and copiotrophic bacteria as rapid responders, which would be reflected in the number of encoded rRNA operon copies and the translational power of the cell (Klappenbach et al., 2000; Dethlefsen and Schmidt, 2007). Both ecological concepts can be applied to marine microbial communities resulting in the distinction of two major trophic strategies that may be related to corresponding super-niches (Polz et al., 2006; Yooseph et al., 2010; Gifford et al., 2012): In summary, typical *r*-strategists (most copiotrophs) are characterized by short periods of fast growth that alternate with long periods of non-growth usually resulting in a low abundance in the environment interrupted by occasional blooms. These bacteria prefer nutrient-rich niches (e.g., detritus) and have a metabolism of low efficiency but high plasticity. In contrast, typical *K*-strategists (most oligotrophs) are highly abundant and have long phases of continuous growth characterized by a constant low growth rate. Niches occupied by these bacteria are usually poor in nutrients leading to substrate limitation and utilization of alternative energy sources. Their metabolism is highly efficient and specialized, but has a low flexibility.

It was previously suggested using two distinct species of marine bacteria, one paradigmatic for a copiotrophic and the other for an oligotrophic life style, that the trophic strategies of

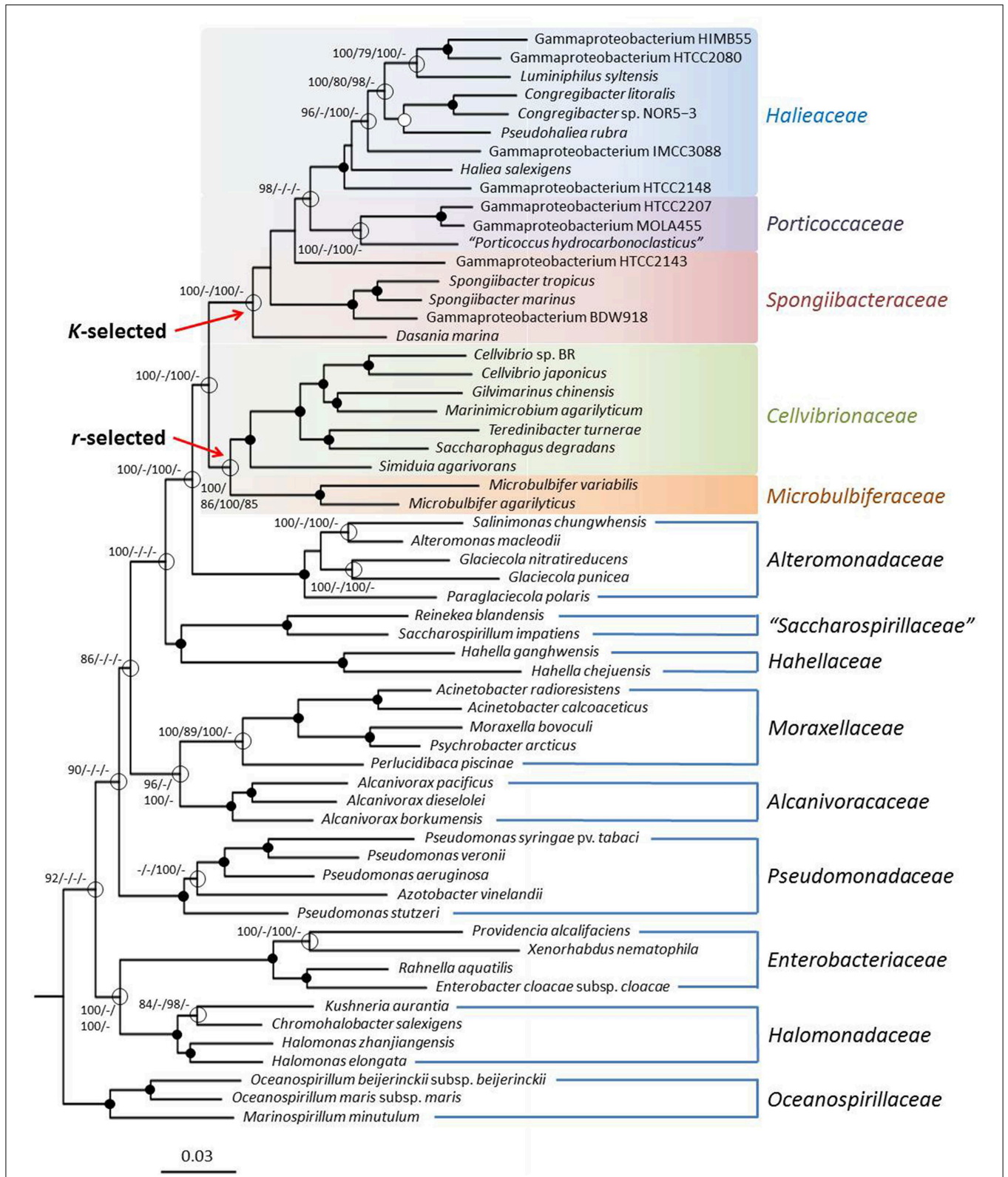


FIGURE 3 | Phylogeny inferred from a gene-content matrix. The tree topology was reconstructed under the maximum likelihood criterion and was rooted with *Magnetococcus marinus* MC-1^T (not shown). Black dots at a distinct node indicate that bootstrap values of 95% or above (percentages of

1000 resamplings) were obtained with four different inference methods, while white dots indicate that values of 95% or above were obtained with three reconstruction methods. Hollow circles indicate that bootstrap values of 75% (Continued)

FIGURE 3 | Continued

or above were obtained with at least one inference method. In such cases values above 75% are given from left to right for the ML gene-content matrix, MP gene-content matrix, ML ortholog-content matrix and MP

ortholog-content matrix analyses. The estimated evolutionary origins of the two different trophic guilds within the *Cellvibrionales* are indicated by red arrows. The branches are scaled in terms of the expected number of substitutions per site.

TABLE 2 | Chemotaxonomic traits of genera with a proposed affiliation to the novel order *Cellvibrionales*.

Genus	No. of species	Major quinone	Major CFA	DNA G + C content (mol%)
Cellvibrionaceae fam. nov.				
<i>Cellvibrio</i> ^{1,2}	8	ND	C _{16:1} ^a , C _{16:0} , C _{18:1}	44–53
<i>Eionea</i> ³	1	Q8	C _{16:1} , C _{18:1} , C _{16:0}	48
<i>Gilvimirinus</i> ⁴	2	Q8	C _{16:1} , C _{16:0} , C _{18:1}	49–52
<i>Maricurvus</i> ⁵	1	Q8	C _{16:1} , C _{18:1} , C _{16:0}	49
<i>Marinimicrobium</i> ⁶	3	Q8	C _{16:0} , C _{19:0} cyclo, C _{16:1}	57–58
<i>Pseudoteredinibacter</i> ⁷	1	Q9	C _{16:1} , C _{17:1} , C _{18:1} , C _{16:0}	52
<i>Saccharophagus</i> ⁷	1	ND	C _{16:0} , C _{17:0} , C _{18:1}	46
<i>Simidua</i> ⁸	4	Q8	C _{16:1} , C _{17:1} , C _{16:0}	50–56
<i>Teredinibacter</i> ⁷	1	ND	C _{18:1} , C _{16:0} , C _{16:1}	49
<i>Umboniiibacter</i> ⁹	1	Q7	C _{17:1} , C _{16:0} , C _{17:0} , C _{16:1}	52
HALIEACEAE fam. nov.				
<i>Haliaea</i> ¹⁰	2	Q8	C _{16:1} , C _{18:1} , C _{17:1}	61–62
<i>Chromatocurvus</i> ¹⁰	1	Q8	C _{17:1} , C _{16:1} , C _{18:1}	63
<i>Congregibacter</i> ¹⁰	1	Q8	C _{16:1} , C _{18:1}	56–58
<i>Halioglobus</i> ¹¹	2	Q8	C _{16:1} , C _{18:1}	59–60
<i>Luminiphilus</i> ¹⁰	1	Q8	C _{16:1} , C _{16:0}	57
<i>Pseudohaliaea</i> ¹⁰	1	Q8	C _{18:1} , C _{16:0} , C _{16:1}	66
MICROBULBIFERACEAE fam. nov.				
<i>Microbulbifer</i> ^{12–16}	19	Q8	iso-C _{15:0} , iso-C _{17:1} , C _{18:1}	49–62
PORTICOCCACEAE fam. nov.				
<i>Porticoccus</i> ¹⁷	1	ND	anteiso-C _{15:0} , anteiso-C _{17:0}	48
SPONGIIBACTERACEAE fam. nov.				
<i>Spongiibacter</i> ¹⁸	2	Q8	C _{17:1} , C _{18:1} , C _{16:1}	58–61
<i>Dasania</i> ¹	1	Q8	C _{16:1} , C _{16:0}	47 ^b
<i>Sinobacterium</i> ¹⁹	1	Q8	C _{16:1} , C _{16:0}	59
<i>Zhongshania</i> ²⁰	4	Q8	C _{16:1} , C _{17:1} , C _{16:0}	52–54

CFA, cellular fatty acid.

^a In the cited literature the summed feature 3 containing the fatty acids C_{16:1}ω7c and/or iso-C_{15:0} 2-OH is given. However, the presence of large amounts of the fatty acid C_{15:0} 2-OH in these bacteria is highly unlikely, so that in this table summed feature 3 was replaced with C_{16:1}.

^b The DNA G + C content value of *Dasania marina* was deduced from the genome sequence, the value given in the literature is 37 mol%.

The superscript numbers indicate the following used references: 1, Lee et al. (2007); 2, Suarez et al. (2014); 3, Urios et al. (2011); 4, Cheng et al. (2015); 5, Iwaki et al. (2012); 6, Yoon et al. (2009); 7, Chen et al. (2011); 8, Park et al. (2014); 9, Romanenko et al. (2010); 10, Spring et al. (2013); 11, Park et al. (2012); 12, Baba et al. (2011); 13, Zhang et al. (2012); 14, Kämpfer et al. (2012); 15, Jeong et al. (2013); 16, Vashist et al. (2013); 17, Oh et al. (2010b); 18, Jang et al. (2011a); 19, Su et al. (2013); 20, Lo et al. (2014).

microorganisms can be inferred from their genome sequences (Lauro et al., 2009). The authors of this study revealed that the genomes of copiotrophic species were often characterized by the overrepresentation of genes associated with the cluster of orthologous group (COG) categories motility (N), defense mechanisms (V), signal transduction mechanisms (T), and transcription (K). In contrast, oligotrophic species were enriched in COGs involved in lipid transport and metabolism (functional category I) and secondary metabolites biosynthesis, transport and catabolism (functional category Q). Accordingly, we performed a COG analysis of both sets of genomes that correspond to distinct ecological guilds in the tree shown in **Figure 3** using tools of the IMG-ER server (Markowitz et al., 2009). As shown

in **Table 3**, we could confirm the predictions made by Lauro et al. (2009) using a comprehensive dataset of whole genomes of environmental important gammaproteobacteria related to the OMG group. We applied this approach also to a set of four genomes affiliated to the distantly related gammaproteobacterial SAR86 clade. The analyzed genomes were however not obtained from pure cultures, but by metagenomic assemblies or single cell sequencing (Dupont et al., 2012). These genomes have a very small size (1.3–1.7 MBp) compared to members of the OMG group indicating a reductive evolution and genome streamlining typical for obligate oligotrophs (Giovannoni et al., 2014). Gene-content analyses revealed a low abundance of the COG categories motility (1%), signal transduction mechanisms

TABLE 3 | Abundance of genes associated with functional COG categories in genome sequenced members of two guilds of the novel order *Cellvibrionales* that are distinguishable by their trophic strategy.

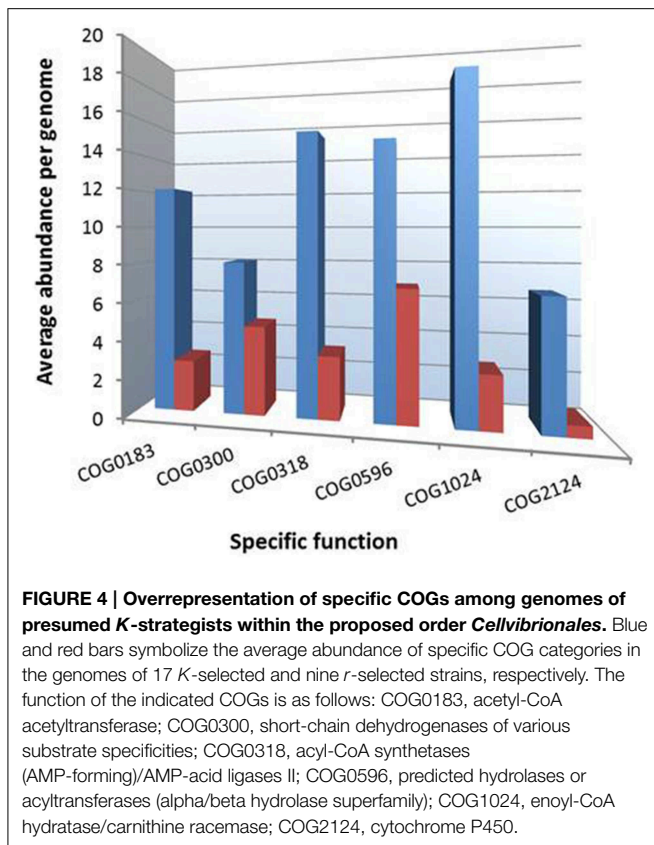
COG category	Percentage of genes associated with functional COG categories ^a	
	<i>K</i> -strategists (17 genomes)	<i>r</i> -Strategists (nine genomes)
Amino acid transport and metabolism	8	7
Carbohydrate transport and metabolism	3 (-)	6 (+)
Cell cycle control, cell division, chromosome partitioning	1	1
Cell motility	2 (-)	3 (+)
Cell wall/membrane/envelope biogenesis	6	6
Coenzyme transport and metabolism	5	5
Defense mechanisms	1 (-)	2 (+)
Energy production and conversion	7	6
Inorganic ion transport and metabolism	6	6
Intracellular trafficking, secretion, and vesicular transport	3	4
Lipid transport and metabolism	7 (+)	3 (-)
Nucleotide transport and metabolism	2	2
Posttranslational modification, protein turnover, chaperones	5	5
Replication, recombination and repair	5	4
Secondary metabolites biosynthesis, transport and catabolism	4 (+)	2 (-)
Signal transduction mechanisms	3 (-)	6 (+)
Transcription	5 (-)	7 (+)
Translation, ribosomal structure and biogenesis	6	6

^a Values correspond to the average percentage of a given COG category per genome. The total for calculating individual percentage values is based on the total number of protein coding genes in the respective annotated genomes. Significant upward and downward deviations from the estimated mean values are labeled with plus or minus signs, respectively.

(2%), and transcription (3%), but a relatively high prevalence of genes involved in lipid metabolism (9%) and secondary metabolites (4%), which suggests their categorization as typical *K*-strategists. In addition, we tried to identify genomic features, which allow a reliable differentiation of both trophic guilds within the *Cellvibrionales*. In **Figure 4** the frequencies of several functional COGs are shown, which differ most significantly among representatives of both trophic guilds. All of these COGs were previously identified in the study of Lauro et al. (2009) and were found to be overrepresented in oligotrophs. Most of these COGs can be associated with the degradation of fatty acids, which may indicate a specialization of marine gammaproteobacterial *K*-strategists on lipids as carbon source due to their higher energy yield compared to carbohydrates (Lauro et al., 2009). In support of this result it was previously reported that a range of fatty acids with various chain lengths like acetate, propionate, butyrate or palmitate can be utilized by several representatives of the OM60/NOR5 clade (Spring et al., 2009, 2013). The largest divergence in the frequency of a distinct function was revealed for cytochrome P450 genes (COG2124) that were found to be 11 times more abundant in the genomes of postulated *K*-strategists than *r*-strategists. P450 enzymes play a role for example in the degradation and detoxification of certain soluble recalcitrant compounds that are used by some oligotrophic marine bacteria as carbon source. This is corroborated by a study that reports utilization of the xenobiotic compound nonylphenol polyethoxylate by members of the OM60/NOR5 clade (Maeda et al., 2005). In general, the combination of COG categories shown in **Figure 4** could be useful for the identification of gammaproteobacteria

representing *K*-strategists in marine metagenomic libraries.

In a recent study by Gifford et al. (2012) the expression patterns of a highly diverse microbial community in a coastal ocean system was studied using a metatranscriptomics approach. Interestingly, their results imply a high degree of niche diversification in seawater, an environment commonly regarded as homogenous. Several of the reported results seem to confirm predictions made by gene-content analyses for the two ecological guilds of *Cellvibrionales*. For instance, transcriptomes of three bins assigned to genome sequences of the putative *r*-strategists *Teredinibacter turnerae*, *Saccharophagus degradans*, and *Cellvibrio japonicus* showed a high prevalence of transcribed ribosomal protein genes (mean value of 14.5%) indicating a high translational power and growth rate, whereas bins assigned to genomes of the highly abundant predicted *K*-strategists *Luminiphilus syltensis* NOR5-1B, HTCC2080, and HTCC2207 had a significantly lower proportion of expressed ribosomal proteins (mean value of 7.6%) indicating a low growth rate. In addition, bins related to the *K*-strategists *Luminiphilus syltensis* NOR5-1B, HTCC2080, and HTCC2207 showed a relatively high expression of genes involved in mixotrophic growth reflecting their effort to continue proliferation under conditions of substrate limitation. The metatranscriptomic data indicate also a metabolism specialized on fatty acid degradation for the bins of NOR5-1B and HTCC2080. A notable discrepancy to our predictions may represent the reported high expression of genes related to motility and carbohydrate utilization allocated to the HTCC2207 bin, however it is unclear if this expression pattern reflects only specific conditions



encountered in the studied site or was even selected by the used sampling method which excluded particles and surface sediments. In this context it should be also noted that in the real world a continuum of ecological strategies exists. Hence, it is unrealistic to expect a clear assignment of all environmental bacteria to one or the other extreme (e.g., Gifford et al., 2012). Nevertheless, gene-content analyses can give a first approximation of the prevailing trophic strategies of bacteria belonging to higher taxonomic ranks, which then can be examined in more detail using metatranscriptomic studies of distinct environments.

In contrast to the observed differences in the genomes of the two trophic guilds of the *Cellvibrionales* there seems to be no obvious genomic divergence between organisms preferring freshwater or marine environments, because as shown in **Figure 3** soil-inhabiting members of the genus *Cellvibrio* cluster together with representatives of several genera bound to marine environments. Consequently, the salinity of the preferred habitat seems to have only little influence on the genetic inventory of these strains.

Conclusions

In this study we demonstrate that the reconstruction of evolutionary relationships among major groups of bacteria requires the analyses of large conserved proteins like RpoB or preferably whole proteomes in order to obtain reliable results, while 16S rRNA sequence analyses alone are not sufficient. Based on our

analyses the phylogenetic structure of an emerging major lineage of ecologically important *Gammaproteobacteria* was revealed. Besides supermatrix-based trees phylogenetic reconstructions based on comparative gene-content analyses were of great value for the identification of two distinct trophic guilds within a monophyletic group of marine bacteria. In this respect, a combination of genome-based classification, gene-content analyses and metatranscriptomic surveys of various environments appears to be the most promising approach to unveil the assumed correlation of the ecological functions of bacteria with their affiliation to higher taxonomic ranks (Philippot et al., 2010). An evaluation of the available phenotypic data of members of the identified phylogenetic clusters suggests the proposal of a novel order within the *Gammaproteobacteria* that comprises five novel families. Formal descriptions of the novel taxa follow below:

Description of *Cellvibrionales* ord. nov.

Cellvibrionales (Cell.vi.bri.o.na'les. N.L. n. *Cellvibrio*, type genus of the order; suff. *-ales*, ending denoting an order; N.L. fem. pl. n. *Cellvibrionales*, the *Cellvibrio* order).

The order belongs to the class *Gammaproteobacteria*. Cells are gram-stain-negative, straight to bent rods, cocci or pleomorphic, non-endospore-forming, motile by means of flagella or non-motile. Aerobic, mesophilic, neutrophilic, chemoorganotrophic, respiratory type of metabolism. Most strains are positive in the oxidase test. Ubiquinone 8 (Q8) is the major respiratory lipoprotein in most species. The G + C content of the genomic DNA of most strains is in the medium range between 44 and 66 mol%. The definition of the order relies mainly on the reconstruction of phylogenetic relationships based on comparative sequence analyses of complete *rpoB* genes or whole proteomes. At present, the order contains the families *Cellvibrionaceae*, *Halieaceae*, *Microbulfiferaceae*, *Porticoccaceae*, and *Spongiibacteraceae*. The type genus of the order is *Cellvibrio* [Blackall et al. 1986 (Parte, 2014)].

Description of *Cellvibrionaceae* fam. nov.

Cellvibrionaceae (Cell.vi.bri.o.na'ce.ae. N.L. n. *Cellvibrio*, type genus of the family; suff. *-aceae*, ending denoting a family; N.L. fem. pl. n. *Cellvibrionaceae*, the *Cellvibrio* family).

The family belongs to the order *Cellvibrionales*, class *Gammaproteobacteria*, and encompasses mainly bacteria isolated from soil, marine environments or invertebrates. The description is as for the order. In addition, the ability to use complex polysaccharides as substrates is a widespread trait among strains belonging to this family. In the cellular fatty acid patterns of most strains unsaturated fatty acids ($C_{16:1}$, $C_{18:1}$, or $C_{17:1}$) are dominating. The G + C content of genomic DNA ranges from 44 to 58 mol%. At present the family comprises the genera *Cellvibrio*, *Eionea*, *Gilvimirinus*, *Maricurvus*, *Marinimicrobium*, *Pseudoteredinibacter*, *Saccharophagus*, *Simidua*, *Teredinibacter*, *Umboniibacter* and the *Candidatus* Endobugula. The type genus of the family is *Cellvibrio* [Blackall et al. 1986 (Parte, 2014)].

Description of *Halieaceae* fam. nov.

Halieaceae (Ha.lie.a'ce.ae. N.L. fem. n. *Haliea*, type genus of the family; suff. *-aceae*, ending to denote a family; N.L. fem. pl. n. *Halieaceae*, the *Haliea* family).

The family belongs to the order *Cellvibrionales*, class *Gammaproteobacteria*, and encompasses mainly bacteria inhabiting marine environments, especially coastal areas. The description is as for the order. In addition, some strains of this family are capable of aerobic photoheterotrophic growth using bacteriochlorophyll *a* and carotenoids for the harvesting of light. Genome analyses indicate that several strains may be able to apply proteorhodopsin for the utilization of light as additional energy source. In the cellular fatty acid patterns of most strains unsaturated fatty acids (C_{16:1} or C_{18:1}) are dominating. The G + C content of genomic DNA ranges from 52 to 66 mol%. Members of this family form a stable clade in reconstructed trees based on 16S rRNA gene sequences, so that currently a reliable affiliation of novel species to this family is possible based on the comparative sequence analyses of complete 16S rRNA genes. At present the family comprises the genera *Chromatocurvus*, *Congregibacter*, *Haliaea*, *Halioglobus*, *Luminiphilus*, and *Pseudohaliaea* as well as some uncharacterized strains belonging to the OM60/NOR5 clade as defined by Yan et al. (2009). The type genus of the family is *Haliaea* [Urios et al. 2008 (Parte, 2014)].

Description of *Microbulbiferaceae* fam. nov.

Microbulbiferaceae (Mi.cro.bul.bi.fer.aceae N.L. masc. n. *Microbulbifer*, type genus of the family; suff. *-aceae*, ending to denote a family; N.L. fem. pl. n. *Microbulbiferaceae*, the *Microbulbifer* family).

The family belongs to the order *Cellvibrionales*, class *Gammaproteobacteria*, and encompasses mainly bacteria isolated from saline sediments or soils. The description is as for the order. In addition, resting coccoid cells are formed in stationary phase. Some strains are able to degrade complex polysaccharides. Branched chain fatty acids (iso-C_{15:0}, iso-C_{17:1}, or iso-C_{17:0}) are dominating in cellular fatty acid patterns. The G + C content of genomic DNA ranges from 49 to 62 mol%. Members of this family form a stable clade in reconstructed trees based on 16S rRNA gene sequences, so that currently a reliable affiliation of novel species to this family is possible based on the comparative sequence analyses of complete 16S rRNA genes. The type and only genus of the family is *Microbulbifer* [González et al. 1997 (Parte, 2014)].

Description of *Porticoccaceae* fam. nov.

Porticoccaceae (Por.ti.coc.ca'ceae. N.L. masc. n. *Porticoccus*, type genus of the family; suff. *-aceae*, ending to denote a family; N.L. fem. pl. n. *Porticoccaceae*, the *Porticoccus* family).

The family belongs to the order *Cellvibrionales*, class *Gammaproteobacteria*, and encompasses mainly bacteria inhabiting marine environments. The description is as for the order. In addition, genome analyses indicate that some strains of this family may apply proteorhodopsin for utilizing light as additional energy source. Branched chain fatty acids (anteiso-C_{15:0} or anteiso-C_{17:0}) are dominating in cellular fatty acid patterns. The G + C content of genomic DNA ranges from 48 to 53 mol%. At present the family only contains the genus *Porticoccus* and some uncharacterized strains belonging to the SAR92 clade as defined by Stingl et al. (2007). The type genus of the family is *Porticoccus* [Oh et al. 2010 (Parte, 2014)].

Description of *Spongiibacteraceae* fam. nov.

Spongiibacteraceae (Spon.gi.i.bac.ter.aceae. N.L. masc. n. *Spongiibacter*, type genus of the family; suff. *-aceae*, ending to denote a family; N.L. fem. pl. n. *Spongiibacteraceae*, the *Spongiibacter* family).

The family belongs to the order *Cellvibrionales*, class *Gammaproteobacteria*, and encompasses mainly bacteria inhabiting marine environments. The description is as for the order. In addition, genome analyses indicate that some strains of this family may be able to use proteorhodopsin for harvesting light as additional energy source. In the cellular fatty acid patterns of most strains unsaturated fatty acids (C_{16:1} or C_{17:1}) are dominating. The G + C content of genomic DNA ranges from 47 to 61 mol%. At present the family comprises the genera *Dasania*, *Sinobacterium*, *Spongiibacter*, *Zhongshania*, "*Oceanicoccus*," and some uncharacterized strains belonging to the BD1-7 clade as defined by Cho and Giovannoni (2004). The type genus of the family is *Spongiibacter* [Graeber et al. 2008 (Parte, 2014)].

Acknowledgments

We are grateful to W. B. Whitman (University of Georgia, USA) and T. Gutierrez (Heriot Watt University, UK) for their allowance to use in this study unpublished genome data of "*Porticoccus hydrocarbonoclasticus*" and to the Leibniz Institute DSMZ for financial support.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2015.00281/abstract>

Supplementary Table S1 | Collection of 16S rRNA gene sequences representing major phylogenetic groups of the class Gammaproteobacteria.

Supplementary Figure S1 | Phylogenetic placement of the known clades of the OMG group. Phylogenetic analyses were based on a comprehensive dataset of almost complete 16S rRNA gene sequences representing mainly type genera within the *Gammaproteobacteria*. The sequence of *Magnetococcus marinus* MC-1^T was used as an outgroup (not shown). Polygons represent groups of two or more sequences. Clades of the OMG group are labeled in red. The phylogenetic lineage corresponding to the proposed order *Cellvibrionales* is shaded in blue. Quotation marks indicate names that are not yet validly published or illegitimate. The tree topology was reconstructed with the neighbor-joining method implemented in the ARB software using the correction of Felsenstein. In addition, trees were reconstructed using the maximum parsimony software of TNT and the RAxML maximum likelihood program. Support of a distinct branching by bootstrap analyses is indicated by symbols. Black dots at a distinct node indicate that bootstrap values of 95% or above (percentages of 1000 resamplings) were obtained with three different reconstruction methods, while white dots indicate that values of 95% or above were obtained with only two reconstruction methods. Hollow circles indicate that bootstrap values of 75% or above were obtained with at least one reconstruction method. In such cases the obtained values are given from left to right for the neighbor-joining, maximum-likelihood and maximum parsimony method. The scale bar represents an estimated sequence divergence.

Supplementary Figure S2 | Subtree of the comprehensive tree shown in Supplementary Figure S1. Only the position of clones or strains affiliated to the proposed order *Cellvibrionales* is displayed. Groups of sequences identified as distinct clades are distinguished by using shades of different colors. Symbols are explained in the legend of **Supplementary Figure S1**.

References

- Abt, B., Han, C., Scheuner, C., Lu, M., Lapidus, A., Nolan, M., et al. (2012). Complete genome sequence of the termite hindgut bacterium *Spirochaeta coccoides* type strain (SPN1^T), reclassification in the genus *Sphaerochaeta* as *Sphaerochaeta coccoides* comb. nov. and emendations of the family *Spirochaetaceae* and the genus *Sphaerochaeta*. *Stand. Genomic Sci.* 6, 194–209. doi: 10.4056/sigs.2796069
- Adékambi, T., Drancourt, M., and Raoult, D. (2009). The *rpoB* gene as a tool for clinical microbiologists. *Trends Microbiol.* 17, 37–45. doi: 10.1016/j.tim.2008.09.008
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Anderson, I., Scheuner, C., Göker, M., Mavromatis, K., Hooper, S. D., Porat, I., et al. (2011). Novel insights into the diversity of catabolic metabolism from ten haloarchaeal genomes. *PLoS ONE* 6:e20237. doi: 10.1371/journal.pone.0020237
- Andrews, J., and Harris, R. (1986). “r- and K-selection and microbial ecology,” in *Advances in Microbial Ecology*, Vol. 9, ed K. C. Marshall (New York, NY: Plenum press), 99–147.
- Baba, A., Miyazaki, M., Nagahama, T., and Nogi, Y. (2011). *Microbulbifer chitini-lyticus* sp. nov. and *Microbulbifer okinawensis* sp. nov., chitin-degrading bacteria isolated from mangrove forests. *Int. J. Syst. Evol. Microbiol.* 61, 2215–2220. doi: 10.1099/ijs.0.024158-0
- Blackall, L. L., Hayward, A. C., and Sl, L. I. (1985). Cellulolytic and dextranolytic Gram-negative bacteria: revival of the genus *Cellvibrio*. *J. Appl. Bacteriol.* 59, 81–97. doi: 10.1111/j.1365-2672.1985.tb01779.x
- Chen, M.-H., Sheu, S.-Y., Arun, A. B., Young, C.-C., Chen, C. A., Wang, J.-T., et al. (2011). *Pseudoteredinibacter isopora* gen. nov., sp. nov., a marine bacterium isolated from the reef-building coral *Isopora palifera*. *Int. J. Syst. Evol. Microbiol.* 61, 1887–1893. doi: 10.1099/ijs.0.024190-0
- Cheng, H., Zhang, S., Huo, Y.-Y., Jiang, X.-W., Zhang, X.-Q., Pan, J., et al. (2015). *Gilvimarinus polysaccharolyticus* sp. nov., an agar-digesting bacterium isolated from seaweed, and emended description of the genus *Gilvimarinus*. *Int. J. Syst. Evol. Microbiol.* 65, 562–569. doi: 10.1099/ijs.0.065078-0
- Cho, J., and Giovannoni, S. J. (2004). Cultivation and growth characteristics of a diverse group of oligotrophic marine gammaproteobacteria. *Appl. Environ. Microbiol.* 70, 432–440. doi: 10.1128/AEM.70.1.432-440.2004
- Courties, A., Riedel, T., Jarek, M., Papadatou, M., Intertaglia, L., Lebaron, P., et al. (2014). Draft genome sequence of the gammaproteobacterial strain MOLA455, a representative of a ubiquitous proteorhodopsin-producing group in the ocean. *Genome Announc.* 2, 2–3. doi: 10.1128/genome.A.01203-13
- Delsuc, F., Brinkmann, H., and Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6, 361–375. doi: 10.1038/nrg1603
- Dethlefsen, L., and Schmidt, T. M. (2007). Performance of the translational apparatus varies with the ecological strategies of bacteria. *J. Bacteriol.* 189, 3237–3245. doi: 10.1128/JB.01686-06
- Dupont, C. L., Rusch, D. B., Yooshep, S., Lombardo, M.-J., Alexander Richter, R., Valas, R., et al. (2012). Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J.* 6, 1186–1199. doi: 10.1038/ismej.2011.189
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Enright, A. J., Van Dongen, S., and Ouzounis, C. A. (2002). An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* 30, 1575–1584. doi: 10.1093/nar/30.7.1575
- Felsenstein, J. (1982). Numerical methods for inferring evolutionary trees. *Q. Rev. Biol.* 57, 379–404. doi: 10.1086/412935
- Gifford, S. M., Sharma, S., Booth, M., and Moran, M. A. (2012). Expression patterns reveal niche diversification in a marine microbial assemblage. *ISME J.* 7, 281–298. doi: 10.1038/ismej.2012.96
- Giovannoni, S. J., Cameron Thrash, J., and Temperton, B. (2014). Implications of streamlining theory for microbial ecology. *ISME J.* 8, 1553–1565. doi: 10.1038/ismej.2014.60
- Gogarten, J. P., Doolittle, W. F., and Lawrence, J. G. (2002). Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol.* 19, 2226–2238. doi: 10.1093/oxfordjournals.molbev.a004046
- Goloboff, P. A., Farris, J. S., and Nixon, K. C. (2008). TNT, a free program for phylogenetic analysis. *Cladistics* 24, 774–786. doi: 10.1111/j.1096-0031.2008.00217.x
- Iwaki, H., Takada, K., and Hasegawa, Y. (2012). *Maricurvus nonylphenolicus* gen. nov., sp. nov., a nonylphenol-degrading bacterium isolated from seawater. *FEMS Microbiol. Lett.* 327, 142–147. doi: 10.1111/j.1574-6968.2011.02471.x
- Jang, G. I., Hwang, C. Y., Choi, H.-G., Kang, S., and Cho, B. C. (2011a). Description of *Spongiibacter borealis* sp. nov., isolated from Arctic seawater, and reclassification of *Melitea salexigens* Urios et al. 2008 as a later heterotypic synonym of *Spongiibacter marinus* Graeber et al. 2008 with emended descriptions of the genus *Spongiibacter* and *Spongiibacter marinus*. *Int. J. Syst. Evol. Microbiol.* 61, 2895–2900. doi: 10.1099/ijs.0.028795-0
- Jang, Y., Oh, H.-M., Kang, I., Lee, K., Yang, S.-J., and Cho, J.-C. (2011b). Genome sequence of strain IMCC3088, a proteorhodopsin-containing marine bacterium belonging to the OM60/NOR5 clade. *J. Bacteriol.* 193, 3415–3416. doi: 10.1128/JB.05111-11
- Jeong, S. H., Yang, S.-H., Jin, H. M., Kim, J. M., Kwon, K. K., and Jeon, C. O. (2013). *Microbulbifer gwangyangensis* sp. nov. and *Microbulbifer pacificus* sp. nov., isolated from marine environments. *Int. J. Syst. Evol. Microbiol.* 63, 1335–1341. doi: 10.1099/ijs.0.042606-0
- Kämpfer, P., Arun, A. B., Young, C.-C., Rekha, P. D., Martin, K., Busse, H.-J., et al. (2012). *Microbulbifer taiwanensis* sp. nov., isolated from coastal soil. *Int. J. Syst. Evol. Microbiol.* 62, 2485–2489. doi: 10.1099/ijs.0.034512-0
- Klappenbach, J. A., Dunbar, J. M., and Schmidt, T. M. (2000). rRNA operon copy number reflects ecological strategies of bacteria. *Appl. Environ. Microbiol.* 66, 1328–1333. doi: 10.1128/AEM.66.4.1328-1333.2000
- Klenk, H.-P., and Göker, M. (2010). En route to a genome-based classification of Archaea and Bacteria? *Syst. Appl. Microbiol.* 33, 175–182. doi: 10.1016/j.syapm.2010.03.003
- Klenk, H. P., and Zillig, W. (1994). DNA-dependent RNA polymerase subunit B as a tool for phylogenetic reconstructions: branching topology of the archaeal domain. *J. Mol. Evol.* 38, 420–432. doi: 10.1007/BF00163158
- Kloesges, T., Popa, O., Martin, W., and Dagan, T. (2011). Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths. *Mol. Biol. Evol.* 28, 1057–1074. doi: 10.1093/molbev/msq297
- Konstantinidis, K. T., and Tiedje, J. M. (2005). Towards a genome-based taxonomy for prokaryotes. *J. Bacteriol.* 187, 6258–6264. doi: 10.1128/JB.187.18.6258-6264.2005
- Kyrpides, N. C., Hugenholtz, P., Eisen, J. A., Woyke, T., Göker, M., Parker, C. T., et al. (2014). Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. *PLoS Biol.* 12:e1001920. doi: 10.1371/journal.pbio.1001920
- Lauro, F. M., McDougald, D., Thomas, T., Williams, T. J., Egan, S., Rice, S., et al. (2009). The genomic basis of trophic strategy in marine bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15527–15533. doi: 10.1073/pnas.0903507106
- Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320. doi: 10.1093/molbev/msn067
- Lee, Y. K., Hong, S. G., Cho, H. H., Cho, K. H., and Lee, H. K. (2007). *Dasania marina* gen. nov., sp. nov., of the order *Pseudomonadales*, isolated from Arctic marine sediment. *J. Microbiol.* 45, 505–509.
- Lerat, E., Daubin, V., and Moran, N. A. (2003). From gene trees to organismal phylogeny in prokaryotes: the case of the γ -proteobacteria. *PLoS Biol.* 1:e19. doi: 10.1371/journal.pbio.0000019
- Li, L., Stoecckert, C. J., and Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178–2189. doi: 10.1101/gr.1224503
- Lo, N., Kang, H. J., and Jeon, C. O. (2014). *Zhongshania aliphaticivorans* sp. nov., an aliphatic hydrocarbon degrading bacterium isolated from marine sediment, and transfer of *Spongiibacter borealis* Jang et al. 2011 to the genus *Zhongshania* as *Zhongshania borealis* comb. nov. *Int. J. Syst. Evol. Microbiol.* 64, 3768–3774. doi: 10.1099/ijs.0.068593-0
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004). ARB: a software environment for sequence data. *Nucleic Acids Res.* 32, 1363–1371. doi: 10.1093/nar/gkh293
- Maeda, T., Hayakawa, K., You, M., Sasaki, M., Yamaji, Y., Furushita, M., et al. (2005). Characteristics of nonylphenol polyethoxylate-degrading bacteria isolated from coastal sediments. *Microbes Environ.* 20, 253–257. doi: 10.1264/jsme2.20.253

- Markowitz, V. M., Mavromatis, K., Ivanova, N. N., Chen, I.-M., Chu, K., and Kyrpides, N. C. (2009). IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 25, 2271–2278. doi: 10.1093/bioinformatics/btp393
- Meier-Kolthoff, J. P., Hahnke, R. L., Petersen, J., Scheuner, C., Michael, V., Fiebig, A., et al. (2014a). Complete genome sequence of DSM 30083^T, the type strain (U5/41^T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand. Genomic Sci.* 9:2. doi: 10.1186/1944-3277-9-2
- Meier-Kolthoff, J. P., Klenk, H. P., and Göker, M. (2014b). Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int. J. Syst. Evol. Microbiol.* 64, 352–356. doi: 10.1099/ijs.0.056994-0
- Meusemann, K., von Reumont, B. M., Simon, S., Roeding, F., Strauss, S., Kück, P., et al. (2010). A phylogenomic approach to resolve the arthropod tree of life. *Mol. Biol. Evol.* 27, 2451–2264. doi: 10.1093/molbev/msq130
- Mollet, C., Drancourt, M., and Raoult, D. (1997). *rpoB* sequence analysis as a novel basis for bacterial identification. *Mol. Microbiol.* 26, 1005–1011. doi: 10.1046/j.1365-2958.1997.6382009.x
- Moore, E. R. B., Tindall, B. J., Martins Dos Santos, V. A. P., Pieper, D. H., Ramos, J.-L., and Palleroni, N. J. (2006). “Nonmedical: *Pseudomonas*,” in *The Prokaryotes*, Vol. 6, eds M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt (New York, NY: Springer New York), 646–703.
- Mulet, M., Lalucat, J., and Garcia-Valdés, E. (2010). DNA sequence-based analysis of the *Pseudomonas* species. *Environ. Microbiol.* 12, 1513–1530. doi: 10.1111/j.1462-2920.2010.02181.x
- Oh, H.-M., Kang, I., Ferriera, S., Giovannoni, S. J., and Cho, J.-C. (2010a). Genome sequence of the oligotrophic marine gammaproteobacterium HTCC2143, isolated from the Oregon Coast. *J. Bacteriol.* 192, 4530–4531. doi: 10.1128/JB.00683-10
- Oh, H.-M., Kim, H., Kim, K.-M., Min, G.-S., and Cho, J.-C. (2010b). *Porticoccus litoralis* gen. nov., sp. nov., a gammaproteobacterium isolated from the Yellow Sea. *Int. J. Syst. Evol. Microbiol.* 60, 727–732. doi: 10.1099/ijs.0.013938-0
- Oyaizu, H., and Komagata, K. (1983). Grouping of *Pseudomonas* species on the basis of cellular fatty acid composition and the quinone system with special reference to the existence of 3-hydroxy fatty acids. *J. Gen. Appl. Microbiol.* 29, 17–40. doi: 10.2323/jgam.29.17
- Park, S., Kim, S.-I., Jung, Y.-T., and Yoon, J.-H. (2014). *Simidiua curdlanivorans* sp. nov., a curdlan-degrading bacterium isolated from the junction between the ocean and a freshwater spring and emended description of the genus *Simidiua*. *Int. J. Syst. Evol. Microbiol.* 64, 3695–3700. doi: 10.1099/ijs.0.065334-0
- Park, S., Yoshizawa, S., Inomata, K., Kogure, K., and Yokota, A. (2012). *Halioglobus japonicus* gen. nov., sp. nov. and *Halioglobus pacificus* sp. nov., members of the class Gammaproteobacteria isolated from seawater. *Int. J. Syst. Evol. Microbiol.* 62, 1784–1789. doi: 10.1099/ijs.0.031443-0
- Parte, A. C. (2014). LPSN—list of prokaryotic names with standing in nomenclature. *Nucleic Acids Res.* 42, D613–D616. doi: 10.1093/nar/gkt1111
- Philippe, H., and Douady, C. J. (2003). Horizontal gene transfer and phylogenetics. *Curr. Opin. Microbiol.* 6, 498–505. doi: 10.1016/j.mib.2003.09.008
- Philippot, L., Andersson, S. G., Battin, T. J., Prosser, J. I., Schimel, J. P., Whitman, W. B., et al. (2010). The ecological coherence of high bacterial taxonomic ranks. *Nat. Rev. Microbiol.* 8, 523–529. doi: 10.1038/nrmicro2367
- Polz, M. F., Hunt, D. E., Preheim, S. P., and Weinreich, D. M. (2006). Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 2009–2021. doi: 10.1098/rstb.2006.1928
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schwier, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219
- Romanenko, L. A., Tanaka, N., and Frolova, G. M. (2010). *Umboniibacter marinipunicus* gen. nov., sp. nov., a marine gammaproteobacterium isolated from the mollusc *Umbonium costatum* from the Sea of Japan. *Int. J. Syst. Evol. Microbiol.* 60, 603–609. doi: 10.1099/ijs.0.010728-0
- Rowland, G. C., Aboshkiwa, M., and Coleman, G. (1993). Comparative sequence analysis and predicted phylogeny of the DNA-dependent RNA polymerase beta subunits of *Staphylococcus aureus* and other eubacteria. *Biochem. Soc. Trans.* 21:40S.
- Ruff, S. E., Probandt, D., Zinkann, A. C., Iversen, M. H., Klaas, C., Würzberg, L., et al. (2014). Indications for algae-degrading benthic microbial communities in deep-sea sediments along the Antarctic Polar Front. *Deep. Res. II Top. Stud. Oceanogr.* 108, 6–16. doi: 10.1016/j.dsr2.2014.05.011
- Scheuner, C., Tindall, B. J., Lu, M., Nolan, M., Lapidus, A., Cheng, J., et al. (2014). Complete genome sequence of *Planctomyces brasiliensis* type strain (DSM 5305^T), phylogenomic analysis and reclassification of *Planctomycetes* including the descriptions of *Gimesia* gen. nov., *Planctopirus* gen. nov. and *Rubinisphaera* gen. nov. and emended descriptions of the order Planctomycetales and the family Planctomycetaceae. *Stand. Genomic Sci.* 9:10. doi: 10.1186/1944-3277-9-10
- Spring, S. (2014). Function and evolution of the Sox multienzyme complex in the marine gammaproteobacterium *Congregibacter litoralis*. *ISRN Microbiol.* 2014:597418. doi: 10.1155/2014/597418
- Spring, S., Fiebig, A., Riedel, T., Göker, M., and Klenk, H.-P. (2014). Genome sequence of gammaproteobacterial *Pseudohalaea rubra* type strain DSM 19751, isolated from coastal seawater of the Mediterranean Sea. *Genome Announc.* 2:e01208-14. doi: 10.1128/genomeA.01208-14
- Spring, S., Lünsdorf, H., Fuchs, B. M., and Tindall, B. J. (2009). The photosynthetic apparatus and its regulation in the aerobic gammaproteobacterium *Congregibacter litoralis* gen. nov., sp. nov. *PLoS ONE* 4:e4866. doi: 10.1371/journal.pone.0004866
- Spring, S., and Riedel, T. (2013). Mixotrophic growth of bacteriochlorophyll a-containing members of the OM60/NOR5 clade of marine gammaproteobacteria is carbon-starvation independent and correlates with the type of carbon source and oxygen availability. *BMC Microbiol.* 13:117. doi: 10.1186/1471-2180-13-117
- Spring, S., Riedel, T., Spröer, C., Yan, S., Harder, J., and Fuchs, B. M. (2013). Taxonomy and evolution of bacteriochlorophyll a-containing members of the OM60/NOR5 clade of marine gammaproteobacteria: description of *Luminiphilus sylvensis* gen. nov., sp. nov., reclassification of *Halaea rubra* as *Pseudohalaea rubra* gen. nov., comb. nov., and emendation of *Chromatococcus halotolerans*. *BMC Microbiol.* 13:118. doi: 10.1186/1471-2180-13-118
- Spring, S., Scheuner, C., Lapidus, A., Lucas, S., Glavina Del Rio, T., Tice, H., et al. (2010). The genome sequence of *Methanohalophilus mahii* SLP^T reveals differences in the energy metabolism among members of the *Methanosarcinaceae* inhabiting freshwater and saline environments. *Archaea* 2010:690737. doi: 10.1155/2010/690737
- Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446
- Stingl, U., Desiderio, R. A., Cho, J.-C., Vergin, K. L., and Giovannoni, S. J. (2007). The SAR92 clade: an abundant coastal clade of culturable marine bacteria possessing proteorhodopsin. *Appl. Environ. Microbiol.* 73, 2290–2296. doi: 10.1128/AEM.02559-06
- Su, J., Zhou, Y., Lai, Q., Li, X., Dong, P., Yang, X., et al. (2013). *Sinobacterium caligoides* gen. nov., sp. nov., a new member of the family Oceanospirillaceae isolated from the South China Sea, and emended description of *Amphritea japonica*. *Int. J. Syst. Evol. Microbiol.* 63, 2095–2100. doi: 10.1099/ijs.0.030965-0
- Suarez, C., Ratering, S., Kramer, I., and Schnell, S. (2014). *Cellvibrio diazotrophicus* sp. nov., a nitrogen-fixing bacteria isolated from the rhizosphere of salt meadow plants and emended description of the genus *Cellvibrio*. *Int. J. Syst. Evol. Microbiol.* 64, 481–486. doi: 10.1099/ijs.0.054817-0
- Swofford, D. L. (2002). *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sunderland, MA: Sinauer Associates.
- Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577. doi: 10.1080/10635150701472164
- Teske, A., Hinrichs, K.-U., Edgcomb, V., de Vera Gomez, A., Kysela, D., Sylva, S. P., et al. (2002). Microbial diversity of hydrothermal sediments in the Guaymas Basin: evidence for anaerobic methanotrophic communities. *Appl. Environ. Microbiol.* 68, 1994–2007. doi: 10.1128/AEM.68.4.1994-2007.2002
- Thompson, J. D., Thierry, J. C., and Poch, O. (2003). RASCAL: rapid scanning and correction of multiple sequence alignments. *Bioinformatics* 19, 1155–1161. doi: 10.1093/bioinformatics/btg133
- Urios, L., Intertaglia, L., Lesongeur, F., and Lebaron, P. (2011). *Eionea nigra* gen. nov., sp. nov., a gammaproteobacterium from the Mediterranean Sea. *Int. J. Syst. Evol. Microbiol.* 61, 1677–1681. doi: 10.1099/ijs.0.023952-0
- Vashist, P., Nogi, Y., Ghadi, S. C., Verma, P., and Shouche, Y. S. (2013). *Microbulbifer mangrovi* sp. nov., a polysaccharide-degrading bacterium isolated from an Indian mangrove. *Int. J. Syst. Evol. Microbiol.* 63, 2532–2537. doi: 10.1099/ijs.0.042978-0

- Williams, K. P., Gillespie, J. J., Sobral, B. W., Nordberg, E. K., Snyder, E. E., Shalom, J. M., et al. (2010). Phylogeny of gammaproteobacteria. *J. Bacteriol.* 192, 2305–2314. doi: 10.1128/JB.01480-09
- Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N. N., et al. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* 462, 1056–1060. doi: 10.1038/nature08656
- Wu, M., and Eisen, J. A. (2008). A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* 9:R151. doi: 10.1186/gb-2008-9-10-r151
- Yamamoto, S., and Harayama, S. (1998). Phylogenetic relationships of *Pseudomonas putida* strains deduced from the nucleotide sequences of *gyrB*, *rpoD* and 16S rRNA genes. *Int. J. Syst. Bacteriol.* 48, 813–819. doi: 10.1099/00207713-48-3-813
- Yan, S., Fuchs, B. M., Lenk, S., Harder, J., Wulf, J., Jiao, N.-Z., et al. (2009). Biogeography and phylogeny of the NOR5/OM60 clade of Gammaproteobacteria. *Syst. Appl. Microbiol.* 32, 124–139. doi: 10.1016/j.syapm.2008.12.001
- Yarza, P., Richter, M., Peplies, J., Euzéby, J., Amann, R., Schleifer, K.-H., et al. (2008). The all-species living tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst. Appl. Microbiol.* 31, 241–250. doi: 10.1016/j.syapm.2008.07.001
- Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K.-H., et al. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* 12, 635–645. doi: 10.1038/nrmicro3330
- Yoon, J.-H., Kang, S.-J., Jung, Y.-T., and Oh, T.-K. (2009). *Marinimicrobium locisalis* sp. nov., isolated from a marine solar saltern, and emended description of the genus *Marinimicrobium*. *Int. J. Syst. Evol. Microbiol.* 59, 2260–2263. doi: 10.1099/ijs.0.008458-0
- Yooséph, S., Neelson, K. H., Rusch, D. B., McCrow, J. P., Dupont, C. L., Kim, M., et al. (2010). Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* 468, 60–66. doi: 10.1038/nature09530
- Zhang, D.-S., Huo, Y.-Y., Xu, X.-W., Wu, Y.-H., Wang, C.-S., Xu, X.-F., et al. (2012). *Microbulbifer marinus* sp. nov. and *Microbulbifer yueqingensis* sp. nov., isolated from marine sediment. *Int. J. Syst. Evol. Microbiol.* 62, 505–510. doi: 10.1099/ijs.0.027714-0

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

Copyright © 2015 Spring, Scheuner, Göker and Klenk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.