Unlocking the Genomic Taxonomy of the Prochlorococcus Collective

Diogo Tschoeke¹ · Vinicius W. Salazar¹ · Livia Vidal¹ · Mariana Campeão¹ · Jean Swings¹,² · Fabiano Thompson¹ · Cristiane Thompson¹

Received: 13 February 2020 / Accepted: 7 May 2020 / Published online: 28 May 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Prochlorococcus is the most abundant photosynthetic prokaryote on our planet. The extensive ecological literature on the Prochlorococcus collective (PC) is based on the assumption that it comprises one single genus comprising the species Prochlorococcus marinus, containing itself a collective of ecotypes. Ecologists adopt the distributed genome hypothesis of an open pan-genome to explain the observed genomic diversity and evolution patterns of the ecotypes within PC. Novel genomic data for the PC prompted us to revisit this group, applying the current methods used in genomic taxonomy. As a result, we were able to distinguish the five genera: Prochlorococcus, Eurycolium, Prolificoccus, Thaumococcus, and Riococcus. The novel genera have distinct genomic and ecological attributes.

Keywords Prochlorococcus collective · Eco-genomics · Genomic taxonomy · Evolution

Introduction

Prochlorococcus cells are the smallest cells that emit red chlorophyll fluorescence, which is how they were discovered approximately 30 years ago [1, 2]. It is the smallest and most abundant photosynthetic microbe on Earth. The oxytrophic free-living marine Prochlorococcus marinus contain divinyl chlorophyll a (chl a2) and both monovinyl and divinyl chlorophyll b (chl b2) as their major photosynthetic pigments rather than chlorophyll a and phycobiliproteins, which are typical of Cyanobacteria, allowing them to adapt to different light intensities [1, 3–6], particularly to capture the blue wavelengths prevailing in oligotrophic waters [7]. Prochlorococcus strains have been cultured and are available in culture collections, but the limited number of cultures does not reflect the true phenotypic and genomic diversity found in the ocean [8]. Prochlorococcus is a smaller, differently pigmented variant of marine Synechococcus [2]. The pioneering work on the eco-genomics of Prochlorococcus was performed by Chisholm’s group at MIT [2, 9–15].

The taxonomy of the Prochlorococcus collective (PC) has been neglected in the last three decades. The designation PC was first coined by Kashtan et al. [11]. This collective comprises the genera Prochlorococcus, Prolificoccus, Eurycolium, and Thaumococcus [16–18], a group of picocyanobacteria that radiated from Synechococcus millions of years ago. We preferer to use the neutral term “collective” here instead of “community,” “complex,” or any other term.

Most literature on PC is based on the assumption that it comprises a single genus, in which the species Prochlorococcus marinus is within a “collective” of ecotypes. The main reason for this concept is that 16S rRNA gene sequence similarities within this group are > 97%. To describe the observed diversity in the PC, ecologists adopt the distributed genome hypothesis of an open bacterial pan-genome (supragenome) [12, 14]. In this hypothesis, the pan-genome is the full complement of genes that each member of the population contributes to and draws genes from. No single isolate contains the full complement of genes, which contributes to high diversity and provides selective advantages [19]. Described in its extreme form, the pan-genome of the bacterial domain is of infinite size, and the bacteria as a whole present an open pan-genome (supragenome), in which the constant “rain” of genetic material in genomes from a cloud of frequently transferred genes enhances the chance of survival of a
species [20]. The PC case illustrates the continuously growing gap between prokaryotic taxonomy as it is practiced today and in-depth eco-genomic studies. Genomic clusters are referred to as ecotypes, ecospecies, sub-ecotypes, clades, populations, subpopulations, OTUs, or CTUs. This informal nomenclature finds its way into databases, adding further nomenclatural chaos to “terrestrial incognitae” [21, 22] comprising these picocyanobacteria.

In their theory of self-amplification and self-organization of the biosphere, Braakman et al. [3] hypothesize that Prochlorococcus clades living in surficial oceanic waters adapted to high light and low nutrient levels have evolved from more ancient Prochlorococcus subpopulations (~800 million years ago) inhabiting deeper oceanic areas, with lower light and higher nutrient concentrations. They refer to these evolutionary paths as “niche constructing adaptive radiations.” Sun and Blanchard [23] hypothesize three phases in the evolutionary history of the PC: genome reduction, genome diversification, and ecological niche specialization, all due to different evolutionary forces. The PC has adapted to live under different light intensities, seawater temperatures, and nutrient levels, and the abundance of the different ecotypes varies markedly along gradients of these parameters [24–30]. However, only high light (HL)–adapted and low light (LL)–adapted ecotypes were initially distinguished [31]. At present, twelve phylogenetic ecotypes are delineated: six HL-adapted ecotypes and six LL-adapted ecotypes [32]. These ecotypes are based on the phylogeny of the 16S/23S rRNA internal transcribed spacer region, differences in behavior towards light, whole-genome analysis, and physiological properties, combined with their environmental distribution [28, 30, 33]. Not only light adaptation but also temperature adaptation have been recognized in the two dominant near-surface HL ecotypes, the HLI ecotype Eurycolium eMIT9312, which both exhibit growth maxima at approximately 25 °C, but Eurycolium eMIT9312 appears to be adapted to thrive at much lower temperatures (approx. 14 °C), whereas Eurycolium eMIT9312 is able to grow at 30 °C [34]. These ecotypes appear to have nearly identical genomes [9], in which the major differences are restricted to five genomic islands, but show marked phenotypic differences.

The relative abundance of cells belonging to the different ecotypes has been estimated in several oceanic regions, revealing patterns that agree for the most part with the HL/LL phenotypes. HL-adapted cells dominate the surface mixed layer, and LL-adapted cells most often dominate deeper waters [32, 35–38]. By combining the physiological features of isolates and ecotype abundance studies in the ocean, it was shown that temperature, in addition to light, is an important determinant of the abundance of these ecotypes in oceans [39, 40]. There is also tremendous variability within ecotypes, particularly with regard to nutrient acquisition [41], susceptibility to predation or phage infection [42], and modes of interaction with other members of the community [43], including interactions with Alteromonas for protection against ROS [44, 45]. Prochlorococcus can use different organic compounds containing key elements to survive in oligotrophic oceans, and can also take up glucose and use it as a source of carbon and energy [46].

Kashtan et al. [11] studied the genomic diversity of single Prochlorococcus cells in samples collected in the Bermuda-Atlantic Time-Series Study (BATS) site in summer, spring, and winter, and hundreds of new Prochlorococcus “subpopulations” with distinct genomic backbones were found. These subpopulations are estimated to have diverged millions of years ago, suggesting ancient, stable niche partitioning followed by selection (not by genetic drift). The authors suggested that the global PC may be conceived as an assortment of thousands of species. They argued that such a large set of coexisting species with distinct genomic backbones is a characteristic feature of free-living prokaryotes with very large population sizes living in highly mixed habitats. In another study by these authors focused on the Pacific Ocean, they concluded that the Atlantic and Pacific Oceans are occupied by different PC populations that experienced separate evolutionary paths for a few million years [13].

From 1970 to the present, the need to classify, name, and identify prokaryotes has been based on a pragmatic polyphasic consensus approach [47]. This non-theory-based approach aims to integrate phenotypic, genotypic, and phylogenetic and ecological characteristics to establish stable and informative classification systems. The central unit is the prokaryotic species. In this procedure, the definition of prokaryotic species is based mainly on DNA–DNA hybridization (DDH) and 16S rRNA sequence similarity: a species is defined as a monophyletic group of strains sharing high phenotypic and genomic similarity (>70% DDH, >98.3% 16S rRNA sequence similarity) [48–52].

Prokaryote taxonomy is becoming more genome based, considering the genomes themselves as the ultimate end products of microbial evolution [50, 53, 54]. Such an approach offers the opportunity to obtain a more natural taxonomic system and allows the definition of consistent and stable monophyletic species and genera with maximal utility for its users [47, 48, 55–75]. Cohan has extensively developed the ecotype or ecospecies theory for prokaryotes, in which “ecotypes are the most newly divergent populations that are distinct ecologically from one another” [76, 77]. In this theory, the splitting of one ecotype into two is the fundamental diversity-creating process of speciation in bacteria. In his stable ecotype model, ecotypes (ecospecies) are long lived, and different ecotypes coexist indefinitely. They are cohesive by virtue of periodic selection events that purge the ecotype of sequence diversity. In this case, the ecotype has all the characteristics of a species [77, 78]. Ecotypes have been discerned among the PC.
The main motivation for the present study arose from the consideration that the taxonomy of Prochlorococcus has been relatively neglected compared with the tremendous advances in marine microbial taxonomy achieved for other groups, e.g., Alteromonas (N = 38 species) [79], Colwellia (N = 21 species) [80], Endozoicomonas (N = 9 species) [81], Marinobacter (N = 45 species) [82], Marinobacterium (N = 17 species) [83], Marinomonas (N = 30 species) [84], Pseudoalteromonas (N = 53 species) [85], Pseudovibrio (N = 6 species) [86], Roseobacter (N = 5 species) [87], Ruegeria (N = 22 species) [88], Shewanella (N = 67 species) [89], Salinispora (N = 3 species) [90], Streptomyces (N = 848 species) [91], Thalassomonas (N = 9 species) [92], and the vibrio clade (N = 199 species, including the genera Vibrio, Photobacterium, Enterovibrio, Aliivibrio, Thaumativibrio, and Grimontia) [93]. Even marine prokaryotes living in extreme oceanic environments such as the OMZ, sulfate-reducing bacteria, N cycle bacteria, and deep sea hydrothermal vent sulfur-oxidizing Sulfurovum (N = 4 species) [94] and Sulforimonas (N = 4 species) [95] have been the subject of advanced taxonomic classification. Nevertheless, even in the well-studied Gram-negative Roseobacter [96] and Gram-positive Streptomyces [97] genera, the paraphyletic nature of these genera leads to noncoherent assignment of scientific names and causes confusion when interpreting their ecology or evolutionary biology.

As new habitats and niches are found for Prochlorococcus, it is important to establish a solid taxonomic system to contribute to the interpretation of the ecology of the PC. Novel PC genomic data allowed us to address the hypothesis that novel genera and species are present within the PC. In this review, we revisit the current taxonomy of the PC and propose the creation of a new genus to encompass the clearly recognized clusters. The novel genera have distinct genomic and ecological attributes.

Methods

A Proposal of a New Taxonogenic Framework for the PC

The availability of Prochlorococcus genomes labelled Prochlorococcus in public databases prompted us to test the hypothesis that the PC comprises a greater number of stable genera and species than have been previously recognized (Table 1 and S1). We first collected all the publicly available genomes designated as Prochlorococcus, and applied the theory-free genomic methodologies and principles discussed previously. The Prochlorococcus genomes came from samples distributed around the world, especially in the Pacific and Atlantic Oceans (Table S1). First, the GC% content of the genomes was determined. A phylogenomic framework was established using genomic signatures, i.e., average amino acid identity (AAI), multilocus sequence analysis (MLSA), and core genome-based tree. We then classified the genomes into different genera and species using relevant AAI and phenotypic information obtained directly from the genomes.

Genome Quality Assessment

A total of 232 genomes and the corresponding metadata were selected from NCBI GenBank [98] because they fit our quality control criteria. Contigs or scaffold files were input into Prodigal software v2.6.3 [55] for gene and protein prediction. To check the quality of the genomes, completeness, contamination, and GC%, the genomes were submitted to CheckM v1.0.11 [56] using default parameters for this estimation. CheckM uses a broader set of marker genes specific to the position of a genome within a reference genome tree. The completeness is measured using a set of marker genes that are expected to be present. This expectation can be defined along an evolutionary gradient, ranging from highly conserved genes to species-specific genes. The genomes used are of high quality and have ≥ 85% of completeness, ≥ 1 Mb of size, and ≤ 5% of contamination (Table S1).

Average Amino Acid Identity Analyses

The average amino acid identity was calculated as described previously for all 232 genomes [57]. All PC genomes were used to calculate the amino acid average identity using the CompareM tool (https://github.com/dparks1134/CompareM). The result was used to plot a dendrogram. The CompareM output was converted to a dissimilarity square matrix. The spatial.distance. squareform function from scipy package v1.1.0 was used in order to obtain a condensed list from the lower triangular elements of the matrix. A complete linkage clustering was performed using cluster.hierarchy. linkage function from scipy with the output of squareform and cityblock metric distance. The figure that represents the linkage matrix was plotted with cluster.hierarchy. dendrogram function from scipy. The intraspecies limit is assumed as ≥ 95%, whereas genera delimitation is assumed as ≥ 70%. Genera were delineated only for those lineages exhibiting at least 3 representative genome sequences.

Phylogenomic Reconstruction

The MLSA approach was based on the concatenated sequences of four housekeeping genes (gyrB, pyrH, recA,
with intraspecies relatedness ranging from > 95 to 100 % MLSA [16]. This was achieved by using tBLASTn [58]. A nucleotide alignment was then made for each of the four genes using MAFFT [59], and these alignments concatenated using MAFFT with default parameters and a custom Perl script. The tree was built from the concatenated alignment using FastTree 2 [60] with default parameters and rooted on *S. elongatus* PCC 6301 using ETE 3 [61]. To build the core genome tree, we used the GToTree package [62] with default parameters. The 232 selected genomes were searched against a Hidden Markov Model of 250 cyanobacterial marker genes using HMMER3 [63]. A concatenated protein alignment from the 250 genes was constructed using Muscle [64] and subsequently trimmed using TrimAl [65]. The alignment was then used to construct a tree using FastTree 2 [60] with default parameters and the pairwise distance matrix using MEGA 6.0 [66]. The tree was rooted on *S. elongatus* PCC 6301 using ETE 3 [61]. Part of the processing was done with GNU parallel [67].

### In Silico Phenotype

The phenotype was predicted utilizing the sequences of the nitrogen-related (*narB, nirA, ntcA, glnB, amt, cynS, ureA, glnA, gisF, gdhA*), phosphorous-related (*phoB, ptrA, phoR*), arsenate-related (*arsC*), iron (*piuC, som*), and photosynthetic pigment-related (*psaAB, pcbABCDEFGH, CP47*, and *katG*) genes of *Prochlorococcus* obtained from UniProt and GenBank as queries for a local blastp [68] search against all PC ORF-predicted genomes. The results were parsed with Python scripts, and the sequences corresponding to the blastp hits were extracted from the respective genomes. The extracted sequences were used to perform a blastp search against the nr database to confirm the functional annotation. Manual verification of these results was performed. The confirmed results were used to build a matrix of presence/absence for the genes related to the indicated proteins in each PC genome.

### Results and Discussion

#### GC Content Reveals Great Genomic Diversity in the PC

One of the most striking findings of our study came from the analysis of GC content, which ranged from 30 to 50.7% (Fig. 1 and Table 1). The distribution of the GC values of all the genomes comprised five groups: (i) the majority (75% of the genomes) exhibited 30 to 33% GC, comprising the genus *Eurycolium*; (ii) a second group (12% of the genomes) exhibited 34 to 35% GC, comprising the genera *Prolificoccus*; (iii) a third group *Prochlorococcus* (4.8% of the genomes) presented 36–37% GC; (iv) a fourth group *Riococcus* (1.4% of the genomes) exhibited 37 to 38% GC; and (v) the fifth group (6.7% of the genomes) exhibited 50 to 51% GC, which was characteristic of the genera *Thaumococcus*. The average genome sizes of the first four GC groups were similar, whereas the genome sizes of the fifth GC group were almost twofold larger. This GC analysis provides a clear indication of the taxonomic width of the PC. Mutational bias is dominated by GC > AT transitions even in bacteria with a high % GC content [69, 70].

![Fig. 1](image-url) Genome clustering in five groups based on GC content and genome size. The colors (green, red, blue, purple, and yellow) represent the five groups. The five different symbols represent the genera where each genome were classified.
Unlocking the Genomic Taxonomy of the Prochlorococcus Collective

CCMP1375T was Prochlorococcus marinus and 99.9%. The closest genus to AAI among PC members varied considerably, between 57 Thaumococcus and Prolificoccus followed by 16) were the genera with the greatest numbers of species of PC (Fig. 2). Eurycolium (112) 180 Riococcus (3) 3 Thaumococcus swingsii
Eurycolium pastoris (16) 25 Prolificoccus Proteus
Thaumococcus
Prochlorococcus

Average Amino Acid Identity of PC

The species and genus delimitations were based on clusters with three or more genomes that shared ≥ 95% or ≥ 70% amino acid average identity (AAI), respectively. Here, we revisit the current taxonomy of the PC and propose the creation of a new genus Riococcus and 128 novel species (N = 137 species) (Table 1 and S1). We confirmed the previous proposal of Walter et al. [18], and we now expand the classification of PC (Fig. 2). Eurycolium (N = 112) and Prolificoccus (N = 16) were the genera with the greatest numbers of species followed by Prochlorococcus (N = 4), Riococcus (N = 3), and Thaumococcus (N = 2) (Fig. 2, Table 1 and S1). The AAI among PC members varied considerably, between 57 and 99%. The closest genus to Prochlorococcus marinus CCMP1375T was Riococcus MIT9211T having 69.4% AAI. The percentages of AAI variability between genera and species were 57 to 69.9% and 70 to 94%, respectively. The percentages of AAI variability within genera and species were > 70 to 94% and 95 to 99.9%, respectively.

Phylogenomic Framework for the PC

To establish the phylogenetic structure of the five PC genera, phylogenetic analysis was performed using a four-locus-based tree (MLSA) and 250 core protein sequences. Bootstrap analysis indicated that most of the branches were highly significant. These analyses have a much higher resolution than the 16S rRNA phylogeny. The PC harbors very similar 16S rRNA gene sequences with similarity ranging from 97 to 99% [2]. The MLSA and the core protein tree showed the relationship of the five genera (Fig. 3). The MLSA and core protein trees were congruent. The MLSA and the core protein sequence identities among PC members varied considerably, from 49 to 100% and from 56 to 100%, respectively. The type species of the genus Prochlorococcus, P. marinus CCMP1375T, showed identity maxima of 80% and 70% MLSA and core protein sequence similarity, respectively, with all other genera. The variability between species was < 95, while that within species ranged from 96 to 100% similarity. The phylogenetic analysis based on MLSA and core protein sequences (Fig. 3) demonstrated congruence between the genetic and ecological features of the five novel genera. For instance, Eurycolium comprises high-light (HLI/II/VI) ecotypes with 30–33% GC contents, whereas the genera Prolificoccus and Prochlorococcus comprise low-light (LLI and LLI/III, respectively) ecotypes with 34–35% and 36–37% GC contents, respectively. Riococcus and Thaumococcus consist of low-light ecotypes (LLIII and LLIV, respectively) with 37 to 38% and 50 to 50.7% GC contents, respectively.

In Silico Phenotype Prediction

Useful phenotypic features that may be used for the identification and classification of the PC were identified, including nitrogen, phosphorous, arsenate, and iron utilization (Table 2). The genus Prochlorococcus does not present some nitrogen-related genes (narB, nirA, cynS, ureA, and gdhA) and phosphorous-related genes (phoB, ptrA, and phoR) but some genomes do present other nitrogen-related genes (ntrA, glnB, amt, glnA, glnF), an arsenate-related gene (arsC), and iron-related genes (pinC and som). Prochlorococcus is differentiated from its closest genus Riococcus by the presence of the arsC gene. The genes related to photosynthetic pigments synthesis (psaAB, pcbABCDEFGH, and CP47) were conserved among the genera (Table S2). All genera present the gene Pro1404 that confers the capability for glucose uptake in PC [46]. Overall, we demonstrated that the methods used (AAI, core protein sequences, and MLSA) correlated and provided significant taxonomic resolution for differentiation of genera and species in PC (Figs. 2 and 3). The proposed genera are in complete agreement with previous literature and support previous theories on the evolution of Prochlorococcus [11, 13, 14, 31, 33, 36, 71]. The PC genomic signatures were more

| Genus (No. of species) | No. of genomes | Type species | Type strain | Range % mol GC | Average length (Mbp) |
|------------------------|----------------|--------------|-------------|----------------|---------------------|
| Eurycolium (112)       | 180            | Eurycolium pastoris | CCMP1986T | 30–33 | 1.5 ± 0.1 |
| Prochlorococcus (04)   | 10             | Prochlorococcus marinus | CCMP1375T | 36–37 | 1.7 ± 0.07 |
| Prolificoccus (16)     | 25             | Prolificoccus Proteus | NATL2AT | 34–35 | 1.7 ± 0.1 |
| Thaumococcus (2)       | 14             | Thaumococcus swingsii | MIT9313T | 50–51 | 2.5 ± 0.06 |
| Riococcus gen. nov. (3)| 3              | Riococcus ceticus sp. nov. | MIT9211T | 37–38 | 1.6 ± 0.1 |

* New taxonomic classification proposed by Walter et al. [18]
* Average genome size was calculated using all genomes from each genus. The standard deviation was represented by ± symbol

Table 1 Proposal of a new PC taxonomic classification
similar between closely related species than between distantly related species. The analysis performed in this study establishes a new taxonomic framework for PC.

**Challenges in the Taxonomy of the Cyanobacteria Phylum and the PC**

Through evolution, *Cyanobacteria* became one of the most diverse and widely distributed prokaryote groups, occupying many niches within terrestrial, planktonic, and benthic habitats. Its long history has resulted in the evolution of broad heterogeneity, comprising unicellular and multicellular, photosynthetic and non-photosynthetic, free-living, symbiotic, toxic, and predatory organisms [72–75]. *Cyanobacteria* have been named according to the Botanical Code [99]. The inclusion of *Cyanobacteria* in taxonomic schemes of Bacteria was only proposed in 1978 by Stanier et al. [100]; over time, bacterial taxonomic names have come into conflict with the botanical nomenclature [101, 102]. More than two decades passed before the Note to General Consideration 5 (1999) was published indicating that *Cyanobacteria* should be included under the rules of the International Committee on Systematic Bacteriology (ICSB)/International Committee on Systematic of Prokaryotes (ICSP) [103–105]. Taxon nomenclature within this group has long been a topic of discussion, but there is currently no consensus [106–109]. As a result, more than 50 genera of *Cyanobacteria* have been described since 2000, and many of them remain unrecognized in the List of Prokaryotic Names with Standing in Nomenclature (LPSN, http://www.bacterio.net) [110] and in databases (e.g., NCBI) or have been the target of high criticism [111, 112].

Cyanobacterial taxonomy is based on morphologic traits and may not reflect the results of phylogenetic analyses [72, 113–118]. The analysis of 16S rRNA gene sequences is useful for charting and characterizing microbial communities [119] but lacks sensitivity to evolutionary changes that occur in association with ecological dynamics, in which microbial diversity is organized by physicochemical parameters [120, 121]. A recent study proposed that there should be 170 genera of *Cyanobacteria* based only on 16S rRNA sequences [119]. Farrant et al. delineated 121 ecologically significant taxonomic units (ESTUs) of *Prochlorococcus* and 15 ESTUs *Synechococcus* in the global ocean using single-copy petB sequences and environmental cues [122]. Although it was assumed that all the disclosed *Prochlorococcus* belonged to a single genus, it was not clear how many phylogenetic groups
Fig. 3 Phylogenetic trees using a 250 core protein sequences and b four housekeeping genes (gyrB, pyrH, recA, rpoB) based on maximum likelihood method. Bootstrap test was conducted using 1000 replicates (red numbers on the branches). The tree outgroup is *Synechococcus elongatus* PCC 6301.
MLSA [16]. Walter et al. [18] further split the same species form monophyletic groups on the basis of identity of GGD, and a maximum GC% range 2%. Strains of share > 95% DNA identity in MLSA, > 95% AAI, > 70% species based on species definition as a group of strains that among the named HL and LL ecotypes was noted by studies. On the other hand, the taxonomic heterogeneity least 12 stable ecotypes have been delineated in ecological species: the type species Prochlorococcus marinus and its type strain MIT 9301 T, and MIT9515 T; (2) the genus Riococcus

Unlocking the Genomic Taxonomy of the PC

Despite the astonishing advances in understanding the ecology of the PC, its taxonomic structure has remained puzzling until recently. The Prochlorococcus collective is thought to present a high degree of panmixis due to horizontal gene transfer and is composed of one genus and one species comprising two subspecies, P. marinus subsp. marinus and P. marinus subsp. pastoris [33, 123–126]. We remark that at least 12 stable ecotypes have been delineated in ecological studies. On the other hand, the taxonomic heterogeneity among the named HL and LL ecotypes was noted by Thompson et al., and the current ecotypes were split into 10 species based on species definition as a group of strains that share > 95% DNA identity in MLSA, > 95% AAI, > 70% identity of GGD, and a maximum GC% range 2%. Strains of the same species form monophyletic groups on the basis of MLSA [16]. Walter et al. [18] further split the Prochlorococcus collective into three new genera: (1) the genus Eurycolium (type species Eurycolium pastoris with the type strain MED4T (=CCMP 1986T), E. chisholmit with its type strain AS9601T, E. ponticus with its type strain MIT9301T, E. nereus with its type strain MIT9302T), E. neptunius with its type strain MIT9312T, E. ponticus with its type strain MIT 9301T, and E. tetissi with its type strain MIT9515T; (2) the genus Prolificoccus, with Prolificoccus proteus as the type species and NATL2AT as its type strain; and (3) the genus Thaumococcus, with Thaumococcus swingsii as the type species and the type strain MIT9313T. The current (type) genus Prochlorococcus contains two species: the type species Prochlorococcus marinus and its type strain SS120T (=CCMP 1375T) and P. ceticus with its type strain MIT9211T. The genus Eurycolium seems to form a large diverse eco-genomic group related to high temperature and oligotrophic environments, whereas the genera Prochlorococcus and Prolificoccus seem to be abundant at low temperature; Thaumococcus appears to thrive at low temperature and in copiotrophic environments [18].

On the basis of the eco-genomic and in silico phenotype evidence gathered amassed here, we propose five new genera within the PC. The taxonomic structure of the PC now comprises the following type species and type strains or type sequences: (1) the genus Prochlorococcus (Pro’chlo ro coe’cous. Gr. pref. pro, before (primitive); Gr. adj. chloros, green; M.L. masc. n. coccus, berry. Primitive green sphere) (the type genus). The type species Prochlorococcus marinus (ma’ri‘nus. marinus L. adj., marine) with its type strain CCMP1375T (=SS120T). The genome of this strain presents a GC content of 36.4%; (2) the genus Prolificoccus (Pro.li.fi.co.ccus. L. prolificus, productive, abundant, numerous; Prolificoccus, referring to an abundant coccus). The type species Prolificoccus proteus (L. n. protein, oprotos, “first,” is an early sea god, the old man of the sea) with its type strain NATL2AT. The genome of this strain presents a GC content of 35.1%; (3) the genus Eurycolium (Eur.y.co.lium. Gr. adj. eury, wide, broad; L. cole, inhabit; Gr. ium, quality or relationship, Eurycolium, referring to the spread inhabiting trait in marine habitats). The types species Eurycolium pastoris (pa.sto ‘ris. Latinized form of Pasteur) with its type strain CCMP1986T (=MED4T). The genome of this strain presents a GC content of 30.8%; (4) the genus Thaumococcus (Thau.mo.co.ccus. Referring to Thaumas, Greek god of the wonders of the sea; N.L. masc. n. coccus (from Gr. Masc. n. kokkos, grain, seed, kernel); N.L. masc. n. Thaumococcus, referring to a coccus living in the sea). The type species Thaumococcus swingsii (L. gen. masc. n. swingsii, of Swings, in honor of the Belgian microbiologist Jean Swings) with its type strain MIT 9313T. The genome of this strain presents a GC content of 50.7%; and (5) the genus Riococcus (Ri.o.co.‘cus. N.L. masc. n. Rio, of Rio de Janeiro, a city in Brazil; N.L. masc. n. kokkos, a grain; N.L. masc. n. Riococcus, referring the city of Rio de Janeiro). The type species Riococcus ceticus (L. n. ceticus, of Cetus, denotes a large fish) with its type strain MIT9211T. The genome of this

| Table 2 Predicted phenotype of new proposed genera of PC |
|---------------------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| **Genus**                        | **Ecotype**      | **narB**         | **nirA**         | **ntcA**         | **glhB**         | **amt**          |
| Eurycolium                      | HL/II/VI         | v                | v                | v                | v                | v                |
| Prochlorococcus                 | LII/II           | -                | v                | v                | v                | -                |
| Prolificoccus                   | LII              | v                | v                | v                | v                | v                |
| Thaumococcus                    | LLIV             | v                | v                | v                | v                | v                |
| Riococcus                       | LL/II/III        | -                | v                | v                | v                | v                |

+ presence; - absent; v, variable. Nitrogen-related genes: narB, nirA, ntcA, glhB, amt, cynS, ureA, glnB, glnF, gdhA. Phosphorous-related genes: phoB, ptrA, phoR. Arsenate-related gene: arsC. Iron-related genes: pucC. som

a Data extracted from Berube et al. [14, 71]
strain presents a GC content of 38%. We suggest that a PC species may be defined as a group of strains that share > 95% DNA identity in core protein sequences, and > 95% AAI. Strains of the same species will form monophyletic groups on the basis of core protein sequences.

Our study further refines previous taxonomic proposals for the Prochlorococcus collective [18]. A recent novel bacterial taxonomy has also been proposed on the basis of phylogenomics [54, 127]. This study proposes the split of Prochlorococcus into 5 distinct genera. GTDB-tk [128] is a quick and easy way to assess how our proposed taxonomy compares with what has been previously published. When we compare our proposal with the GTDB proposal we have the following correspondence: Prochlorococcus = Prochlorococcus, Eurycolium = Prochlorococcus_A, Prolificoccus = Prochlorococcus_B, Thaumococcus = Prochlorococcus_C, and Riococcus = Prochlorococcus_D, suggesting an excellent agreement between these different studies. These studies provide insights into the evolutionary relationship among the different genera and species, as well in silico phenotypes of PC.

Conclusions

The PC represents an excellent model for further theory development in ecology, evolution, phylogeny, and taxonomy. The PC refers to a group of picocyanobacteria that radiated from Synechococcus during millions of years of evolution in the oceans. The taxonomic analysis of 232 genomes revealed at least 5 new genera in addition to the original genus, Prochlorococcus [129]. As we delineated genera only for those groups with at least three available genome sequences, we anticipate that new genera will be included in our new proposal when more representative genomes are added to groups with only 1 or 2 genomes. On average, 20% of the genes of PC genomes have unknown functions and await further characterization. The pan-genome of the PC is estimated to include over 80,000 genes [14, 32].

The PC has evolved in multiple ecologic niches in the global ocean and participates in relevant biogeochemical cycles, such as the recently postulated parasitic arsenic cycle among the different genera and species, as well in silico phenotypes of PC.

Acknowledgments The authors thank CNPq, CAPES and FAPERJ.

References

1. Chisholm SW, Frankel SL, Goericke R et al (1992) Prochlorococcus marinus nov. gen. nov. sp.: an oxyphototrophic marine prokaryote containing divinyl chlorophyll a and b. Arch Microbiol 297–300
2. Chisholm SW (2017) Prochlorococcus. Curr Biol 27:R447–R448. https://doi.org/10.1016/j.cub.2017.02.043
3. Braakman R, Follows MJ, Chisholm SW (2017) Metabolic evolution and the self-organization of ecosystems. Proc Natl Acad Sci 114:E3091–E3100. https://doi.org/10.1073/pnas.1619573114
4. Ting CS, Ramsey ME, Wang YL, Jon E, Durham T (2009) Minimal genomes, maximal productivity: comparative genomics of the photosystem and light-harvesting complexes in the marine cyanobacterium, Prochlorococcus. Photosynth Res 101:1–19
5. Scanlan DJ, Ostrowski M, Mazar D, Dufresne A, Garzarek L, Hess WR, Post AF, Hagemann M, Paulsen I, Partensky F (2009) Ecological genomics of marine picocyanobacteria. Microbiol Mol Biol Rev 73:249–299. https://doi.org/10.1128/MMBR.00035-08
6. Ralf G, Repeta DJ (1992) The pigments of Prochlorococcus marinus: the presence of divinylchlorophyll a and b in a marine prokaryote. Limnol Oceanogr 37:425–433
7. Partensky F, Garzarek L (2010) Prochlorococcus: advantages and limits of minimalism. Ann Rev Mar Sci 2:305–331
8. Mühling M (2012) On the culture-independent assessment of the diversity and distribution of Prochlorococcus. Environ Microbiol 14:567–579
9. Coleman ML, Sullivan MB, Martiny AC et al (2006) Genomic islands and the ecology and evolution of Prochlorococcus. Science (80–) 311:1768–1770. https://doi.org/10.1126/science.1122050
10. Coleman ML, Chisholm SW (2007). Code and context: Prochlorococcus as a model for cross-scale biology. 15:15–407. https://doi.org/10.1016/j.tim.2007.07.001
11. Kashtan N, Roggensack SE, Rodrigue S et al (2014) Single-cell genomics reveals hundreds of coexisting subpopulations in wild Prochlorococcus. Science (80–) 344:416–420. https://doi.org/10.1126/science.1248575
12. Biller SJ, Berube PM, Berta-Thompson JW, Kelly L, Roggensack SE, Awad L, Roache-Johnson KH, Ding H, Giovannoni SJ, Rocap G, Moore LR, Chisholm SW (2014) Genomes of diverse isolates of the marine cyanobacterium Prochlorococcus. Sci Data 1:140034
13. Kashtan N, Roggensack SE, Berta-Thompson JW, Grinberg M, Stepanauskas R, Chisholm SW (2017) Fundamental differences in diversity and genomic population structure between Atlantic and Pacific Prochlorococcus. ISME J 11:1997–2011. https://doi.org/10.1038/ismej.2017.64
14. Berube PM, Biller SJ, Hackl T, Hogle SL, Satinsky BM, Becker JW, Braakman R, Collins SB, Kelly L, Berta-Thompson J, Coe A, Bergauer K, Bouman HA, Brownie TJ, de Corte D, Hassler C, Hulata Y, Jacquot JE, Maas EW, Reinthaler T, Stotes E, Yokokawa T, Lindell D, Stepanauskas R, Chisholm SW (2018) Data descriptor: single cell genomes of Prochlorococcus, Synechococcus, and sympatric microbes from diverse marine environments. Sci Data 5:1–11. https://doi.org/10.1038/sdata.2018.154
15. Chisholm SW (2012) The life and times of the ocean’s smallest photosynthetic cell. In: Kolter R, Maloy S (eds) Microbes and environments. Sci Data 5:1–11. https://doi.org/10.1038/sdata.2018.154
16. Thompson CC, Silva GGZ, Vieira NM, Edwards R, Vicente ACP, Thompson FL (2013) Genomic taxonomy of the genus
Prochlorococcus. Microb Ecol 66:752–762. https://doi.org/10.1007/s00248-013-0270-8

17. Coutinho F, Tschoeke DA, Thompson F, Thompson C (2016) Comparative genomics of Synechococcus and proposal of the new genus Parasynechococcus. PeerJ 4:e1522. https://doi.org/10.7717/peerj.1522

18. Walter JM, Coutinho FH, Dutilh BE, Swings J, Thompson FL, Thompson CC (2017) Ecogenomics and taxonomy of Cyanobacteria phylum. Front Microbiol 8. https://doi.org/10.3389/fmicb.2017.02132

19. Tettelin H, Riley D, Cattuto C, Medini D (2008) Comparative genomics: the bacterial pan-genome. Curr Opin Microbiol 11:472–477

20. Lapierre P, Gogarten JP (2009) Estimating the size of the bacterial pan-genome. Trends Genet 25:107–110

21. Rosselló-Móra R, Whitman WB (2019) Dialogue on the nomenclature and classification of prokaryotes. Syst Appl Microbiol 42:5–14. https://doi.org/10.1016/j.syapm.2018.07.002

22. De Vos P, Thompson F, Thompson C, Swings J (2017) A flavor of prokaryotic taxonomy: systematics revised. Microbial resources: from functional existence in nature to applications. Elsevier Inc., pp 29–44

23. Sun Z, Blanchard JL (2014) Strong genome-wide selection early in the evolution of Prochlorococcus resulted in a reduced genome through the loss of a large number of small effect genes. PLoS One 9:e88837

24. Johnson ZI, Zinser ER, Coe A et al (2006) Niche partitioning among Prochlorococcus ecotypes along ocean-scale environmental gradients. Science (80-) 311:1737–1740

25. Garczarek L, Dufresne A, Rousvoal S, West NJ, Mazard S, Marie D, Claustre HÅC, Raimbault P, Post AF, Scanlan DJ, Partensky FÅO© (2007) High vertical and low horizontal diversity of Prochlorococcus ecotypes in the Mediterranean Sea in summer. FEMS Microbiol Ecol 60:189–206. https://doi.org/10.1111/j.1574-6941.2007.00297.x

26. Rusch DB, Martiny AC, Dupont CL, Halpern AL, Venter JC (2010) Characterization of Prochlorococcus clades from iron-depleted oceanic regions. Proc Natl Acad Sci 107:16184–16189

27. Larkin AA, Blinébry SK, Howes C, Lin Y, Loftus SE, Schmaus CA, Zinser ER, Johnson ZI (2016) Niche partitioning and biogeography of high light adapted Prochlorococcus across taxonomic ranks in the North Pacific. ISME J 10:1555–1567

28. West NJ, Lebaron P, Strutton PG, Suzuki MT (2011) A novel clade of Prochlorococcus found in high nutrient low chlorophyll waters in the South and Equatorial Pacific Ocean. ISME J 5:933–944

29. Huang S, Wilhelms SR, Harvey HR, Taylor K, Jiao N, Chen F (2012) Novel lineages of prochlorococcus and synechococcus in the global oceans. ISME J 6:285–297. https://doi.org/10.1038/ismej.2011.106

30. Kent AG, Dupont CL, Yoosop S, Martiny AC (2016) Global biogeography of Prochlorococcus genome diversity in the surface ocean. ISME J 10:1856–1865. https://doi.org/10.1038/ismej.2015.265

31. Moore LR, Chisholm SW (1999) Photophysiology of the marine cyanobacterium Prochlorococcus: ecotypic differences among cultured isolates. Limnol Oceanogr 44:628–638

32. Biller SJ, Berube PM, Lindell D, Chisholm SW (2015) Prochlorococcus: the structure and function of collective diversity. Nat Rev Microbiol 13:13–27. https://doi.org/10.1038/nrmicro3378

33. Rocap G, Lamerin FW, Lamerin J, Malfatti S, Chain P, Ahlgren NA, Arellano A, Coleman M, Hauser L, Hess WR, Johnson ZI, Land M, Lindell D, Post AF, Regala W, Shah M, Shaw SL, Steglich C, Sullivan MB, Ting CS, Tolonen A, Webb EA, Zinser ER, Chisholm SW (2003) Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation. Nature 424:1042–1047

34. Chandler JW, Lin Y, Gainer PJ, Post AF, Johnson ZI, Zinser ER (2016) Variable but persistent coexistence of Prochlorococcus ecotypes along temperature gradients in the ocean’s surface mixed layer. Environ Microbiol Rep 8:272–284

35. Binder BJ, Chisholm SW, Olson RJ, Frankel SL, Worden AZ (1996) Dynamics of picophytoplankton, ultraphytoplankton and bacteria in the central equatorial Pacific. Deep Sea Res Part II Top Stud Oceanogr 43:907–931

36. Ferris MJ, Palenik B (1998) Niche adaptation in ocean cyanobacteria. Nature 396:226–228

37. García-Fernández JM, Hess WR, Hournard J, Partensky F (1999) Expression of the psbA gene in the marine oxyphotobacteria Prochlorococcus spp. Arch Biochem Biophys 359:17–23

38. Fuhrman JA, McCallum K, Davis AA (1993) Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans. Appl Environ Microbiol 59:1294–1302

39. Martiny AC, Coleman ML, Chisholm SW (2006) Phosphate acquisition genes in Prochlorococcus ecotypes: evidence for genome-wide adaptation. Proc Natl Acad Sci 103:12552–12557

40. Martiny AC, Huang Y, Li W (2009) Occurrence of phosphate acquisition genes in Prochlorococcus cells from different ocean regions. Environ Microbiol 11:1340–1347

41. Moore LR, Post AF, Rocap G, Chisholm SW (2002) Utilization of different nitrogen sources by the marine cyanobacteria Prochlorococcus and Synechococcus. Limnol Oceanogr 47:989–996

42. Sullivan MB, Waterbury JB, Chisholm SW (2003) Cyanophages infecting the oceanic cyanobacterium Prochlorococcus. Nature 424:1047–1051

43. Sher D, Thompson JW, Kashtan N, Croul L, Chisholm SW (2011) Response of Prochlorococcus ecotypes to co-culture with diverse marine bacteria. ISME J 5:1125–1132

44. Biller SJ, Coe A, Chisholm SW (2016) Torn apart and reunited: impact of a heterotroph on the transcriptome of Prochlorococcus. ISME J 10:2831–2843

45. Biller SJ, Berube PM, Dooley K, Williams M, Satinsky BM, Hackl T, Hogle SL, Coe A, Bergauer K, Bouman HA, Browning TJ, de Corte D, Hassler C, Hulston D, Jacquot JE, Maas EW, Reinthaler T, Sintes E, Yokokawa T, Chisholm SW (2018) Data descriptor: marine microbial metagenomes sampled across space and time. Sci Data 5:1–7. https://doi.org/10.1038/sdata.2018.176

46. Muñoz-marín MC, Gómez-baena G, López-lozano A et al (2020) Mixotrophy in marine picocyanobacteria: use of organic compounds by Prochlorococcus and Synechococcus. ISME J. 14:1065–1073. https://doi.org/10.1038/s41396-020-0603-9

47. Vandinamme P, Pot B, Gilliss M et al (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Mol Biol Rev 60:407–438

48. Wayne LG, Brenner DJ, Colwell RR et al (1987) Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. Int J Syst Bacterial 47:463–464. https://doi.org/10.1099/00207713-37-4-463

49. Stackebrandt E, Frederiksen W, Garrity GM, Grimont PAD, Kämpfer P, Maiden MCJ, Nesme X, Rosselló-Mora R, Swings J, Trüper HG, Vauterin L, Ward AC, Whitman WB (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. Int J Syst Evol Microbiol 52:1043–1047

50. Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, Stackebrandt E, de Peer YV, Vandamme P, Thompson FL, Swings J (2005) Defining prokaryotic species: Reevaluating prokaryotic species. Nat Rev Microbiol 3:733–739
51. Thompson CC, Chimetto L, Edwards RA, Swings J, Stackebrandt E, Thompson FL (2013) Microbial genomic taxonomy. BMC Genomics 14:913. https://doi.org/10.1186/1471-2164-14-913

52. Thompson CC, Amoral GR, Campeão M, Edwards RA, Polz MF, Dutilh BE, Ussery DW, Sawabe T, Swings J, Thompson FL (2015) Microbial taxonomy in the post-genomic era: rebuilding from scratch? Arch Microbiol 197:359–370

53. Castelle CJ, Banfield JF (2018) Major new microbial groups expand diversity and alter our understanding of the tree of life. Cell 172:1181–1197

54. Parks DH, Chuvchenko M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996–1004. https://doi.org/10.1038/nbt.4229

55. Hyatt D, Chen G-L, LoCasto PF et al (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf 11:119

56. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055

57. Konstantinidis KT, Tiedje JM (2005) Towards a genome-based taxonomy for prokaryotes. J Bacteriol 187:6258–6264

58. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421

59. Kato K, Asiminos G, Toh H (2009) Multiple alignment of DNA sequences with MAFFT. Methods Mol Biol. https://doi.org/10.1007/978-1-59745-251-9_3

60. Price MN, Dehal PS, Arkin AP (2010) FastTree 2 - approximately maximum-likelihood trees for large alignments. PLoS One 5: e9490. https://doi.org/10.1371/journal.pone.0009490

61. Huerta-Cepas J, Serra F, Bork P (2016) ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. Mol Biol Evol. 33:1635–1638. https://doi.org/10.1093/molbev/msw046

62. Lee MD, Ponty Y (2019) GToTree: A user-friendly workflow for phylogenome analyses. Bioinformatics. 35:4162–4164. https://doi.org/10.1093/bioinformatics/btz888

63. Eddy SR (2011) Accelerated profile HMM searches. PLoS Comput Biol. 7:e1002195. https://doi.org/10.1371/journal.pcbi.1002195

64. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1797–1800. https://doi.org/10.1093/nar/gkh340

65. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25:1972–1973. https://doi.org/10.1093/bioinformatics/btp348

66. Tamura K, Stecher G, Peterson D et al (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. https://doi.org/10.1093/molbev/msr067

67. Tange O (2011) GNU parallel: the command-line power tool. https://doi.org/10.5281/zenodo.16303

68. Altschul SF, Madden TL, Schäffer AA et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

69. Hildebrand F, Meyer A, Eyre-Walker A (2010) Evidence of selection upon genomic GC-content in bacteria. PLoS Genet 6: e1001107

70. Hershberg R, Petrov DA (2010) Evidence that mutation is universally biased towards AT in bacteria. PLoS Genet 6:e1001115

71. Berube PM, Rasmussen A, Braakman R, Stepanauskas R, Chisholm SW (2019) Emergence of trait variability through the lens of nitrogen assimilation in Prochlorococcus. Elife 8:1–28. https://doi.org/10.7554/elife.41043

72. Schirmeister BE, Antonelli A, Bagheri HC (2011) The origin of multicellularity in cyanobacteria. BMC Evol Biol 11:45

73. Di Rienzi SC, Sharon I, Wrighton KC et al (2013) The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. Elife 2:e01022

74. Soo RM, Skennerton CT, Sekiguchi Y, Imelfort M, Paech SJ, Dennis PG, Steen JA, Parks DH, Tyson GW, Hugenholtz P (2014) An expanded genomic representation of the phylum Cyanobacteria. Genome Biol Evol 6:1031–1045

75. Soo RM (2015) In search of non-photosynthetic Cyanobacteria. Univ Queensl

76. Cohan FM (2002) What are bacterial species? Annu Rev Microbiol 56:457–487

77. Cohan FM (2016) Bacterial speciation: genetic sweeps in bacterial species. Curr Biol 26:R112–R115. https://doi.org/10.1016/j.cub.2015.10.022

78. Cohan FM, Kopac SM (2017) A theory-based pragmatism for discovering and classifying newly divergent species of bacterial pathogens. Genet Epidemiol 41:175–189. https://doi.org/10.1002/j.1097-4636.2016.01180.x

79. Baumann L, Baumann P, Mandel M et al (1972) Taxonomy of aerobic marine. Microbiology 110:402–429

80. Deming JW, Somers LK, Straube WL, Swartz DG, Macdonell MT (1988) Isolation of an obligately barophilic bacterium and description of a new genus, Colwellia gen. nov. Syst Appl Microbiol 10:152–160. https://doi.org/10.1016/S0722-2020(88)80030-4

81. Kurahashi M, Yokota A (2007) Endozoicoomarinas elysiosa gen. nov., sp. nov., a γ-proteobacterium isolated from the sea slug Elysia omata. Syst Appl Microbiol 30:202–206. https://doi.org/10.1016/j.syapm.2006.07.003

82. GAUTHIER MJ, LAFAY B, CHRISTEN BR et al (1992) Marinobacter hydrocarbonoclasticus gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. Int J Syst Bacteriol 42:568–576. https://doi.org/10.1099/ijss.0.349.644

83. GONZALEZ JM, MAYER F, MORAN MA et al (2009) Microbulbifer hydrolyticus gen. nov., sp. nov., and Marinobacterium georgiense gen. nov., sp. nov., two marine bacteria from a lignin-rich pulp mill waste enrichment community. Int J Syst Bacteriol 49:369–376. https://doi.org/10.1099/ijs.0.002771-74.2-369

84. LANDSCHOOT A VAN, LEY J DE (1983) Intra- and intergeneric similarities of the rRNA cistrons of Alteromonas, Murimonas (gen. nov.) and some other Gram-negative bacteria. J Gen Microbiol 129:3057–3074

85. Gauthier G, Gauthier M, R C (1995) Phylogenetic analysis of the genera Alnteromonas, Shewanella, and Moritella using genes coding for small-subunit RNA sequences and division of the genus Alnteromonas into two genera, Alnteromonas (Emended) and Pseudalteromonas gen. nov., and proposal of Tw. Int J Syst Bacteriol 45:755–761

86. Liew WY, Te Lin Y, Jean WD et al (2004) Pseudovibrio denitrificans gen. nov., sp. nov., a marine, facultatively anaerobic, fermentative bacterium capable of denitrification. Int J Syst Evol Microbiol 54: 2307–2312. https://doi.org/10.1099/ijsem.0.63107-0

87. Shibata T (1991) Roseobacter litoralis gen. nov., sp. nov., and Roseobacter denitrificans sp. nov., aerobic pink-pigmented bacterium which contain bacteriochlorophyll a. Syst Appl Microbiol 14: 2307–2312. https://doi.org/10.1099/ijs.0.63107-0

88. Uchino Y, Hirata A, Yokota A, Sugiyama J (2005) Micrococcus albus gen. nov., sp. nov., a Nobutobacter isolated from the sea slug Elysia omata. Syst Appl Microbiol 28:164–170. https://doi.org/10.1086/429281

89. Ruegeria gelatinovora comb. nov., Ruegeria algicola comb. nov., Ruegeria atlantica gen. nov., comb. nov., Ruegeria atlantica gen. nov., nom. rev., Ruegeria phyllosphaerae gen. nov., sp. nov., Ruegeria gelatinovora comb. nov., Ruegeria algicola comb. nov., and A. J Gen Appl Microbiol 44:201–210. https://doi.org/10.2323/jgam.44.201
Makarova KS, Ostrowski M, Oztas S, Robert C, Rogozin IB, Scanlan DJ, de Marsac NT, Weissenbach J, Wincker P, Wolf YI, Hess WR (2003) Genome sequence of the cyanobacterium Prochlorococcus marinus SS120, a nearly minimal oxyphototrophic genome. Proc Natl Acad Sci 100:10020–10025

124. Kettler GC, Martiny AC, Huang K, Zucker J, Coleman ML, Rodrigue S, Chen F, Lapidus A, Ferriera S, Johnson J, Steglich C, Church GM, Richardson P, Chisholm SW (2007) Patterns and implications of gene gain and loss in the evolution of Prochlorococcus. PLoS Genet 3:2515–2528. https://doi.org/10.1371/journal.pgen.0030231

125. Zinser ER, Lindell D, Johnson ZI, Futschik ME, Steglich C, Coleman ML, Wright MA, Rector T, Steen R, McNulty N, Thompson LR, Chisholm SW (2009) Choreography of the transcriptome, photophysiology, and cell cycle of a minimal photoautotroph, Prochlorococcus. PLoS One 4:e5135

126. García-Fernandez JM, Marsac NT, Diez J (2004) Streamlined regulation and gene loss as adaptive mechanisms in prochlorococcus for optimized nitrogen utilization in oligotrophic environments. Microbiol Mol Biol Rev 68:630–638. https://doi.org/10.1128/MMBR.68.4.630

127. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P (2020) A complete domain-to-species taxonomy for Bacteria and Archaea. Nat Biotechnol. https://doi.org/10.1038/s41587-020-0501-8

128. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH (2019) GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. https://doi.org/10.1093/bioinformatics/btz848

129. Chisholm SW, Frankel SL, Goericke R, Olson RJ, Palenik B, Waterbury JB, West-Johnsrud L, Zettler ER (1992) Prochlorococcus marinus nov. gen. nov. sp.: an oxyphototrophic marine prokaryote containing divinyl chlorophyll a and b. Arch Microbiol 157:297–300

130. Giovannoni SJ, Halsey KH, Saw J, Muslin O, Suffridge CP, Sun J, Lee CP, Moore ER, Temperton B, Noell SE (2019) A parasitic arsenic cycle that shuttles energy from phytoplankton to heterotrophic bacterioplankton. MBio 10

131. Saunders JK, Rocap G (2016) Genomic potential for arsenic efflux and methylation varies among global Prochlorococcus populations. ISME J 10:197–209. https://doi.org/10.1038/ismej.2015.85