Research Article

Prochlorococcus and *Synechococcus* marine cyanobacteria:
a scientometrics review

Caio Cesar-Ribeiro[1,2,3] , **Cristiane Santos Barbosa[1]**
Vanessa Terra[1,4] **& Nédia de Castilhos Ghisi[1]**
[1]Pós-graduação em Biologia Molecular, Universidade Tecnológica Federal do Paraná
Curitiba, Paraná, Brasil
[2]Pós-graduação em Evolução e Diversidade, Universidade Federal do ABC, Santo André, Brasil
[3]NEPEA, Universidade Estadual Paulista, São Vicente, São Paulo, Brasil
[4]Universidade Federal de Uberlândia, Monte Carmelo, Minas Gerais, Brasil
Corresponding author: Caio Cesar-Ribeiro (caiocribeiro@hotmail.com)

ABSTRACT. *Prochlorococcus* and *Synechococcus* are the dominant primary producers in the ocean and account for a significant percentage of oceanic carbon fixation. This review summarizes the available data in the literature (1992-2021; n = 480) using CiteSpace software to provide insights into the specific characteristics of *Prochlorococcus* and *Synechococcus* genes that scientists can use for establishing future research directions. From 1992 to 2021, the most frequent and with higher centrality keyword was "multiple evolutionary origin" (0.43), followed by flow cytometry (0.40), because this period was the main method used to identify cyanobacteria. In 2001-2010, were "sequence" and "gene", both with 0.13 of centrality; this can be explained by the period when gene sequencing was developed. In the last 10 years period 2011-2021, were diversity (0.17); marine *Synechococcus* (0.10); *Prochlorococcus* ecotype (0.08); gene expression (0.14). Additionally, the migration characteristics of flow cytometry, genes identification methods, genomes, and ecotypes of cyanobacteria, indicate how future studies should focus in 1) molecular ecology: how genetic variability among different populations affects their adaptation to different ocean environments, 2) physiology: how this phytoplankton respond to changes in nutrient availability and solar radiation, 3) interactions with other living beings: how they interact with other organisms, such as bacteria, zooplankton, and fish, and how these interactions affect the dynamics of the ocean ecosystem, and 4) modeling: computational models are being developed to predict how the distribution will change in the future as a result of climate change and human use of the oceans.

Keywords: *Prochlorococcus*; *Synechococcus*; CiteSpace; genome; cyanobacteria; scientometrics

INTRODUCTION

Prochlorococcus contains some of the smallest photosynthetic organisms known to science (Farrant et al. 2016). *Prochlorococcus* taxa are found throughout the euphotic zone and are some of the ocean's most abundant life forms. They are adapted to oligotrophic conditions and inhabit areas with reduced light intensities by increasing their pigment concentration. Studies have shown that *Prochlorococcus* is an abundant and important primary producer in the open ocean, accounting for over 82% of the North Pacific Subtropical Ocean (Liu et al. 1997). Meanwhile, *Synechococcus* ecotypes differ from typical prokaryotes due to their more complex cell cycle, which results in higher growth rates (Lalli & Parsons 1997). *Synechococcus* taxa are abundant in temperate and tropical euphotic zones and play a crucial role in ocean primary productivity. Research has revealed that *Synechococcus* is the most abundant genus of bacteria and dominant in terms of biomass and productivity in the northern Atlantic, tropical, and subtropical regions (Olson et al. 1990, Veldhuis & Kraay 2000).

_____

Corresponding editor: Sonia Soto-Rodríguez

Prochlorococcus and *Synechococcus* are among the most abundant taxa in the collective photosynthetic activity of cyanobacteria and occupy deeper areas of the water column (Veldhuis & Kraay 2000). These two genera are generally found together, with their depth distributions differing in absolute concentrations and vertical extension, except in homogeneous water columns (Li et al. 1992, Campbell & Vaulot 1993, Veldhuis & Kraay 1993, 2000, Cesar-Ribeiro et al. 2020).

Prochlorococcus and *Synechococcus* are usually calculated separately for estimating phytoplankton biomass because they can be easily distinguished from pico-eukaryotes (Li et al. 1993, Campbell et al. 1994). The most abundant prokaryotic group in marine environments is heterotrophic bacteria, which generally dominates the size range of 0.2 to 2 µm in cell number and biomass (Li et al. 1993). However, the taxonomic differentiation of this group is limited due to the absence of pigments (Giovannoni et al. 1990). Still, recent advances in metagenomics have made it possible to identify different taxa and even approximate their abundance.

The genera *Prochlorococcus* and *Synechococcus* are estimated to be responsible for approximately 25% of the ocean's net primary productivity. Their abundance and global distribution make *Prochlorococcus* and *Synechococcus* essential components of marine food webs, primarily supplying fixed carbon to higher trophic levels (Flombaum et al. 2013).

Prochlorococcus is the dominant phototroph in oligotrophic subtropical gyres, some of the largest biomes on Earth (Berube et al. 2018). Despite having the smallest genomes of known oxygenic phototrophs, the global diversity of *Prochlorococcus* harbors an immense number of protein-encoding genes. Estimates based on 41 cultivated isolate genomes suggest that the *Prochlorococcus* pan-genome, the complete set of genes found in all *Prochlorococcus*, contains over 80,000 distinct genes, many of which likely play a role in local environmental adaptation (Berube et al. 2018).

Prochlorococcus contains the smallest photosynthetic organisms ever identified, adapted to the oligotrophic ocean, and inhabit areas with reduced light intensities by increasing their pigment concentration (Farrant et al. 2016). The species in this genus have efficient carbon concentration mechanisms and high CO_2 fixation rates by photosynthetic pigments (Farrant et al. 2016). *Prochlorococcus* marinus is an abundant primary producer in the open ocean, reaching over 82% of the North Pacific Subtropical Ocean (Farrant et al. 2016). On the other hand, species of the genus

Synechococcus differ from typical prokaryotes due to their more elaborate cell cycle and higher growth rates (Grébert et al. 2018). These species are abundant in temperate and tropical euphotic zones and crucial to ocean primary productivity (Doré et al. 2020).

Studies have shown that *Prochlorococcus* and *Synechococcus* generally co-occur in the ocean, with their depth distributions differing in absolute concentrations and vertical extension (Farrant et al. 2016). These two genera are among the most abundant taxa in the collective photosynthetic activity of cyanobacteria and occupy deeper areas of the water column (Farrant et al. 2016). Marine cyanobacteria within the genera *Prochlorococcus* and *Synechococcus* are estimated to be responsible for roughly 25% of ocean net primary productivity, performing vital functions at the base of marine food webs (Farrant et al. 2016).

Recent advances in the genomics of single cells have uncovered previously unknown marine microbial phyla and functions and have revealed the existence of new clades with distinct ecological and physiological adaptations and high degrees of genomic and functional diversity among *Prochlorococcus* cells (Doré et al. 2020). The study of *Prochlorococcus* and *Synechococcus* research using CiteSpace software aims to provide a comprehensive and systematic scientometrics review by summarizing general characteristics of publication output, evaluating the performance of countries, institutions, and journals, and identifying emerging hotspots and the intellectual structure of *Prochlorococcus* and *Synechococcus* genes (Farrant et al. 2016).

Based on previous studies, we adopted CiteSpace software to perform a visual bibliometric analysis of *Prochlorococcus* and *Synechococcus* research in 1992-2021. This study aims to comprehensively and systematically provide a scientometrics review. More specifically, our analysis aims to 1) summarize general characteristics of publication output and main subject categories; 2) evaluate the performance of countries, institutions, and journals; and 3) identify the emerging hotspots and the evolution of *Prochlorococcus* and *Synechococcus* genes. Previous studies have yet to analyze its research corpus to such depth to include aspects such as co-citation clusters, keywords, or research clusters. To bridge these gaps in extant literature by undertaking an in-depth scientometrics review of global sustainability, we used CiteSpace software to conduct a visual scientometrics analysis. The most apparent advantage of CiteSpace is that it allows scholars to investigate specific research areas by analyzing the citations, co-citations, and geographical

distribution, thus drawing a beneficial conclusion. Studies of *Prochlorococcus* and *Synechococcus* are important in understanding the significance of this phytoplankton for the planet's functioning and taking measures to protect the oceans and preserve the health of our planet. These two types of phytoplankton can be found in all the world's oceans and play a crucial role in regulating the climate and the global carbon cycle, as they are the main producers of oxygen on Earth.

MATERIALS AND METHODS

Data were obtained from the Web of Science (WoS) database of Clarivate Analytics (Chen 2014). The searched terms included topic search (TS): *Prochlorococcus* or *Synechococcus*. This database is regarded as the most important and frequently used scientific database in most fields. The records, including titles, abstracts, and cited references, were then exported to CiteSpace for subsequent analysis.

CiteSpace provides various functions for facilitating the understanding and interpretation of network patterns, including identifying the major topic finding hotspots and areas and labeling clusters with terms from selected literature. Scienciometry can help identify trends and patterns in research on these cyanobacteria, which can be useful in guiding future research and promoting the development of discoveries. Analysis of collaboration networks among researchers and institutions can help identify new collaboration opportunities, while analysis of research topics can help identify knowledge gaps and areas that deserve more attention. In addition, analysis of bibliographic references helps identify the main contributions to advancing knowledge in these areas, which can guide future research and promote the development of new approaches and methodologies. With this information, researchers can identify areas that deserve more attention and focus, as well as approaches and techniques that have been most effective in generating new knowledge about these important marine cyanobacteria. Scientometrics analysis methods CiteSpace is a Java-based scientific visualization software package used for analyzing and visualizing co-citation networks developed by Chen (2006). In CiteSpace, countries, contributing authors, and journals were mapped; this can help visualize and identify contributing factors for the evolution of the hospitality knowledge map. Hot research topics and frontline research in hospitality were identified based on the frequency of famous keywords used in journal articles. A node represents one item in viewing the maps, such as a journal, keyword, and reference.

One of the key functions of CiteSpace is to identify major topics and hotspots in a given field of research, done by analyzing the frequency of keywords and co-occurrence of terms within a selected literature set. CiteSpace can then visualize the network of relationships between the different terms, with the size of each node representing its frequency of occurrence and the thickness of the edges representing the strength of the co-occurrence relationship. Once the network has been generated, CiteSpace allows users to label clusters with terms from the selected literature, which involves selecting a term that represents the central theme of a cluster and then automatically labeling all other nodes in the cluster with related terms from the literature set, making it easier for users to understand the major themes and topics that are emerging from the analysis. Another important feature of CiteSpace is the ability to identify areas of the network that are most influential or central to the field of research by analyzing the centrality of nodes and edges within the network, using measures such as betweenness centrality and degree centrality. By identifying the most central nodes and edges, researchers can gain insight into the key players and relationships within the field and the most important research topics and themes (Chen 2006).

The search generated 1201 references between 1992 and 2021. Later, refining manually analyzed each publication by reading titles and abstracts to select only studies involving "genes." A total of 721 publications were excluded due to duplicated citations or because they were unrelated to genes, genomic, or even genetic; they were related to these cyanobacteria ecology, distribution, and identification (480 selected). Document type, publication year, source and impact factor, category, country, and author were identified. The impact factor was obtained from Journal Citation Reports (JCR) (clarivate.com/products/web-of-science). The results were analyzed using Microsoft Office Excel, Mapchart, and CiteSpace software (Chen 2014).

We use the centrality of CiteSpace because of the comparison between data, which means each document's importance in a co-citing network can be partially evaluated by the indicator betweenness centrality (Liu et al. 2017). Centrality is defined in the following equation:

$$\text{Centrality}\left(\text{node i}\right) = \sum_{i \neq j \neq k} \frac{\rho_{jk}\,(i)}{\rho_{jk}} \qquad (1)$$

where ρ_{jk} represents the number of shortest paths between node j and node k, and ρ_{jk} (i) is the number of those paths that pass-through node i.

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart demonstrates the steps of identification, selection, eligibility, and inclusion of articles in this study; a total of 721 citations were excluded due to duplicated citations and publications that did not were related to genetic (Fig. 1).

RESULTS

An increase was observed in yearly publications, from one in 1992 to 33 in 2009 and 2015. These publications were classified into nine document groups: article, procedure paper, abstract, review, letter, note, editorial material, book review, and others. Articles were predominant, comprising 93.33% of the total. The only language was English (100%).

The articles in our refined dataset were cited 29,572 times, averaging 61.60 citations per item. The H-index is the maximum value of H for which an author or area has published H papers cited at least H times. Our dataset had 16 papers with at least 61 citations each (Hirsch 2005). The dispersion graph represents the mean of publications and citations per year (Fig. 2). It shows that since 2002 there has been a considerable and growing increase in publications and citations regarding *Prochlorococcus* and *Synechococcus* genetics. In the period analyzed, no correlation was observed between citations and publications (Pearson's correlation: $R^2 = 0.30$; $P > 0.05$). There may be a correlation between posts and citations in academic publications, as the visibility and dissemination of research findings can influence their impact and the extent to which other researchers cite them. Furthermore, any correlation between posts and citations may be time-lagged, as research can take some time to gain visibility and recognition within the academic community. It is also possible that the relationship between posts and citations is not linear and that there may be diminishing returns to increased visibility and dissemination of research findings.

Information analysis based on the origin of the authors of the publications can be observed (Fig. 3a). The publication leader was the USA, followed by Germany, France, and China. In Figure 3b, information on the number of citations and cooperation networks between countries is summarized, showing the USA as the country that stands out in marine cyanobacteria studies.

The letter font size in Figure 3b is proportional to the publication volume, while the thickness of the connecting lines between countries demonstrates the intensity of cooperation. The thicker the connection line, the higher the frequency of cooperation between two or more countries. The formation of clusters of countries separated by different colors can be seen in Figure 3b. The numbers represent the citation counts of each country. The top-ranked item by citation counts is in order: USA, France, Germany, and China. Figure 3b also presents the centrality of some countries (purple halo). Centrality is an important measure that varies from 0 to 1 (Chen 2014), representing the influence of a country. In this study, the USA has the most expressive centrality (0.85), followed by France (0.2), Germany and Italy (0.15), and England and Spain (0.14).

Moreover, some countries presented citation bursts. A citation burst indicates a most active region (in this case), which can last for multiple years and a single year. A citation burst proves that a publication/country is associated with a surge of citations. In other words, the item has attracted extraordinary attention from the scientific community (Chen 2014). The top-ranked item in bursts is France (1995), with bursts of 7.13. The 2nd was Canada (2005) (5.15), the 3rd was Germany (1998) (4.19), and the 4th was China (2018) (3.16).

After analysis in the CiteSpace software, the institutions with more frequency of publications were Massachusetts Institute of Technology (MIT), Centre National de la Recherche Scientifique (CNRS), University of Warwick, and Université Pierre et Marie Curie (now part of Sorbonne Université); according to the centrality, the institutions with the highest values were: MIT (0.21), University of Warwick (0.19), CNRS (0.16), and the University of California, San Diego (0.13) (Fig. 4).

The principal authors that published more articles with higher citations are presented in Figure 5, and the keywords grouped in clusters related to these publications: cluster #0 marine *Synechococcus* is related to the authors D.J. Scanlan and W.R. Hess; the cluster #2 phage genome related to the author A.C. Martiny and S.W. Chisholm and the cluster #4 microbial communities related to O. Beja and B. Palenik; this analysis had a high Silhouette 0.9846. The silhouette coefficient greater than 0.5 indicates reasonable data clustering (Kaufman & Rousseeuw 1990). The silhouette coefficients of all five significant clusters exceed 0.9, which shows reasonable clustering results. The authors with the most robust Citation Bursts were in order: Chisholm (1992) - 9.70; Palenik & Koke (1995) - 8.04; and Hess (1995) - 6.12.
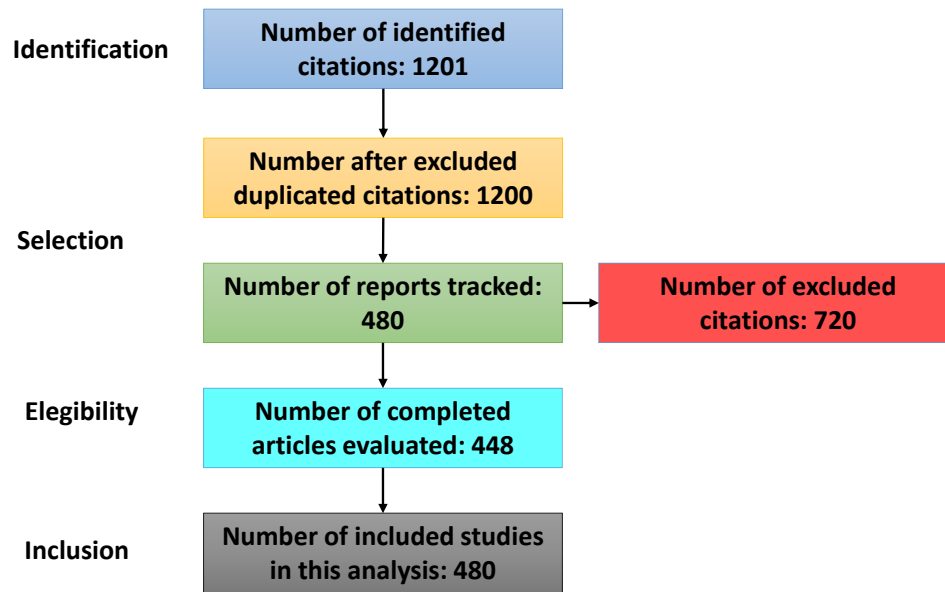
Figure 1. PRISMA flowchart, showing the flow of data collection for inclusion in scientometrics.
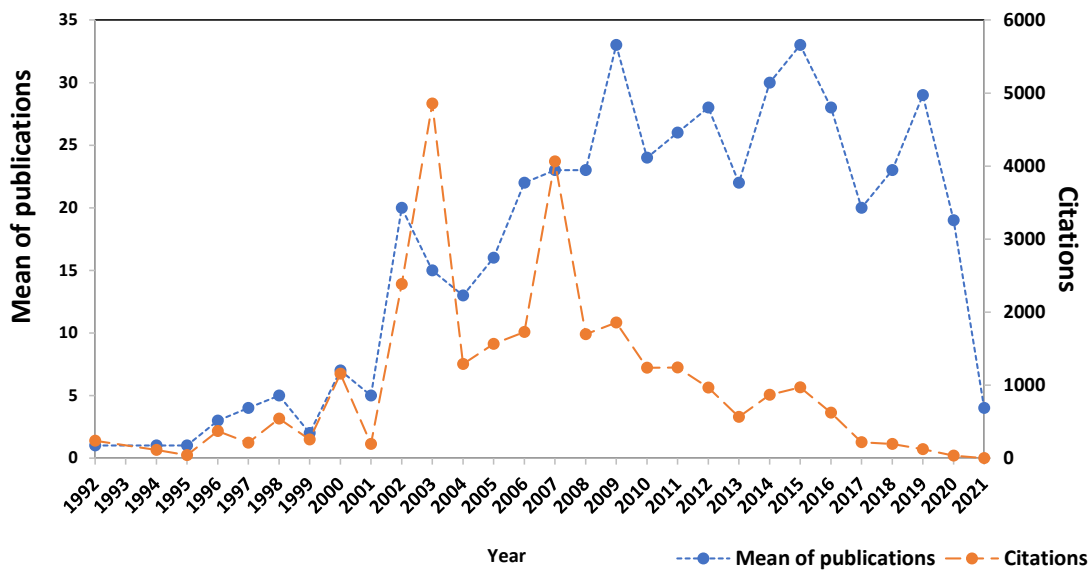

Figure 2. Cumulative distribution of citations and publications between 1992 and 2021.

The leading areas that published more articles with higher citations are presented (Fig. 6a), and the keywords grouped in clusters related to these publications: cluster #0 and #2 genome analysis is related to the Journal Environmental Sciences and Ecology; Biotechnology & Applied Microbiology, Genetics & Heredity, and Virology; the cluster #1 genome reduction related to the Journal Biochemistry & Molecular Biology, and the cluster #3 arctic ocean related to Engineering, Environmental Engineering; this analysis had a high Silhouette 0.9452, demon-

strating that the representation is significant. The journals with the most substantial Citation Bursts were in order: Plant Sciences (1995) - 6.99; Biotechnology & Applied Microbiology (1997) - 5.93; Genetics & Heredity (2007) - 4.04, Evolutionary Biology (2009) - 5.42, and Microbiology (2017) - 3.9.

Figure 6b represents the frequency (%) and centrality of publications by area. The journals with the highest frequency of publications were also the same with the highest centrality, and they are in order (frequency-centrality): Microbiology (230-0.21), Envi-
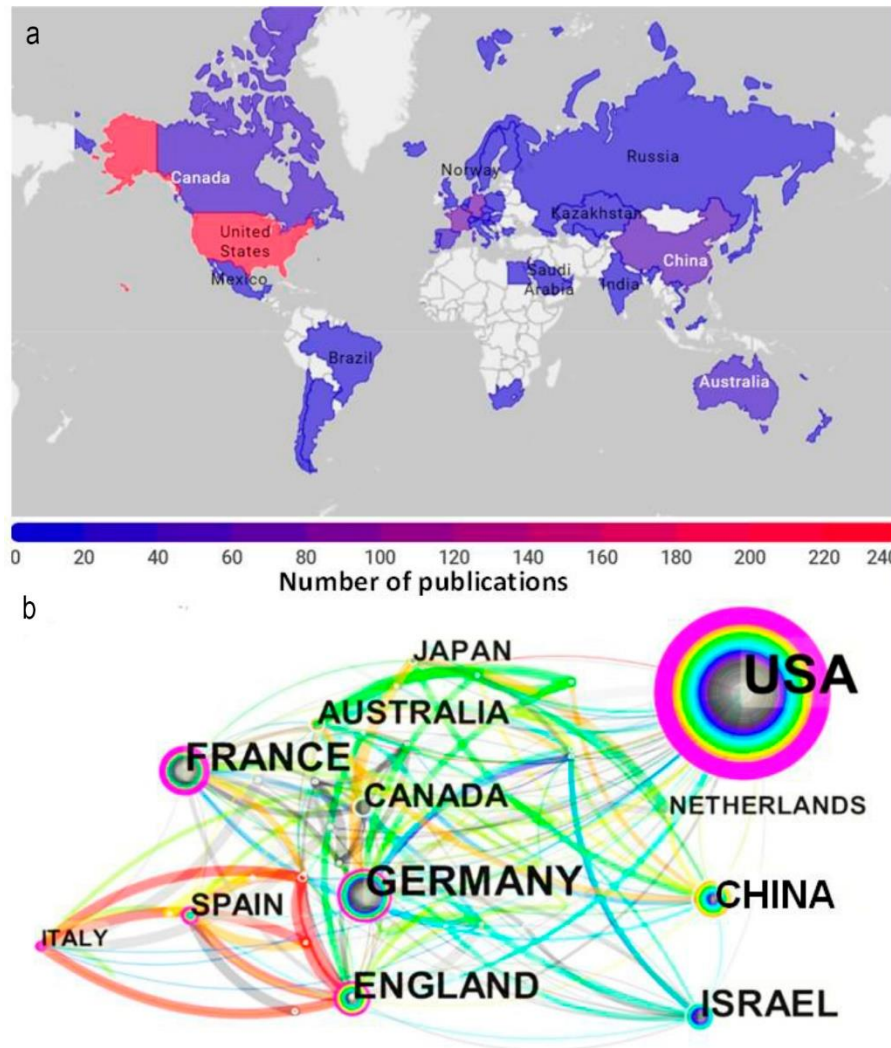
Figure 3. a) Localization of scientific papers published in peer-reviewed journals about genes of *Prochlorococcus* and *Synechococcus*. The geographic distribution of the reviewed papers is related to the color of each country, proportional to the publication number. b) The cooperation network of the producing countries.

ronmental Sciences & Ecology (88-0.19), Science & Technology (70-0.16), and Multidisciplinary Sciences (70-0.13).

Table 1 expresses the publications with the highest number of citations between 1992-2021 regarding the genetics of *Prochlorococcus* or *Synechococcus,* and some authors of these publications were also the same with the most substantial Citation Burst expressed (Fig. 5).

DISCUSSION

The evolution of the Citation Burst scores from areas related to plant sciences to microbiology may reflect the changing interests of researchers and the growth of microbiology as a field. As more research is conducted on microorganisms, including cyanobacteria *Prochlorococcus* and *Synechococcus,* in marine environments, publications related to microbiology may be more likely to have high Citation Burst scores. It could be due to the increasing importance of microbiology in research areas such as ecology, environmental science, and biotechnology. The type of study could also influence the journal selection and the Citation Burst score. Similarly, the study focuses on the ecology or biogeochemistry of these cyanobacteria in marine environments. In that case, it may be more likely to be published in a journal such as Limnology and Oceanography or Environmental Microbiology, which may have a higher impact factor in environmental science.
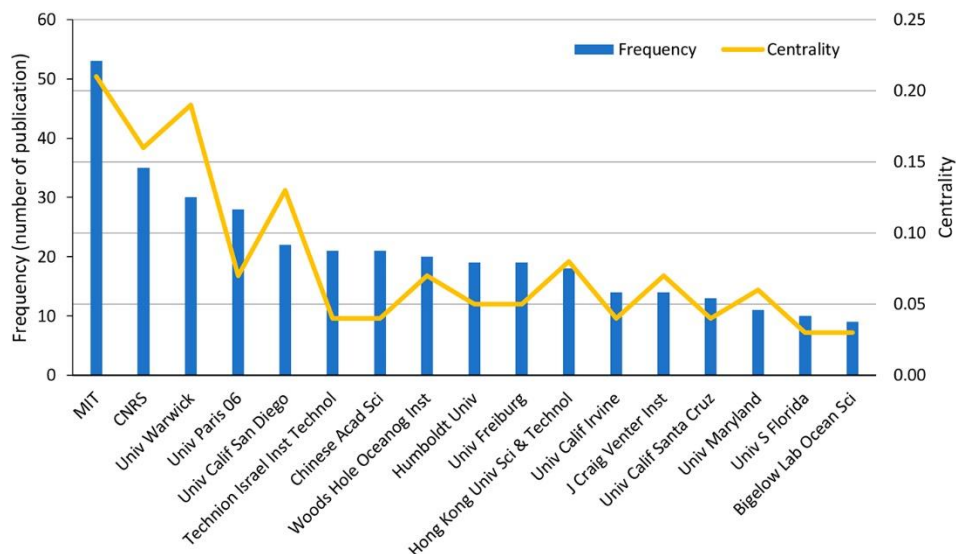
Figure 4. Frequency (number of publications) and centrality of publications per institution. MIT: Massachusetts Institute of Technology, CNRS: Centre National de la Recherche Scientifique.

Table 1. Publications regarding the genetics of *Prochlorococcus or Synechococcus* with the highest number of citations between 1992-2021.

	Title	Author and year	Journal	Citations (n)
1	Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*	Kettler et al. (2007)	Plos Genetics	2737
2	Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation.	Rocap et al. (2003)	Nature	1655
	The genome of a motile marine *Synechococcus*	Palenik et al. (2003)	Nature	935
3	Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5 '-nuclease assays	Suzuki et al. (2000)	Appl. Environ. Microbiology	823
5	Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome	Dufresne et al. (2003)	PNAS	764
6	CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity, and evolution	Badger & Price (2003)	J. Exp. Botany	513
7	Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences	Rocap et al. (2002)	Appl. Environ. Microbiology	400
8	Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*	Moore et al. (2002)	Limnology & Oceanography	383
9	Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*	Sullivan et al. (2003)	Nature	379
10	Transfer of photosynthesis genes to and from *Prochlorococcus* viruses	Lindell et al. (2004)	PNAS	361

Based on the analysis of the occurrence of a subject category, it is possible to describe which area is highlighted in the genetics of *Prochlorococcus* or *Synechococcus*. In this study, the most recurring areas for analysis of the characteristics of categories listed by WoS were chosen and analyzed using CiteSpace.

The main keywords used in the publications and the themes in clusters are presented (Fig. 7). The keyword "cyanobacteria" were related to the clusters #0 *Prochlorococcus* ecotype, #2 thylakoid membrane, and #4 genome analysis, the keyword "gene" was related to #8 molecular component and #9 rRNA gene sequence
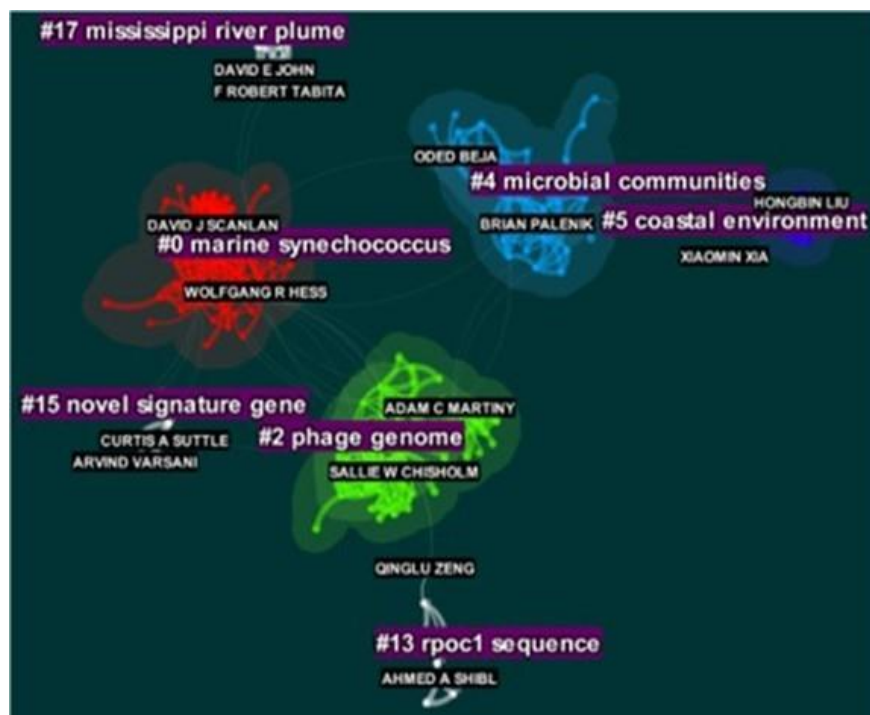
Figure 5. CiteSpace representation of the principal authors and the keywords used in the publications.

analysis, the word "*Synechococcus*" was related to #1 lake Kinneret and #3 *Prochlorococcus* strain, and the word "phytoplankton" related to #5 single *psbA* gene, #1 lake Kinneret and #6 ribulose bisphosphate carboxylase gene expression.

Keywords can be regarded as the soul of an article, and a keyword co-word analysis can be used to detect research topics, analyze research hotspots, and monitor the research frontier transitions of a specific knowledge domain. CiteSpace conducted the keyword co-word analysis related to *Prochlorococcus* and *Synechococcus*. The evolution of the keywords in the publications in the periods of A (1992-2000), B (2001-2010), and C (2011-2021) are shown (Fig. 8). Future studies on *Prochlorococcus* and *Synechococcus* should aim to Research in this area can help improve our understanding of the distribution and abundance of these cyanobacteria in different environments. Genes identification methods are also an important area of research, as they can help identify and characterize genes important for the survival and adaptation of these cyanobacteria in different environments. Understanding the functions and regulation of these genes can provide insight into the mechanisms of evolution and adaptation in *Prochlorococcus* and *Synechococcus*. Finally, future studies should focus on cyanobacteria's genomic diversity and clades. *Prochlorococcus* and

build on the constant expansion and development of research on these cyanobacteria over the past 30 years. To guide future research, Table 2 provides the top 15 keywords regarding frequency and centrality in *Prochlorococcus* and *Synechococcus*, which can help identify the most important and influential research topics and themes. Keywords for future studies should be selected based on frequency and centrality to ensure they are both widely used and highly influential within the field. Future studies should focus on several key areas, including flow cytometry's migration characteristics, gene identification methods, and cyanobacteria's genomic diversity and clades. Flow cytometry is a powerful technique to study *Prochlorococcus* and *Synechococcus* migration pattern*s* in marine ecosystems.

Synechococcus are known to have high genomic diversity, which is related to their diverse adaptation strategies and ecological niches. Research in this area can help improve our understanding of the genetic basis of this diversity and how it relates to the distribution, ecology, and evolution of these cyanobacteria in different environments.

The most frequent and with higher centrality keyword in period A (1992-2000) was "multiple evolutionary origins" (centrality - 0.43). The phylogenetic analyses of cyanobacteria have gained in quantity over the past 20 years (Giovannoni et al. 1988,
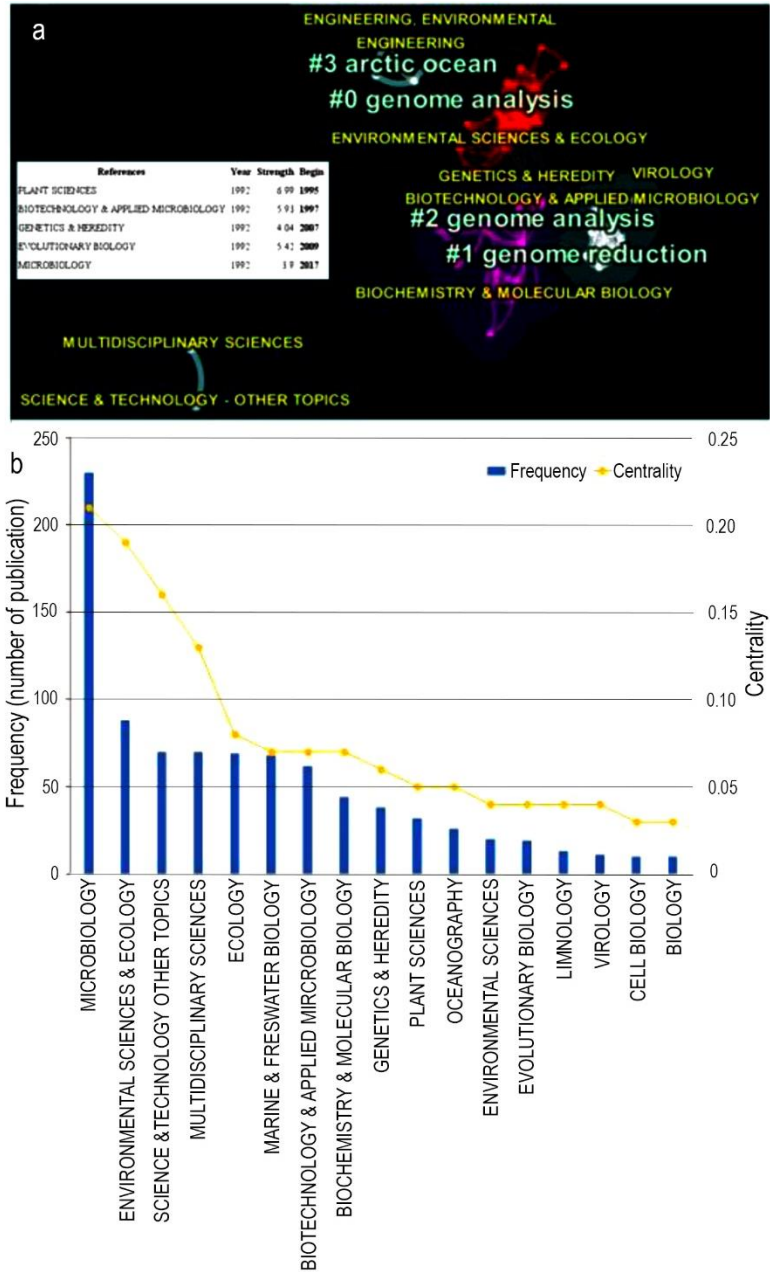
Figure 6. a) CiteSpace representation of the leading journals and the keywords used in the publications and strongest Citation Burst, b) frequency (number of publications) and centrality of publications by a journal.
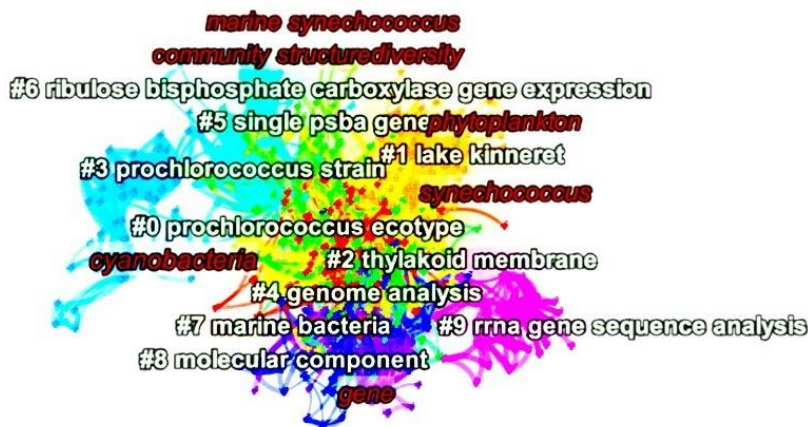

Figure 7. CiteSpace representation of the main keywords used in the publications and the themes in clusters.
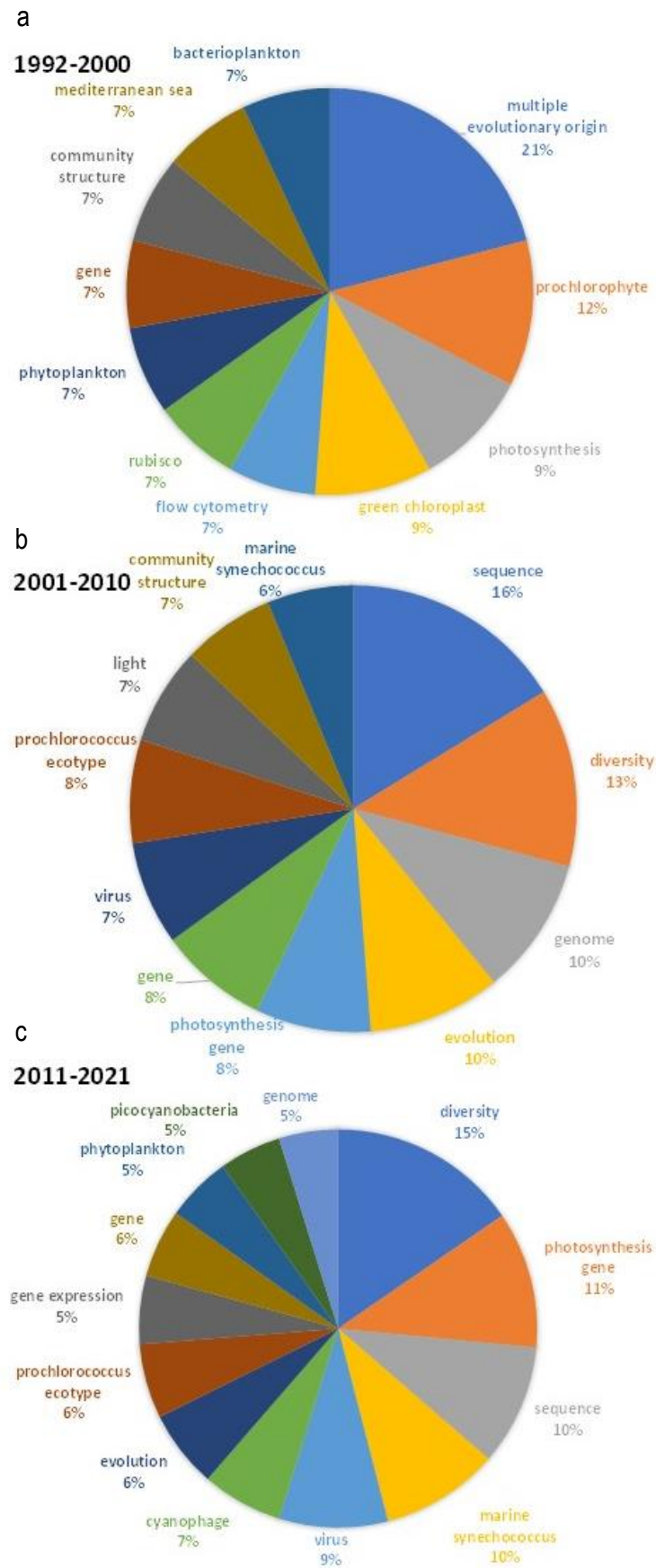
a

1992-2000

b

2001-2010

c

2011-2021

Figure 8. Evolution of the keywords in the publications in the periods a) 1992-2000, b) 2001-2010, and c) 2011-2021.

Table 2. Top 15 keywords related to *Prochlorococcus and Synechococcus* for 1992 to 2021.

Period	Keyword	Frequency	Centrality
A - 1992-2000	multiple evolutionary origins	9	0.43
	prochlorophyte	5	0.16
	photosynthesis	4	0.04
	green chloroplast	4	0.33
	flow cytometry	3	0.40
	rubisco	3	0.00
	phytoplankton	3	0.12
	gene	3	0.12
	community structure	3	0.11
	Mediterranean Sea	3	0.03
	bacterioplankton	3	0.00
	mRNA	2	0.05
B - 2001-2010	sequence	39	0.13
	diversity	31	0.11
	genome	24	0.05
	evolution	23	0.05
	photosynthesis gene	20	0.02
	gene	19	0.13
	virus	18	0.04
	Prochlorococcus ecotype	18	0.07
	light	17	0.04
	community structure	16	0.12
	marine *Synechococcus*	15	0.07
	population	14	0.03
	phytoplankton	13	0.07
	phycoerythrin	12	0.04
	16s ribosomal RNA	11	0.05
C - 2011-2021	diversity	61	0.17
	photosynthesis gene	44	0.05
	sequence	39	0.05
	marine *Synechococcus*	38	0.10
	virus	35	0.06
	cyanophage	26	0.07
	evolution	25	0.06
	Prochlorococcus ecotype	24	0.08
	gene expression	22	0.14
	gene	22	0.08
	phytoplankton	21	0.09
	picocyanobacteria	20	0.02
	genome	19	0.04
	light	18	0.07
	ecotype	18	0.03

Tomitani et al. 2006). Analyses assessing the characteristics of cyanobacterial ancestors provide essential information on the history of cyanobacteria and the evolution of life forms (Swingley et al. 2008). The second word with higher centrality is "flow cytometry" (0.40) because this period was the primary method used to identify *Prochlorococcus* and

Synechococcus. When *Synechococcus* cells are excited by the wavelength of blue light (and to a lesser degree UV - from 405 to 488 nm), phycoerythrin emits a strong orange fluorescence that can be registered separately from the red fluorescence of their chlorophyll. As *Synechococcus* are significantly larger than *Prochlorococcus* (diameters: about 1 and 0.6 µm,

respectively), these two cell types can be discriminated by their scatter signals in flow cytometry related to their size (Marie et al. 1997).

The period B (2001-2010), the keywords with higher centrality were "sequence" and "gene," both with 0.13 of centrality; this can be explained by the period when gene sequencing was developed, and the scientific community discovered the genome of different species. The DNA analysis and the genomic development make identifying cyanobacteria genes for phylogenetic analyses difficult and require an ideal case comparison of complete genome data. Genomic studies in cyanobacteria have emphasized that marine species of *Synechococcus* and *Prochlorococcus* are a particularly well-studied group (Rocap et al. 2003, Dufresne et al. 2008, Scanlan et al. 2009), as shown in Table 2. According to Table 1 was the period with the publications more cited due to gene sequencing (Suzuki et al. 2000, Moore et al. 2002, Rocap et al. 2002, 2003, Badger & Price 2003, Dufresne et al. 2003, Palenik et al. 2003, Sullivan et al. 2003, Lindell et al. 2004, Kettler et al. 2007).

In the last 10 years, period C (2011-2021), the keywords with higher centrality were diversity (0.17), marine *Synechococcus* (0.10), *Prochlorococcus* ecotype (0.08), gene expression (0.14), phytoplankton (0.09), and light (0.07). The most important activities regarding these cyanobacteria were the concern with the diversity and phytoplankton ecology and the discovery of ecotypes in *Prochlorococcus*. The main distinction between ecotypes in *Prochlorococcus* refers to adaptations to low and high light [low light (LL) and high light (HL)] (Campbell & Vaulot 1993), and the physiological differentiation into two HL clades and 4 LL clades allow them to occupy a wide range of luminosity spectrum and nutrient availability at low and mid-latitudes (Partensky et al. 1999, Kettler et al. 2007); the biogeography of *Synechococcus* genetic diversity is less clear. The current notion is that several ecotypes (I, IV, and perhaps several others) are abundant in cooler, nutrient-rich waters. Likewise, several ecotypes (classes II, III, V, VI, and VII) are frequent but variable in tropical and subtropical waters (Kent et al. 2018).

CONCLUSIONS

This study offers a comprehensive scientometrics review of the genetics of cyanobacteria. In addition, we examined the genetics of *Prochlorococcus* and *Synechococcus* characteristics from multiple perspectives: characteristics of publication outputs, cooperation among countries and authors, and co-occurrence analysis of subject categories and keywords. Over the past 29 years, there were 480 journal articles related to this field, and the number of articles increased steadily, which indicated that researchers paid increased attention to this field and that the attention was not evenly distributed around the world but was mainly concentrated in North America, Europe, and Asia. Meanwhile, international collabo-rations among North America, Europe, and Asia were relatively intimate. Furthermore, the USA was the most active contributor to most publications and cooperation among all the countries. In addition, this research is a multifaceted and multidisciplinary field and covers a wide range of interests, from microbiology, environ-mental sciences & ecology, science & technology, and multidisci-plinary sciences.

Additionally, the migration of keywords from "multiple evolutionary origins" and "flow cytometry" to "sequence" and "gene" to "diversity"; "marine *Synechococcus*" and "*Prochlorococcus* ecotype" indicate how future studies should focus, showing that the concern with phytoplankton diversity, ecotypes of cyanobacteria are the new way followed by the scientists regarding these cyanobacteria responsible for a considerable amount of atmospheric oxygen. In the future, the studies will focus on molecular ecology: scientists are studying how genetic variability among different populations of *Prochlorococcus* and *Synechococcus* affects their adaptation to different ocean environments; physiology: researchers are investigating how this phytoplankton respond to changes in nutrient availability and solar radiation, as well as how they are affected by climate change and changes in seawater quality; interactions with other organisms: scientists are studying how *Prochlorococcus* and *Synechococcus* interact with other organisms, such as bacteria, zooplankton, and fish, and how these interactions affect the dynamics of the ocean ecosystem; and modeling: computational models are being developed to predict how the distribution of *Prochlorococcus* and *Synechococcus* will change in the future as a result of climate change and human use of the oceans. Studies of *Prochlorococcus* and *Synechococcus* are important because they help us understand the significance of these microorganisms for the planet's functioning and allow us to take measures to protect the oceans and preserve the health of our planet.

ACKNOWLEDGMENTS

of Graduate Education. Multiuser Core Laboratory of Biological Analysis and Molecular Biology (BioMol) at Universidade Tecnológica Federal do Paraná (UTFPR), *Campus* Dois Vizinhos.

REFERENCES

Badger, M.R. & Price, G.D. 2003. CO2 concentrating mechanisms in cyanobacteria: molecular components, their diversity, and evolution. Journal of Experimental Botany, 54383: 609-622.

Berube, P.M., Biller, S.J., Hackl, T., Hogle, S.L., Satinsky, B.M., Becker, J.W. & Chisholm, S.W. 2018. Single-cell genomes of *Prochlorococcus, Synechococcus*, and sympatric microbes from diverse marine environments. Scientific Data, 51: 1-11.

Campbell, L. & Vaulot, D. 1993. Photosynthetic picoplankton community structure in the subtropical North Pacific Ocean near Hawaii Station Aloha. Deep-Sea Research Part I: Oceanographic Research Papers, 40: 2043-2060.

Campbell, L., Nolla, H.A. & Vaulot, D. 1994. The importance of *Prochlorococcus* to community structure in the central North Pacific Ocean. Limnology and Oceanography, 39: 954-961.

Cesar-Ribeiro, C., Piedras, F.R., da Cunha, L.C., de Lima, D.T., Pinho, L.Q. & Moser, G.A.O. 2020. Is oligotrophy an equalizing factor driving microplankton species functional diversity within Agulhas rings? Frontiers in Marine Science 7: 599185.

Chen, C. 2006. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. Journal of the American Society for information Science and Technology, 57: 359-377.

Chen, C. 2014. The citespace manual. College of Computing and Informatics, 1: 1-84.

Chisholm, S.W. 1992. Phytoplankton size. Primary productivity and biogeochemical cycles in the sea. Springer, Berlin, pp. 213-237.

Doré, H., Farrant, G.K., Guyet, U., Haguait, J., Humily, F., Ratin, M., et al. 2020. Evolutionary mechanisms of long-term genome diversification associated with niche partitioning in marine picocyanobacteria. Frontiers in Microbiology, 11: 567431.

Dufresne, A., Ostrowski, M., Scanlan, D.J., Garczarek, L., Mazard, S., Palenik, B.P. & Partensky, F. 2008. Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. Genome Biology, 9: 1-16.

Dufresne, A., Salanoubat, M., Partensky, F., Artiguenave, F., Axmann, I.M., Barbe, V., et al. 2003. Genome

sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome. Proceedings of the National Academy of Sciences, 100: 10020-10025.

Farrant, G.K., Doré, H., Cornejo-Castillo, F.M., Partensky, F., Ratin, M., Ostrowski, M., et al. 2016. Delineating ecologically significant taxonomic units from global patterns of marine picocyanobacteria. Proceedings of the National Academy of Sciences, 113: E3365-E3374.

Flombaum, P., Gallegos, J.L., Gordillo, R.A., Rincón, J., Zabala, L.L., Jiao, N., et al. 2013. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. Proceedings of the National Academy of Sciences, 110: 9824-9829.

Giovannoni, S.J., DeLong, E.F., Olsen, G.J. & Pace, N.R. 1988. Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. Journal of Bacteriology, 170: 720-726.

Giovannoni, S.J., Delong, E.F., Schmidt, T.M. & Pace, N.R. 1990. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Applied and Environmental Microbiology, 56: 2572-2575.

Grébert, T., Doré, H., Partensky, F., Farrant, G.K., Boss, E.S., Picheral, M., et al. 2018. Light color acclimation is a key process in the global ocean distribution of *Synechococcus cyanobacteria*. Proceedings of the National Academy of Sciences, 115: 2010-2019.

Hess, W.R., Weihe, A., Goër, S.L.D., Partensky, F. & Vaulot, D. 1995. Characterization of the single psbA gene of *Prochlorococcus marinus* CCMP 1375 (Prochlorophyta). Plant Molecular Biology, 27: 1189-1196.

Hirsch, J.E. 2005. An index to quantify an individual's scientific research output. Proceedings of the National academy of Sciences, 102: 16569-16572.

Kaufman, L. & Rousseeuw, P.J. 1990. Finding groups in data. John Wiley & Sons, New Jersey.

Kent, A.G., Baer, S.E., Mouginot, C., Huang, J.S., Larkin, A.A., Lomas, M.W. & Martiny, A.C. 2018. Parallel phylogeography of *Prochlorococcus* and *Synechococcus*. The Isme Journal, 13: 430-441.

Kettler, G.C., Martiny, A.C., Huang, K., Zucker, J., Coleman, M.L., Rodrigue, S., et al. 2007. Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. Plos Genetics, 3: e231.

Lalli, C.M. & Parsons, T.R. 1997. Biological oceanography: an introduction. Butterworth-Heinemann, Oxford.

Li, W.K.W., Lewis, M.R. & Lister, A. 1992 Low cytometric detection of prochlorophytes and cyano-

bacteria in the Gulf of Policastro. Archiv für Hydrobiologie, 124: 309-316.

Li, W.K.W., Rao, D.V., Harrison, W.G., Smith, J.C., Cullen, L.L., Irwin, B. & Platt, T. 1993. Autotrophic picoplankton in the tropical ocean. Science, 219: 292-295.

Lindell, D., Sullivan, M.B., Johnson, Z.I., Tolonen, A.C., Rohwer, F. & Chisholm, S.W. 2004. Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. Proceedings of the National Academy of Sciences, 101: 11013-11018.

Liu, H., Nolla, H.A. & Campbell, L. 1997. *Prochlorococcus* growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. Aquatic Microbial Ecology, 12: 39-47.

Liu, B., Li, Z., Chen, X., Huang, Y. & Liu, X. 2017. Recognition and vulnerability analysis of key nodes in power grid based on complex network centrality. IEEE Transactions on Circuits and Systems II: Express Briefs, 65: 346-350.

Marie, D., Partensky, F., Jacquet, S. & Vaulot, D. 1997. Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I applied and environmental. Microbiology, 63: 186-193.

Moore, L.R., Post, A.F., Rocap, G. & Chisholm, S.W. 2002. Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. Limnology and Oceanography, 47: 989-996.

Olson, R.J., Chisholm, S.W., Zettler, E.R., Altabet, M.A. & Dusenberry, J.A. 1990 Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. Deep-Sea Research Part A: Oceanographic Research Papers, 37: 1033-1051.

Palenik, B. & Koke, J.A. 1995. Characterization of a nitrogen-regulated protein identified by cell surface biotinylation of a marine phytoplankton. Applied and Environmental Microbiology, 61: 3311-3315.

Palenik, B., Brahamsha, B., Larimer, F.W., Land, M., Hauser, L., Chain, P., et al. 2003. The genome of a motile marine *Synechococcus*. Nature, 424: 1037-1042.

Partensky, F., Blanchot, J. & Vaulot, D. 1999. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. Bulletin de l'Institut Océanographique, 19: 457-475.

Rocap, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N.A., et al. 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. Nature, 424: 1042-1047.

Rocap, G., Distel, D.L., Waterbury, J.B. & Chisholm, S.W. 2002. Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. Applied and Environmental Microbiology, 683: 1180-1191.

Scanlan, D.J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W.R. & Partensky, F. 2009. Ecological genomics of marine picocyanobacteria. Microbiology and Molecular Biology Reviews, 73: 249-299.

Sullivan, M.B., Waterbury, J.B. & Chisholm, S.W. 2003. Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. Nature, 424: 1047-1051.

Suzuki, M.T., Taylor, L.T. & DeLong, E.F. 2000. Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5′-nuclease assays. Applied and Environmental Microbiology, 66: 4605-4614.

Swingley, W.D., Blankenship, R.E. & Raymond, J. 2008. Integrating Markov clustering and molecular phylogenetics to reconstruct the cyanobacterial species tree from conserved protein families. Molecular Biology and Evolution, 25: 643-654.

Tomitani, A., Knoll, A.H., Cavanaugh, C.M. & Ohno, T. 2006. The evolutionary diversification of cyanobacteria: molecular-phylogenetic and paleontological perspectives. Proceedings of the National Academy of Sciences, 103: 5442-5447.

Veldhuis, M.J.W. & Kraay, G.W. 1993. Cell abundance and fluorescence of picoplankton in relation to growth irradiance and nitrogen availability in the Red Sea Netherlands. Journal of Sea Research, 31: 135-145.

Veldhuis, M.J.W. & Kraay, G.W. 2000. Application of flow cytometry in marine phytoplankton research: current applications and future perspectives. Scientia Marina, 64: 121-134.