Metadata of the article that will be visualized in OnlineFirst

ArticleTitle	Advances of high-throughput sequencing for unraveling biotechnological potential of microalgal-bacterial communities	
Article Sub-Title		
Article CopyRight	The Author(s), under exclusive licence to Springer Nature B.V. (This will be the copyright line in the final PDF)	
Journal Name	Journal of Applied Phycology	
Corresponding Author	FamilyName	**Solovchenko**
	Particle	
	Given Name	**Alexei E.**
	Suffix	
	Division	Department of Bioengineering, Faculty of Biology
	Organization	Lomonosov Moscow State University
	Address	Moscow, 119234 GSP-1, Russia
	Phone	
	Fax	
	Email	solovchenkoae@my.msu.ru
	URL	
	ORCID	
Author	FamilyName	**Zaytsev**
	Particle	
	Given Name	**Petr A.**
	Suffix	
	Division	Department of Bioengineering, Faculty of Biology
	Organization	Lomonosov Moscow State University
	Address	Moscow, 119234 GSP-1, Russia
	Phone	
	Fax	
	Email	
	URL	
	ORCID	
Author	FamilyName	**Rodin**
	Particle	
	Given Name	**Vladimir A.**
	Suffix	
	Division	Department of Chemistry of Natural Compounds, Faculty of Chemistry
	Organization	Lomonosov Moscow State University
	Address	Moscow, 119234, Russia
	Phone	
	Fax	
	Email	
	URL	
	ORCID	
Author	FamilyName	**Zaytseva**
	Particle	
	Given Name	**Anna A.**
	Suffix	
	Division	Department of Bioengineering, Faculty of Biology

	Organization	Lomonosov Moscow State University
	Address	Moscow, 119234 GSP-1, Russia
	Phone	
	Fax	
	Email	
	URL	
	ORCID	
Author	FamilyName	**Zvereva**
	Particle	
	Given Name	**Maria I.**
	Suffix	
	Division	Department of Chemistry of Natural Compounds, Faculty of Chemistry
	Organization	Lomonosov Moscow State University
	Address	Moscow, 119234, Russia
	Phone	
	Fax	
	Email	
	URL	
	ORCID	

Schedule	Received	31 Dec 2023
	Revised	28 Apr 2024
	Accepted	1 May 2024

Abstract	Although established biotechnological applications of microalgae e.g., the production of high-value metabolites is based on axenic cultures, exploitation of the mutualistic consortia of microalgae and bacteria quickly comes to foreground, especially in bioremediation and wastewater treatment. This trend shifts the focus from genomic research of certain microalgal species to metagenomic studies of interactions between microalgae and bacteria in natural communities and in artificial consortia. Dissection of the genetic determinants of the robustness and productivity of the consortia become a hot research direction, too. Admirable contribution to this topic had been made by high-throughput sequencing (HTS), while recent breakthrough in this field was entailed by the advent and rapid development of the 3rd generation nanopore sequencing which becomes increasingly accurate while providing unprecedented sequencing performance. Recent progress of the Oxford Nanopore Technologies (ONT) enabled both classical metagenomic analysis of microalgal-bacterial communities based on whole metagenome sequencing as well as taxonomic and genetic profiling based on the amplicon sequencing. The parallel emergence of novel bioinformatic algorithms for processing the metagenomic datasets opened new opportunities for the analysis of structure and physiology of microalgal-bacterial communities. From the practical perspective, the new HTS techniques became a time- and labor-savers in discovery of new microalgae with a high potential for the accumulation of valuable metabolites, biodegradation of hazardous micropollutants, and biosequestration of nutrients from waste streams. Search for prokaryotic species boosting the biotechnological potential of eukaryotic microalgae via mutualistic interactions with them is another important goal. The insights from the both short-read and long-read metagenomics will form a solid foundation for the rational design of microalgal-bacterial consortia for biotechnology. In this review, we briefly outline the benefits of the long-read sequencing for structural and functional investigation of algal-bacterial consortia and summarize recent reports on using this approach for achieving the biotechnology-related goals.

Keywords (separated by '-')	HTS - Nanopore - Amplicon sequencing - Microalgae - Metagenome - Metabarcoding - Profiling - Functional prediction
Footnote Information	The online version contains supplementary material available at https://doi.org/10.1007/s10811-024-03267-1 .

Journal of Applied Phycology
https://doi.org/10.1007/s10811-024-03267-1

REVIEW

Advances of high-throughput sequencing for unraveling biotechnological potential of microalgal-bacterial communities

Petr A. Zaytsev[1] · Vladimir A. Rodin[2] · Anna A. Zaytseva[1] · Maria I. Zvereva[2] · Alexei E. Solovchenko[1]

Received: 31 December 2023 / Revised: 28 April 2024 / Accepted: 1 May 2024

Abstract

Although established biotechnological applications of microalgae e.g., the production of high-value metabolites is based on axenic cultures, exploitation of the mutualistic consortia of microalgae and bacteria quickly comes to foreground, especially in bioremediation and wastewater treatment. This trend shifts the focus from genomic research of certain microalgal species to metagenomic studies of interactions between microalgae and bacteria in natural communities and in artificial consortia. Dissection of the genetic determinants of the robustness and productivity of the consortia become a hot research direction, too. Admirable contribution to this topic had been made by high-throughput sequencing (HTS), while recent breakthrough in this field was entailed by the advent and rapid development of the 3rd generation nanopore sequencing which becomes increasingly accurate while providing unprecedented sequencing performance. Recent progress of the Oxford Nanopore Technologies (ONT) enabled both classical metagenomic analysis of microalgal-bacterial communities based on whole metagenome sequencing as well as taxonomic and genetic profiling based on the amplicon sequencing. The parallel emergence of novel bioinformatic algorithms for processing the metagenomic datasets opened new opportunities for the analysis of structure and physiology of microalgal-bacterial communities. From the practical perspective, the new HTS techniques became a time- and labor-savers in discovery of new microalgae with a high potential for the accumulation of valuable metabolites, biodegradation of hazardous micropollutants, and biosequestration of nutrients from waste streams. Search for prokaryotic species boosting the biotechnological potential of eukaryotic microalgae via mutualistic interactions with them is another important goal. The insights from the both short-read and long-read metagenomics will form a solid foundation for the rational design of microalgal-bacterial consortia for biotechnology. In this review, we briefly outline the benefits of the long-read sequencing for structural and functional investigation of algal-bacterial consortia and summarize recent reports on using this approach for achieving the biotechnology-related goals.

Keywords HTS · Nanopore · Amplicon sequencing · Microalgae · Metagenome · Metabarcoding · Profiling · Functional prediction

Introduction: Microalgal consortia as a promising vehicle for biotechnology

In nature, microalgae exist within microbial communities with other microbial species including diverse fungi, bacteria, and/or archaea. In these communities, microalgae become engaged in a complex network of interactions with their partner species represented mostly by bacteria, implemented as trophic exchange and/or chemical signaling. There is ever increasing evidence of the correlation between composition and activity of the bacterial component of the consortium and the physiological condition of the microalgae. This evidence suggests that the interactions between the microalgae and the bacteria can be significant for the

✉ Alexei E. Solovchenko
solovchenkoae@my.msu.ru

[1] Department of Bioengineering, Faculty of Biology, Lomonosov Moscow State University, Moscow 119234 GSP-1, Russia

[2] Department of Chemistry of Natural Compounds, Faculty of Chemistry, Lomonosov Moscow State University, Moscow 119234, Russia

consortium itself. It can also affect practically relevant characteristics such as cell division rate, its biochemical composition and excretion of assorted compounds (Danish-Daniel et al. 2023; Li et al. 2023b).

There are species and whole taxa of microalgae whose microbiomes are of a considerable interest due to their high biotechnological potential or even a possible threat to human health and economics (Kublanovskaya et al. 2020b; Danish-Daniel et al. 2023). Among the most conspicuous forms of microalgal-bacterial interactions, and hence most studied so far, is the formation of complex structures such as floccules and biofilms or their biomimetic analogs—photogranules (Trebuch et al. 2020, 2023). They frequently include prokaryotic oxygenic phototrophs—cyanobacteria (Kublanovskaya et al. 2019, 2020a).

A crucial role in the formation and evolution of microalgal-bacterial consortia is played by the phycosphere. This term was coined to denote a spatial zone in close proximity to the microalgal cell surface characterized by the presence of superficial structures of microalgal cells as well as by gradients of chemical and physical parameters making the phycosphere especially favorable for other organisms. In other words, microalgae acts as ecosystem engineers or, in terms of ecology, edificator of the microbial community formed around its cells. Eventually, the phycosphere becomes inhabited by microorganisms engaged in diverse (mostly symbiotic) interactions with the basibiont (the microalga) and between themselves (Fig. 1).

These interactions can be significant from the practical standpoint (Seymour et al. 2017). The most known is successful application of mixed cultures of microalgae with plant growth-promoting bacteria (PGPB) for soil remediation and biofertilization (Gonzalez and Bashan 2000; de-Bashan et al. 2021; Gonzalez-Gonzalez and de-Bashan 2023). The co-culture improves soil health and stimulates crop plants productivity by synthesizing a broad spectrum of bioactive molecules (de-Bashan et al. 2004; 2021) including the phytohormone analogs excreted by representatives of *Chlorella*, *Scenedesmus*, and *Chlamydomonas*.

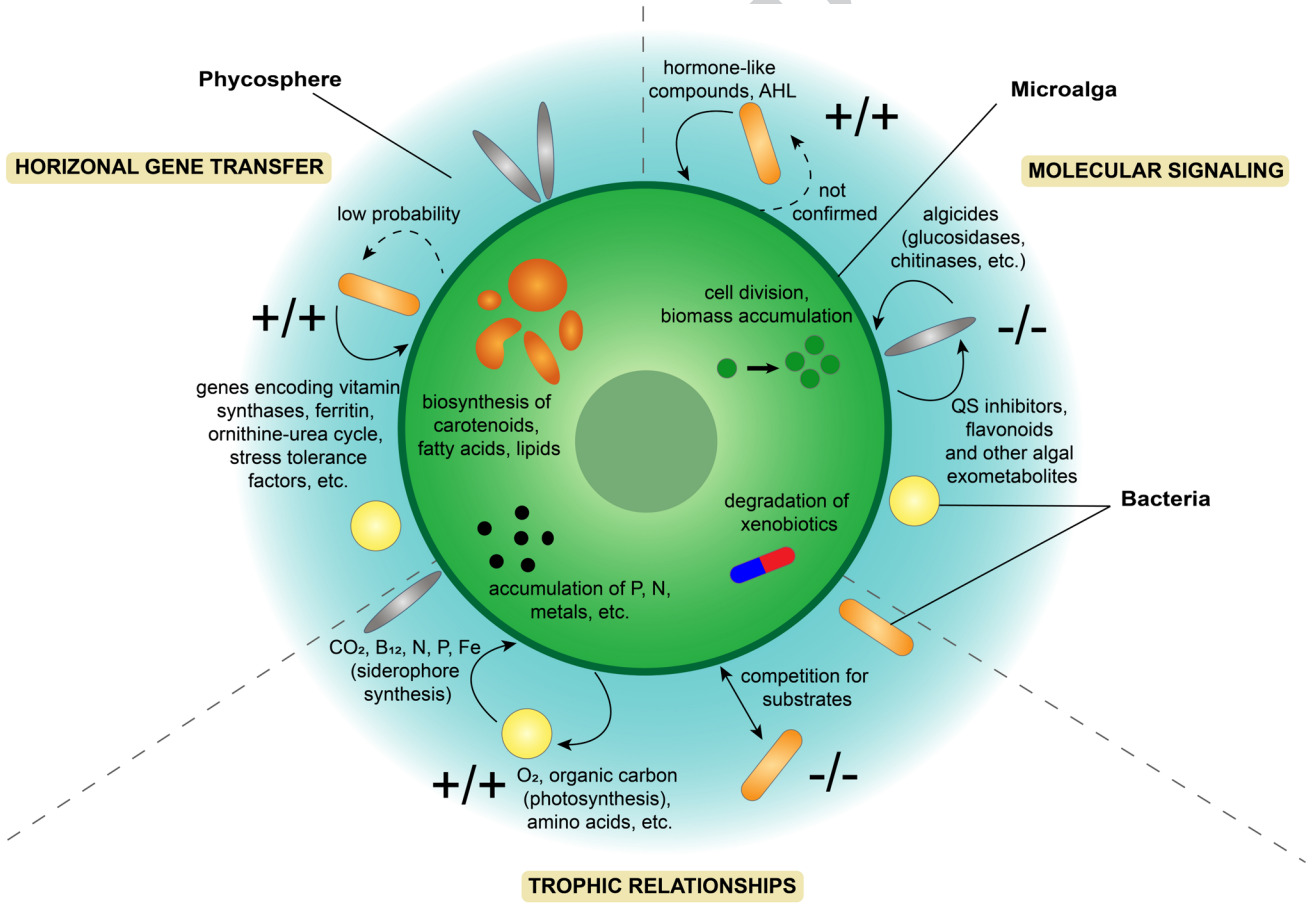


Fig. 1 Schematic representation of the phycosphere formed around a microalgal cell, its bacterial inhabitants, and the processes within. The interactions are divided into three main categories shown in the scheme. Thin black arrows show directions of the interactions, dashed arrow denote unconfirmed interactions. The "+" and "–" signs denote positive and negative interaction types, respectively. The biotechnologically relevant processes affected by the microalgal-bacterial interactions are listed inside the microalgal cell silhouette

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

Cyanobacteria also fix nitrogen (Llamas et al. 2023) and make it, together with phosphorus, more bioavailable to crop plants with participation of microorganisms from the genera *Azospirillum*, *Azotobacter* and other diazotrophic cyanobacteria (Scognamiglio et al. 2021; Solomon et al. 2023). Consortia of microalgae and plant growth-promoting bacteria (PGPB) boost growth and pathogen resistance of important vegetable crops including tomato, onion, and cucumber by stimulating their nitrogen uptake and producing bioactive polysaccharides (Kang et al. 2021). With cyanobacteria added to a microalgae-PGPB consortium, a robust synthetic consortium is formed which can serve as efficient biofertilizer (Sadvakasova et al. 2023). A similar result could be achieved by co-immobilization of microalgae-PGPB consortia on alginate and/or chitosan beads (Gonzalez and Bashan 2000).

Another major application field for microalgal-bacterial consortia is the biotreatment of wastewater by biosequestration of nutrients, decomposition of bulk pollutants, and biodegradation of hazardous micropollutants. Common issues of the microalgae-based solutions for environmental applications including their stability and sustained efficiency under fluctuating environmental conditions and wastewater composition, as well as economic viability can be, in principle, addressed by appropriate microalgal-bacteria consortia (Saravanan et al. 2021).

Bacteria from certain taxa, frequently belonging to PGPB as well, also exert stimulatory effects on microalgal growth and productivity. In analogy with PGPB, those bacteria were named microalgal growth-promoting bacteria (MGPB). Supplementation of MGPB to axenic cultures of microalgae from the genera *Chlorella*, *Chlamydomonas*, and *Euglena* frequently used in wastewater treatment increase biomass accumulation and the treatment efficiency (Toyama et al. 2018).

The most robust form of the algal-bacterial consortia in the wastewater treatment systems are algal-bacteria biofilms (Clagnan et al. 2023). Tehes biofilms can be formed with participation of quorum-sensing mechanisms orchestrating the microalgal-bacterial interactions to attract the MGPB to populate the niches formed around the photoautotrophic cells (Qixin et al. 2022). The MGPB can either stimulate the growth of microalga by supplying them with essential co-factors and vitamins (Shetty et al. 2019; Iqbal et al. 2022) or perform enzymatic hydrolysis of the microalgal cell wall increasing the product yield in case of valuable metabolite production (Carrillo-Reyes et al. 2016). Increasing the bioavailability of nitrogen by bacteria in wastewater sludge communities facilitates accumulation of microalgal biomass (Leong et al. 2020) and, in certain cases, lipid productivity of species from the genera *Chlorella*, *Chlorococcum, Scenedesmus*, and *Nannochloropsis* (Koreivienė et al. 2014; Arutselvan et al. 2021; Upadhyay et al. 2021). Future breakthroughs in wastewater treatment are expected from application of multi-omics approach and high-throughput methods for screening for selection and/or design of even more robust and productive consortia (Patel et al. 2017; Padmaperuma et al. 2018; Nagarajan et al. 2022).

Clearly, the environmental and agricultural applications of microalgae are about "xenic" cultures and consortia. Moreover, the advent of molecular methods of culture purity control revealed that many microalgal cultures that passed conventional axenicity tests appeared to be not really axenic and harbored other (non-cultivable) microorganisms. Interestingly, using strictly axenic cultures in microalgal biotechnology was frequently complicated by deterioration of culture vigor and productivity, let alone the costs of axenicity maintenance at large scale (Patel et al. 2017; Padmaperuma et al. 2018).

These circumstances have focused interest to the consortia themselves and methods of their investigation and engineering. It became clear that engineering of the phycosphere aimed at to populating it with desirable MGPB would ensure a kind of division of labor between the components of the consortium for avoiding metabolic overload, enhanced biomass accumulation, balancing the growth by quorum sensing mechanisms, and increase of nutrient availability for microalgae (Park et al. 2017; Patel et al. 2017; González-González and de-Bashan 2021). Specific examples include significant increase of chlorophyll, lipid, and carotenoid content in co-cultures of microalga with the bacteria that are frequently found in microalgal core microbiome such as *Paracoccus haeundaensis – Lactobacillus fermentum*, *Characium* sp. – *Pseudomonas composti*, *Tetradesmus obliquus* and *Coelastrella* sp. – *Variovorax paradoxus* (Berthold et al. 2019; Choi et al. 2021; Perera et al. 2021).

Of special interest is boosting the productivity of the biotechnologically important microalgae such as *Haematococcus lacustris* without resorting to their genome modification. Solving this problem would make the natural astaxanthin from microalgae much more competitive than it is now. Thus, *H. lacustris* has shown more than two-fold increase in its major secondary carotenoid astaxanthin yield in co-culture with the bacteria *Sphingomonas hankookensis* or *Paenarthrobacter ureafaciens*, or the fungus *Simplicillium lanosoniveum* (Lee et al. 2022). Co-culturing of microalgae with certain yeast species results in beneficial cross-feeding that either increases the rate of carbon dioxide assimilation or enables the utilization of organic carbon sources for higher biomass accumulation (Cheirsilp et al. 2012; Wang et al. 2016; Gao et al. 2023b). Such co-cultures are designed by high-throughput screening of suitable auxotrophs among microalga, bacteria, and fungi to arrange the most efficient trophic interactions (Saleski et al. 2019).

Taking a closer look on the publication landscape related with microalgal genomics, one might notice that

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

molecular biology methods have become widespread within that field (Fig. 2, see also online Supplementary). However, the topic of high-throughput sequencing (HTS) is still underrepresented for microalgal biotechnology. But even then, the initial focus of sequencing techniques on microalgal metabolism and culturing has been gradually shifting towards the genomics approach of microalgal-bacterial community investigation. As a result, HTS methods are becoming tightly related to environmental research, such as ecological monitoring of phytoplankton (including notorious algal blooms and eutrophication), aquaculture, and wastewater treatment. Today we see the emerging understanding of the importance of HTS for monitoring of microbial and microalgal diversity, as well as for estimation of microalga-bacterial consortia functional potential. At the same time, it became clear that solving these problems demands new experimental methods and data processing algorithms in metagenomics. Below, we attempt to outline the importance of long-read sequencing for getting insights into the structure, functioning, and biotechnological potential of algal-bacterial consortia. *Pro et contra* of the mainstream sequencing technologies will be discussed with an emphasis of nanopore sequencing represented by Oxford Nanopore Technology (ONT). Special attention will be given to novel algorithms developed for gaining actionable insights from the data output of ONT sequencing platform. The review covers the reports (Fig. 2, see also online Supplementary) on the successful applications of HTS in the field of microalgal ecology and microalgal-bacterial interactions in the context of biotechnology. Additionally, the amount and specificity of the long read-based metagenomics is considered.

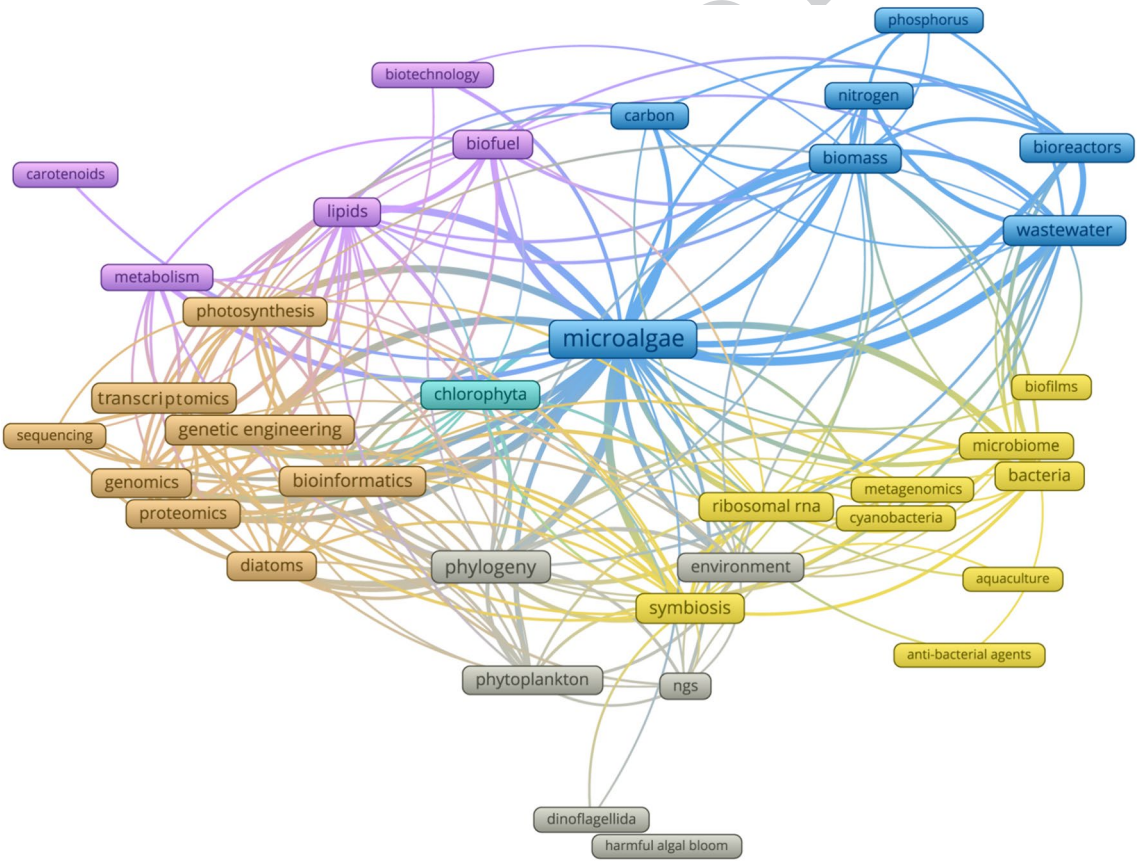


Fig. 2 Co-occurrence map of the keywords in publications related to microalga genomics. The largest number of edges are within 'molecular methods' cluster (brown) and to 'symbiosis' node of 'metagenomics' cluster (yellow), stating the increasing attention to microalga-bacteria interactions in molecular systems biology field. At the same time, the nodes 'metagenomics' and 'microbiome' have few and thin edges with 'bioproduction' and 'microalga technology' clusters (purple and blue, consequently), which highlights future potential of metagenomic studies for practical application of alga. The initial set of titles, abstracts, and keywords of 767 research articles was collected from PubMed and analyzed in VOSviewer 1.6.20 (only the keywords which occurred 10 times and more were taken, clustering resolution = 1.3, min. strength = 8)

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

Metagenomics in microalgal research

Metagenomics is a key to knowledge of the microbial universe

Metagenomic approach to investigation of microbial communities evolved in the last two decades. It became a powerful tool for studies of the microbiomes of soil, marine and freshwater sediments, and planktonic communities, as well as microbiomes of animals and plants. This approach also proliferated into diverse practical applications such as environmental monitoring, control of food quality and fermentation, medical research, and wastewater treatment. Recently, researchers started to use metagenomics to dissect microalgae-based communities from various biotopes from active sludge of wastewater treatment plants to photobioreactors. An illustrious example is comprised by culture crash "forensics" (Lane et al. 2013).

Metagenomic approach becomes increasingly widespread in the studies of microbial communities while the classical methods that are based on isolation and cultivation are giving up their positions since the latter are (i) labour- and time-intensive and (ii) suffer from a high organism-dependent bias. An important advantage of metagenomics is its potential to reveal hidden microbial diversity represented by uncultured species. This is especially relevant to bacterial symbionts of microalgae in natural and artificial systems.

Generally, metagenomic studies of alga-bacterial communities aim to answer three practical questions:

1. What organisms form the community (which taxa do they belong to)?

This question is solved using molecular identifiers or barcodes uniquely identifying organisms at different levels of taxonomy. According to the principle of DNA barcoding, sets of genomic loci are selected to ensure the desired level of identification accuracy for bacterial and microalgal strains. While the 16S rRNA gene locus is usually sufficient for identification of the most of heterotrophic bacterial species in the microalgal phycosphere (Lebonah et al. 2014), reliable identification of oxygenic phototrophs requires a more extended set of loci. Thus, for eukaryotic microalgal nuclear genes (18S rRNA, *nu*ITS1, and *nu*ITS2), chloroplastic genes (*rbcL*, *tufA*, and cp23S), as well as mitochondrial cytochrome *c* oxidase subunit I (COI) gene are used in most situations (Hadi et al. 2016; Zou et al. 2016; Ballesteros et al. 2021). Among those, the *tufA* gene encoding a plastidial elongation factor currently is the most promising marker capable of resolving lower taxa within the class Chlorophyceae (Vieira et al. 2016). For identification of Cyanophyta, the 16S rRNA gene and ITS between 16S and 23S rRNA genes, functional *rbcL* or *nif* genes, and a subunit of RNA polymerase (*rpoB*/*C*/*D* genes) are commonly used (Mishra 2020; Ballesteros et al. 2021). The CBOL (Consortium for the Barcoding of Life) recommends the consequent application of at least two markers for reliable identification of microalgal taxa (Pawlowski et al. 2012).

2. What is the potential functional profile (ecological function) of the community?

The possible physiological and other features of a community are defined by list of functional orthologs represented in the genomes of species forming this community. A more or less specific set of genetic determinants can be compiled for any major phenotypical trait expressed at the level of community. Typical examples include (but not limited to) nitrification (*amo*, *nxr*, *hao*, etc.), denitrification (*nap, nar, and nirS,* etc.), and uptake of phosphate with its subsequent accumulation in form of polyphosphate (*pho* genes, *PSR1, PTC1, ppk, ppk2,* etc.) (Wang et al. 2023; Xiong et al. 2023). Following the concept of reverse ecology, such gene sets might be the basis of metabolic reconstruction of an entire microalga-bacterial community (Cao et al. 2016). This approach might also reveal a lot of information about the biotechnological potential of a community which might be useful e.g. for *in silico* pre-screening.

3. What are the possible interactions between the organisms forming the community?

Answering this question requires study of the genes responsible for different modes of communication between microalga and bacteria in the phycosphere, from trophic interactions to chemical signaling based on specific molecules (see e.g., Fig. 1). Excellent examples of the latter are quorum sensing substances, phytohormones, algicides, growth inhibitors, and extracellular enzymes that modulate their activity in the medium (Dow 2021; Astafyeva et al. 2022; Santo et al. 2022). Another route of interactions within the community is horizontal gene transfer between species—the phenomenon noticed in natural microalga-microbial communities under selective pressure of hazardous micropollutants (Liu et al. 2022; Li et al. 2023a).

HTS in microbial community research: pro et contra

Since the advent of the first method for sequencing nucleic acids, this approach has evolved dramatically yielding three generations of sequencing with distinct advantages and drawbacks (Table 1; Sanger and Coulson 1975; Slatko et al. 2018). Each method has its own unique characteristics and

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

Table 1 Three generations of nucleic acid sequencing methods generally applied in microalgal-bacterial community research

Generation	Sequencing method –Platform	Underlying principle	Advantages	Drawbacks	Remarks
First generation sequencimg	Chain termination method (Sanger sequencing) – Agilent Bioanalyzer	Synthesis of a new DNA chain, a complementary matrix chain, with the inclusion of labeled nucleotides stopping the synthesis process	High accuracy	Low throughput, high costs, inefficient for sequencing large genomes and metagenomes	Suitable for: sequencing of individual genes (short fragments), verification of the results of other methods, identification of axenic cultures
Second/Next generation sequencing (SGS/NGS), High-throughput sequencing* (HTS), Sequencing by synthesis	Pyrosequencing –Roche/454	Measurement of the released pyrophosphate during the incorporation of nucleotides into a growing DNA chain	Longer reads	Errors in sequencing homopolymer regions, higher error in bases, sensitive to DNA quality	Good for amplicon sequencing, suitable for detecting variations
	Dye sequencing –Illumina	Sequencing by synthesis (incorporation of designated nucleotides and fluorescence detection)	High throughput, low cost of reading, high accuracy	Shorter reads (up to 250 b.p., difficulties in resolving long repeats	A mainstream platform for genomics and transcriptomics of axenic cultures
	Semiconductor sequencing – Ion Torrent	Sequencing by synthesis (detection of hydrogen ions released during incorporation of deoxyribonucleotide triphosphates)	High throughput, low cost of reading	Reads up to 400 b.p. (errors in sequencing homopolymers)	Suitable for microbial genome and transcriptome sequencing, targeted sequencing
Third generation sequencing (TGS),	Single-molecule real-time sequencing (SMRT) – PacBio	Sequencing by replication with labeled nucleotides	Long reads	Low throughput, relatively high costs and relatively scarce availability worldwide	Fits well for closing gaps in reference assemblies and characterization of structural variation in genomes
	Nanopore sequencing – Oxford Nanopore Technolog (ONT)	Recording of conductivity changes during movement of nucleic acid molecule through a nanopore	Long reads (tens thousands b.p.), portability, high throughput, low cost of reading	Relatively low accuracy	A promising platform for metagenomics and transcriptomics of xenic cultures and communities from environmental applications

*Depending on a particular ONT product the nanopore sequencing can also be considered as HTS

the method of choice depends on specific goals and requirements of the study. Instead of reviewing technical details of each method (for those, we refer the reader to recent overviews: (Mardis 2017; Slatko et al. 2018)), here we will highlight their scope and applicability for investigation of microalgal-bacterial communities with emphasis on the most recent long read-based technologies.

The third-generation sequencing is distinguished by the ability to read long DNA sequences. This method is represented on the market by single-molecule real-time sequencing technology (SMRT) from PacBio and by solutions from Oxford Nanopore Technology (ONT) company. Admittedly, PacBio has proven itself as the most powerful sequencing method existing up to now, due to the optimal ratio of potential read length to sequencing precision. Still, it has not become a mainstream technology, particularly due to its high cost (almost twice as much compared to other TGS) that leads to lower availability (Athanasopoulou et al. 2022). However, the second technology is becoming more widespread due to the opposite trends. In the process of nanopore sequencing, individual single-stranded DNA or RNA molecules pass through the nanopore causing changes in electrical conductivity in the nanopore unit. The corresponding changes in the electrical signal are recorded, and the nucleotide sequence is inferred from these records. Further details on the principles and technical implementation of this method can be found in (Kasianowicz et al. 1996; Stoddart et al. 2009).

A key advantage of TGS is the length of reading, which can reach 100 thousand b.p. which is unattainable for other sequencing platforms. Another decisive advantage of TGS is its ability to perform single-molecule sequencing without the need to average signals from a group of molecules making the results more accurate. Combination of these advantages makes the sequencers capable of precise analyzing of long repeats and GC-rich regions of the genome, unlike the alternative technologies (Jain et al. 2018). Furthermore, nanopore sequencing, provided that a sufficient amount of genomic DNA (above 200 fmol) was extracted from the biological sample, makes it possible to omit the amplification step in the sample preparation routine, while PacBio requires pre-amplification for some purposes (Athanasopoulou et al. 2022). The lack of the PCR amplification step lowers the risk of a bias due to selective enrichment of certain parts of the genome and shortens the sample preparation time, saving reagents and making the whole process more portable. These circumstances and the small size of nanopore sequencers makes them extremely mobile, so the whole sequencing can be carried out as a kitchen-table effort (Edwards et al. 2016).

At the same time, this technique from the very beginning suffered from a greater (relative to other modern sequencing approaches) number of reading errors (determining specific nucleotides in particular position with thin the nucleic acid sequence). Admittedly, this problem is increasingly mitigated every year. Thus, although almost 40% of reading errors were reported in 2015 (Laver et al. 2015), in 2018 the error rate has been reduced to 0-10% thanks to improved data processing and sample preparation (Jain et al. 2018), and in 2022 a solution was announced to increase the reading accuracy to 99.9%. Using the last modification ONT a systematic study of *de novo* genome assembly with control the quality of assembled genomes as well as reads by the ability to reproduce SNVs and deletion of gene found in alternative experiments for the same samples the application only ONT technology for *de novo* genome assembly was proved (Khrenova et al. 2022). As a result, nanopore sequencing now has a broad range of applications including genomics, epigenomics, metagenomics, and RNA research. It is widely used in life sciences research, medicine, agriculture, and other fields where genome sequencing or nucleotide sequence analysis is required (Zhang et al. 2022; Badger et al. 2023; Mastrorosa et al. 2023).

One of the distinguishing features of ONT is the direct nucleic acids sequencing ability, which opens new opportunities for high accuracy transcriptomics, including identification of novel isoforms and detection of full-length RNA (Athanasopoulou et al. 2022). On the other hand, the accurate differential analysis based on long-read sequencing data may require higher throughput via generation of cDNA library, which in case of nanopore sequencing still provides an advantage over existing methods by reading full-length isoforms and avoiding (or at least reducing) additional bioinformatics step to assemble reads into transcripts. However, this potential has not been fully leveraged due to the limitations of current long-read assembly methods and underdeveloped short-read data integration approaches. Unevenly low coverage when using short-read technologies leads to the splitting of one transcript into several transcripts or incorrect definition of ends and, as a result, to errors in the assessment of differential gene expression. Conversely, long-read sequencing libraries lack depth of coverage and suffer from artifacts in cDNA-based methods, leading to erroneous assembly and quantification of transcripts. To overcome these problems, a hybrid assembly approach (short and long reads together) is used, which dramatically increases the sensitivity and accuracy of full-length transcript assembly on the correct strand and improves the detection of biological features of the transcriptome (Kainth et al. 2023). When alternative splicing has a significant contribution to transcriptomic variation, ONT protocols have been shown to be superior to short-read sequencing protocols in terms of transcriptome assembly and the risk of false positives due to unambiguous mapping of reads to transcripts (Engelhard et al. 2023).

Both PacBio and ONT are suitable for implementation of two main strategies of metagenomic studies: whole

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

metagenome sequencing and amplicon sequencing of a specific loci either for identification of the microbes and/or revealing their functional potential (Athanasopoulou et al. 2022; Kim et al. 2022). Thus, in 16S-based studies, PacBio and ONT allow the creation of primers covering the entire 16S10 gene or even entire ribosomal operons, increasing dramatically the resolution of the taxonomic assignment i.e., the number of precisely distinguishable species (Kerkhof et al. 2017; Tedersoo et al. 2018). Reading the whole metagenome leads to minimal bias in species composition and amount. At the same time, amplicon sequencing of DNA-barcodes (or metabarcoding), e.g., 16S rRNA, its internal transcribed spacer (ITS), *rbcL* etc., offers a cheaper alternative which features a higher throughput but is potentially prone to bias due to the presence of amplification step (see above).

It is well known that "traditional" short-read sequencing technologies cannot reliably resolve repeats and duplicated regions of the genome, so their using for taxonomical assignment and genome assembling of closely related species is complicated (Ashton et al. 2015), while heterogeneity inherent in the metagenome might lead to incorrect assembly between species. In case of metabarcoding, the *16S rRNA* gene sequence harboring a combination of conservative and highly variable regions allows for precise species identification, but limitations of the short-read technologies (NGS, Table 1) prevent them from covering a sufficiently long part of this gene to provide species-level resolution (Shin et al. 2016).

Nowadays, TGS (mostly nanopore sequencing) has secured its place in the array of methods for studies of microbial communities offering distinct advantages for metagenomics. Despite some admirable results produced by PacBio technology in assembling whole genomes of microorganisms, including microalgae (Luo et al. 2018; Maeda et al. 2019; Gao et al. 2023a), there are few works dedicated to PacBio evaluation of microbial communities (Tedersoo et al. 2018; Gueidan et al. 2019; Kim et al. 2022). Therefore, we shall consider below the specific applications of the nanopore technology (solely or in combination with short-read methods) for scrutinizing the microalgal community structure and functional profile.

Studying the whole metagenome of microalgal communities with HTS technologies

The whole metagenome sequencing (WMS) approach stands as the golden standard for metagenomic studies of various sample types harboring microalgal-bacterial consortia, mostly due to the large amount of sequence data enabling thorough analysis of the consortia. That includes precise taxonomical identification of eukaryotic and prokaryotic species forming a community, confirming the presence of diverse functional genes sets, search for new efficient and stable enzymes and reconstruction of metagenome-scale metabolic models (Belcour et al. 2020; Zorrilla et al. 2021; Kuppa Baskaran et al. 2023). Further insights can be obtained by investigating raw metagenome reads or scaffolds, for example from phylotyping based on straightforward count in alignment-free algorithms (Inskeep et al. 2013; Patil and McHardy 2013), more precise taxonomical identification by BLAST or another sequence comparison tool such as implemented in MEGAN or TAXAssign algorithm (Huson et al. 2007; Inskeep et al. 2013), or classification based on the species-level genome bins e.g., with MetaPhlan 4 algorithm (Ljaz and Quince 2013; Blanco-Míguez et al. 2023).

The most popular approach relies on pre-assembled genomes from the metagenome (MAG) for prokaryotic and eukaryotic species, which however might be limited by insufficient coverage of taxa and quality of the assemblies (Yang et al. 2021). Though application of both mentioned approaches is better adopted for prokaryotic species, there is an emerging trend in algorithm development for eukaryotic microorganisms, including microalgae. Such tools as EukRep and Tiara utilize machine learning and deep learning methods to classify read subsets that are related to a microalga (or even its plastids and mitochondria) in a whole metagenome, then extract and assemble them (West et al. 2018; Karlicki et al. 2021). Completeness and contamination are two main characteristics of MAGs, which are estimated by single-copy marker gene analysis (SCMG). For prokaryotic MAGs, the CheckM algorithm is widely used and shows good performance, while quality check of eukaryotic MAGs is a challenge, it is however reached by using a defined set of eukaryotic SCMG (BUSCO and CEGMA algorithms) or dynamic selection of an appropriate SCMG set for improved evaluation e.g., with EukCC algorithm (Saary et al. 2020).

Though the short-read WMS inherently provides excessive metagenome coverage, its results are still limited by the read length. Confident assignment of the metagenomic reads to a specific taxon by comparison with known DNA barcodes or reference genomes requires longer sequences than obtainable with currently available NGS platforms (Table 1). The robustness of genus or species identification within the WMS data can be improved either by assembly of short reads or by application of longer reads (Tran and Phan 2020; Pessi et al. 2023). In some cases, workable DNA-barcode loci can be difficult to assemble from short reads due to their highly conserved sequences making the taxonomical assignment of MAGs challenging. As an example of such case, Pessi et al. (2023) reported that among 37 MAGs, obtained from 17 cyanobacterial mats from polar regions, only one included a complete sequence of 16S rRNA gene, therefore it was impossible to map most of the MAGs to a 16S rRNA sequences database. Since the step of assembly

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

is not required for processing of the output of long-read sequencing by ONT, it can be directly used for easy on-site taxonomical classification. The efficiency of this approach is additionally boosted by developing frameworks for rapid classification, like System for Mobile Analysis in Real-Time of Environment (SMARTEn), which is implemented in Coriolis – a mobile metagenomic classification tool (Mikalsen and Zola 2023).

WMS allows investigation of the microbial species in different natural and artificially created biotopes, from natural habitats to laboratory and industrial cultures. One of the most valuable outputs of WMS of natural communities is the information about the genetic diversity of microalgae and evaluation of their physiological potential. This direction is highly contributed by large international projects aiming at collecting metagenomic samples from wide geographical area covering a lot of diverse habitats. These are represented by Tara Oceans Expedition, Microbial Atlas, etc. which have produced a large amount of data for metagenomic mining (Delmont et al. 2022). More advanced sample collection techniques, like targeting the layers of water column with the maximum chlorophyll *a* concentration or filtering the cells by their size, help to narrow the microbial diversity of a sample and thus further improve metagenomic algorithms output (Yergeau et al. 2017; Delmont et al. 2022; Duncan et al. 2022). This enables study the genetic variability of a particular microalgal species, such as the chlorophyte *Bathycoccus prasinos*—a dominating member of marine eukaryotic picoplankton.

On the practical side, functional analysis of the MAGs showed amino acids content shift among polar populations of microalgae, which explains adaptation to the changes in temperatures (Duncan et al. 2022). Studying the functional landscape of eukaryotic and prokaryotic MAGs in picoplankton also allows prediction of microbiome succession, including such crucial events such as microalgal blooms (Kavagutti et al. 2023). The same approach can be used for revealing the functional potential of microalgal species discovered within metagenomes for the destruction of hazardous micropollutants by search for the relevant metabolic pathways. Examples include plastic biodegradation by adhesion on cell surface with following enzymatic hydrolysis; this process is extensively studied with the focus on the enzymes polyethylene hydrolase (PETase) and mono(2-hydroxyethyl) terephthalic acid hydrolase (MHETase) (Chia et al. 2020). Other examples include heavy metal phycoremediation by their uptake by and enzymatic reduction (e.g., by chromium reductase) in the microalgal cells (Priya et al. 2022), and xenobiotics degradation (Cheng et al. 2021; Ovis-Sánchez et al. 2023; Vasilieva et al. 2023) e.g., by nitrilase (Vingiani et al. 2019).

Though known sets of genes in metagenome can be detected by targeted PCR-analysis with degenerative primers (Gulvik et al. 2012), the results of this approach might be compromised. One of the reasons is functional redundancy—presence of alternative pathways of similar function in the community (Graham et al. 2015), another one is the functional divergence of orthologs within a species (Ma et al. 2021). Therefore, WMS remains a powerful approach for estimating the efficiency and stability of microalgal communities under particular conditions as well as for bioprospecting of promising strains from e.g., wastewater stabilization/oxidation ponds or other polluted areas (Chia et al. 2020; Jankowski et al. 2022; Nagarajan et al. 2022). The investigation of the genetic variation landscape for microalgal and cyanobacterial species is a promising way to mine new homologs of biotechnologically valuable enzymes or alterations in biosynthetic pathways. A pangenomic analysis of databases-retrieved *Nannochloropsis* species genomes revealed length and sequence variations between photosystems I and II genes (*psaB*, *J*, *L*, and *psbH*, *Y*, *N*, *I*, *T*), energy conservation genes (*atpH*, *G*, *E*), as well as loss of the acetohydroxyacid synthase negative feedback regulation gene (*ilvH*) in branched chain amino acids pathway, that indicated its alternative regulation (Starkenburg et al. 2014).

Pangenomic studies have demonstrate that transposable elements are as important for the phenotype of algae as single nucleotide polymorphism (SNP), indicating the importance of sequencing method precision (Carrier et al. 2024). While deep shotgun NGS sequencing provides good nucleotide resolution, the accuracy of the assembly can be greatly enhanced by joint application of the genome-wide chromosome conformation capture (Hi-C) method with nanopore sequencing of long reads (Pan et al. 2023; Carrier et al. 2024). Being originally developed for chromatin-DNA interaction studies within a given eukaryotic species, Hi-C showed great potential in reconstruction of high-quality MAGs from microbial communities (called metaHi-C), as the capturing technique artificially gathers DNA molecules within each organism and thus improves metagenomic binning procedure (Beitel et al. 2014). The combination of short-read NGS, long-read nanopore sequencing and metaHi-C opens up the opportunity for pangenomic analysis within a certain microalga-bacterial community and the development of metapangenomic approach (Delmont and Eren 2018).

Getting insights into interactions within microalgal-bacterial communities with HTS

The WMS is also a powerful tool for studying the interactions within microalga-bacteria communities. On one hand, it relies on search for a specific set of genes encoding pathways for molecular signaling and/or trophic substrate exchange. Trophic relationships can be revealed starting already from elemental metabolism level, by classification

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

of the relevant genes found in the MAGs as related, e.g., to phosphorus, nitrogen, or sulfur fluxes between microalgae and bacteria in a community (Saini et al. 2023). The keen attention to this kind of studies is due to importance of microalga-bacterial consortia for nutrient biosequestration from wastewater ponds, marine sediments, and biofertilizer-treated soils (Vučić and Müller 2021), where both sides can affect phosphorus accessibility for each other by enzymatic solubilization by bacteria (Dong et al. 2022) or pH modulation by the microalgae. The balance in flux between carbon, oxygen and nitrogen is crucial for the aerobic enhanced biological phosphorus removal (EBPR) process in microalga-bacteria biofilms during wastewater treatment (Mohamed et al. 2021).

Another well-known mode of interaction between microalgae and bacteria is syntrophy, where bacterial organisms produce the vitamins biotin, cobalamin and thiamin, for which most of microalgae are auxotrophic and require them for e.g., lipid biosynthesis (Wirth et al. 2020). Comparison of metabolic potential and substrate spectrum can also reveal the spatial interaction within the cyanosphere (cyanobacterial analogue of phycosphere), where filamentous cyanobacteria (representatives of *Lyngbya*, *Planktothricoides*, *Pseudochroococcus* and other genera) are able to build extracellular polymeric substance (EPS) of polysaccharide mucilage, which is then inhabited by heterotrophic bacteria capable of its partial degradation and utilization in catabolic reactions (Halary et al. 2022). Besides that, a more specific interaction way exists in a form of signaling molecules exchange within such consortia: phytohormones are produced by bacteria with either stimulating or suppressing mode for microalgae (for example most known L-amino oxidase manages conversion of L-tryptophan to indole-3-acetic acid) (Wang et al. 2021; Mars Brisbin et al. 2022), algicides that cause microalgal cell damage (Jia et al. 2023), and other quorum sensing agents with wide spectrum of impacts on photosynthetic cells (Dow 2021). One can do WMS data mining not only for the biosynthetic pathways for these operating molecules, but also for the related molecular transporters, like ABC-transporters (Krohn-Molt et al. 2017; Li et al. 2022). While solid evidence of chemical interaction between microalgae and bacteria usually requires integration with other omics (ideally proteomics and metabolomics methods), WMS provides firm background for genome-centric approach in such studies (Krohn-Molt et al. 2017).

An interesting and promising approach to investigate microalgal-bacterial interactions is one based on hologenome concept. The phycosphere can be considered as a classical holobiont—metaorganism, where certain bacteria persist and co-evolve with microalgae acting as the ecosystem engineer (edificator). That co-evolution might be revealed by comparative genomics through searching for phylosymbiotic signals (correlation in divergence) in phylogeny of both host and symbiont, codivergence of dominant microbiome groups with a host, and metabolic complementary (Cooke et al. 2019). The phycosphere is known to be highly dynamic system responding to biotic and abiotic factors and featuring the hologenome evolution mechanisms: amplification or reduction of bacterial partners, acquisition of new bacteria, and horizontal gene transfer (HGT) (Rosenberg and Zilber-Rosenberg 2018). Though HGT between eukaryotic and prokaryotic species faces many obstacles based on difference in genome structure and mechanisms, it has been shown that the gene flow from bacteria to microalga does exist (Li et al. 2023a). It is most evident for the antibiotic resistance genes (ARG) transfer in environments with high evolutionary pressure, such as anthropogenically polluted sites, making it resonable to propose a concept of 'PollutantBiome' as a special case of hologenome (Ashraf et al. 2023; Li et al. 2023a).

Investigation of the hologenome structure via comparative genomics requires low contamination values of MAGs, since presence of heterogenous reads in the final sequence leads to severe misinterpretation. Thus, long-read nanopore sequencing with the following polishing by NGS short reads is the best technique for revealing the status quo for holobiont and symbiont, as nanopore-produced long contigs reduce the probability of interspecies read contamination, while short reads increase consensus accuracy and enable analysis of SNP variants (Sauvage et al. 2019). In addition, long reads can be efficiently sorted not only by species of origin, but also by assignment to specific compartments within cells. The heteroplasmy and genetic variation of organellar genomes (nuclear, plastid, mitochondrial) of cellular endosymbionts can provide proof of gene transfer and metabolic complementarity between the microalgae holobiont and the symbionts (Sauvage et al. 2019).

Hologenome studies can be greatly enhanced by nanopore long-reads supported metaHi-C approach and opens new horizons for HGT studies, by making it possible to capture DNA–DNA interaction between host genome and mobile genetic elements (plasmids, viral loci, etc.) (Bickhart et al. 2022). The recently developed MetaCC algorithm has been shown to be a powerful tool for MAG reconstruction and plasmids search in complex microbial communities hybrid assembly of long and short reads (Du and Sun 2023). However, the holistic approach for investigation of microalga-bacteria communities currently remains underrepresented and still needs to be developed and critically reviewed.

Advantages of long-read HTS for taxonomical profiling of microalgal-bacterial communities

Opposite to the WMS, taxonomic profiling of microbial communities is based on amplicon sequencing of genetic barcodes, specifically determining taxonomical assignment

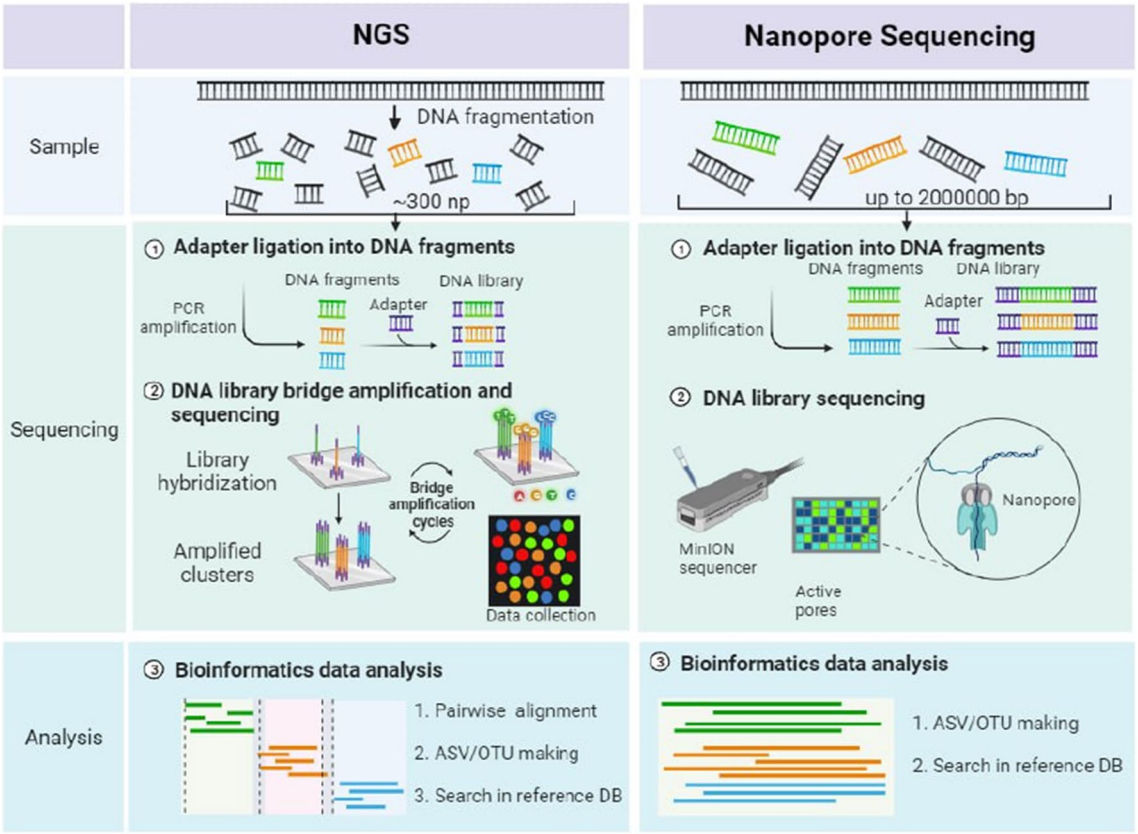
Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

of microorganisms. The variety of metabarcoding methods mainly depend on loci that are used for each particular group of organisms, with the main criteria of conservativity within the taxon and variability between taxa. Thus, the ribosomal operon is widely used for bacteria identification, since 16S and 23S rRNA genes, combined with ITS provides strain-level resolution. Recently Pushpakumara et al. (2023) have demonstrated the high potential of the 16S rRNA gene metabarcoding for analysis of microalgal-bacterial communities revealing previously unknown associations between microorganisms. The identification of eukaryotic microalgae usually requires other genetic barcodes, such as 18S rRNA gene, its ITS regions, or more specific *rbcL* and *tufA*. Metabarcoding based on functional *rbcL* and *tufa* genes has several advantages over ribosomal loci, which are increased richness of a studied communities, and identification of haplotypes presence and microevolution via population genetic approach (Sauvage et al. 2016; Turk Dermastia et al. 2023). The second becomes available due to high resolution of identification provided by such barcodes, though it requires accurately considering possible errors and correction strategies. 16S and 23S rRNA genes are also applied for microalgae identification as plastid and mitochondrial ribosomal loci, which can be applied simultaneously to identify both components of microalgal-bacterial communities (Kezlya et al. 2023). At the same time, the presence of the plastid or mitochondrial ribosomal loci reads reduces community sample richness and affects diversity index estimation, and therefore is considered as unwelcomed contamination (Thomas et al. 2020). Both experimental techniques, such as physical removal of eukaryotic DNA (Demkina et al. 2023) and optimization of bacteria-specific primers for ribosomal operons, have been evaluated recently to obtain pure prokaryotic profiles (Thomas et al. 2020) as well as training bioinformatic classifiers on chloroplast-derived datasets, such as QIIME2 naïve Bayes tool trained on PhytoREF database (Bonfantine et al. 2021).

Until recently, the DNA metabarcoding method was firmly based on short-read sequencing on the NGS platforms. Though widely spread and routine, it possesses severe drawbacks for studying the microbial communities of microalgae cultures and natural samples. The main and crucial drawback is that short read length limits taxonomical resolution. While the general rule states that ribosome small subunit rRNA gene and its ITS is required for strain identification, the NGS platforms of sequence-by-synthesis method has a limitation of maximum 300–500 b.p. (in case of pyrosequencing) and 150–300 b.p. (in case of Illumina), which allows reading of only part of the barcode. The V3-V4 regions of 16S rRNA gene is the most popular variant for microbiome profiling, though other regions, such as V2-V3 are shown to be more specific and provide higher taxa resolution (Bukin et al. 2019). Even then, the drawback lies in the interplay between resolution and richness of the community, as the increased specificity leads to the loss of particular groups of organisms. The rapid recent development of long read TGS technology enables full length barcode reading and thus removes the taxonomical resolution issues (Fig. 3) (Kerkhof et al. 2017; Portik et al. 2022). However, a one should carefully consider choice of sequencing platform for such purpose. Despite obvious advantages of long over short reads for barcode sequencing, either throughput or accuracy of sequencing itself can suffer in such race, which affects taxa identification. While PacBio can provide very accurate results at a low throughput, Oxford Nanopore products have increased throughput (especially with PromethION) but it is notorious for low accuracy of basecalling. Comparison of simulation results for different platforms showed that 50% exceed of sequencing launch capacity for Illumina over Oxford Nanopore can provide maximum accuracy of read classification and taxa identification (Pearman et al. 2020). Currently, there are many research directions of how to improve the accuracy of nanopore sequencing basecalling: by improving the technology itself through cross membrane voltage varying, by implementing other amplification strategies (such as The Rolling Circle Amplification to Concatemeric Consensus (R2C2) method), or by training basecaller models on specific datasets (Volden et al. 2018; Noakes et al. 2019; Ferguson et al. 2022). The last can be performed on species-specific datasets to improve minor taxa identification in environmental samples (Ciuffreda et al. 2021). It should be mentioned, that PacBio is considered as a useful and robust sequencing method for metabarcoding of relatively species-poor communities while targeting large regions of SSU (around 2500–3000 b.p.) of microeukaryotes (Tedersoo et al. 2018; Gueidan et al. 2019).

Both experimental data and bioinformatics simulations prove that long read barcode sequences also contribute to greater richness of a studied community (Jamy et al. 2020; Lemoinne et al. 2023). Nanopore sequencing showed high potential of finding up to twice more hidden species compared to Illumina short read (Huggins et al. 2022; Lemoinne et al. 2023; Szoboszlay et al. 2023). This was shown to be especially useful for marine ecosystems, which usually possesses high richness, such as marine biofilms (Wang et al. 2022). Long read taxonomic profiling research on algal-bacterial communities of *Ulva* species has shown the decrease of microbiome richness but increase of relative abundance of MGPB *Sulfitobacter* and *Roseobacter* when passing from marine environmental samples to laboratory cultures (van der Loos et al. 2021). Nanopore sequencing has been demonstrated as a useful tool for investigation of the interactions within microalgal natural communities, such as harmful blooms of dinoflagellates. Sequencing of long ribosomal genes cluster cassette more than 3 kb long harboring 18S, ITS and partial 28S rRNA genes enabled identification of

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024



AQ2 **Fig. 3** Comparison of short read NGS and long read nanopore sequencing in application for taxonomic profiling of microbial communities. The genetic barcodes molecules from different species colored in green, orange, and blue. Nanopore sequencing technology enables reading whole unfragmented loci of genetic barcodes, also with only one PCR procedure during library preparation, thus contributing to lower amplification bias

a nearly complete list of species, including the toxic microalgae *Alexandrium*, *Gonyaulax*, *Prorocentrum,* and *Lingulodinium* (Hatfield et al. 2020). Studying the prokaryotic components of natural dinoflagellate communities by nanopore sequencing revealed associations between particular microalgal species and bacteria clades, such as *Alexandrium tamarense* and *Roseobacter* bacteria (Shin et al. 2018). The research authors propose that growth of *A. tamarense* can be promoted by sulfonate, which is produced by *Roseovarius* genus bacteria with Sox multienzyme complex (Shin et al. 2018).

Another issue to be kept in sight for metabarcoding is quantitative bias as a result of uneven amplification occurring for different barcode sequences (Pawluczyk et al. 2015). Though targeting amplicons with conserved priming sites or application of degenerate primers slightly improves in that situation, they still cannot overcome another bias coming from various gene copy number in genomes (Krehenwinkel et al. 2017). In case of nanopore sequencing, lack of a DNA synthesis step during the sequencing step improves amplification bias for species abundance but does not remove it completely (Fig. 3) (Huggins et al. 2022). Application of optimized primers set for target barcode amplification can drastically improves PCR bias, as well as new possible selection and amplification strategies to create barcode libraries (Matsuo et al. 2021).

Despite difficulties faced by the research community in application of long reads in metabarcoding method, nanopore sequencing is shown to be an extremely useful tool for quick (within 24 hours), cost-efficient and research-friendly technology for taxonomical identification in microalgal communities and in revealing microalgal-bacterial interactions (van der Loos et al. 2021). This is highly supported by development of new bioinformatic analysis pipelines enabling real-time identification and richness analysis of nanopore sequenced 16S rRNA gene long reads – such as the NanoR-Tax pipeline (Rodríguez-Pérez et al. 2022).

Augmenting functional annotation of microalgal communities with advantageous HTS

The biology of microalgal-bacterial consortia has a severe lack of understanding of the functional genetic landscape underlying interactions between these organisms. Even

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

though an emerging trend towards microalgal metagenomics enriches us with MAGs and other genomic information, we are far from its complete functional annotation and thus prediction of a role of a particular organism in a community. Classical workaround is complementing the genomic data with transcriptome—the approach successfully tested for microbial communities, including those sampled from the environment (Wang et al. 2020). This can be implemented within integrative omics pipelines and algorithms (like Galaxy) to create fully annotated metabolic networks of a particular MAG from a community (Schiml et al. 2023). Integration of metagenomics with metatranscriptomics (and full way down to other omics methods) enables investigation of complex interplay between abiotic factors (illumination, biogenic elements, etc.) and microalgal response in aquatic biomes, as well as microbial interactions within microalgal biofilms (Krohn-Molt et al. 2017; Trench-Fiol and Fink 2020). Recent advances in nanopore sequencing of both RT-PCR amplicons and direct RNA opened a way for unbiased and full-length transcripts reading for complex environmental communities, such as soil (Salzberg 2019; Poursalavati et al. 2023). Although this approach requires particular caution when handling RNA from samples of complex chemical mixtures and thus is hardly feasible in the field, it holds promise for simultaneous taxonomical identification and functional profiling of microbial communities with defined pipelines (Poursalavati et al. 2023). By accumulating a sufficient amount of accurate and complete metatranscriptomic data from known conditions the further reverse predictions of functional profile of a community can be made from similar environmental contexts and taxonomic profile only (Krinos et al. 2023).

The golden dream of microalgal communities' researchers is an implementation of prediction algorithms based on taxonomical profile data to reveal functional potential of a community. Among the most popular are PICRUST(2) and Tax4Fun(2) whose main principle is comparison of OTU/ASV against the reference databases consisted of assembled metagenomes with functional annotation (Liu et al. 2020). Though highly reference-depended, not taking into account the true physiological state of the cell as well as genome context, these algorithms were welcomed in studies of species interactions within a microalgal consortium including searches for potential N and/or P recovery bacteria for soil health mitigation or waste treatment (Zarezadeh et al. 2019). Besides the trophic interactions, this approach can reveal signaling cross-talk between algicidal bacteria species and microphototrophs (Le et al. 2022). Though not yet adjusted for these algorithms, the long read metabarcoding data produced by nanopore sequencing can dramatically improve the accuracy of such functional prediction, as species or strain-level information narrows down the functional landscape even within one taxon.

Increasing accuracy of species identification together with capability of capturing the community richness can greatly contribute development of Microbial Genome-Wide Association Studies (mGWAS) – an approach aiming for detection of genetic variants and genes responsible for specific phenotypic features (Power et al. 2017; San et al. 2020). Nanopore sequencing can provide post-GWAS fine-mapping of determined candidate loci for their further investigation and application (Magdy et al. 2020). For microalgal-bacterial consortia studies these can be genes encoding antimicrobial or algicidal agents, growth-promoting factors, phytohormones, or members of biogenic element conversion cycles. At the end of the day, such "environmental GWAS" ("eGWAS") can serve the great deal for microalga biotechnology by highlighting those genetic variants (strains) that might be useful for target process as a part of bioengineered consortia.

Concluding remarks and outlook

Admittedly, short-read NGS/SGS is now a mainstream platform for sequencing of genomes and transcriptomes as well as for providing support to other "omics" studies in microalgae. As such, the short-read sequencing has provided a plethora of invaluable insights into different aspects of microalgal biology, also crucial for microalgal biotechnology applications. Now we see that long-read sequencing platforms, especially nanopore-based sequencing technology, confidently enters the stage of algal research. This is especially true for metagenomic studies of microbial communities harboring microalgae as the edificator and other microorganisms contributing to the robustness, productivity, and biotechnological versatility of the whole community.

At the current level of sequencing technology evolution, both metagenomic strategies can by implemented with either short-read NGS or long-read nanopore sequencing. Still, it becomes increasingly obvious that the latter has distinct advantages that warrant its increasing application in this field (although the most fruitful approach is that employing both platforms). The number of publications dedicated to microalgal communities studied with well-established short read sequencing exponentially increased over last 15 years, as the number of the papers on microalgal genomics (Fig. 4).

The most promising directions of the metagenomic studies of microalgae include: 1) ecological monitoring of harmful microalgal blooms that cause economical and health treats to human activities; 2) mining of microalgal and/or associated bacterial strains for bioprospecting of biosynthetic pathways of valuable molecules (carotenoids, fatty acids, bioactive compounds); 3) strategies development for rational design of microalga-bacteria consortia for wastewater treatment, micropollutants biodegradation and enhanced bioproduct

Journal : Large 10811 | Article No : 3267 | Pages : 19 | MS Code : 3267 | Dispatch : 6-5-2024

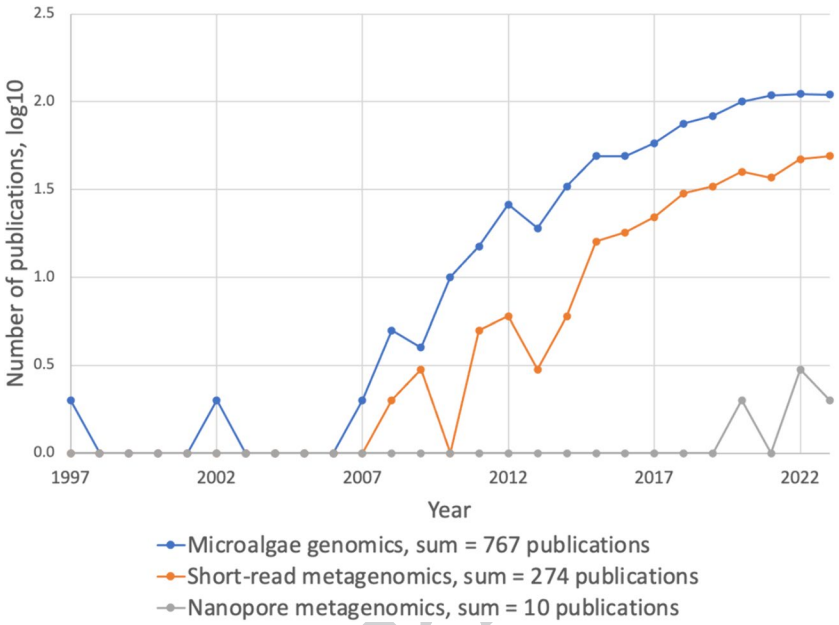


Fig. 4 Dynamics of microalga-bacteria metagenomics-related publications. Calculations were based on PubMed database but excluded reviews and editorial notes. For the specific query terms see the Online Supplementary materials

production. None of these is reachable without the information about taxonomical structure and functional potential of communities, which can be easily obtained from HTS data, especially with rapid development of nanopore sequencing.

Systematic reports on nanopore-based studies of microalga metagenomes have started to emerge only recently, so one can anticipate a boom in this field in the next few years. To keep up with this trend, one should realize the tremendous potential of the long-read sequencing technologies for studies of the biology of microalgae. Therefore, it is important to highlight the benefits of the long-read sequencing for revealing taxonomic structure, genetic diversity, and functional potential of microalgae-based communities for biotechnological applications. We hope that the present review makes a good step in this direction.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10811-024-03267-1.

Author's contributions P.Z., M.Z. and A.S. wrote the main manuscript text, A.Z. and V.R. prepared the figures. All authors reviewed the manuscript.

Funding This work was supported by the project of the Interdisciplinary Scientific and Educational School of Moscow State University "Molecular technologies of living systems and synthetic biology" (#23-Sh04-45).

Availability of data and materials Not applicable.

Declarations

Competing interests A.S. is an Associate Editor for Journal of Applied Phycology and the peer-review process for this article was independently handled by another member of the journal editorial board.

Ethical Approval Not applicable.

Conflict of Interest A.S. is an Associate Editor for Journal of Applied Phycology and the peer-review process for this article was independently handled by another member of the journal editorial board.

References

Arutselvan C, Narchonai G, Pugazhendhi A, LewisOscar F, Thajuddin N (2021) Evaluation of microalgal strains and microalgal consortium for higher lipid productivity and rich fatty acid profile towards sustainable biodiesel production. Bioresour Technol 339:125524

Ashraf N, Ahmad F, Lu Y (2023) Synergy between microalgae and microbiome in polluted waters. Trends Microbiol 31:9–21

Ashton PM, Nair S, Dallman T, Rubino S, Rabsch W, Mwaigwisya S, WainO'grady JJ (2015) MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. Nat Biotechnol 33:296–300

Astafyeva Y, Gurschke M, Qi M, Bergmann L, Indenbirken D, de Grahl I, Katzowitsch E, Reumann S, Hanelt D, Alawi M, Streit WR (2022) Microalgae and bacteria interaction—evidence for division of diligence in the alga microbiota. Microbiol Spectr 10:e0063322

Athanasopoulou K, Boti MA, Adamopoulos PG, Skourou PC, Scorilas A (2022) Third-generation sequencing: the spearhead towards the radical transformation of modern genomics. Life 12:30

Badger JH, Giordano R, Zimin A, Wappel R, Eskipehlivan SM, Muller S, Donthu R, Soto-Adames F, Vieira P, Zasada I, Goodwin S (2023) Direct sequencing of insect symbionts via nanopore adaptive sampling. Curr Opin Insect Sci 61:101135

Ballesteros I, Terán P, Guamán-Burneo C, González N, Cruz A, Castillejo P (2021) DNA barcoding approach to characterize microalgae isolated from freshwater systems in Ecuador. Neotrop Biodiversity 7:170–183

Beitel CW, Froenicke L, Lang JM, Korf IF, Michelmore RW, Eisen JA, Darling AE (2014) Strain-and plasmid-level deconvolution of a

synthetic metagenome by sequencing proximity ligation products. PeerJ 2:e415
Belcour A, Frioux C, Aite M, Bretaudeau A, Hildebrand F, Siegel A (2020) Metage2Metabo, microbiota-scale metabolic complementarity for the identification of key species. eLife 9:e61968
Berthold DE, Shetty KG, Jayachandran K, Laughinghouse HD, Gantar M (2019) Enhancing algal biomass and lipid production through bacterial co-culture. Biomass Bioenergy 122:280–289
Bickhart DM, Kolmogorov M, Tseng E, Portik DM, Korobeynikov A, Tolstoganov I, Uritskiy G, Liachko I, Sullivan ST, Shin SB, Zorea A (2022) Generating lineage-resolved, complete metagenome-assembled genomes from complex microbial communities. Nature Biotech 40:711–719
Blanco-Míguez A, Beghini F, Cumbo F, McIver LJ, Thompson KN, Zolfo M, Manghi P, Dubois L, Huang KD, Thomas AM, Nickols WA (2023) Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nature Biotech 41:1633–1644
Bonfantine KL, Trevathan-Tackett SM, Matthews TG, Neckovic A, Gan HM (2021) Dumpster diving for diatom plastid 16S rRNA genes. PeerJ 9:e11576
Bukin YS, Galachyants YP, Morozov IV, Bukin SV, Zakharenko AS, Zemskaya TI (2019) The effect of 16S rRNA region choice on bacterial community metabarcoding results. Sci Data 6:190007
Cao Y, Wang Y, Zheng X, Li F, Bo X (2016) RevEcoR: an R package for the reverse ecology analysis of microbiomes. BMC Bioinformat 17:294
Carrier G, Berthelier J, Maupetit A, Nicolau E, Marbouty M, Schreiber N, Charrier A, Carcopino C, Leroi L, Saint-Jean B (2024) Genetic and phenotypic intra-species diversity of alga *Tisochrysis lutea* reveals original genetic structure and domestication potential. Eur J Phycol 59:94–111
Carrillo-Reyes J, Barragán-Trinidad M, Buitrón G (2016) Biological pretreatments of microalgal biomass for gaseous biofuel production and the potential use of rumen microorganisms: a review. Algal Res 18:341–351
Cheirsilp B, Kitcha S, Torpee S (2012) Co-culture of an oleaginous yeast *Rhodotorula glutinis* and a microalga *Chlorella vulgaris* for biomass and lipid production using pure and crude glycerol as a sole carbon source. Ann Microbiol 62:987–993
Cheng X, Xu J, Smith G, Nirmalakhandan N, Zhang Y (2021) Metagenomic profiling of antibiotic resistance and virulence removal: activated sludge vs. algal wastewater treatment system. J Env Manage 295:113129
Chia WY, Ying Tang DY, Khoo KS, Kay Lup AN, Chew KW (2020) Nature's fight against plastic pollution: algae for plastic biodegradation and bioplastics production. Environ Sci and Ecotechnol 4:100065
Choi SS, Seo YB, Nam S-W, Kim G-D (2021) Enhanced production of astaxanthin by co-culture of *Paracoccus haeundaensis* and lactic acid bacteria. Front Mar Sci 7:
Ciuffreda L, Rodríguez-Pérez H, Flores C (2021) Nanopore sequencing and its application to the study of microbial communities. Comp Struct Biotech J 19:1497–1511
Clagnan E, Dell'Orto M, Štěrbová K, Grivalský T, Manoel JA, Masojídek J, D'Imporzano G, Acién-Fernández FG, Adani F (2023) Impact of photobioreactor design on microalgae-bacteria communities grown on wastewater: Differences between thin-layer cascade and thin-layer raceway ponds. Bioresour Technol 374:128781
Cooke I, Mead O, Whalen C, Boote C, Moya A, Ying H, Robbins S, Strugnell JM, Darling A, Miller D, Voolstra CR (2019) Molecular techniques and their limitations shape our view of the holobiont. Zoology 137:125695
Danish-Daniel M, Noor MEM, Chui-Pin L, Po-Teen L, Tan MP, Sung YY, Usup G (2023) 16S-based metagenomics analysis of microbial communities associated with toxic and non-toxic dinoflagellate cultures. In: AIP Conference Proceedings 2738
de Bashan LE, Hernandez J-P, Morey T, Bashan Y (2004) Microalgae growth-promoting bacteria as "helpers" for microalgae: a novel approach for removing ammonium and phosphorus from municipal wastewater. Water Res 38:466–474
de Bashan LE, Magallon-Servin P, Lopez BR, Nannipieri P (2021) Biological activities affect the dynamic of P in dryland soils. Biol Fertil Soils 58:105–119
Delmont TO, Eren AM (2018) Linking pangenomes and metagenomes: the *Prochlorococcus* metapangenome. PeerJ 6:e4320
Delmont TO, Gaia M, Hinsinger DD, Frémont P, Vanni C, Fernandez-Guerra A, Eren AM, Kourlaiev A, d'Agata L, Clayssen Q, Villar E (2022) Functional repertoire convergence of distantly related eukaryotic plankton lineages abundant in the sunlit ocean. Cell Genomics 2:100123
Demkina A, Slonova D, Mamontov V, Konovalova O, Yurikova D, Rogozhin V, Belova V, Korostin D, Sutormin D, Severinov K, Isaev A (2023) Benchmarking DNA isolation methods for marine metagenomics. Sci Rep 13:22138
Dong H, Liu W, Zhang H, Zheng X, Duan H, Zhou L, Xu T, Ruan R (2022) Improvement of phosphate solubilizing bacteria *Paenibacillus xylanexedens* on the growth of *Chlorella pyrenoidosa* and wastewater treatment in attached cultivation. Chemosphere 306:135604
Dow L (2021) How do quorum-sensing signals mediate algae-bacteria interactions?. Microorganisms 9:
Du Y, Sun F (2023) MetaCC allows scalable and integrative analyses of both long-read and short-read metagenomic Hi-C data. Nature Commun 14:6231
Duncan A, Barry K, Daum C, Eloe-Fadrosh E, Roux S, Schmidt K, Tringe SG, Valentin KU, Varghese N, Salamov A, Grigoriev IV (2022) Metagenome-assembled genomes of phytoplankton microbiomes from the Arctic and Atlantic Oceans. Microbiome 10:67
Edwards A, Debbonaire AR, Sattler B, Mur L, Hodson AJ (2016) Extreme metagenomics using nanopore DNA sequencing: a field report from Svalbard, 78 N. BioRxiv 10:073965
Engelhard CA, Khani S, Derdak S, Bilban M, Kornfeld JW (2023) Nanopore sequencing unveils the complexity of the cold-activated murine brown adipose tissue transcriptome. iScience 26:107190
Ferguson S, McLay T, Andrew RL, Bruhl JJ, Schwessinger B, Borevitz J, Jones A (2022) Species-specific basecallers improve actual accuracy of nanopore sequencing in plants. Plant Methods 18:137
Gao B, Xu M, Shan D, Zhang C, Yang Y, Dong Z, Zhang H, Han B, Huang L, Zhang C (2023a) The genomes of *Vischeria oleaginous* microalgae shed light on the molecular basis of hyperaccumulation of lipids. BMC Biol 21:133
Gao H, Wang H, Zhang Y, Wang Y, Liu G, Zhao Q, Yu Z, Xin F, Zhang W (2023b) Design and optimization of artificial light-driven microbial consortia for the sustainable growth and biosynthesis of 2-phenylethanol. Chem Eng J 466:143050
Gonzalez LE, Bashan Y (2000) Increased growth of the microalga *Chlorella vulgaris* when coimmobilized and cocultured in alginate beads with the plant-growth-promoting bacterium *Azospirillum brasilense*. Appl Environ Microbiol 66:1527–1531
González-González LM, de Bashan LE (2021) Toward the enhancement of microalgal metabolite production through microalgae–bacteria consortia. Biology 10:282
Gonzalez-Gonzalez LM, de-Bashan LE (2023) The Potential of microalgae–bacteria consortia to restore degraded soils. Biology 12:
Graham LE, Wilcox LW, Knack JJ (2015) Why we need more algal metagenomes. J Phycol 51:1029–1036

AQ3

Gueidan C, Elix JA, McCarthy PM, Roux C, Mallen-Cooper M, Kantvilas G (2019) PacBio amplicon sequencing for metabarcoding of mixed DNA samples from lichen herbarium specimens. MycoKeys 53:73–91

Gulvik CA, Effler TC, Wilhelm SW, Buchan A (2012) De-MetaST-BLAST: a tool for the validation of degenerate primer sets and data mining of publicly available metagenomes. PLoS One 7:e50362

Hadi SI, Santana H, Brunale PP, Gomes TG, Oliveira MD, Matthiensen A, Oliveira ME, Silva FC, Brasil BS (2016) DNA barcoding green microalgae isolated from neotropical inland waters. PLoS One 11:e0149284

Halary S, Duperron S, Demay J, Duval C, Hamlaoui S, Piquet B, Reinhardt A, Bernard C, Marie B (2022) Metagenome-based exploration of bacterial communities associated with cyanobacteria strains isolated from thermal muds. Microorganisms 10:2337

Hatfield RG, Batista F, Bean T, Fonseca V, Santos A, Turner A, Lewis A, Dean K, Martinez-Urtaza J (2020) The application of nanopore sequencing technology to the study of dinoflagellates: a proof of concept study for rapid sequence-based discrimination of potentially harmful algae. Front Microbiol 11:00844

Huggins LG, Colella V, Atapattu U, Koehler AV, Traub RJ (2022) Nanopore sequencing using the full-length 16s rrna gene for detection of blood-borne bacteria in dogs reveals a novel species of hemotropic mycoplasma. Microbiol Spectr 10:e03088-03022

Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data. Genome Res 17:377–386

Inskeep WP, Jay ZJ, Herrgard MJ, Kozubal MA, Rusch DB, Tringe SG, Macur RE, Jennings RD, Boyd ES, Spear JR, Roberto FF. (2013) Phylogenetic and functional analysis of metagenome sequence from high-temperature archaeal habitats demonstrate linkages between metabolic potential and geochemistry. Front Microbiol 4:

Iqbal K, Saxena A, Pande P, Tiwari A, Chandra Joshi N, Varma A, Mishra A (2022) Microalgae-bacterial granular consortium: striding towards sustainable production of biohydrogen coupled with wastewater treatment. Bioresour Technol 354:127203

Jain M, Koren S, Miga KH, Quick J, Rand AC, Sasani TA, Tyson JR, Beggs AD, Dilthey AT, Fiddes IT, Malla S (2018) Nanopore sequencing and assembly of a human genome with ultra-long reads. Nature Biotech 36:338–345

Jamy M, Foster R, Barbera P, Czech L, Kozlov A, Stamatakis A, Bending G, Hilton S, Bass D, Burki F (2020) Long-read metabarcoding of the eukaryotic rDNA operon to phylogenetically and taxonomically resolve environmental diversity. Mol Ecol Resour 20:429–443

Jankowski P, Gan J, Le T, McKennitt M, Garcia A, Yanaç K, Yuan Q, Uyaguari-Diaz M (2022) Metagenomic community composition and resistome analysis in a full-scale cold climate wastewater treatment plant. Env Microbiome 17:3

Jia Y, Lu J, Wang M, Qin W, Chen B, Xu H, Ma Z (2023) Algicidal bacteria in phycosphere regulate free-living *Symbiodinium* fate via triggering oxidative stress and photosynthetic system damage. Ecotoxicol Env Safety 263:115369

Kainth AS, Haddad GA, Hall JM, Ruthenburg AJ (2023) Merging short and stranded long reads improves transcript assembly. PLoS Comput Biol 19:e1011576

Kang Y, Kim M, Shim C, Bae S, Jang S (2021) Potential of algae–bacteria synergistic effects on vegetable production. Front Plant Sci 12:

Karlicki M, Antonowicz S, Karnkowska A (2021) Tiara: deep learning-based classification system for eukaryotic sequences. Bioinformatics 38:344–350

Kasianowicz JJ, Brandin E, Branton D, Deamer DW (1996) Characterization of individual polynucleotide molecules using a membrane channel. Proc Nat Acad Sci 93:13770–13773

Kavagutti VS, Bulzu PA, Chiriac CM, Salcher MM, Mukherjee I, Shabarova T, Grujčić V, Mehrshad M, Kasalický V, Andrei AS, Jezberová J (2023) High-resolution metagenomic reconstruction of the freshwater spring bloom. Microbiome 11:15

Kerkhof LJ, Dillon KP, Häggblom MM, McGuinness LR (2017) Profiling bacterial communities by MinION sequencing of ribosomal operons. Microbiome 5:116

Kezlya E, Tseplik N, Kulikovskiy M (2023) Genetic markers for metabarcoding of freshwater microalgae: review. Biology (Basel) 12:

Khrenova MG, Panova TV, Rodin VA, Kryakvin MA, Lukyanov DA, Osterman IA, Zvereva MI (2022) Nanopore sequencing for *de novo* bacterial genome assembly and search for single-nucleotide polymorphism. Int J Mol Sci 23:8569

Kim CY, Ma J, Lee I (2022) HiFi metagenomic sequencing enables assembly of accurate and complete genomes from human gut microbiota. Nature Commun 13:6367

Koreivienė J, Valčiukas R, Karosienė J, Baltrėnas P (2014) Testing of *Chlorella*/*Scenedesmus* microalgae consortia for remediation of wastewater, CO_2 mitigation and algae biomass feasibility for lipid production. J Env Eng Landscape Manage 22:105–114

Krehenwinkel H, Wolf M, Lim JY, Rominger AJ, Simison WB, Gillespie RG (2017) Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. Sci Rep 7:17668

Krinos AI, Cohen NR, Follows MJ, Alexander H (2023) Reverse engineering environmental metatranscriptomes clarifies best practices for eukaryotic assembly. BMC Bioinf 24:74

Krohn-Molt I, Alawi M, Förstner KU, Wiegandt A, Burkhardt L, Indenbirken D, Thieß M, Grundhoff A, Kehr J, Tholey A, Streit WR (2017) Insights into microalga and bacteria interactions of selected phycosphere biofilms using metagenomic, transcriptomic, and proteomic approaches. Front Microbiol 8:01941

Kublanovskaya A, Chekanov K, Solovchenko A, Lobakova E (2019) Cyanobacterial diversity in the algal–bacterial consortia from Subarctic regions: new insights from the rock baths at White Sea Coast. Hydrobiologia 830:17–31

Kublanovskaya A, Baulina O, Chekanov K, Lobakova E (2020a) The microalga *Haematococcus lacustris* (Chlorophyceae) forms natural biofilms in supralittoral White Sea coastal rock ponds. Planta 252:37

Kublanovskaya A, Solovchenko A, Fedorenko T, Chekanov K, Lobakova E (2020b) Natural communities of carotenogenic chlorophyte Haematococcus lacustris and bacteria from the White Sea coastal rock ponds. Microb Ecol 79:785–800

Kuppa Baskaran DK, Umale S, Zhou Z, Raman K, Anantharaman K (2023) Metagenome-based metabolic modelling predicts unique microbial interactions in deep-sea hydrothermal plume microbiomes. ISME Commun 3:42

Lane TW, Lane PD, Koh C-Y, Carney LT, Solberg OD (2013) Pond Crash forensics: microbiome analysis and field diagnostics. Sandia National Lab. (SNL-CA), Livermore, CA (United States)

Laver T, Harrison J, O'Neill PA, Moore K, Farbos A, Paszkiewicz K, Studholme DJ (2015) Assessing the performance of the Oxford Nanopore Technologies MinION. Biomolec Detect Quantif 3:1–8

Le VV, Ko S-R, Kang M, Park C-Y, Lee S-A, Oh H-M, Ahn C-Y (2022) The cyanobactericidal bacterium *Paucibacter aquatile* DH15 caused the decline of *Microcystis* and aquatic microbial community succession: A mesocosm study. Environ Pollut 311:119849

Lebonah DE, Dileep A, Chandrasekhar K, Sreevani S, Sreedevi B, Pramoda Kumari J (2014) DNA barcoding on bacteria: a review. Adv Biol 2014:541787

Lee S-A, Kim M, Esterhuizen M, Van Le V, Kang M, Ko SR, Oh HM, Kim YJ, Ahn CY (2022) An acceleration of carotenoid production and growth of *Haematococcus lacustris* induced by host-microbiota network interaction. Microbiol Res 262:127097

Lemoinne A, Dirberg G, Georges M, Robinet T (2023) Fine-scale congruence in bacterial community structure from marine sediments sequenced by short-reads on Illumina and long-reads on Nanopore. bioRxiv:2023.2006.2006.541006

Leong WH, Kiatkittipong K, Kiatkittipong W, Cheng YW, Lam MK, Shamsuddin R, Mohamad M, Lim JW (2020) Comparative performances of microalgal-bacterial co-cultivation to bioremediate synthetic and municipal wastewaters whilst producing biodiesel sustainably. Processes 8:1427

Li X, Li X, Yang X, Lan C, Huang Y, Jia B (2022) Identification and characterization of ATP-binding cassette transporters in *Chlamydomonas reinhardtii*. Mar Drugs 20:20100603

Li S, Li X, Chang H, Zhong N, Ren N, Ho S-H (2023a) Comprehensive insights into antibiotic resistance gene migration in microalgal-bacterial consortia: mechanisms, factors, and perspectives. Sci Total Environ 901:166029

Li S, Peng H, Shi X, Gu Q, Shen Z, Wang M (2023b) Significant effects of associated microorganisms on the community of photosynthetic picoeukaryotes. Microb Ecol 85:1164–1178

Liu Y-X, Qin Y, Chen T, Lu M, Qian X, Guo X, Bai Y (2020) A practical guide to amplicon and metagenomic analysis of microbiome data. Protein Cell 12:315–330

Liu W, Huang W, Cao Z, Ji Y, Liu D, Huang W, Zhu Y, Lei Z (2022) Microalgae simultaneously promote antibiotic removal and antibiotic resistance genes/bacteria attenuation in algal-bacterial granular sludge system. J Hazard Mat 438:129286

Ljaz U, Quince C (2013) TAXAassign v0.4. https://github.com/umerijaz/TAXAassign. Accessed November 15 2023

Llamas A, Leon-Miranda E, Tejada-Jimenez M (2023) Microalgal and nitrogen-fixing bacterial consortia: from interaction to biotechnological potential. Plants 12:2476

Luo Q, Bian C, Tao M, Huang Y, Zheng Y, Lv Y, Li J, Wang C, You X, Jia B, Xu J (2018) Genome and transcriptome sequencing of the astaxanthin-producing green microalga, *Haematococcus pluvialis*. Genome Biol Evol 11:166–173

Ma H, Wu X, Wei Z, Zhao L, Li Z, Liang Q, Zheng J, Wang Y, Li Y, Huang L, Hu Q (2021) Functional divergence of diacylglycerol acyltransferases in the unicellular green alga *Haematococcus pluvialis*. J Exp Bot 72:510–524

Maeda Y, Nojima D, Sakurai M, Nomaguchi T, Ichikawa M, Ishizuka Y, Tanaka T (2019) Genome analysis and genetic transformation of a water surface-floating microalga Chlorococcum sp. FFG039. Sci Reps 9:11200

Magdy T, Kuo HH, Burridge PW (2020) Precise and cost-effective nanopore sequencing for post-GWAS fine-mapping and causal variant identification. iScience 23:100971

Mardis ER (2017) DNA sequencing technologies: 2006–2016. Nature Protocols 12:213–218

Mars Brisbin M, Mitarai S, Saito MA, Alexander H (2022) Microbiomes of bloom-forming *Phaeocystis* algae are stable and consistently recruited, with both symbiotic and opportunistic modes. ISME J 16:2255–2264

Mastrorosa FK, Miller DE, Eichler EE (2023) Applications of long-read sequencing to Mendelian genetics. Genome Med 15:42

Matsuo Y, Komiya S, Yasumizu Y, Yasuoka Y, Mizushima K, Takagi T, Kryukov K, Fukuda A, Morimoto Y, Naito Y, Okada H (2021) Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MinION™ nanopore sequencing confers species-level resolution. BMC Microbiol 21:35

Mikalsen AJ, Zola J (2023) Coriolis: enabling metagenomic classification on lightweight mobile devices. Bioinformatics 39:i66–i75

Mishra S (2020) Cyanobacterial imprints in diversity and phylogeny. In: Singh PK, Kumar A, Singh VK, Shrivastava AK (eds) Advances in Cyanobacterial Biology. Academic Press, New York, pp 1–15

Mohamed AYA, Welles L, Siggins A, Healy MG, Brdjanovic D, Rada-Ariza AM, Lopez-Vazquez CM (2021) Effects of substrate stress and light intensity on enhanced biological phosphorus removal in a photo-activated sludge system. Water Res 189:116606

Nagarajan D, Lee D-J, Varjani S, Lam SS, Allakhverdiev SI, Chang J-S (2022) Microalgae-based wastewater treatment – Microalgae-bacteria consortia, multi-omics approaches and algal stress response. Sci Total Environ 845:157110

Noakes MT, Brinkerhoff H, Laszlo AH, Derrington IM, Langford KW, Mount JW, Bowman JL, Baker KS, Doering KM, Tickman BI, Gundlach JH (2019) Increasing the accuracy of nanopore DNA sequencing using a time-varying cross membrane voltage. Nature Biotech 37:651–656

Ovis-Sánchez JO, Perera-Pérez VD, Buitrón G, Quintela-Baluja M, Graham DW, Morales-Espinosa R, Carrillo-Reyes J (2023) Exploring resistomes and microbiomes in pilot-scale microalgae-bacteria wastewater treatment systems for use in low-resource settings. Sci Total Environ 882:163545

Padmaperuma G, Kapoore RV, Gilmour DJ, Vaidyanathan S (2018) Microbial consortia: a critical look at microalgae co-cultures for enhanced biomanufacturing. Crit Rev Biotech 38:690–703

Pan R, Hu H, Xiao Y, Xu L, Xu Y, Ouyang K, Li C, He T, Zhang W (2023) High-quality wild barley genome assemblies and annotation with nanopore long reads and Hi-C sequencing data. Sci Data 10:535

Park J, Park BS, Wang P, Patidar SK, Kim JH, Kim SH, Han MS (2017) Phycospheric native bacteria Pelagibaca bermudensis and Stappia sp. ameliorate biomass productivity of Tetraselmis striata (KCTC1432BP) in co-cultivation system through mutualistic interaction. Front Plant Sci 8:289

Patel VK, Sahoo NK, Patel AK, Rout PK, Naik SN, Kalra A (2017) Exploring microalgae consortia for biomass production: a synthetic ecological engineering approach towards sustainable production of biofuel feedstock. In: Gupta SK, Malik A, Bux F (eds) Algal Biofuels: Recent Advances and Future Prospects. Springer, Cham, pp 109–126

Patil KR, McHardy AC (2013) Alignment-Free genome tree inference by learning group-specific distance metrics. Genome Bio Evol 5:1470–1484

Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, Bowser SS, Cepicka I, Decelle J, Dunthorn M, Fiore-Donno AM (2012) CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. PLoS Biol 10:e1001419

Pawluczyk M, Weiss J, Links MG, Egaña Aranguren M, Wilkinson MD, Egea-Cortines M (2015) Quantitative evaluation of bias in PCR amplification and next-generation sequencing derived from metabarcoding samples. Anal Bioanal Chem 407:1841–1848

Pearman WS, Freed NE, Silander OK (2020) Testing the advantages and disadvantages of short- and long- read eukaryotic metagenomics using simulated reads. BMC Bioinformat 21:220

Perera IA, Abinandan S, Subashchandrabose SR, Venkateswarlu K, Naidu R, Megharaj M (2021) Microalgal–bacterial consortia unveil distinct physiological changes to facilitate growth of microalgae. FEMS Microbiol Ecol 97:fiab012

Pessi IS, Popin RV, Durieu B, Lara Y, Tytgat B, Savaglia V, Roncero-Ramos B, Hultman J, Verleyen E, Vyverman W, Wilmotte A (2023) Novel diversity of polar Cyanobacteria revealed by genome-resolved metagenomics. Microbial Genomics 9:1056

Portik DM, Brown CT, Pierce-Ward NT (2022) Evaluation of taxonomic classification and profiling methods for long-read shotgun metagenomic sequencing datasets. BMC Bioinf 23:541

Poursalavati A, Javaran VJ, Laforest-Lapointe I, Fall ML (2023) Soil metatranscriptomics: an improved RNA extraction method toward functional analysis using Nanopore direct RNA sequencing. Phytobiomes J 7:42–54

Power RA, Parkhill J, de Oliveira T (2017) Microbial genome-wide association studies: lessons from human GWAS. Nature Rev Genet 18:41–50

Priya AK, Jalil AA, Vadivel S, Dutta K, Rajendran S, Fujii M, Soto-Moscoso M (2022) Heavy metal remediation from wastewater using microalgae: Recent advances and future trends. Chemosphere 305:135375

Pushpakumara B, Tandon K, Willis A, Verbruggen H (2023) Unravelling microalgal-bacterial interactions in aquatic ecosystems through 16S rRNA gene-based co-occurrence networks. Sci Rep 13:2743

Qixin L, Xuan F, Zhiya S, Wenxin S, Shuo W, Ji L (2022) Enhanced wastewater treatment performance by understanding the interaction between algae and bacteria based on quorum sensing. Bioresour Technol 354:127161

Rodríguez-Pérez H, Ciuffreda L, Flores C (2022) NanoRTax, a real-time pipeline for taxonomic and diversity analysis of nanopore 16S rRNA amplicon sequencing data. Computat Struct Biotech J 20:5350–5354

Rosenberg E, Zilber-Rosenberg I (2018) The hologenome concept of evolution after 10 years. Microbiome 6:78

Saary P, Mitchell AL, Finn RD (2020) Estimating the quality of eukaryotic genomes recovered from metagenomic analysis with EukCC. Genome Biol 21:244

Sadvakasova AK, Bauenova MO, Kossalbayev BD, Zayadan BK, Huang Z, Wang J, Balouch H, Alharby HF, Chang JS, Allakhverdiev SI (2023) Synthetic algocyanobacterial consortium as an alternative to chemical fertilizers. Env Res 233:116418

Saini JS, Manni M, Hassler C, Cable RN, Duhaime MB, Zdobnov EM (2023) Genomic insights into the coupling of a *Chlorella*-like microeukaryote and sulfur bacteria in the chemocline of permanently stratified Lake Cadagno. ISME J 17:903–915

Saleski TE, Kerner AR, Chung MT, Jackman CM, Khasbaatar A, Kurabayashi K, Lin XN (2019) Syntrophic co-culture amplification of production phenotype for high-throughput screening of microbial strain libraries. Metabol Eng 54:232–243

Salzberg SL (2019) Next-generation genome annotation: we still struggle to get it right. Genome Biol 20:92

San JE, Baichoo S, Kanzi A, Moosa Y, Lessells R, Fonseca V, Mogaka J, Power R, de Oliveira T (2020) Current affairs of microbial genome-wide association studies: approaches, bottlenecks and analytical pitfalls. Front Microbiol 10:3119

Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Molec Bioly 94:441–448

Santo ÉDE, Ishii M, Pinto UM, Matsudo MC, Carvalho JCMd (2022) Obtaining Bioproducts from the studies of signals and interactions between microalgae and bacteria. Microorganisms 10:2029

Saravanan A, Kumar PS, Varjani S, Jeevanantham S, Yaashikaa PR, Thamarai P, Abirami B, George CS (2021) A review on algal-bacterial symbiotic system for effective treatment of wastewater. Chemosphere 271:129540

Sauvage T, Schmidt WE, Suda S, Fredericq S (2016) A metabarcoding framework for facilitated survey of endolithic phototrophs with tufA. BMC Ecol 16:8

Sauvage T, Schmidt WE, Yoon HS, Paul VJ, Fredericq S (2019) Promising prospects of nanopore sequencing for algal hologenomics and structural variation discovery. BMC Genom 20:850

Schiml VC, Delogu F, Kumar P, Kunath B, Batut B, Mehta S, Johnson JE, Grüning B, Pope PB, Jagtap PD, Griffin TJ (2023) Integrative meta-omics in Galaxy and beyond. Environ Microbiome 18:56

Scognamiglio V, Giardi MT, Zappi D, Touloupakis E, Antonacci A (2021) Photoautotrophs–bacteria co-cultures: advances, challenges and applications. Materials 14:3027

Seymour JR, Amin SA, Raina J-B, Stocker R (2017) Zooming in on the phycosphere: the ecological interface for phytoplankton–bacteria relationships. Nature Microbiol 2:17065

Shetty P, Boboescu IZ, Pap B, Wirth R, Kovács KL, Bíró T, Futó Z, White RA III, Maróti G (2019) Exploitation of algal-bacterial consortia in combined biohydrogen generation and wastewater treatment. Front Energy Res 7:0052

Shin J, Lee S, Go MJ, Lee SY, Kim SC, Lee CH, Cho BK (2016) Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing. Sci Rep 6:29681

Shin H, Lee E, Shin J, Ko SR, Oh HS, Ahn CY, Oh HM, Cho BK, Cho S (2018) Elucidation of the bacterial communities associated with the harmful microalgae *Alexandrium tamarense* and *Cochlodinium polykrikoides* using nanopore sequencing. Sci Rep 8:5323

Slatko BE, Gardner AF, Ausubel FM (2018) Overview of next-generation sequencing technologies. Curr Protocols Mol Biol 122:e59

Solomon W, Mutum L, Janda T, Molnár Z (2023) Potential benefit of microalgae and their interaction with bacteria to sustainable crop production. Plant Growth Regul 101:53–65

Starkenburg SR, Kwon KJ, Jha RK, McKay C, Jacobs M, Chertkov O, Twary S, Rocap G, Cattolico RA (2014) A pangenomic analysis of the *Nannochloropsis* organellar genomes reveals novel genetic variations in key metabolic genes. BMC Genom 15:212

Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H (2009) Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. Proc Nat Acad Sci 106:7702–7707

Szoboszlay M, Schramm L, Pinzauti D, Scerri J, Sandionigi A, Biazzo M (2023) Nanopore Is preferable over Illumina for 16S amplicon sequencing of the gut microbiota when species-level taxonomic classification, accurate estimation of richness, or focus on rare taxa is required. Microorganisms 11:

Tedersoo L, Tooming-Klunderud A, Anslan S (2018) PacBio metabarcoding of Fungi and other eukaryotes: errors, biases and perspectives. New Phytol 217:1370–1385

Thomas F, Dittami SM, Brunet M, Le Duff N, Tanguy G, Leblanc C, Gobet A (2020) Evaluation of a new primer combination to minimize plastid contamination in 16S rDNA metabarcoding analyses of alga-associated bacterial communities. Env Microbiol Rep 12:30–37

Toyama T, Kasuya M, Hanaoka T, Kobayashi N, Tanaka Y, Inoue D, Sei K, Morikawa M, Mori K (2018) Growth promotion of three microalgae, *Chlamydomonas reinhardtii*, *Chlorella vulgaris* and *Euglena gracilis*, by in situ indigenous bacteria in wastewater effluent. Biotechnol Biofuels 11:176

Tran Q, Phan V (2020) Assembling reads improves taxonomic classification of species. Genes 11:946

Trebuch LM, Oyserman BO, Janssen M, Wijffels RH, Vet LEM, Fernandes TV (2020) Impact of hydraulic retention time on community assembly and function of photogranules for wastewater treatment. Water Res

Trebuch LM, Sohier J, Altenburg S, Oyserman BO, Pronk M, Janssen M, Vet LE, Wijffels RH, Fernandes TV (2023) Enhancing phosphorus removal of photogranules by incorporating polyphosphate accumulating organisms. Water Res

Trench-Fiol S, Fink P (2020) Metatranscriptomics from a small aquatic system: microeukaryotic community functions through the diurnal cycle. Front Microbiol 11:

Turk Dermastia T, Vascotto I, Francé J, Stanković D, Mozetič P (2023) Evaluation of the *rbcL* marker for metabarcoding of marine diatoms and inference of population structure of selected genera. Front Microbiol 14:

Upadhyay AK, Singh R, Singh DV, Singh L, Singh DP (2021) Microalgal consortia technology: A novel and sustainable approach of

resource reutilization, waste management and lipid production. Env Technol Innovat 23:101600
van der Loos LM, D'Hondt S, Willems A, De Clerck O (2021) Characterizing algal microbiomes using long-read nanopore sequencing. Algal Res 59:102456
Vasilieva S, Lukyanov A, Antipova C, Grigoriev T, Lobakova E, Chivkunova O, Scherbakov P, Zaytsev P, Gorelova O, Fedorenko T, Kochkin D (2023) Interactive effects of ceftriaxone and chitosan immobilization on the production of arachidonic acid by and the microbiome of the chlorophyte *Lobosphaera* sp. IPPAS C-2047. Int J Mol Sciences 24:
Vieira HH, Bagatini IL, Guinart CM, Vieira AAH (2016) *tufA* gene as molecular marker for freshwater Chlorophyceae. Algae 31:155–165
Vingiani GM, De Luca P, Ianora A, Dobson ADW, Lauritano C (2019) Microalgal enzymes with biotechnological applications. Mar Drugs 17:459
Volden R, Palmer T, Byrne A, Cole C, Schmitz RJ, Green RE, Vollmers C (2018) Improving nanopore read accuracy with the R2C2 method enables the sequencing of highly multiplexed full-length single-cell cDNA. Proc Nat Acad Sci 115:9726–9731
Vučić V, Müller S (2021) New developments in biological phosphorus accessibility and recovery approaches from soil and waste streams. Eng Life Sci 21:77–86
Wang R, Tian Y, Xue S, Zhang D, Zhang Q, Wu X, Kong D, Cong W (2016) Enhanced microalgal biomass and lipid production via co-culture of *Scenedesmus obliquus* and *Candida tropicalis* in an autotrophic system. J Chem Technol Biotechnol 91:1387–1396
Wang Y, Hu Y, Gao GF (2020) Combining metagenomics and metatranscriptomics to study human, animal and environmental resistomes. Med Microecol 3:100014
Wang C, Qi M, Guo J, Zhou C, Yan X, Ruan R, Cheng P (2021) The active phytohormone in microalgae: the characteristics, efficient detection, and their adversity resistance applications. Molecules 27:
Wang S, Su X, Cui H, Wang M, Hu X, Ding W, Zhang W (2022) Microbial richness of marine biofilms revealed by sequencing full-length 16S rRNA genes. Genes 13:1050
Wang L, Jia X, Xu L, Yu J, Ren S, Yang Y, Wang K, López-Arredondo D, Herrera-Estrella L, Lambers H, Yi K (2023) Engineering microalgae for water phosphorus recovery to close the phosphorus cycle. Plant Biotech J 21:1373–1382
West PT, Probst AJ, Grigoriev IV, Thomas BC, Banfield JF (2018) Genome-reconstruction for eukaryotes from complex natural microbial communities. Genome Res 28:569–580
Wirth R, Pap B, Böjti T, Shetty P, Lakatos G, Bagi Z, Kovács KL, Maróti G (2020) *Chlorella vulgaris* and its phycosphere in wastewater: microalgae-bacteria interactions during nutrient removal. Front Bioeng Biotech 8:557572
Xiong W, Wang S, Jin Y, Wu Z, Liu D, Su H (2023) Insights into nitrogen and phosphorus metabolic mechanisms of algal-bacterial aerobic granular sludge via metagenomics: Performance, microbial community and functional genes. Bioresour Technol 369:128442
Yang C, Chowdhury D, Zhang Z, Cheung WK, Lu A, Bian Z, Zhang L (2021) A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data. Computat Struct Biotech J 19:6301–6314
Yergeau E, Michel C, Tremblay J, Niemi A, King TL, Wyglinski J, Lee K, Greer CW (2017) Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian. Arctic Sci Rep 7:42242
Zarezadeh S, Moheimani NR, Jenkins SN, Hülsen T, Riahi H, Mickan BS (2019) Microalgae and phototrophic purple bacteria for nutrient recovery from agri-industrial effluents: influences on plant growth, rhizosphere bacteria, and putative carbon- and nitrogen-cycling genes. Front Plant Sci 10:1193
Zhang LL, Zhang C, Peng JP (2022) Application of nanopore sequencing technology in the clinical diagnosis of infectious diseases. Biomed Environ Sci 35:381–392
Zorrilla F, Buric F, Patil KR, Zelezniak A (2021) metaGEM: reconstruction of genome scale metabolic models directly from metagenomes. Nucleic Acids Res 49:e126
Zou S, Fei C, Wang C, Gao Z, Bao Y, He M, Wang C (2016) How DNA barcoding can be more effective in microalgae identification: a case of cryptic diversity revelation in *Scenedesmus* (Chlorophyceae). Sci Rep 6:36822

Journal:	**10811**
Article:	**3267**

Author Query Form

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query	Details Required	Author's Response
AQ1	Please check if the authorgroup and affiliations are captured/presented correctly.	
AQ2	Figure 3 contains poor quality of text inside the artwork. Please do not re-use the file that we have rejected or attempt to increase its resolution and re-save.	
AQ3	Please provide complete bibliographic details of this reference Choi et al., 2021,Danish-Daniel et al., 2023, Dow, 2021, Gonzalez-Gonzalez and de-Bashan, 2023, Inskeep et al., 2013, Kang et al., 2021, Kezlya et al., 2023, Lane et al., 2013, Lemoinne et al., 2023, Ljaz and Quince, 2013, Szoboszlay et al., 2023, Trebuch et al., 2020, 2023, Trench-Fiol and Fink, 2020, Turk Dermastia et al., 2023, Vasilieva et al., 2023 and Wang et al., 2021.	