



Article

# Obtaining Genome Sequences of Mutualistic Bacteria in Single *Microcystis* Colonies

Jing Tu <sup>1,\*</sup> , Liang Chen <sup>1</sup>, Shen Gao <sup>1</sup>, Junyi Zhang <sup>1,2</sup>, Changwei Bi <sup>1</sup> , Yuhan Tao <sup>1</sup>, Na Lu <sup>1</sup> and Zuhong Lu <sup>1,\*</sup>

<sup>1</sup> State Key Lab of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China; tongust@163.com (L.C.); gaoshen12345gg@hotmail.com (S.G.); blocksharon@163.com (J.Z.); bichwei@163.com (C.B.); 220174561@seu.edu.cn (Y.T.); nlu@seu.edu.cn (N.L.)

<sup>2</sup> Wuxi Environmental Monitoring Center, Wuxi 210096, China

\* Correspondence: jtu@seu.edu.cn (J.T.); zhlu@seu.edu.cn (Z.L.); Tel.: +86-25-8379-2396 (J.T.); +86-25-8379-3779 (Z.L.)

Received: 22 September 2019; Accepted: 10 October 2019; Published: 11 October 2019



**Abstract:** Cells of *Microcystis* are associated with heterotrophic bacteria and organized in colonies in natural environment, which are basic elements in the mass occurrence of cyanobacterial species. Analyzing these colonies by using metagenomics is helpful to understand species composition and relationship. Meanwhile, the difference in population abundance among *Microcystis* colonies could be used to recover genome bins from metagenome assemblies. Herein, we designed a pipeline to obtain high-quality genomes of mutualistic bacteria from single natural *Microcystis* colonies. Single colonies were lysed, and then amplified by using multiple displacement amplification to overcome the DNA quantity limit. A two-step assembly was performed after sequencing and scaffolds were grouped into putative bins based on their differential-coverage among species. We analyzed six natural colonies of three prevailing *Microcystis* species from Lake Taihu. Clustering results proved that colonies of the same species were similar in the microbial community composition. Eight putative population genome bins with wide bacterial diversity and different GC content were identified based on coverage difference among colonies. At the phylum level, proteobacteria was the most abundant besides cyanobacteria. Six of the population bins were further refined into nearly complete genomes (completeness > 90%).

**Keywords:** cyanobacterial blooms; *Microcystis*; metagenomics; single *Microcystis* colony; mutualistic bacteria

## 1. Introduction

Cyanobacterial blooms occur as a result of massive proliferation of cyanobacteria [1], and have become a worldwide environmental problem [2]. Among several cyanobacterial genera, the unicellular, worldwide distributed cyanobacterium *Microcystis* is the most representative genus of bloom-forming cyanobacteria [3,4].

In natural environment, *Microcystis* cells are associated with heterotrophic bacteria and organized in colonies which vary in sizes and shapes [5]. The rise and accumulation of large *Microcystis* colonies is the immediate cause of abrupt occurrence of *Microcystis* blooms [6]. Lab culture experiments support that the natural heterotrophic bacterial community have a role in the development of *Microcystis* blooms in natural water [7]. To improve our understanding of the relationship within *Microcystis* colonies, genomes of representative heterotrophic bacteria are desired to be obtained. Systematically sequencing genomes from pure cultures is an effective method to obtain bacteria genomes. However, it is challenging to disassociate and purify these heterotrophic bacteria from *Microcystis* colonies and only

a small portion of them is representable by cultivated organisms [8]. By directly extracting bulk DNA from environmental samples and sequencing, metagenomics has overcome the disassociation and cultivation obstacles [9]. High-quality genomes of dominant populations have been assembled from relatively low-complexity communities using metagenomics [10,11]. Individual genomes can be recovered by grouping sequences of particular individuals from a background of community sequencing data [12]. The applications of metagenomics in cyanobacterial blooms [2,13–15] have revealed useful information on the microbial communities. However, the genomes of individual members are still not well assembled because the microbial communities of *Microcystis* blooms are high complex microorganism mixtures in natural freshwater. These mixtures contain different types of *Microcystis* colonies and some other planktonic microbial communities.

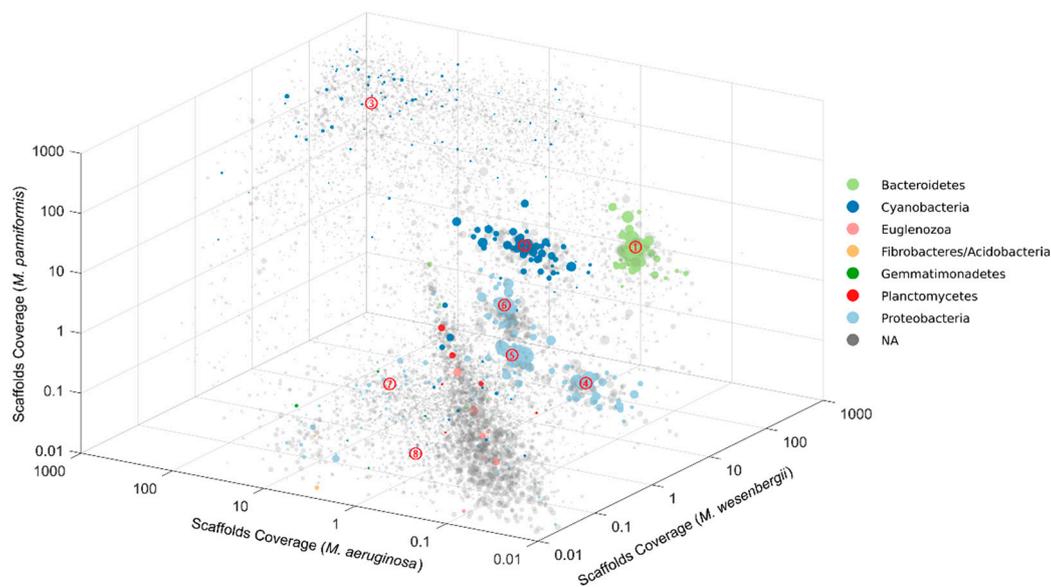
As the basic element in the mass occurrence of cyanobacterial species, single *Microcystis* colony is less complex in species composition, is an ideal sample for metagenomic studies, and has been used in microcystin gene analysis [16,17] or genetic study [18]. *Microcystis* colonies differ in morphological characteristics and can be divided into different morphospecies [19]. Each type of *Microcystis* colony is diverse in microbial community and has different metabolic pathways among the members. Many researches attempted to understand the morphological variation among *Microcystis* in several directions [20–25]. However, because the DNA quantity is quite limited in a single *Microcystis* colony, it is difficult to perform efficient metagenome analysis. This bottleneck can be bypassed by culturing individual type of *Microcystis* colonies in lab. Xie et al. cultured *Microcystis wesenbergii* T100 in lab for bulk DNA extraction and downstream analysis, and the genomes of eight individual mutualistic bacteria were reconstructed from metagenomic assemblies [26]. However, some colonial *Microcystis* stains lose their mucilage and convert into individual cells gradually after some generations in lab culture [23]. The lab-maintained colonies are less complex and their microorganic compositions are unlike the ones in nature [7,26]. Single natural *Microcystis* colonies are more suitable for metagenomic studies to obtain the genome sequences of associated bacteria from the bacterial communities in cyanobacterial blooms. Meanwhile, the difference in population abundance among *Microcystis* colonies can be used to recover genome bins from metagenome assemblies.

Herein, we designed a pipeline to obtain genome sequences of mutualistic bacteria in single *Microcystis* colonies, which lysed single colonies and amplified the metagenomes before library construction, and grouped sequences of metagenome based on their differential-coverage among three prevailing species of *Microcystis* colonies. Six related metagenome sequencing libraries were generated by disassociating single *Microcystis* colonies from a mixed sample, following by multiple displacement amplification (MDA) to overcome the DNA quantity limit. The assembled metagenome scaffolds were grouped in a three-dimension space based on sequencing depth. We obtained eight population genome bins of the cohabitation bacteria which differed in abundance among *M. aeruginosa*, *M. wesenbergii* and *M. panniformis* colonies, and six of them were further refined into high-quality genomes.

## 2. Results

### 2.1. Binning of Single *Microcystis* Colony Metagenomes

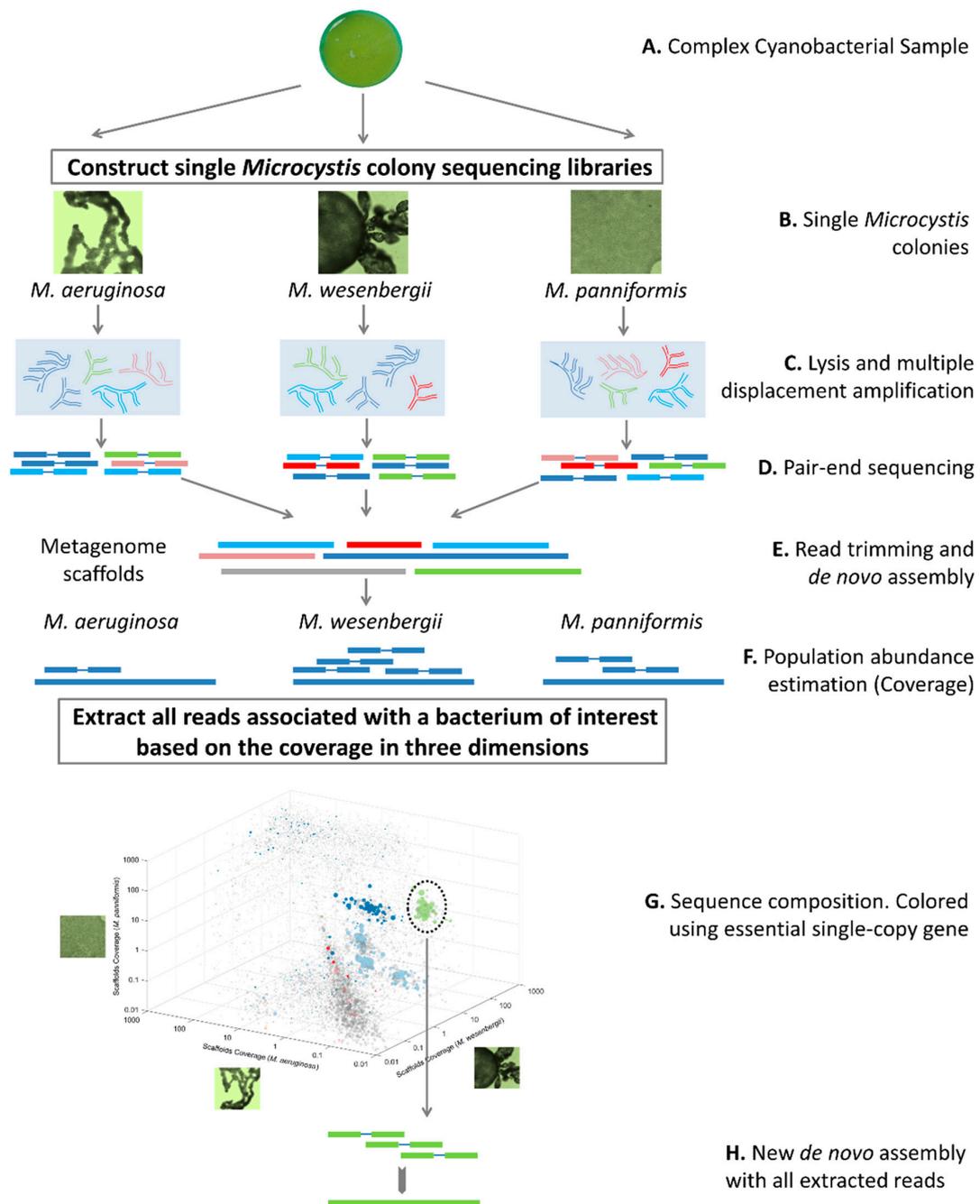
We deeply sequenced DNA from single *Microcystis* colonies of three species (*M. aeruginosa*, *M. wesenbergii* and *M. panniformis*) collected from Lake Taihu, the third largest freshwater lake in China, which suffers annual *Microcystis* blooms. A total of 87 Gbp of paired-end sequencing data were produced and were assembled into a total of 105 Mbp of scaffolds ranging in size from 1 Kbp to 198 Kbp (Table S1). Due to the abundance difference among species, populations were differentially represented in each data set. We firstly binned the scaffolds into population genomes by plotting all scaffolds in three-dimension (3D) based on their coverage in different *Microcystis* colonies (Figure 1 and Figure S2). Some subsets of spots were further binned based on the differences in nucleotide composition. In total, eight putative population genome bins with wide bacterial diversity and big range of GC content were identified, and six of them were further refined into high-quality genomes.



**Figure 1.** Sequence 3D coverage binning of metagenome scaffolds from the single *Microcystis* colonies. Plots represent scaffolds, the size indicates the length of scaffolds and the color shows the phylum-level taxonomic affiliation. Gray spots are scaffolds which cannot be assigned to the essential genes of the 8 most abundant phyla. Clusters of same colored spots represent potential genome bins, and are indicated by numbers in the center (Table 2). Spots in similar coverage pattern, such as around diagonal, were extracted and further binned based on the differences in nucleotide composition.

## 2.2. Pipeline of Obtaining Genomes in Single *Microcystis* Colonies

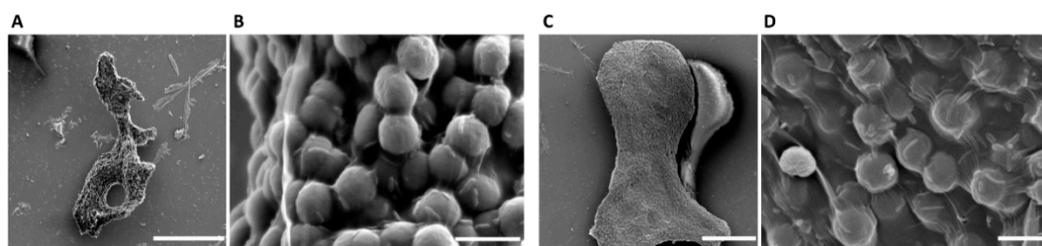
The overall pipeline of obtaining genomes in single *Microcystis* colonies is illustrated in Figure 2. Briefly, colonies of *Microcystis* were separated from a cyanobacterial bloom sample collected from Lake Taihu, and were classified under light microscope based their species features. *Microcystis* colonies which were classified as *M. aeruginosa*, *M. wesenbergii* and *M. panniformis* were selected for metagenome analysis. These three species are dominated species in Lake Taihu. Two parallel colonies of each species were used for following analysis. Due to limit of DNA quantity in a single *Microcystis* colony, MDA, which is superior in genome coverage, amplification accuracy and efficiency [27,28], was used to amplify the lysed single colonies. Sequencing libraries were constructed using the amplified nucleotides and the six single colonies were sequenced separately. All the paired-end sequence data of the six samples were used unitedly in de novo assembly. After the evaluating parallel colonies, reads from the first one of two parallel colonies of each species were used for mapping to assembled scaffolds, and the mapping results showed that populations were differentially represented due to their diverse abundance in specific species. The binning of the scaffolds was firstly facilitated by plotting the three coverage against each other in the 3D space. Some putative genome bins were also purified based on their GC content and using principal component analysis of tetranucleotide frequencies. After that, all the sub-network of scaffolds associated with the grouped bins were extracted. To improve the assembly, all relevant reads of a genome bin were extracted for reassembly, including the paired-end reads which both ends mapped to the scaffolds and those which only one end mapped to the scaffolds. The extracted reads were used for a new de novo assembly.



**Figure 2.** Overview of the pipeline to obtain population genomes from single *Microcystis* colony metagenomes. A. The complex cyanobacterial sample was collected. B. Colonies of three dominated *Microcystis* species (*M. aeruginosa*, *M. wesenbergii* and *M. panniformis*) were separated and classified based on their species features. C. Colonies were lysed and amplified by using MDA. D. Paired-end sequencing libraries were constructed and the metagenomes were sequenced separately. E. All data of metagenomes were used unitedly in *de novo* assembly after read trimming. F. The population abundance of each assembled scaffold was estimated based on its coverage in each metagenome. G. The binning of the scaffolds was facilitated by plotting coverages against each other. The colored spots are scaffolds which can be assigned to the essential genes of the 8 most abundant phyla. H. The subsets of scaffolds associated with the grouped bins were extracted. The paired-end reads which at least one end mapped to the scaffolds were extracted and used for a new *de novo* assembly.

### 2.3. Colony Classification and Sample Preparation

Colonies of *Microcystis* were classified and selected under light microscope and further imaged under scanning electron microscope (SEM) (Figure 3 and Figure S1). *Microcystis* frequently forms large colonies with mucilage and heterotrophic bacteria. Colonies of different species have specific morphological characteristics. The colonies of *M. aeruginosa* form in irregular shape, and the mucilage is frequently broken or perforated, making the population into a network of branches or panes (Figure 3A and Figure S1A). In *M. aeruginosa* colonies, numerous bacteria, which are usually rods, were observed on the surface or embedded in the polysaccharide shell under SEM (Figure 3B). The mucilage of *M. wesenbergii* has clear boundary and is transparent. There are fewer *microcystis* cells of in a *M. wesenbergii* colony, which are sparsely, arranged randomly and formed in hollow population (Figure S1B). Due to the abundance of mucilage and the hollow population, dehydration in the sample preparation of SEM changed the morphology of *M. wesenbergii* colonies, preventing them from being identified under SEM. The colonies of *M. panniformis* are usually flat, with irregular edges and cells arranged densely (Figure 3C and Figure S1C). In SEM micrographs of *M. panniformis* colonies, numerous rod-like bacteria were also observed on the surface or embedded in the polysaccharide shell (Figure 3D).



**Figure 3.** The SEM micrographs of the single *Microcystis* colonies. (A) Global view of a *M. aeruginosa* colony. (B) Local view of the selected *M. aeruginosa* colony. Bacilliform bacteria were seen on the surface of the polysaccharide shell. (C) Global view of a *M. panniformis* colony. (D) Local view of the selected *M. panniformis* colony. Bacilliform bacteria were also seen on the surface of the polysaccharide shell. Scale bars in A and C, 100  $\mu\text{m}$ . Scale bars in B and D, 5  $\mu\text{m}$ .

Colonies of different species were classified under morphological characteristics. Under SEM, plenty of adherent bacteria were observed obviously on the surface of polysaccharide shell of these natural colonies, which revealed that the mutualistic system of *Microcystis* colonies in natural environment were composed of algal cells and abundant heterotrophic bacteria.

### 2.4. Mapping Results

After sequencing, 45 to 166 million reads were generated for each colony (Table 1). Six metagenomes were aligned to NIES843 [3] and FACHB-1757 [29] separately. The mapping results are shown in Table 1. At least 67.87% of the two genomes were covered at least five folds using 45 million reads (Table 1). The high coverage ratio of all metagenomes using both two references revealed that the microbial communities of these metagenomes were mainly composed of *Microcystis* cells. At same coverage, higher coverage ratios were got for each metagenome using FACHB-1757 as reference. The two metagenomes of *M. panniformis* colony both covered over 95% of the FACHB-1757 genome for at least five folds. As a contrast, none of the metagenomes covered over 85% of the NIE-843 genome at same coverage. Similar coverage ratios within species revealed a well consistency between duplicates.

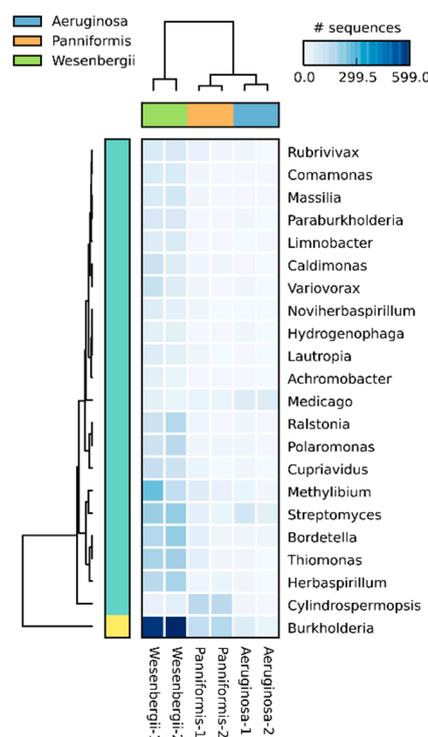
**Table 1.** Data summary and mapping results.

	Aeruginosa-1	Aeruginosa-2	Wesenbergii-1	Wesenbergii-2	Panniformis-1	Panniformis-2
Total reads (million)	94.68	77.94	45.23	165.66	106.81	91.77
Coverage Ratio of NIES-843*	67.87%	69.71%	83.83%	78.41%	80.56%	75.70%
Coverage Ratio of FACHB-1757 *	68.20%	70.35%	87.09%	94.30%	99.89%	95.32%

\* With 45 million reads in 5× coverage.

## 2.5. Clustering Results

After aligning to non-redundant protein database, the relative abundance of orders belonging to Cyanobacterioidetes were calculated and shown in Figure S3. Taxonomic analysis was performed subsequently using the alignment files. Genera with differential abundance were extracted for clustering analysis (Figure 4). Obviously, metagenomes of the same species show a significantly high similarity. The lineage of each species is distinct from the lineage of the other two species. Relatively, the two metagenomes of *M. panniformis* are closer to *M. aeruginosa* than that of *M. wesenbergii*. The clustering results revealed that single colonies of the same *microcystis* species are similar in the microbial community composition. The relatively close distance between single colonies of *M. panniformis* and *M. aeruginosa* is in accord with their similarity in morphological characteristics.



**Figure 4.** Clustering result of taxonomic analysis results at genus level. Based on the taxonomic analysis results, six samples were clustered at genus level. Metagenomes of the same species showed to be closest with each other. Relatively, the two metagenomes of *M. panniformis* are closer to *M. aeruginosa* than that of *M. wesenbergii*.

## 2.6. Assembly and Reads Binning

After quality control and reads trimming, 84 Gbp paired-end sequencing data were used for de novo assembly (Table S1). A total of 105 Mbp of scaffolds with N50 for 4935 bp were assembled using SPAdes. The assembled scaffolds captured 49.83% of all sequenced reads.

To evaluate to bias induced by MDA, a bulk sample was obtained and 6.17 Gbp of paired-end sequencing data were produced. The metagenome of the bulk sample was assembled under the same condition. We calculated the percentage of scaffold with specific GC content in both

amplified metagenome and bulk metagenome (Table S2). The results proved that MDA prefers to amplify fragments with low GC content rather than with high GC content. Therefore, we calculated standardization factors for each GC content to normalize the relative abundance.

The binning was based on the different coverage of scaffolds in the data set of each species. Due to the high similarity in microbial composition, only the first metagenomes of each species were used for binning. In the 3D space of coverage, scaffolds clustered together represent putative population genome bins. Further refinement was facilitated for the bins which contained multiple species. In total, eight population bins were identified with GC content ranging from 42.03% to 68.52% and represented with bacterial diversity (five phyla) (Table S3). Six of the population bins were assembled into high-quality genomes (Table 2) which included many rare members with the lowest being 0.126% in relative abundance. In comparison, binning based on GC content and overall coverage which is classically used in metagenomics identified five population bins, and these bins contained more duplicated essential genes which represented multiple species (Figure S4, Table S4). Population bins with similar GC content and overall coverage cannot be separated using classical pipeline. For example, group E in Supplementary Figure 4 which contained 80 duplicated essential genes is the mixture of group 4 (67.60%, GC content) and group 6 (68.52%, GC content). At the phylum level, four of the eight groups were proteobacteria, two were cyanobacteria, one was bacteroidetes, one was gemmatimonadetes and one was the group of euglenozoa. Proteobacteria was discovered to be the most abundant phylum in cyanobacterial water bloom samples, while bacteroidetes was another important phylum [30]. Proteobacteria was also reported to be the most abundant phylum in Lake Taihu [14]. Because of some putative bins, group 3, group 7, and group 8 do not stand out in Figure 1, we highlighted the selected scaffolds of each group in separated figures (Figure S5).

Except for group 3, the genome completeness of the other five groups were estimated to be over 90% based on the number of essential genes, and four of them were over 95% completed (Table 2). The completeness validated using CheckM was in accordance with the results based on essential genes (Table S5). These five genomes have few duplicated essential genes, suggesting that each group had little contamination. However, the contamination validated by CheckM was relatively higher and diverse among genomes (Table S5). The inconsistency in contamination validation might due to the different evaluation criterions used. Different methods had been reported to generate different contamination level [31]. All these results indicated that five pure and nearly completed genomes were obtained through the differential-coverage binning of single *Microcystis* colonies. Group 3 was the most abundant group among the eight discovered groups, and ought to be the most abundant group in all populations in the metagenome according to its quite high relative abundance (15.34%).

At the genus level, group 3 was from *Microcystis*. A lot number of duplicated essential genes of group 3 indicated the genome was not well separated from one the others. During assembly, metagenomes of single colonies of *M. aeruginosa*, *M. wesenbergii* and *M. panniformis* were used unitedly. Group 3 is part of the chimeric genome of *M. aeruginosa*, *M. wesenbergii* and *M. panniformis*.

**Table 2.** Assembly statistics for the 6 representative genomes.

Figure ID	No. Contigs	Total Length (bp)	GC (%)	No. Essential Genes	No. Duplicated Essential Genes	Standardized Relative Abundance* (%)				Genus	Phylogenetic Affiliation
						<i>aeruginosa</i>	<i>wesenbergii</i>	<i>paniformis</i>	Overall		
1	149	3,487,294	50.16	101/105	1	0.004	2.250	0.116	1.004	<i>Chryseolinea</i>	Bacteroidetes
2	309	4,301,792	42.03	102/106	5	0.190	0.528	0.095	0.443	<i>Pseudanabaena</i>	Cyanobacterioidetes
3	3434	10,835,247	43.20	81/106	40	9.356	23.222	8.004	15.344	<i>Microcystis</i>	Cyanobacteroidetes
4	160	2,863,377	67.60	98/106	5	0.003	0.567	0.064	0.376	<i>Brevundimonas</i>	Proteobacteria
5	39	1,954,923	54.00	102/106	1	0.007	0.435	0.004	0.126	<i>Cupriavidus</i>	Proteobacteria
6	219	3,678,702	68.52	103/106	2	0.040	0.774	0.094	0.842	<i>Burkholderia</i>	Proteobacteria

\* Standardized relative abundance was calculated as the standardized percentage of reads of a genome bin in the total number of reads or the number of reads in a specific metagenome.

Group 1 was from *Chryseolinea*, a genus of *Cytophagia*. The GC content of the assembled genome (50.2%) is close to the reported species (49.9%) of *Chryseolinea* which was isolated from soil [32]. Bacteria belonging to *Cytophagia* have been reported to be isolated from *Microcystis* colonies [33,34]. Members of another genus in *Cytophagia*, *Cytophage* act as cyanobacteriolytic bacteria [35], which are able to lyse the cyanobacteria as a food resource. Therefore, the high relative abundance of group 1 in *M. wesenbergii* colony compared with the other two species might be a reason of the sparseness structure. Group 2 was from *Pseudanabaena*, another genus of *Cyanobacteria*. As autotrophic organism, *Pseudanabaena mucicola* can often be found associated to *Microcystis* colonies [36–38]. The co-cultivation of *Pseudanabaena* with *Microcystis* strains in the lab revealed strong interaction between these two cyanobacteria, and evidenced that the growth of *Pseudanabaena* in natural systems may cause either *Microcystis* lysis, or colony sedimentation [39]. Group 4 was from *Brevundimonas*, a genus of *Alphaproteobacteria*. *Brevundimonas* has been reported to be a frequenter of *Microcystis* colony associated bacterial in different locations [40]. Group 5 and 6 were from two genera of the same family (*Burkholderiaceae*) in *Betaproteobacteria*, *Cupriavidus* and *Burkholderia*. Two bacteria in family *Burkholderiaceae* were binned from the metagenome of a lab cultured *Microcystis* strain, *M. wesenbergii* T100 [26]. *Burkholderiales* have been evidenced to be one of the dominated orders in buoyant cyanobacteria [41]. Group 7 was from *Roseomonas*, another genus of *Alphaproteobacteria*, and was also demonstrated to be one of the dominant genera during the algal bloom [42]. Group 8 from *Gemmatimonas* was the lowest abundant population among the eight assembled genomes. *Gemmatimonas* is frequently associated with cyanobacterial colonies [43,44].

### 3. Discussion

*Microcystis* colonies are basic units of cyanobacterial blooms. The colonies are constituted of *Microcystis* cells and heterotrophic bacteria, which vary in size and shape [5]. Morphological characteristics of cells and colonies are considered as species definition of *Microcystis*. The mass *Microcystis* bloom sample is a colony mixture with high complexity. Single colonies in nature are less complex in bacterial composition and more suitable for metagenomic study to reveal the natural mutualistic microsystem. Morphological studies in advance confirmed that colonies were consisted of *Microcystis* cells and numerous heterotrophic bacteria (Figure 3 and Figure S1). After which we constructed a comprehensive pipeline to construct and analyze the metagenomes of these natural mutualistic microsystem.

Due to the trace amount of nucleic acid within a single colony, extracting DNAs step by step is not practicable. Most nucleic acids of the single colony might be lost during extraction and external contaminants might be brought in. Therefore, we lysed single colonies and digested proteins and RNAs in tube. Nucleic acids were amplified using  $\Phi$ 29 DNA polymerase and random hexamer primers. This isothermal amplification has been extensively applied due to the high accuracy and great efficiency [28]. Metagenomes obtained in our study demonstrated that DNAs within single colonies were well amplified using MDA. However, amplification biases among DNA fragments and in GC content have been reported to be induced by MDA [45,46]. To promote the amplification uniformity and to recover more populations within single colonies, improved amplification approach such as micro-channel MDA ( $\mu$ cMDA) [47,48] or droplet MDA are desired [49,50].

Mapping results revealed that the nucleic acids in single *Microcystis* colonies were well recovered by our experimental pipeline. The six metagenomes mapped at least 67.87% of the two reference genomes. Higher coverage of all metagenomes using FACHB-1757 as reference indicated that geographical location is one of the dominant factors in genome similarity. FACHB-1757 is a strain of *M. panniformis* isolated from Lake Taihu in August 2011 [29]. The two metagenomes of *M. panniformis* colony both covered over 95% of the FACHB-1757 genome for at least five folds, indicating that colonies of the same *Microcystis* species at the same geographical location have the highest genome similarity. Metagenomes of the same species have similar coverage ratio regardless of which reference genome was used. Taxonomic analysis of metagenomes after alignment revealed similar results. All these

results supported that colonies of the same species are not only consistent in genome, but also similar in microbial community composition.

Metagenomes are complex genome mixtures of individuals. The assembly of metagenomes will result in a large set of independent contigs that are not easily aggregated into individuals [51]. The process which is referred as binning is applied to reduce the complexity of metagenomics data by grouping similar assembled contigs together and annotated to individual genomes [9,12]. Genomic features, such as GC content and tetranucleotide frequency are commonly used in assembly [52]. However, only a few organisms with distinct or even extreme features in base composition could be grouped in complex communities. Herein, we applied an alternative approach to separate organisms by using abundance differences of contigs among samples. This approach is referred as differential-coverage binning, which had been applied in recover genomes from samples processed using two different DNA extraction methods [53] and time series community samples [54]. Colony samples of different *Microcystis* species are the exactly parallel samples for differential-coverage binning. Therefore, we separated population bins by using the differential-coverage of contigs among colonies. In the 3D space of coverage, scaffolds clustered together represent putative population genome bins (Figure 1 and Figure S3). Although scaffolds could be clustered in the 2D space of coverage (Figure S6), 3D space provides a superior binning, in which scaffolds within a putative bin could be screened in the third dimension. Because scaffolds from different microorganisms in one putative bin seriously influence the following assembly result, a purified putative bin means a higher quality genome. The six high-quality genomes assembled demonstrated the high efficiency of our pipeline.

In total, eight population bins were identified which are greatly varied in GC content and abundance, and six of which were assembled into high-quality genomes. All these organisms are the component of the mutualistic system within single *Microcystis* colonies. The constitution of bacteria at the phylum level is consistent with the previous study [14].

## 4. Material and Methods

### 4.1. Sample Collection, Colony Classification, and Observation

A cyanobacterial bloom sample was collected from the water surface using a plastic bucket in Meiliang Bay of Lake Taihu in August 2015, and transported to the laboratory in 50 mL centrifuge tubes. Assemblages were rinsed with 0.02 M phosphate-buffered saline (PBS) to remove free-living bacteria. A bulk sample was obtained to evaluate and standardize the bias of MDA. Centrifuge tubes, culture dishes, and other materials were sterilized and dried before use. Single *Microcystis* colonies were classified based on species features under an optical microscope. *M. aeruginosa* and *M. wesenbergii* colonies were identified referring to Yu et al. [55]. The single colonies which had flattened irregular monolayers with small holes (in old colonies) were identified as *M. panniformis* [29] (Figure S1).

Colonies for SEM observation were prepared as follows. Colonies sample in 1 mL lake water were selected and washed with 0.02 M PBS, and then fixed with 2.5% glutaraldehyde at room temperature for 10 h, and post-fixed in 1% osmium tetroxide at 4 °C for 30 min, followed by rinsing with 0.02 M PBS. After that, the samples were dehydrated by ethanol series (50%, 70%, 95%, and 100%), and dried. Finally, the samples were coated with gold for analysis. The prepared samples were examined under an SEM (S-3000N, Hitachi, Japan).

### 4.2. Colony Lysis and DNA Amplification

The selected single *Microcystis* colonies were rinsed with 0.02 M PBS for three times and with 100 µL TES (50 mM Tris, 1 mM EDTA, 100 mM NaCl) for one time. After removing the supernatant, 5 µL TES and 5 µL lysozyme (10 mg/mL, cat. 62970, Sigma, St. Louis, MO, USA) was added to each colony and the tubes were incubated at 37 °C for 10 h. Subsequently, 0.2 µL proteinase K solution (10 mg/mL, cat. P6556, Sigma) was added to each colony to digest proteins by incubating at 50 °C for 3 h and was inactivated at 95 °C for 10 min. After that, RNAase A (cat. 19101, Qiagen, Hilden,

Germany) for 0.1  $\mu$ L was added to each tube to digest RNAs by incubating at 37 °C for 30 min and was inactivated at 95 °C for 10 min. DNA of the bulk cyanobacterial bloom sample were extracted using E.Z.N.A.<sup>®</sup> Water DNA Kit (cat. D5525, OMEGA, Norcross, GA, USA).

The 90  $\mu$ L MDA reaction buffer was added to each lysed colony with a final concentration of 1  $\times$   $\Phi$ 29 buffer (cat. M0269, NEB, Ipswich, MA, USA), 50  $\mu$ M N6 primer with two phosphorothioate bonds at the 3' end (Genscript, Nanjing, China), 1 mM dNTP (cat. N0447, Ipswich, MA, NEB) and 1 mg/mL BSA (cat. B9001, Ipswich, MA, NEB). To anneal the random hexamers to single colony DNA, we heated the reaction system at 95 °C for 10 min, and immediately put on ice for at least 30 min. Then 20 units of  $\Phi$ 29 DNA polymerase (cat. M0269, Ipswich, MA, NEB) were added to each tube with brief centrifuging. MDA reactions were carried out at 30 °C for 10 h and terminated at 65 °C for 10 min.

#### 4.3. Library Construction and Sequencing

Barcoded paired-end sequencing libraries for 500–700 bp in length were constructed with 2  $\mu$ g amplified DNA for each colony, and 2  $\mu$ g DNA for bulk sample according to the manufacturer's protocol (Illumina, San Diego, CA, USA). The resulting libraries were size-checked by Agilent 2100 Bioanalyzer system (Agilent, Santa Clara, CA, USA). The libraries were then subjected to paired-end sequencing on Illumina X10 platform (Illumina, San Diego, CA, USA) with the 150 bp read option, according to manufacturer's instructions.

Raw sequence data of the six samples reported in this study have been submitted to the National Center for Biotechnology Information (NCBI) Sequence Read Archive under accessions SRX4337635, SRX4337636, SRX4337637, SRX4337638, SRX4337639, and SRX4337640.

#### 4.4. Read Trimming, Alignment

The bioinformatics analysis strategy used to obtain genomes of mutualistic bacteria in single *Microcystis* colonies is detailed in Figure 2. Sequencing reads in FASTQ format were trimmed with trimmomatic-0.36 [56] under following steps and parameters: removing sequencing adapters, cutting the start or end of a read if the quality score is below 3, scanning the reads with a 4-base wide sliding window and cutting the windows whose average quality score per base are below 20, and dropping the reads whose length are below 50 bp.

The trimmed reads were aligned to NIES843 [3] and FACHB-1757 [29] using bwa mem [57]. The ratio of aligned reads and coverage of reference were calculated for each metagenome.

#### 4.5. Taxonomic Analysis and Clustering

The trimmed reads of each metagenome were aligned to non-redundant protein database (nr) [58] using DIAMOND-0.7.11 [59]. Alignment files were loaded into MEGAN5 [60] to perform taxonomic analysis, and taxonomic information at genus level was extracted. Six metagenomes were aligned separately in this step and clustered by STAMP-2.1.3 [61].

#### 4.6. De Novo Metagenome Assembly and Scaffold Coverage

The filtered reads were assembled using SPADes-3.9.0 [62], with a k-mer of 101 and a minimum scaffold length of 1 kbp. The six metagenomes were assembled unitedly.

Reads were mapped to scaffolds using bwa mem [57] and the scaffold coverage information was reported by qualimap-2.1.1 (<http://qualimap.bioinfo.cipf.es/>). Relative metagenome abundance of each genome bin was calculated as the percentage of metagenome reads mapping to the individual genomes compared to the total number of metagenome reads.

The number of scaffold of specific GC content was counted for both amplified metagenome and bulk metagenome to calculate standardization factor of relative abundance.

#### 4.7. Conserved Marker Genes Identification

Firstly, open reading frames were predicted in the assembled scaffolds using the metagenome version of Prodigal [63]. Secondly, proteins were identified using HMMER3 (<http://hmmer.janelia.org/>) with the default settings by searching a set of 107 HMMs of essential single-copy genes [64] against the predicted open reading frames, except the trusted cutoff was used (-cut\_tc). Using BLASTP, the identified proteins were taxonomic classified against the RefSeq (version 84) protein database with a maximum e-value cutoff of 1e-5. Lastly, from the BLAST.xml output file, class level taxonomic assignments were extracted using MEGAN5 [60].

#### 4.8. Sequence Binning

Sequence binning was facilitated by using the metagenomes generated from single *Microcystis* colonies of three species. The first metagenomes of the two duplicates of each species were selected for differential-coverage binning. Difference in colony-specific bacteria abundance among the three single *Microcystis* colonies resulted in different relative abundance of the populations in the three metagenomes. By independently mapping the reads from each selected metagenome to the assembled scaffolds, coverages of scaffolds in the three metagenomes were calculated. Distinct grouping was identified by plotting the three coverage against one the others in R. The grouped scaffolds suggest that they own similar coverage in all three metagenomes, indicating that they are from the same species. Initial groups were selected by coloring scaffolds according to their taxonomic affiliation as determined above. For some groups which contain a lot of duplicated genes, tetranucleotide frequencies and GC content were used in the next step to purify these initial groups.

For comparison, scaffolds were also binned based on their GC content and overall coverage.

#### 4.9. Re-Assembly

Re-assembly was performed by extracting reads mapping to scaffolds of a particular bin and all associated paired-end reads. The extracted reads were used for a new de novo assembly using SPADes-3.9.0 [62] with a k-mer of 101. The expected coverage and coverage cutoff parameters were adjusted individually by running multiple assemblies. To validate the assemblies, CheckM was used to assess genome quality [65].

### 5. Conclusions

In this work, we proposed a universal metagenome construction and analysis pipeline for uncultured single *Microcystis* colonies, the basic mutualistic units in cyanobacterial blooms. By the isothermal DNA amplification and differential-coverage binning of scaffolds, six high-quality genomes (completeness > 90%) were assembled, including some rare members (< 1%). The analysis of wild single *Microcystis* colonies improved our understanding of the mutualistic relationship within *Microcystis* colonies and the mechanism of cyanobacterial blooms.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/20/20/5047/s1>.

**Author Contributions:** Conceptualization, J.T. and Z.L.; methodology, S.G. and J.Z.; software, L.C., Y.T. and C.B.; data curation, N.L.; Writing—Original Draft preparation, J.T.; Writing—Review and Editing, J.T. and Z.L.; supervision, Z.L.; project administration, J.T.; funding acquisition, J.T. and J.Z.

**Funding:** This research was funded by “National Natural Science Foundation of China, grant number 61571121 and 31600096” and “Fundamental Research Funds for the Central Universities of China, grant number 2242019K41038”.

**Acknowledgments:** We would like to thank Chunpeng He for the technical assistance in scanning electron microscope imaging.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

MDA	multiple displacement amplification
3D	three-dimension
SEM	scanning electron microscope
μcMDA	micro-channel MDA
PBS	phosphate-buffered saline
NCBI	National Center for Biotechnology Information

## References

1. Blaha, L.; Babica, P.; Marsalek, B. Toxins produced in cyanobacterial water blooms-toxicity and risks. *Interdiscip. Toxicol.* **2009**, *2*, 36–41. [[CrossRef](#)] [[PubMed](#)]
2. Pope, P.B.; Patel, B.K. Metagenomic analysis of a freshwater toxic cyanobacteria bloom. *FEMS Microbiol. Ecol.* **2008**, *64*, 9–27. [[CrossRef](#)] [[PubMed](#)]
3. Kaneko, T.; Nakajima, N.; Okamoto, S.; Suzuki, I.; Tanabe, Y.; Tamaoki, M.; Nakamura, Y.; Kasai, F.; Watanabe, A.; Kawashima, K.; et al. Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* nies-843. *DNA Res.* **2007**, *14*, 247–256. [[CrossRef](#)] [[PubMed](#)]
4. Mur, L.R. Some aspects of the ecophysiology of cyanobacteria. *Ann. Microbiol.* **1983**, *134*, 61–72. [[CrossRef](#)]
5. Frangeul, L.; Quillardet, P.; Castets, A.M.; Humbert, J.F.; Matthijs, H.C.; Cortez, D.; Tolonen, A.; Zhang, C.C.; Gribaldo, S.; Kehr, J.C.; et al. Highly plastic genome of *Microcystis aeruginosa* pcc 7806, a ubiquitous toxic freshwater cyanobacterium. *BMC Genom.* **2008**, *9*, 274. [[CrossRef](#)]
6. Zhu, W.; Li, M.; Luo, Y.G.; Dai, X.X.; Guo, L.L.; Xiao, M.; Huang, J.Y.; Tan, X. Vertical distribution of *Microcystis* colony size in lake taihu: Its role in algal blooms. *J. Gt. Lakes Res.* **2014**, *40*, 949–955. [[CrossRef](#)]
7. Shen, H.; Niu, Y.; Xie, P.; Tao, M.; Yang, X. Morphological and physiological changes in *Microcystis aeruginosa* as a result of interactions with heterotrophic bacteria. *Freshw. Biol.* **2011**, *56*, 1065–1080. [[CrossRef](#)]
8. Hugenholtz, P. Exploring prokaryotic diversity in the genomic era. *Genome Biol.* **2002**, *3*, reviews0003.0001-0003.0008. [[CrossRef](#)]
9. Kunin, V.; Copeland, A.; Lapidus, A.; Mavromatis, K.; Hugenholtz, P. A bioinformatician's guide to metagenomics. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 557–578. [[CrossRef](#)] [[PubMed](#)]
10. Tyson, G.W.; Jarrod, C.; Philip, H.; Allen, E.E.; Ram, R.J.; Richardson, P.M.; Solovyev, V.V.; Rubin, E.M.; Rokhsar, D.S.; Banfield, J.F. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **2004**, *428*, 37–43. [[CrossRef](#)] [[PubMed](#)]
11. Héctor, G.M.; Natalia, I.; Victor, K.; Falk, W.; Barry, K.W.; Mchardy, A.C.; Christine, Y.; Shaomei, H.; Salamov, A.A.; Ernest, S. Metagenomic analysis of two enhanced biological phosphorus removal (ebpr) sludge communities. *Nat. Biotechnol.* **2006**, *24*, 1263–1269.
12. Mande, S.S.; Monzoorul Haque, M.; Tarini Shankar, G. Classification of metagenomic sequences: Methods and challenges. *Brief Bioinform.* **2012**, *13*, 669–681. [[CrossRef](#)] [[PubMed](#)]
13. Eiler, A.; Heinrich, F.; Bertilsson, S. Coherent dynamics and association networks among lake bacterioplankton taxa. *ISME J.* **2012**, *6*, 330–342. [[CrossRef](#)] [[PubMed](#)]
14. Steffen, M.M.; Li, Z.; Effler, T.C.; Hauser, L.J.; Boyer, G.L.; Wilhelm, S.W. Comparative metagenomics of toxic freshwater cyanobacteria bloom communities on two continents. *PLoS ONE* **2012**, *7*, e44002. [[CrossRef](#)] [[PubMed](#)]
15. Mou, X.Z.; Jacob, J.S.; Lu, X.X.; Robbins, S.; Sun, S.L.; Ortiz, J.D. Diversity and distribution of free-living and particle-associated bacterioplankton in sandusky bay and adjacent waters of lake erie western basin. *J. Gt. Lakes Res.* **2013**, *39*, 352–357. [[CrossRef](#)]
16. Kurmayer, R.; Dittmann, E.; Fastner, J.; Chorus, I. Diversity of microcystin genes within a population of the toxic cyanobacterium *Microcystis* spp. in lake wannsee (berlin, germany). *Microb Ecol.* **2002**, *43*, 107–118. [[CrossRef](#)] [[PubMed](#)]
17. Via-Ordorika, L.; Fastner, J.; Kurmayer, R.; Hisbergues, M.; Dittmann, E.; Komarek, J.; Erhard, M.; Chorus, I. Distribution of microcystin-producing and non-microcystin-producing *Microcystis* sp. in european freshwater bodies: Detection of microcystins and microcystin genes in individual colonies. *Syst. Appl. Microbiol.* **2004**, *27*, 592–602. [[CrossRef](#)]

18. Lara, Y.; Lambion, A.; Menzel, D.; Codd, G.A.; Wilmotte, A. A cultivation-independent approach for the genetic and cyanotoxin characterization of colonial cyanobacteria. *Aquat. Microb. Ecol.* **2013**, *69*, 135–143. [[CrossRef](#)]
19. Otsuka, S.; Suda, S.; Li, R.; Matsumoto, S.; Watanabe, M.M. Morphological variability of colonies of *Microcystis* morphospecies in culture. *J. Gen. Appl. Microbiol.* **2000**, *46*, 39–50. [[CrossRef](#)]
20. Doers, M.P.; Parker, D.L. Properties of *Microcystis-Aeruginosa* and *m flos-aquae* (cyanophyta) in culture: Taxonomic implications. *J. Phycol.* **1988**, *24*, 502–508.
21. Parker, D.L. Improved procedures for the cloning and purification of *Microcystis*-cultures (cyanophyta). *J. Phycol.* **1982**, *18*, 471–477. [[CrossRef](#)]
22. Parker, D.L.; Mihalick, J.E.; Plude, J.L.; Plude, M.J.; Clark, T.P.; Egan, L.; Flom, J.J.; Rai, L.C.; Kumar, H.D. Sorption of metals by extracellular polymers from the cyanobacterium *Microcystis aeruginosa* f. *Flos-aquae* strain c3-40. *J. Appl. Phycol.* **2000**, *12*, 219–224. [[CrossRef](#)]
23. Plude, J.L.; Parker, D.L.; Schommer, O.J.; Timmerman, R.J.; Hagstrom, S.A.; Joers, J.M.; Hnasko, R. Chemical characterization of polysaccharide from the slime layer of the cyanobacterium *Microcystis-flos-aquae*-c3-40. *Appl. Environ. Microbiol.* **1991**, *57*, 1696–1700. [[PubMed](#)]
24. Kurmayer, R.; Christiansen, G.; Chorus, I. The abundance of microcystin-producing genotypes correlates positively with colony size in *Microcystis* sp. and determines its microcystin net production in lake wannsee. *Appl. Environ. Microbiol.* **2003**, *69*, 787–795. [[CrossRef](#)] [[PubMed](#)]
25. Rohrlack, T.; Henning, M.; Kohl, J.G. Mechanisms of the inhibitory effect of the cyanobacterium *Microcystis aeruginosa* on daphnia galeata's ingestion rate. *J. Plankton Res.* **1999**, *21*, 1489–1500. [[CrossRef](#)]
26. Xie, M.; Ren, M.; Yang, C.; Yi, H.; Li, Z.; Li, T.; Zhao, J. Metagenomic analysis reveals symbiotic relationship among bacteria in *Microcystis*-dominated community. *Front. Microbiol.* **2016**, *7*, 56. [[CrossRef](#)] [[PubMed](#)]
27. Dean, F.B.; Hosono, S.; Fang, L.H.; Wu, X.H.; Faruqi, A.F.; Bray-Ward, P.; Sun, Z.Y.; Zong, Q.L.; Du, Y.F.; Du, J.; et al. Comprehensive human genome amplification using multiple displacement amplification. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5261–5266. [[CrossRef](#)] [[PubMed](#)]
28. Hosono, S.; Faruqi, A.F.; Dean, F.B.; Du, Y.F.; Sun, Z.Y.; Wu, X.H.; Du, J.; Kingsmore, S.F.; Egholm, M.; Lasken, R.S. Unbiased whole-genome amplification directly from clinical samples. *Genome Res.* **2003**, *13*, 954–964. [[CrossRef](#)] [[PubMed](#)]
29. Zhang, J.Y.; Guan, R.; Zhang, H.J.; Li, H.; Xiao, P.; Yu, G.L.; Du, L.; Cao, D.M.; Zhu, B.C.; Li, R.H.; et al. Complete genome sequence and genomic characterization of *Microcystis panniformis* fachb 1757 by third-generation sequencing. *Stand. Genom. Sci.* **2016**, *11*, 11. [[CrossRef](#)]
30. Berg, K.A.; Lyra, C.; Sivonen, K.; Paulin, L.; Suomalainen, S.; Tuomi, P.; Rapala, J. High diversity of cultivable heterotrophic bacteria in association with cyanobacterial water blooms. *ISME J.* **2009**, *3*, 314–325. [[CrossRef](#)]
31. Cornet, L.; Meunier, L.; Van Vlierberghe, M.; Leonard, R.R.; Durieu, B.; Lara, Y.; Misztak, A.; Sirjacobs, D.; Javaux, E.J.; Philippe, H.; et al. Consensus assessment of the contamination level of publicly available cyanobacterial genomes. *PLoS ONE* **2018**, *13*, e0200323. [[CrossRef](#)] [[PubMed](#)]
32. Kim, J.-J.; Alkawally, M.; Brady, A.L.; Rijpstra, W.I.C.; Sinninghe Damsté, J.S.; Dunfield, P.F. *Chryseolinea serpens* gen. Nov., sp. Nov., a member of the phylum bacteroidetes isolated from soil. *Int. J. Syst. Evol. Microbiol.* **2013**, *63*, 654–660.
33. Shi, L.M.; Huang, Y.X.; Zhang, M.; Shi, X.L.; Cai, Y.F.; Gao, S.L.; Tang, X.M.; Chen, F.Z.; Lu, Y.P.; Kong, F.X. Large buoyant particles dominated by cyanobacterial colonies harbor distinct bacterial communities from small suspended particles and free-living bacteria in the water column. *Microbiologyopen* **2018**, *7*, e00608. [[CrossRef](#)] [[PubMed](#)]
34. Fan, Q.; Xiao, H.J.; Qiang, W.U.; Wang, S.J.; Pengfu, L.I. Characterization of epiphytic bacteria associated with colonial *Microcystis*. *J. Lake Sci.* **2017**, *29*, 617–624.
35. Yamamoto, Y.; Niizuma, S.; Kuroda, N.; Sakamoto, M. Occurrence of heterotrophic bacteria causing lysis of cyanobacteria in a eutrophic lake. *JPN J. Phycol.* **1993**, *41*, 46–48.
36. Sedmak, B.; Kosi, G. Microcystins in slovene freshwaters (central europe)—First report. *Nat. Toxins* **1997**, *5*, 64–73. [[CrossRef](#)]
37. Vasconcelos, V.; Morais, J.; Vale, M. Microcystins and cyanobacteria trends in a 14 year monitoring of a temperate eutrophic reservoir (aguieira, portugal). *J. Environ. Monit.* **2011**, *13*, 668–672. [[CrossRef](#)] [[PubMed](#)]

38. Yarmoshenko, L.P.; Kureyshevich, A.V.; Yakushin, V.M. *Microcystis botrys* and *lemmermanniella flexa*—New species of cyanoprokaryota for the flora of ukraine in phytoplankton of the kanev reservoir. *Hydrobiol. J.* **2013**, *49*, 39–44. [[CrossRef](#)]
39. Agha, R.; del Mar Labrador, M.; de los Ríos, A.; Quesada, A. Selectivity and detrimental effects of epiphytic *pseudanabaena* on *Microcystis* colonies. *Hydrobiologia* **2016**, *777*, 139–148. [[CrossRef](#)]
40. Yang, C.; Wang, Q.; Simon, P.N.; Liu, J.; Liu, L.; Dai, X.; Zhang, X.; Kuang, J.; Igarashi, Y.; Pan, X.; et al. Distinct network interactions in particle-associated and free-living bacterial communities during a *Microcystis aeruginosa* bloom in a plateau lake. *Front. Microbiol.* **2017**, *8*, 1202. [[CrossRef](#)]
41. Shi, L.M.; Huang, Y.X.; Zhang, M.; Yu, Y.; Lu, Y.P.; Kong, F.X. Bacterial community dynamics and functional variation during the long-term decomposition of cyanobacterial blooms in-vitro. *Sci. Total Environ.* **2017**, *598*, 77–86. [[CrossRef](#)] [[PubMed](#)]
42. Zhang, H.; Jia, J.; Chen, S.; Huang, T.; Wang, Y.; Zhao, Z.; Feng, J.; Hao, H.; Li, S.; Ma, X. Dynamics of bacterial and fungal communities during the outbreak and decline of an algal bloom in a drinking water reservoir. *Int. J. Environ. Res. Public Health* **2018**, *15*, 361. [[CrossRef](#)] [[PubMed](#)]
43. Coman, C.; Bica, A.; Druga, B.; Barbu-Tudoran, L.; Dragos, N. Methodological constraints in the molecular biodiversity study of a thermomineral spring cyanobacterial mat: A case study. *Antonie Van Leeuwenhoek* **2011**, *99*, 271–281. [[CrossRef](#)] [[PubMed](#)]
44. Shi, L.; Cai, Y.; Kong, F.; Yu, Y. Specific association between bacteria and buoyant *Microcystis* colonies compared with other bulk bacterial communities in the eutrophic lake taihu, china. *Environ. Microbiol. Rep.* **2012**, *4*, 669–678.
45. Spits, C.; Le Caignec, C.; de Rycke, M.; Van Haute, L.; van Steirteghem, A.; Liebaers, I.; Sermon, K. Optimization and evaluation of single-cell whole, genome multiple displacement amplification. *Hum. Mutat.* **2006**, *27*, 496–503. [[CrossRef](#)]
46. Pinard, R.; de Winter, A.; Sarkis, G.J.; Gerstein, M.B.; Tartaro, K.R.; Plant, R.N.; Egholm, M.; Rothberg, J.M.; Leamon, J.H. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. *BMC Genom.* **2006**, *7*, 216. [[CrossRef](#)] [[PubMed](#)]
47. Li, J.J.; Shi, X.L.; Qiao, Y.; Chen, L.; Duan, M.Q.; Hou, Y.; Ge, Q.Y.; Tao, Y.H.; Tu, J.; Lu, Z.H. 1d-reactor decentralized mda for uniform and accurate whole genome amplification. *Anal. Chem.* **2017**, *89*, 10147–10152. [[CrossRef](#)]
48. Li, J.J.; Lu, N.; Tao, Y.H.; Duan, M.Q.; Qiao, Y.; Xu, Y.M.; Ge, Q.Y.; Bi, C.W.; Fu, J.Y.; Tu, J.; et al. Accurate and sensitive single-cell-level detection of copy number variations by micro-channel multiple displacement amplification (mu cmda). *Nanoscale* **2018**, *10*, 17933–17941. [[CrossRef](#)]
49. Sidore, A.M.; Lan, F.; Lim, S.W.; Abate, A.R. Enhanced sequencing coverage with digital droplet multiple displacement amplification. *Nucleic Acids Res.* **2016**, *44*, 66. [[CrossRef](#)]
50. Rhee, M.; Light, Y.K.; Meagher, R.J.; Singh, A.K. Digital droplet multiple displacement amplification (ddmda) for whole genome sequencing of limited DNA samples. *PLoS ONE* **2016**, *11*, e0153699. [[CrossRef](#)]
51. Brinig, M.M.; Lepp, P.W.; Ouverney, C.C.; Armitage, G.C.; Relman, D.A. Prevalence of bacteria of division tm7 in human subgingival plaque and their association with disease. *Appl. Environ. Microbiol.* **2003**, *69*, 1687–1694. [[CrossRef](#)] [[PubMed](#)]
52. Marcy, Y.; Ouverney, C.; Bik, E.M.; Losekann, T.; Ivanova, N.; Martin, H.G.; Szeto, E.; Platt, D.; Hugenholtz, P.; Relman, D.A.; et al. Dissecting biological “dark matter” with single-cell genetic analysis of rare and uncultivated tm7 microbes from the human mouth. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11889–11894. [[CrossRef](#)] [[PubMed](#)]
53. Albertsen, M.; Hugenholtz, P.; Skarshewski, A.; Nielsen, K.L.; Tyson, G.W.; Nielsen, P.H. Genome sequences of rare, uncultured bacteria obtained by differential-coverage binning of multiple metagenomes. *Nat. Biotechnol.* **2013**, *31*, 533–538. [[CrossRef](#)] [[PubMed](#)]
54. Sharon, I.; Morowitz, M.J.; Thomas, B.C.; Costello, E.K.; Relman, D.A.; Banfield, J.F. Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res.* **2013**, *23*, 111–120. [[CrossRef](#)] [[PubMed](#)]
55. Yu, G.; Song, L.; Li, R. Taxonomic notes on water bloom forming *Microcystis* species (cyanophyta) from china—An example from samples of the dianchi lake. *Acta Phytotaxon. Sin.* **2007**, *45*, 727–741. [[CrossRef](#)]
56. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]

57. Li, H.; Durbin, R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
58. Pruitt, K.D.; Tatusova, T.; Maglott, D.R. Ncbi reference sequences (refseq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **2007**, *35*, 61–65. [[CrossRef](#)]
59. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using diamond. *Nat. Methods* **2015**, *12*, 59–60. [[CrossRef](#)]
60. Huson, D.H.; Beier, S.; Flade, I.; Gorska, A.; El-Hadidi, M.; Mitra, S.; Ruscheweyh, H.J.; Tappu, R. Megan community edition-interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* **2016**, *12*, e1004957. [[CrossRef](#)]
61. Parks, D.H.; Tyson, G.W.; Hugenholtz, P.; Beiko, R.G. Stamp: Statistical analysis of taxonomic and functional profiles. *Bioinformatics* **2014**, *30*, 3123–3124. [[CrossRef](#)] [[PubMed](#)]
62. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pribelski, A.D.; et al. Spades: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [[CrossRef](#)] [[PubMed](#)]
63. Hyatt, D.; LoCascio, P.F.; Hauser, L.J.; Uberbacher, E.C. Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* **2012**, *28*, 2223–2230. [[CrossRef](#)] [[PubMed](#)]
64. Dupont, C.L.; Rusch, D.B.; Yooseph, S.; Lombardo, M.J.; Richter, R.A.; Valas, R.; Novotny, M.; Yee-Greenbaum, J.; Selengut, J.D.; Haft, D.H.; et al. Genomic insights to sar86, an abundant and uncultivated marine bacterial lineage. *ISME J.* **2012**, *6*, 1186–1199. [[CrossRef](#)] [[PubMed](#)]
65. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. Checkm: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).