

OPEN

# Structural and functional microbial diversity along a eutrophication gradient of interconnected lakes undergoing anthropopressure

Bartosz Kiersztyn, Ryszard Chróst, Tomasz Kaliński, Waldemar Siuda, Aleksandra Bukowska , Grzegorz Kowalczyk & Karolina Grabowska

We present the results of an analysis of the 16S rRNA-based taxonomical structure of bacteria together with an analysis of carbon source utilization ability using EcoPlate (Biolog, USA) metabolic fingerprinting assessment against the backdrop of physicochemical parameters in fifteen interconnected lakes. The lakes exhibit a wide spectrum of trophic gradients and undergo different intensities of anthropopressure. Sequences of V3–V4 16S rRNA genes binned by taxonomic assignment to family indicated that bacterial communities in the highly eutrophicated lakes were distinctly different from the bacterial communities in the meso-eutrophic lakes (ANOSIM  $r = 0.99$ ,  $p = 0.0002$ ) and were characterized by higher richness and more diverse taxonomical structure. Representatives of the *Actinobacteria*, *Proteobacteria*, *Cyanobacteria*, *Planctomycetes*, *Verrucomicrobia*, *Bacteroides* phyla predominated. In most cases their relative abundance was significantly correlated with lake trophic state. We found no similar clear relationship of community-level physiological profiling with lake trophic state. However, we found some significant links between the taxonomic and metabolic structure of the microbes in the studied lakes (Mantel's correlation  $r = 0.22$ ,  $p = 0.006$ ). The carbon source utilization ability of the studied microorganisms was affected not only by the taxonomic groups present in the lakes but also by various characteristics like a high  $\text{PO}_4^{3-}$  concentration inhibiting the utilization of phosphorylated carbon.

Rapid anthropogenic eutrophication of natural freshwater environments has been an increasing problem in the last few decades<sup>1</sup>. Eutrophication may lead to serious changes in microbial community composition, affecting the functioning of the microbial loop and thus the entire aquatic food web. Simultaneous monitoring of phylogenetic and functional (metabolic) diversity is crucial for understanding human influences on the microbial community through the acceleration of eutrophication processes. Over the past 20 years, many studies have been carried out on the diversity of bacteria in lakes with different trophic states<sup>2–8</sup>. Authors often present results on geographically isolated lakes located in different regions of the world with different catchment areas and, as a consequence, different physicochemical conditions. This prevents the direct comparison of the impacts of the trophic state on the microorganisms living in these separated ecosystems. However, an example of an exception is the work of Zwirgmaier *et al.*<sup>9</sup>, which described the seasonal changes in bacterial diversity in 19 connected pre-Alpine, hard water, alkaline lakes based on pyrosequencing of 16S rRNA genes. Nevertheless, despite the years of research focusing on lake bacteria taxonomic composition, there is relatively little information regarding the connection of the phylogenetic structure of bacteria with their metabolic functions<sup>10</sup>. Metatranscriptomics provides new possibilities for the analysis of potentially active metabolic pathways; nevertheless, it does not allow the determination of the actual rates of specific metabolic processes. For example, it does not directly take into account the competitive inhibition of previously produced enzymes or the influence of environmental conditions (temperature, pH, oxido-redox potential) on metabolic pathways. Therefore, it seems to be important to look for a link between the phylogenetic potential of microorganisms and the various directly measured metabolic process indicators. Using massive parallel sequencing of 16S rRNA genes<sup>11</sup> together with metabolic fingerprinting, based on, for example,

Microbial Ecology and Environmental Biotechnology Department, Institute of Botany, Faculty of Biology, University of Warsaw; Biological and Chemical Research Centre, Warszawa, Poland. Correspondence and requests for materials should be addressed to B.K. (email: [kiersztyn@biol.uw.edu.pl](mailto:kiersztyn@biol.uw.edu.pl))

Biolog EcoPlate methodology<sup>12–17</sup>, it is becoming possible, to some extent, to make the connection between structure and function and to answer the question ‘who is doing what?’, at least at the community level. Answering this question is not simple because the relationships between structure and function are complex and subject to the impact of various additional conditions<sup>18</sup>. For example, Dickerson and Williams<sup>19</sup> showed, using Biolog EcoPlate methodology, that the metabolic profiles of microorganisms from two lakes, meso- and oligotrophic, were different from those in a eutrophic lake. At the same time, the eutrophic lake showed a larger representation of *Flavobacteriia* and *Gammaproteobacteria*, but it was impossible to create a clear picture of the relationships between microorganism phylogeny and physiology. Understanding these complex relationships and the way they are influenced by human activity disturbing the physicochemical conditions of reservoirs and their trophic status is crucial for understanding the changes in biochemical cycles as a consequence of anthropopressure.

We present an analysis of the community-level physiological profiling and phylogenetic diversity of the microbial community in 15 interconnected lakes located in the temperate climate zone of north-eastern Poland. These lakes belong to the Great Masurian Lakes System (GMLS), which was formed by ice sheets during the Pomeranian phase of the glaciation in the late Pleistocene<sup>20</sup>. The GMLS serves as a useful model for comparative studies of lakes with different degrees of eutrophication resulting from anthropogenic stress. In addition, the lakes in the GMLS are homogenous in terms of their time of origin, geological location, catchment characteristics and climate influence. Different levels of anthropogenic pressure on lakes during the past hundred years<sup>21</sup> have resulted in varying degrees of eutrophication in the Masurian lakes. In particular, stronger anthropopressure in the southern part of the GMLS has resulted in a higher trophic state in the southern part of the GMLS compared to the northern lakes<sup>22</sup>. An exception is Śniardwy lake, which is located at the southern end of the system. Śniardwy lake is a shallow, large lake that, despite its location and strong anthropopressure, remains in a meso-eutrophic state. We devote special attention to this particular lake because, in some respects, it combines the features of the meso-eutrophic northern and eutrophic southern lakes in the GMLS.

In this paper, we present a parallel assessment of the physiological and functional diversity in the GMLS along a eutrophication gradient in the near-shore littoral zones of the lakes. We address two main research hypotheses: (1) the taxonomic and metabolic diversities of microorganisms depend strongly on the trophic state of the studied lakes, and (2) differences in the metabolic profiles of the microorganisms are determined by the taxonomic composition—there is a strong relationship between structure and function.

## Results

The TSI values calculated for the sampling sites are shown in Table 1. The studied lakes were characterized by a wide spectrum of TSI values from mesotrophy to advanced eutrophy. For further analysis, the lakes were divided into two groups according to their trophic status: meso-eutrophic (ME), consisting of seven lakes (TSI from 40 to 53) and eutrophic (E), consisting of eight lakes (TSI from 56 to 60). We assessed the characteristics of the sampling sites taking into consideration the basic physicochemical characteristics. The following parameters were measured: concentration of dissolved organic carbon (DOC),  $P-PO_4^{3-}$ , total phosphorus (TP), total nitrogen (TN), dissolved oxygen, turbidity, chlorophyll *a*, conductivity, pH, and temperature (Table 1). The results of the PCA (Fig. 1a,b) indicate that the lakes grouped according to the degree of eutrophication to E and ME were significantly different in terms of their various physicochemical conditions (Fig. 1a). Focusing on supplementary values, we found that higher TSI values were related to bacterial (LB) and heterotrophic nanoflagellate (HNF) abundances (Fig. 1b).

The percentages of the dominant phyla of the domain *Bacteria* in each lake are shown in Fig. 2. Only the phyla whose relative percentage exceeded 1% were taken into account as individual units. We found three dominant phyla - *Actinobacteria*, *Proteobacteria* and *Cyanobacteria*. The most abundant phylum was *Actinobacteria*, contributing from 20% in highly eutrophicated lake Beldany to more than 40% in less eutrophicated lakes (Przystań, Śniardwy, Boczne). We found a significant positive correlation between mean TSI values and the relative abundance of the phyla *Planctomycetes*, *Verrucomicrobia*, *Cyanobacteria* and *Chloroflexi* and a significant negative correlation between TSI and microorganisms belonging to the phyla *Actinobacteria* and *Bacteroides*. This indicated that the dominant taxon might be affected by water quality parameters related to the trophic state of the sampling area.

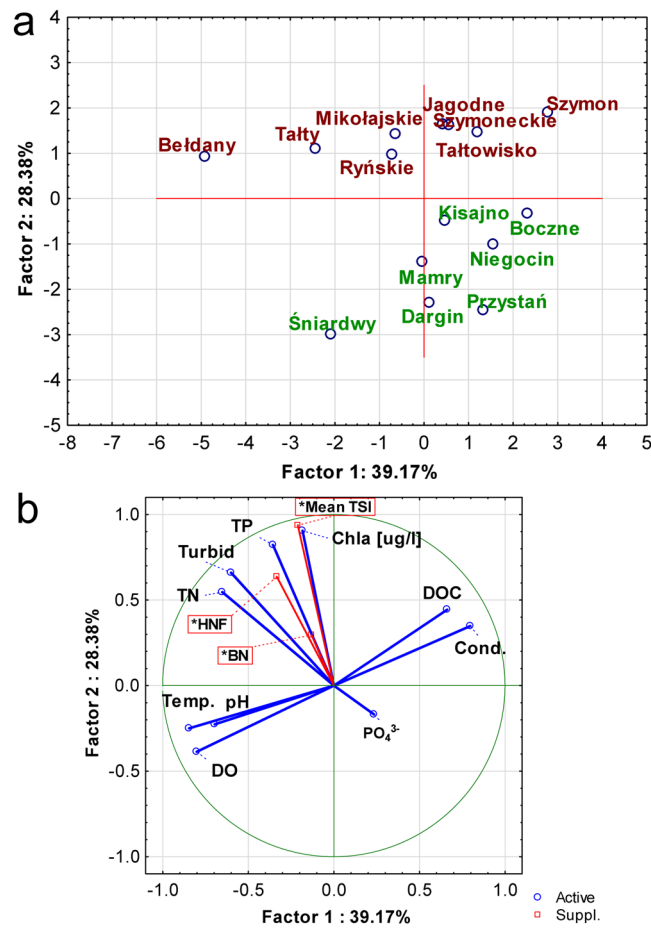
The Bray-Curtis-based NMDS analysis of sequence data, binned by taxonomic assignment to family, indicated that the bacterial communities living in the eutrophic lakes (E) were significantly distinct from the bacterial communities living in the meso-eutrophic (ME) lakes (Fig. 3a). The difference between eutrophic and meso-eutrophic lakes was statistically significant (ANOSIM  $r = 0.99$ ,  $p = 0.0002$ ). After excluding 25 families belonging to the phylum *Cyanobacteria* from the analysis, we obtained similar results. The differences between ME and E lakes were even more clear (ANOSIM  $r = 0.998$ ,  $p = 0.0001$ , lower NMDS stress, Fig. 3b). To determine which taxonomic groups are mainly responsible for the observed differences between eutrophic and meso-eutrophic lakes, the SIMPER test was used. The taxonomic groups mostly responsible for the observed differences between ME and E lakes are presented in Table 2. The microorganisms belonging to the families presented in Table 2 were responsible for more than 78% of the differences observed between eutrophic and meso-eutrophic lakes. The four bacterial families with the highest influence are the *ACK-M1* family of phylum *Actinobacteria*, *Pelagibacteraceae* of the phylum *Proteobacteria*, *Chthoniobacteraceae* of the phylum *Verrucomicrobia* and *Pseudanabaenaceae* of the phylum *Cyanobacteria*.

We found that at the family level, the OTU richness and diversity were positively correlated with the mean TSI of the studied lakes. A higher trophic state of lakes was generally associated with both increased taxonomical richness and a higher Simpson's diversity index, indicating a more equal distribution of detected families (Table 3).

The average relative respiration of carbon sources assigned to the basic groups of compounds is shown in Fig. 4. We did not find a strong relationship between the respiration ability of microorganisms in association with 31 different carbon sources and the trophic status of the studied lakes. The Bray-Curtis-based NMDS analysis

	LAKE	Przystań	Mamry	Dargin	Kisajno	Niegocin
Area [ha]		115	2 504	3 030	1 896	2 600
Max depth [m]		22.8	43.8	37.6	25	39.7
Coordinates of sampling location		54.207241, 21.657892	54.157234, 21.723144	54.145939, 21.731939	54.042000, 21.738594	54.009215, 21.738598
Mean TSI		42 ± 2	40 ± 9	43 ± 1	53 ± 3	49 ± 3
Trophic state		ME	ME	ME	ME	ME
Temp [°C]		21.35 ± 0.06	21.68 ± 0.141	22.50 ± 0.02	21.64 ± 1.56	22.38 ± 0.74
Cond [µScm <sup>2</sup> ]		297.00 ± 0.0	295.00 ± 0.0	302.00 ± 0.0	324.00 ± 6.0	370.00 ± 1.4
Oxygen [mgL <sup>-1</sup> ]		9.77 ± 0.03	10.05 ± 0.03	10.19 ± 0.0	10.27 ± 1.08	10.47 ± 0.26
pH		8.38 ± 0.01	8.32 ± 0.05	8.50 ± 0.0	8.40 ± 0.09	8.45 ± 0.0
Turbidity [NTU]		0.03 ± 0.06	4.35 ± 5.3	0.30 ± 0.14	0.87 ± 0.50	1.20 ± 0.14
Chl a [µgL <sup>-1</sup> ]		3.77 ± 0.40	0.94 ± 1.13	3.37 ± 0.57	12.26 ± 1.69	5.29 ± 0.28
TP [µgL <sup>-1</sup> ]		15.05 ± 0.53	15.67 ± 1.34	15.67 ± 0.53	30.16 ± 1.54	26.97 ± 0.64
P-PO <sub>4</sub> <sup>3-</sup> [µgL <sup>-1</sup> ]		3.68 ± 0.35	3.13 ± 0.21	3.48 ± 0.62	3.58 ± 0.30	14.26 ± 0.92
TN [µgL <sup>-1</sup> ]		367 ± 58	633 ± 115	533 ± 152	650 ± 71	600 ± 0
DOC [mgL <sup>-1</sup> ]		8.43 ± 0.07	8.09 ± 0.07	9.11 ± 0.08	9.23 ± 0.1	11.26 ± 0.03
HNF [cellmL <sup>-1</sup> ]		3.92 ± 0.77 × 10 <sup>3</sup>	2.10 ± 0.76 × 10 <sup>4</sup>	1.47 ± 0.2 × 10 <sup>4</sup>	2.69 ± 0.44 × 10 <sup>4</sup>	2.13 ± 0.32 × 10 <sup>4</sup>
Bacteria Number [cellmL <sup>-1</sup> ]		3.9 ± 0.39 × 10 <sup>6</sup>	5.2 ± 0.21 × 10 <sup>6</sup>	4.2 ± 0.41 × 10 <sup>6</sup>	4.9 ± 0.31 × 10 <sup>6</sup>	4.9 ± 0.24 × 10 <sup>6</sup>
	LAKE	Boczne	Jagodne	Szymoneckie	Szymon	Tałowisko
Area [ha]		183	420	523	154	327
Max depth [m]		25	37.4	28.5	2.9	39.5
Coordinates of sampling location		53.967407, 21.758265	53.945283, 21.721688	53.918687, 21.697991	53.891046, 21.633682	53.880376, 21.560412
Mean TSI		50 ± 4	58 ± 3	59 ± 2	58 ± 1	56 ± 3
Trophic state		ME	E	E	E	E
Temp [°C]		20.95 ± 0.01	22.03 ± 0.03	21.65 ± 0.01	20.87 ± 0.01	21.48 ± 0.91
Cond [µScm <sup>2</sup> ]		373.00 ± 0.0	354.00 ± 0.0	354.00 ± 0.0	354.50 ± 0.71	361.00 ± 5.2
Oxygen [mgL <sup>-1</sup> ]		10.05 ± 0.01	10.08 ± 0.08	9.55 ± 0.01	7.61 ± 0.03	9.82 ± 0.16
pH		8.44 ± 0.06	8.44 ± 0.0	8.42 ± 0.01	8.02 ± 0.01	8.31 ± 0.04
Turbidity [NTU]		1.57 ± 0.31	3.77 ± 0.11	3.60 ± 0.28	2.95 ± 0.35	1.90 ± 0.1
Chl a [µgL <sup>-1</sup> ]		7.93 ± 0.15	23.23 ± 1.56	24.07 ± 0.42	18.11 ± 0.64	19.05 ± 0.40
TP [µgL <sup>-1</sup> ]		31.80 ± 1.52	35.19 ± 0.18	40.13 ± 0.36	37.35 ± 1.25	30.67 ± 0.71
P-PO <sub>4</sub> <sup>3-</sup> [µgL <sup>-1</sup> ]		6.17 ± 0.46	3.73 ± 0.21	3.43 ± 0.21	3.38 ± 0.46	2.78 ± 0.17
TN [µgL <sup>-1</sup> ]		433 ± 58	700 ± 100	700 ± 100	667 ± 153	867 ± 153
DOC [mgL <sup>-1</sup> ]		11.12 ± 0.29	11.99 ± 0.08	10.08 ± 0.05	10.28 ± 0.03	11.52 ± 0.11
HNF [cellmL <sup>-1</sup> ]		2.60 ± 0.70 × 10 <sup>4</sup>	2.54 ± 0.37 × 10 <sup>4</sup>	3.83 ± 0.54 × 10 <sup>4</sup>	1.15 ± 0.14 × 10 <sup>4</sup>	3.93 ± 0.53 × 10 <sup>4</sup>
Bacteria Number [cellmL <sup>-1</sup> ]		5.0 ± 0.53 × 10 <sup>6</sup>	4.3 ± 0.24 × 10 <sup>6</sup>	4.2 ± 0.33 × 10 <sup>6</sup>	6.4 ± 0.21 × 10 <sup>6</sup>	3.8 ± 0.29 × 10 <sup>6</sup>
	LAKE	Tały	Ryńskie	Mikołajskie	Beldany	Śniardwy
Area [ha]		1 160	671	498	941	11 340
Max depth [m]		50.8	20.2	25.9	46	23.4
Coordinates of sampling location		53.813409, 21.567998	53.935961, 21.544402	53.801142, 21.570925	53.686344, 21.578000	53.766621, 21.842127
Mean TSI		59 ± 4	60 ± 1	59 ± 2	60 ± 1	42 ± 4
Trophic state		E	E	E	E	ME
Temp [°C]		22.66 ± 0.07	22.59 ± 0.11	21.72 ± 0.86	23.50 ± 0.11	23.58 ± 0.13
Cond [µScm <sup>2</sup> ]		306.33 ± 0.58	309.33 ± 0.58	310.33 ± 5.0	256.00 ± 0.0	291.36 ± 1.4
Oxygen [mgL <sup>-1</sup> ]		11.52 ± 0.50	9.90 ± 0.04	10.12 ± 0.94	12.65 ± 0.05	13.58 ± 0.4
pH		8.54 ± 0.03	8.40 ± 0.0	8.40 ± 0.11	8.63 ± 0.01	8.46 ± 0.02
Turbidity [NTU]		6.17 ± 0.38	4.23 ± 0.06	4.80 ± 3.0	7.03 ± 0.15	0.45 ± 0.11
Chl a [µgL <sup>-1</sup> ]		12.30 ± 5.54	20.99 ± 0.68	20.86 ± 3.00	23.74 ± 0.93	2.05 ± 1.58
TP [µgL <sup>-1</sup> ]		58.32 ± 0.64	43.42 ± 0.82	48.86 ± 0.36	45.57 ± 1.41	16.28 ± 0.31
P-PO <sub>4</sub> <sup>3-</sup> [µgL <sup>-1</sup> ]		4.48 ± 0.0	3.38 ± 0.17	4.48 ± 0.30	3.58 ± 0.30	3.78 ± 0.17
TN [µgL <sup>-1</sup> ]		800 ± 58	633 ± 0	700 ± 100	1000 ± 100	700 ± 0
DOC [mgL <sup>-1</sup> ]		9.05 ± 0.07	9.5 ± 0.02	9.15 ± 0.11	7.55 ± 0.07	8.02 ± 0.05
HNF [cellmL <sup>-1</sup> ]		2.42 ± 0.75 × 10 <sup>4</sup>	3.96 ± 1.2 × 10 <sup>4</sup>	2.35 ± 0.36 × 10 <sup>4</sup>	4.28 ± 0.41 × 10 <sup>4</sup>	9.21 ± 2.8 × 10 <sup>3</sup>
Bacteria Number [cellmL <sup>-1</sup> ]		5.5 ± 0.2 × 10 <sup>6</sup>	5.6 ± 0.12 × 10 <sup>6</sup>	5.6 ± 0.4 × 10 <sup>6</sup>	5.2 ± 0.35 × 10 <sup>6</sup>	4.9 ± 0.51 × 10 <sup>6</sup>

**Table 1.** The basic characteristics of the studied lakes. ± Are the standard deviation of mean values. Temp: temperature; Cond: special conductivity; Oxygen: Dissolved oxygen concentration; Turbidity: nephelometric turbidity; Chl a: chlorophyll a concentration; TP: total phosphorus; P-PO<sub>4</sub><sup>3-</sup>: dissolved orthophosphates; TN: total nitrogen; DOC: dissolved organic carbon; HNF: number of heterotrophic nanoflagellates; Bacterial number: number of DAPI-stained bacteria; Trophic state E: Lake belongs to the group of eutrophic lakes; Trophic state ME: Lake belongs to the group of meso-eutrophic lakes; Mean TSI: Mean Trophic State Index calculated on the basis of the concentration of chlorophyll a, total phosphorus and the visibility of the Secchi disc.



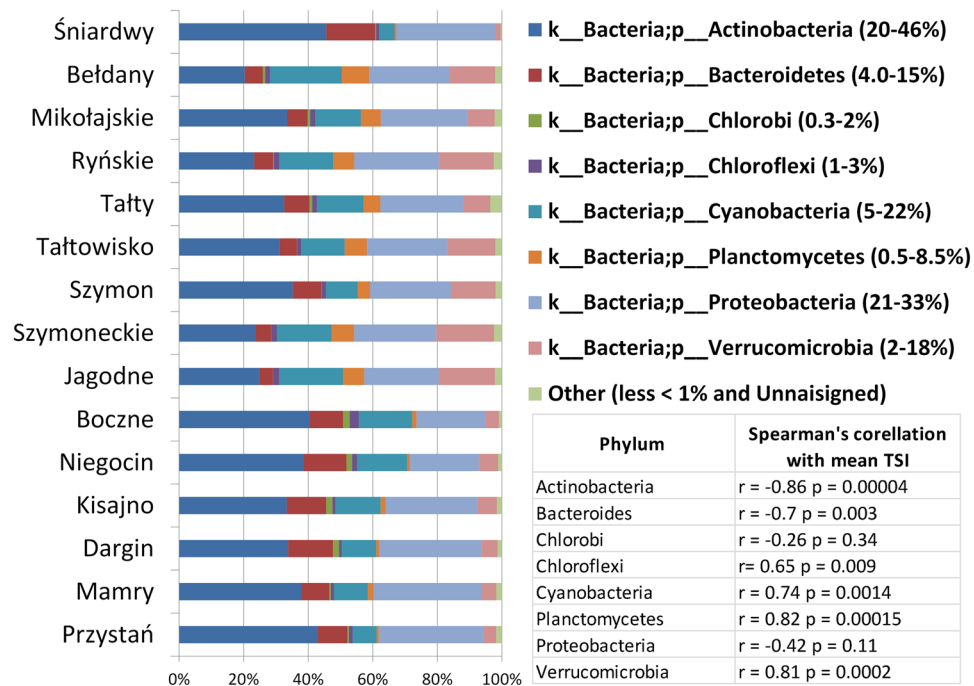
**Figure 1.** (a) PCA-based grouping of lakes according to physicochemical parameters. The lakes marked in red are eutrophic (E), and lakes marked in green are meso-eutrophic (ME). (b) Principal component analysis (PCA) results. Projection of variables on the factor planes of PCA analyses of physicochemical variables. The variables in red squares are supplementary and have no influence on the principal components of the analysis. Cond.: conductivity [ $\mu\text{scm}^{-2}$ ], DOC: dissolved organic carbon [ $\text{mgL}^{-1}$ ], Chl a: chlorophyll a concentration [ $\mu\text{gL}^{-1}$ ], HNF: heterotrophic nanoflagellates [ $\text{cellmL}^{-1}$ ], TN: total nitrogen concentration [ $\mu\text{gL}^{-1}$ ], Turbid: nephelometric turbidity [NTU], Temp.: temperature [ $^{\circ}\text{C}$ ], DO: dissolved oxygen concentration [ $\text{mg O}_2\text{L}^{-1}$ ], TP: total phosphorus concentration [ $\mu\text{gL}^{-1}$ ],  $\text{PO}_4^{3-}$ : orthophosphate concentration [ $\mu\text{gL}^{-1}$ ], mean trophic state index, BN: bacteria number [ $\text{cellmL}^{-1}$ ].

did not lead to an obvious grouping of the eutrophic and meso-eutrophic lakes (Fig. 5), as was the case with the taxonomic data (Fig. 3a). However, there were weak signs of grouping the lakes into eutrophic and mesotrophic based on metabolic diversity (ANOSIM  $r = 0.25$ ,  $p = 0.003$ ). We found that only the relative respiration rates of complex carbon sources were significantly negatively correlated with the mean TSI ( $r = -0.54$ ,  $p = 0.03$ ). Other carbon sources were not significantly correlated with the TSI values. It is interesting that, in the case of lake Niegocin, there is no participation of the respiration of phosphorylated carbon substrates in the total respiration of the analysed carbon sources. In Niegocin lake, we observed an exceptionally high level of  $\text{P-PO}_4^{3-}$  ( $14.3 \mu\text{gL}^{-1}$ , Table 1). In general, we found a significant negative logarithmic relationship between the phosphate-carbon respiration rate and the  $\text{P-PO}_4^{3-}$  concentration ( $y = -2.733\ln(x) + 11.802$ ,  $r^2 = 0.8836$ ,  $p < 0.05$ ). We did not find a significant correlation between the TSI and the number of used carbon sources (richness equivalent) or Simpson's diversity index (Table 3).

To investigate the relationship between the Bray-Curtis distance matrices created on the basis of relative phylogenetic and metabolic differences, we used the Mantel test of matrix correlations. We found that the two matrices were positively and significantly correlated ( $r = 0.22$ ,  $p = 0.006$ ), but the correlation was not high.

## Discussion

Carlson's TSI index<sup>23</sup> allows the numeric classification of lake trophic state. Due to the relatively high heterogeneity of lake water and the variability in measurable TSI on a yearly basis, we included the temporal TSI values characterizing the sampling sites during summer stratification in our analysis (Table 1). However, the values obtained generally correspond with the TSI values characterizing the studied lakes over several years<sup>21,24,25</sup>. The TSI value alone provides general information about the fertility of a reservoir and allows for a general assessment

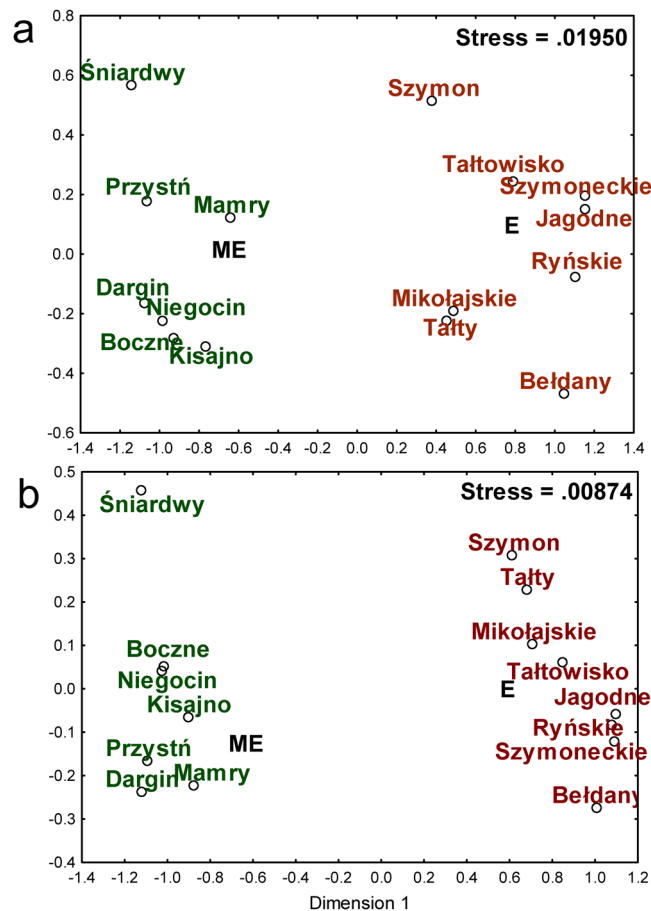


**Figure 2.** Relative sequence abundance of dominant bacterial phyla associated with different lakes. Only phyla whose relative share exceeded 1% were taken into account as individual units. Spearman correlations between the TSI and microorganisms belonging to the dominant phyla were found.

of its physicochemical properties by providing information on the amount and quality of dissolved organic matter, the number of bacteria<sup>24</sup> and the penetration of light into deeper parts of the lake water body, among other variables. The classification of lakes into two eutrophic and meso-eutrophic groups was reflected in the classification of lakes based on their physicochemical diversity (Fig. 1a). It is interesting that despite the geographical location of the southern lakes and the strong anthropopressure, the sampling area of Lake Śniardwy remained meso-eutrophic<sup>22</sup>, and its physicochemical characteristics were similar to those characterizing the northern, less eutrophicated lakes (based on the PCA). Based on the description of the studied lakes, it can be concluded that the numbers of HNF and bacteria remain positively related with characteristics strongly associated with a high trophic state, which is in accordance with the literature<sup>24,26</sup>. This paper poses the question of whether the phylogenetic structure of microorganisms belonging to domain *Bacteria* inhabiting the group of eutrophic lakes differs from the structure of bacteria inhabiting the meso-eutrophic lakes. The obtained data support hypothesis that the taxonomic diversities of microorganisms depend strongly on the trophic state of the studied lakes. Obviously, Illumina sequencing does not allow for a fully quantitative evaluation of the frequencies of individual taxonomic groups based on 16S rRNA genes amplification due to, for example, differing numbers of 16S rRNA coding DNA copies in individual bacterial species or some PCR amplification bias. However, the highly significant difference between the relative occurrence of bacterial families in ME and E lakes (ANOSIM  $r = 0.99$ ,  $p = 0.0002$ ) leads to the conclusion that the meso-eutrophic lakes were inhabited by different groups of bacteria than the eutrophic ones. Differences in the taxonomic structure of microorganisms from domain *Bacteria* in meso- and eutrophic reservoirs can partially result from differences in the presence of cyanobacteria. As widely described<sup>27–29</sup> the distribution of dominant bloom-forming cyanobacteria is usually strongly related to the geology of the catchments, nutrient loading (especially the phosphorus load), temperature and the general trophic state of lakes. We also found a positive correlation between the trophic state and the frequency of *Cyanobacteria* occurrence (Fig. 2). However, we did not find that *Cyanobacteria* significantly affected the differences in the taxonomic composition of *Bacteria* in the studied lakes. The exclusion of *Cyanobacteria* did not eliminate the differences between the two groups of lakes and even made them more pronounced (higher  $r^2$  and lower  $p$  in the ANOSIM test). We used the Bray-Curtis distance for the NMDS grouping of the studied lakes. On the basis of the SIMPER test, we found that the differences between the E and EM lakes were mostly due to the participation of representative of families belonging to the phyla *Actinobacteria*, *Proteobacteria*, *Verrucomicrobia*, *Cyanobacteria*, and *Planctomycetes* (Table 3). The proportions of the relative abundance of bacteria at the phylum level were similar to the results obtained by Liu *et al.*<sup>30</sup>, who analysed the biogeography of the abundance of bacterioplankton in the lakes and reservoirs of China.

As mentioned above, we found a strong positive correlation between the TSI and *Cyanobacteria* occurrence, which has been documented for lakes by other authors<sup>29</sup>. Differences in the trophic state were also reflected in the dominance of other phyla. We found a statistically significant negative correlation between the frequency of *Actinobacteria* and the TSI of the examined lakes and a high relative abundance of *Actinobacteria* compared to the that of the other detected phyla, particularly in the meso-eutrophic lakes. *Actinobacteria*, historically





**Figure 3.** (a) The graphical results of the Bray-Curtis-based NMDS analysis of sequence data, binned by taxonomic assignment to family. The whole domain Bacteria was taken into consideration. The lakes marked in red are eutrophic (E), and lakes marked in green are meso-eutrophic (ME). (b) The graphical results of the Bray-Curtis-based NMDS analysis of sequence data, binned by taxonomic assignment to family, after excluding 25 families belonging to phylum Cyanobacteria from the analysis. The lakes marked in red are eutrophic (E), and lakes marked in green are meso-eutrophic (ME).

classified as mostly soil bacteria<sup>31</sup>, are frequently present in freshwater lentic systems<sup>9,32</sup> and are often associated with oligotrophic ecosystems<sup>3</sup>. *Actinobacteria* is often the numerically dominant phylum in lake epilimnions<sup>30,33</sup>. Due to the production of additional pigments and high G + C content, they are relatively resistant to UV light, which easily penetrates deep into oligotrophic lakes<sup>10</sup>. *Planctomycetes* and *Verrucomicrobia*, on the other hand, showed a positive correlation with the TSI in our study. In the past, *Planctomycetes* and *Verrucomicrobia*, which are included in the Planctomycetes-Verrucomicrobia-Chlamydiae (PVC) superphylum, were underestimated due to the small number of deposited sequences in reference databases. Currently, the increase in the number of sequences of *Planctomycetes* and *Verrucomicrobia* deposited in the databases allows a better estimation of their role in aquatic ecosystems. The significant participation of *Planctomycetes* and *Verrucomicrobia* in the phyla found in our studies is in line with this trend. Recently, *Verrucomicrobia* have been shown to be ubiquitous in freshwater and exhibit a cosmopolitan distribution in lakes and rivers. Their abundance is usually between 1% and 6% of the total microbial community<sup>34,35</sup>. In our study, the mean participation of *Verrucomicrobia* was slightly higher and amounted to 9.5% (min. 1.5, max. 17.8). There is limited information about the physiology of *Verrucomicrobia* and *Planctomycetes*. It is known that the abundance of some members of the phylum *Verrucomicrobia* is favoured by high nutrient availabilities or cyanobacterial blooms<sup>36,37</sup>. The efficient use of glycolate by representatives of *Verrucomicrobia* is also noted<sup>38</sup>. Glycolate is secreted during photosynthesis by phytoplankton, occurring in higher numbers in eutrophic reservoirs (higher chlorophyll a concentrations). This corresponds to the positive correlation we found between the occurrence of *Verrucomicrobia* and the trophic status of the lakes. The *Proteobacteria* phylum did not show any significant correlation with the TSI because members of this phylum belong to a very physiologically diverse group, such as those in the *Alphaproteobacteria* or *Gammaproteobacteria* classes. *Alphaproteobacteria* often dominate in oligotrophic reservoirs and are known to be competitive under conditions of low nutrient/substrate availability<sup>10</sup>. In our study, the frequency of the occurrence of microorganisms belonging to the *Alphaproteobacteria* class showed a significant negative correlation with the TSI (Spearman correlation  $r = -0.7$ ,  $p = 0.003$ ). In addition, this class includes *Pelagibacteraceae*, which significantly contributed (SIMPER test, contrib. 12.3%) to the observed difference between the ME and E lakes (Table 3). On the

Family	Contrib. %	Cumulative %
p_Actinobacteria; c_Actinobacteria; o_Actinomycetales; f_ACK-M1	13.42	13.42
p_Proteobacteria; c_Alphaproteobacteria; o_Rickettsiales; f_Pelagibacteraceae	12.35	25.77
p_Verrucomicrobia; c_Spartobacteria; o_Chthoniobacterales; f_Chthoniobacteraceae	9.503	35.27
p_Cyanobacteria; c_Synechococcophycidae; o_Pseudanabaenales; f_Pseudanabaenaceae	5.066	40.34
p_Verrucomicrobia; c_[Methylacidiphilae]; o_Methylacidiphilales; f_LD19	4.11	44.45
p_Planctomycetes; c_Planctomycetia; o_Pirellulales; f_Pirellulaceae	3.072	47.52
p_Cyanobacteria; c_Synechococcophycidae; o_Synechococcales; f_Synechococcaceae	3.025	50.55
p_Proteobacteria; c_Betaproteobacteria; o_Burkholderiales; f_Comamonadaceae	2.648	53.19
p_Verrucomicrobia; c_Opitutae; o_Cerasicoccales; f_Cerasicoccaceae	2.52	55.71
p_Bacteroidetes; c_Flavobacteriia; o_Flavobacteriales; f_Cryomorphaceae	2.296	58.01
p_Cyanobacteria; c_; o_; f_	2.104	60.11
p_Actinobacteria; c_Actinobacteria; o_Actinomycetales; f_Mycobacteriaceae	2.043	62.16
p_Planctomycetes; c_Phycisphaerae; o_Phycisphaerales; f_	1.861	64.02
p_Bacteroidetes; c_Cytophagia; o_Cytophagales; f_Cytophagaceae	1.593	65.61
p_Bacteroidetes; c_Flavobacteriia; o_Flavobacteriales; f_Flavobacteriaceae	1.559	67.17
p_Actinobacteria; c_Acidimicrobiia; o_Acidimicrobiales; f_C111	1.476	68.65
p_Chloroflexi; c_Anaerolineae; o_Caldilineales; f_Caldilineaceae	1.417	70.06
p_Cyanobacteria; c_Oscillatoriophycidae; o_Oscillatoriales; f_Phormidiaceae	1.397	71.46
Unassigned	1.319	72.78
p_Cyanobacteria; c_Nostocophycidae; o_Nostocales; f_Nostocaceae	1.226	74
p_Chlorobi; c_OPB56; o_; f_	1.199	75.2
p_Cyanobacteria; c_Oscillatoriophycidae; o_Chroococcales; f_Microcystaceae	1.19	76.39
p_Actinobacteria; c_Actinobacteria; o_Actinomycetales; f_	1.162	77.56
p_Bacteroidetes; c_[Saprosirae]; o_[Saprosirales]; f_Chitinophagaceae	1.022	78.58

**Table 2.** The SIMPER test results of the contribution and cumulative percentage of families mainly responsible for the observed differences between groups of eutrophic (E; consisting of eight lakes, TSI from 56 to 60) and mesotrophic-eutrophic (ME; consisting of seven lakes, TSI from 40 to 53) lakes.

other hand, *Gammaproteobacteria* (and among them many potentially pathogenic bacteria, such as *Aeromonas* spp. or *Legionella* spp.) usually require high concentrations of organic matter<sup>10</sup>. Gammaproteobacteria grew fast when nitrogen and phosphorus (elements occurring in high concentrations in eutrophic lakes) were added to enclosures<sup>39</sup>. Generally, *Alphaproteobacteria* (mostly *Pelagibacteraceae*, from 3% to 22% in our study) and Betaproteobacteria (for example, *Comamonadaceae*—in our study, their contribution was from 2% to 11%) have been regarded as some of the most important aquatic microorganisms on our planet. In our study, the abundance of *Pelagibacteraceae* was from 21% to 33%, which corresponded well with data obtained by other researchers examining various lake ecosystems<sup>9,30,40</sup>.

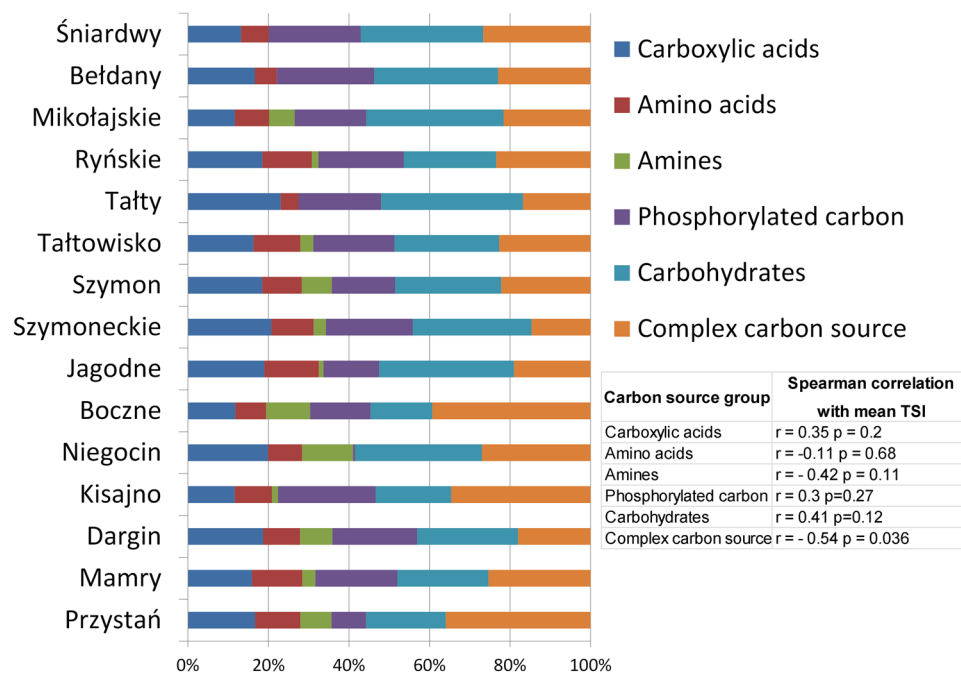
Biodiversity influences the way in which ecosystems function. We observed a significant positive correlation between the trophic status of lakes and the taxonomic richness of microorganisms and their even distribution. We found not only more taxa in the higher eutrophicated lakes but also higher values of Simpson's diversity index in those lakes. This corresponds with results obtained by Jankowski *et al.*<sup>41</sup>, who showed that the bacterial community composition changed and was richer and more heterogeneous within lakes as the trophic status increased. As suggested by Bell *et al.*<sup>42</sup>, the greater diversity of bacteria in eutrophic lakes may be associated with their higher overall productivity and contribute to bacterial services. The same situation is observed in marine environments, where primary productivity drives greater bacterial richness<sup>43</sup>. It should be emphasized that the studied lakes were not hyper-eutrophic. Under hyper-eutrophy, biodiversity loss may be observed, especially in the case of phototrophic microorganisms. Nevertheless, as reported by Crossetti *et al.*<sup>44</sup>, the loss of biodiversity following trophic change is not a single dimension of a single factor but rather a result of a combination of factors related to larger levels of biomass in the reservoir.

In contrast to results obtained by Zwirgmaier *et al.*<sup>9</sup>, who showed a spatial distribution in the Osterseen Lake District, it seems interesting that in our study, the differences in the taxonomic structure of the *Bacteria* domain were associated with the trophic state rather than the geographical location of the lakes. We found some similarities that may result from the proximity of the sampling points (e.g., lakes Tały and Mikołajskie), but the mesotrophic part of Śniardwy lake (to which water flows from Mikołajskie lake) has been classified to the group of northern lakes, which are distant from lake Śniardwy. This suggests that trophic characteristics rather than geographical proximity determine the taxonomic structure of bacteria in the studied ecosystem.

The situation is not as clear in the case of community-level physiological profiling. To obtain a community-level physiological profile, we used the commercially available Biolog Corporation's EcoPlate test, enabling the creation of a metabolic fingerprint of the bacterial community. The most commonly used EcoPlate procedure is based on the colorimetric determination of the final concentration of the respiratory product after 48–76 hours. However, it has some limitations, which should be mentioned here. First, the tetrazolium chloride used in the test may lead to a selective reduction in the intensity of respiration of various groups of microorganisms after a longer incubation

Lake	Phylogenetic Richness	Phylogenetic Simpson 1-D index	Number of used Carbon source	Used Carbon source Simpson 1-D index
Przystań	105	0.5733	30	0.937
Mamry	127	0.5889	25	0.9458
Dargin	108	0.5963	25	0.9277
Kisajno	122	0.6453	27	0.9264
Niegocin	110	0.633	24	0.9247
Boczne	129	0.6191	22	0.9283
Jagodne	140	0.6951	28	0.9368
Szymoneckie	137	0.707	28	0.9391
Szymon	135	0.6719	31	0.9456
Tałowisko	138	0.6791	30	0.9439
Tały	132	0.6733	18	0.9126
Ryńskie	123	0.6979	25	0.9399
Mikołajskie	136	0.6702	21	0.9302
Beldany	132	0.7058	17	0.9234
Śniardwy	90	0.555	17	0.8976
<i>Spearman correlation with TSI</i>	<i>r = 0.61 p = 0.01</i>	<i>r = 0.91 p &lt; 0.0005</i>	<i>r = -0.1 p = 0.7</i>	<i>r = -0.03 p = 0.9</i>

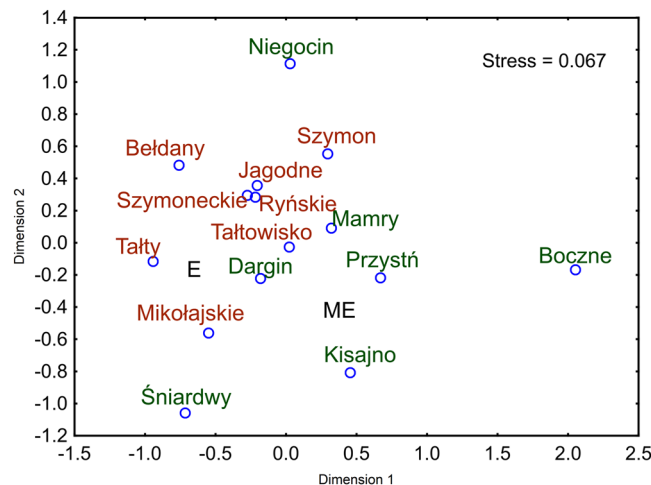
**Table 3.** Richness (number of taxa or used carbon sources) and Simpson's diversity index characterizing the studied lakes. In the case of taxonomic diversity at the family level, OTU richness and diversity were positively correlated with the mean TSI of the studied lakes. In the case of metabolic fingerprinting, we did not find a significant correlation between the TSI and the number of bacterial carbon sources (richness equivalent) or Simpson's diversity index.



**Figure 4.** The average relative respiration of carbon sources assigned to the basic groups of compounds. Only the relative respiration rates of complex carbon sources were significantly negatively correlated with the mean TSI.

period. The accumulation of secondary metabolites and the decrease in the substrate concentration also contribute to this phenomenon. To limit this effect, we determined not only the final concentration of the colorimetric product of respiration but also the maximum rate of respiration (maximal increasing rate of the production of the colorimetric product of respiration) observed throughout the whole 72-hour incubation period. Differences in the obtained values for each of the substrates were used to calculate the Bray-Curtis similarity matrix for the studied lakes.





**Figure 5.** Bray-Curtis-based NMDS analysis of the relative respiration of 31 different carbon sources by bacteria in the studied lakes. The lakes marked in red are eutrophic (E), and lakes marked in green are meso-eutrophic (ME).

The EcoPlate-based results of metabolic profiling were not as strongly associated with the trophic state of lakes as was the case with phylogenetic classification. However, we also observed a tendency towards the separate clustering of microorganisms originating from meso-eutrophic and eutrophic lakes (ANOSIM  $r = 0.25$ ,  $p = 0.003$ ). A similar EcoPlate system was used by Dickerson and Williams<sup>19</sup> to test whether three Florida lakes of different trophic levels could be distinguished by their functional profiles. These authors showed that the metabolic profiles of the two mesotrophic lakes were more similar to one another than to that of the eutrophic lake.

We did not find a significant correlation between the TSI and the number of carbon sources usable for microorganisms from lakes (richness) or between the TSI and Simpson's diversity index. In other words, an increase in the trophic state did not contribute to an increase in the potential of the bacterial community to use a larger number of more diverse carbon sources. We also found no significant correlation (based on Spearman correlation coefficients) and no significant linear or non-linear relationships between richness and Simpson's diversity indices for metabolic and taxonomic structures. Our results do not allow to strongly confirm or reject the concept of microbial functional redundancy described by Delgado-Baquerizo *et al.*<sup>45</sup> who, in model experiments, observed no functional redundancy in a majority of analyzed cases.

However, while the experimentally induced microbial biodiversity decline analyzed by Delgado-Baquerizo *et al.*<sup>45</sup> was very rapid, our studied lakes have been undergoing eutrophication for many years. The microbial communities in each lake are a part of a complex web of relationship between various components of the microbial loop. Relatively slow changes in biodiversity may have less influence on functional diversity than rapid changes, under which the lack of redundancy might be particularly dangerous to environmental functioning. The long term adaptation of microbial communities to environmental characteristics and available food sources is important for stable environmental homeostasis. For example, in eutrophicated lakes the excretion of organic matter by phytoplankton or degradation of dead phytoplankton cells might be a primary food source for chemoorganotrophic bacteria. In mesotrophic lakes the lack of intense primary production forces bacteria to use many different sources of carbon.

In our research, only in the case of complex carbon sources was there a significant negative correlation between the rate of use of this group of C compounds and the TSI. The utilization of complex carbon sources is typical for oligotrophic lakes, where low primary production forces microorganisms to use allochthonous organic matter. Such organic matter is often built with a long polymer chain with a high C:N ratio and is difficult to degrade<sup>46</sup>. The more efficient respiration of complex carbon sources in meso-eutrophic lakes that we observed may indicate the adaptation of microorganisms in less eutrophicated reservoirs to the degradation of this type of compound. Most of the literature is focused on pelagial zones; however, this zone of lakes may differ from the littoral zone in terms of bacterial activity and bacterial diversity<sup>47</sup>. Samples for our study were taken from littoral zones. The near-shore, ecotone and littoral zones of lakes are productive and potentially very diverse, where allochthonous matter, often of anthropogenic origin (including xenobiotics), may be a significant source of carbon. The inflow of this type of compound can have a significant impact on the metabolism of microorganisms that occupy the littoral zone of lakes.

Due to the high plasticity of chemoorganotrophic bacteria, their metabolic profiles are shaped by the presence of carbon sources in the environment, not just by the taxonomic composition. The presence of inhibitors may also have an impact. For example, in Lake Niegocin, a lack of phosphate-carbon utilization was observed in the EcoPlate-based metabolic profile. At the same time, abnormally high concentrations of orthophosphates (likely as a result of land inflow of anthropogenic origin) were found. Assimilation of phosphorylated carbon sources by microorganisms is not possible without their prior dephosphorylation. To acquire phosphorus from organic phosphorylated compounds, microorganisms require active alkaline phosphatases<sup>48</sup>. Because orthophosphates inhibit the activity of phosphatases<sup>49</sup>, in the case of Lake Niegocin, we did not observe the respiration of this type of substrate. We found a negative logarithmic relationship between the  $P-PO_4^{3-}$  concentration and respiration

for all lakes we examined. This confirms that the results of the EcoPlate analysis were influenced by inhibitors present in water, making it impossible to precisely characterize the full metabolic potential of the bacteria. At the same time, our data indicate that the EcoPlate test allows the assessment of the current potential of the bacteria. Therefore, the EcoPlate-based metabolic fingerprint should be interpreted as a result of the metabolic capacity of the microorganism community, its taxonomic composition and environmental characteristics (availability of carbon sources, adaptation of bacteria to carbon sources and the presence of inhibitors). Despite these reservations, the Mantel correlation of the Bray-Curtis distance matrices calculated for the community-level physiological profiling distance and taxonomic diversity gave significant results. Although the value of the correlation was rather low ( $r = 0.22$ , Mantel test), it remained statistically significant ( $p = 0.006$ ). The Mantel test results together with the results of the NMPS analysis allow us to conclude that there is some relationship between the taxonomic structure of microorganisms and their ability to decompose the carbon sources in the studied lakes; however, this relationship may be influenced by various environmental factors. Additionally, functional metabolic potentials may be strongly affected not only by bacteria-bacteria interactions and environmental conditions<sup>50</sup> but also by heterotrophic bacteria-cyanobacteria<sup>36</sup> or bacteria-eukaryotic algae interactions<sup>51</sup>.

## Conclusions

The Great Masurian Lake System can be considered to be a testing ground that allows us to study the impact of the anthropogenic eutrophication process on the functional and structural diversity of aquatic bacteria. Our results highlight the dependence of the taxonomic composition on lake trophic state and suggest that accelerated anthropogenic eutrophication may lead to a significant transformation of the taxonomic structure of microorganisms from the *Bacteria* domain. This is due to the changes in the physicochemical conditions prevailing in the lakes undergoing anthropopressure, among other reasons. The increase in eutrophication in the studied lakes led to an increase in bacterial family richness and the evenness of their distribution. The change in the taxonomic structure resulting from the change in trophic state did not clearly follow the change in the metabolic structure reflected in the differences in the preference for the use of various carbon sources by microorganisms. The data, however, indicate the existence of a structure-function relationship. Although this relationship is not strong, it reflects the high plasticity of the metabolism of microorganisms depending not only on the taxonomic composition but also on the environmental conditions and, in the context of nutrition, the available resources of labile organic matter. This plasticity is one of the reasons why, despite increasing research, the relationship between bacterial biodiversity and functioning remains poorly understood. The presented study concerns only a small part of the huge population of bacteria living in lakes, mainly the aerobic bacteria inhabiting surface waters of the littoral zone during summer stratification. To what extent the observed relationship between the trophic state of lakes and the functioning of bacteria from the littoral zone is representative of the entire lake bacteriome requires further research.

## Methods

The research area comprised 15 lakes belonging to the Great Masurian Lake System (Fig. 6). The GMLS is a chain of lakes located in the Great Masurian Lakes region in north-eastern Poland. All lakes of the GMLS are connected by natural or artificial channels, and create a broad, long chain that is unique in Europe. The northern meso-eutrophic part of the GMLS includes the lakes Przystań, Mamry, Dargin, Łabap, Kisajno and Niegocin, whereas the southern, mostly eutrophic lakes are Boczne, Jagodne, Szymoneckie, Szymon, Kotek, Tałtowisko, Ryńskie, Tały, Mikołajskie, Beldany and the meso-eutrophic lake Śniardwy (Fig. 6). Detailed information about the characteristics of the lakes is shown in Table 1. Samples were taken in July 2016 from the surface water of the littoral zones during the summer thermal stratification. The approximate sampling sites are shown in Fig. 6. The samples were taken in 3 repetitions within the radius of 100 m from the central points of sampling area (geographical coordinates in Table 1). Each sample constituted a mixture (v/v) of water sampled from three depths: 1, 2 and 3 metres, taken at all lakes except Mamry, Niegocin, Szymoneckie, Szymon and Śniardwy, where samples were taken at 1 m and 2 m depth because the water was shallower than 3 metres at the sampling sites. The water samples were collected in sterile 5 l polyethylene containers. The samples were subjected to further laboratory analyses within 2–4 h.

The physicochemical parameters (temperature, conductivity, chlorophyll autofluorescence, pH, oxygen concentration, turbidity) were determined *in situ* using a multiparametric YSI 6000 probe (Yellow Spring, YSI Inc./Xylem Inc. USA). Parameters were measured at every 20 cm depth, in each sampling point. The results are presented as the average for the analysed water column. Dissolved organic carbon (DOC) was determined with a Shimadzu TOC 5050 carbon analyser.

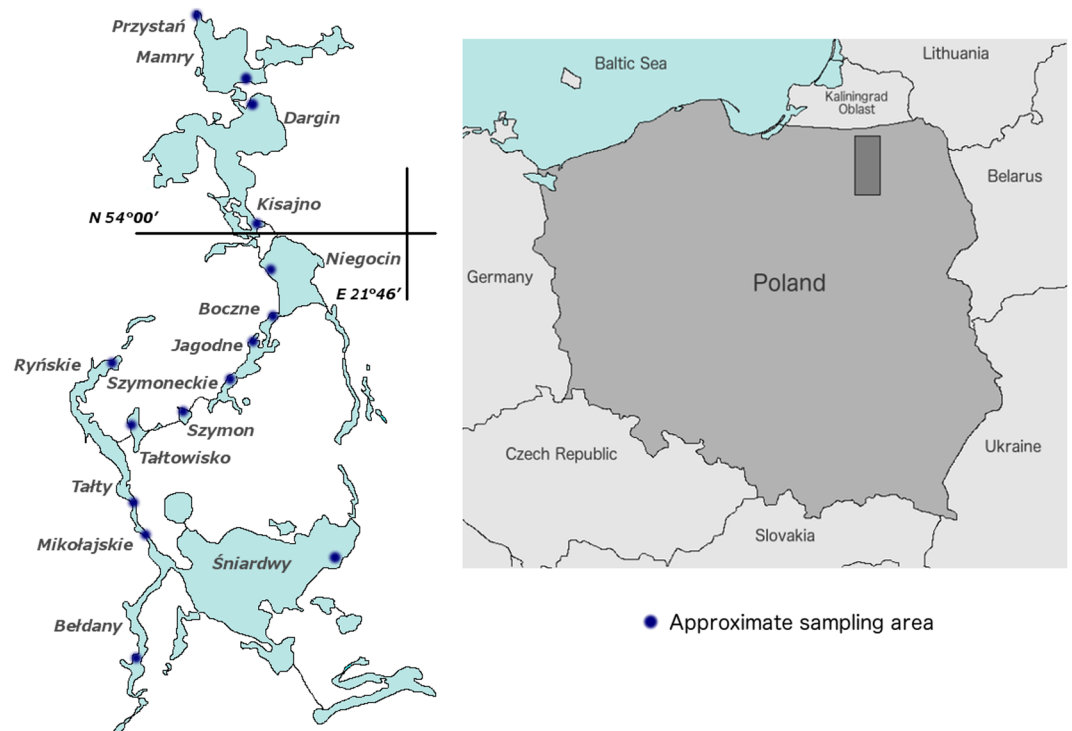
The number of DAPI (4,6-diamidino-2-phenylindole) -stainable bacteria was determined by the direct counting of cells collected on 0.2 µm black polycarbonate membrane filters (Millipore) using epifluorescence microscopy<sup>52</sup>. A computer image analysing system composed of a Nikon epifluorescence E450 microscope, Nikon Digital Camera DXM 1200F and NIS elements software (Nikon) were used. HNF (heterotrophic nanoflagellates) were counted under Nikon epifluorescence E450 microscope (Nikon)<sup>53</sup>.

The total phosphorus (TP) and phosphate ( $\text{PO}_4^{3-}$ ) concentrations were determined using the molybdate method using a Shimadzu UV-VIS 1201 spectrophotometer<sup>54</sup>.

Total nitrogen was determined using commercially certified Merck-Millipore cell tests (Spectroquant Nitrogen (Total) Cell Test, 114537) according to the manufacturer's instructions using Merck Spectroquant Pharo 300 spectrophotometer.

The chlorophyll *a* (Chl *a*) concentration was determined fluorimetrically using a TD-700 fluorometer (Turner)<sup>55</sup>.

The mean TSI (Trophic State Index) was defined as the arithmetic mean of the TSI calculated on the basis of the concentration of chlorophyll *a*, total phosphorus and the visibility of the Secchi disc<sup>23</sup>.



**Figure 6.** Locations of sampling sites within the Great Masurian Lakes System (GMLS).

**Taxonomic diversity of the bacteria domain.** DNA from the samples was isolated after the suspension of microorganisms on the surface of 0.2  $\mu\text{m}$  sterile polycarbonate filters (Merck Millipore). From each lake, exactly 5  $\times$  200 ml of water was filtered on Nalgene (Nalgene®) filtration equipment and immediately frozen in sterile 2 ml Eppendorf vials (each filter separately) at  $-30^\circ\text{C}$ . The DNA from the shredded filters was isolated by spin-column-based DNA extraction using the GeneMATRIX DNA Purification Kit (EurX, Gdańsk). The isolated purified DNA was stored at  $-80^\circ\text{C}$  for further analysis.

The phylogenetic analysis of the bacterial community was performed using Illumina sequencing<sup>56</sup>. For the sequencing, the 16S rRNA genes, V3–V4 hypervariable regions (amplicons of approximately 459 bp) were selected. PCR amplification was carried out using Q5 Hot Start High-Fidelity 2X Master Mix using reaction conditions as recommended by the manufacturer ( $95^\circ\text{C}$  for 3 minutes, 25 cycles of  $95^\circ\text{C}$  for 30 seconds,  $55^\circ\text{C}$  for 30 seconds,  $72^\circ\text{C}$  for 30 seconds and, after the last cycle,  $72^\circ\text{C}$  for 5 minutes) with region-specific (341F and 785R)<sup>57</sup> primers that include the Illumina flowcell adapter sequences. The primer sequence was as follows: forward primer: 5' CGGGNGGCWGCAG 3', reverse primer: 5' GACTACHVGGGTATCTAATCC 3'. The Illumina overhang adapter sequences added to the locus-specific-sequences were as follows: forward overhang: 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG (locus-specific sequence), reverse overhang: 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG (locus-specific sequence).

Sequencing took place on an Illumina MiSeq sequencer using the MiSeq Reagent Kit v2 and 2  $\times$  250 bp protocol. Initial automatic data analysis was carried out on the MiSeq system using the MiSeq Reporter software (MSR) v2.6, which was used for demultiplexing the data and raw FASTQ file generation. The bioinformatic analysis resulting in the classification of reads up to the species level was carried out using the QIIME software package and the Greengenes v13\_8 reference sequence database. The analysis consisted of the following stages:

(1) Removal of adapter sequences and quality filtering (Phred quality score  $<20$ , minimum length 30 – Phytonproject – cutadapt); (2) Pair-end read merging (ea-utils in FASTQ-join in QIIME). (3) Clustering – uclust; (4) Chimeric sequence removal – ChimeraSlayer; (5) Taxonomic analysis of OTU clusters – (uclust implemented using the assign\_taxonomy.py command in QIIME).

For the Bray-Curtis similarity matrices, the relative abundance of OTUs at the family level expressed as a percentage was used. These values were calculated as the sequence count for each identified group divided by the total sequence counts<sup>58</sup>. For the correlation with TSI, the relative abundance of each group was calculated in the same way at the phylum level.

**Metabolic diversity of microbial communities.** The Biolog EcoPlate method was used to determine the community-level physiological profiles<sup>19</sup> with the modifications described below. EcoPlate (Biolog, USA) is used to measure the ability of the bacterial community to utilize carbon substrates. An EcoPlate is a 96-well microplate composed of triplicates of 31 response wells with different carbon sources and control wells. The following carbon sources were used: carbohydrates—D-cellobiose,  $\alpha$ -D-lactose,  $\beta$ -methyl-D-glucoside, D-xylose, erythritol, D-mannitol, N-acetyl-D-glucosamine, D-galactonic acid  $\gamma$ -lactone; phosphorylated carbons—glucose-1-phosphate, D,L- $\alpha$ -glycerol phosphate; amines—phenylethylamine, putrescine; carboxylic

acids—D-glucosaminic acid, D-galacturonic acid,  $\gamma$ -hydroxybutyric acid, itaconic acid,  $\alpha$ -ketobutyric acid, D-malic acid, pyruvic acid methyl ester, 2-hydroxy benzoic acid, 4-hydroxy benzoic acid; complex carbon—Tween 40, Tween 80,  $\alpha$ -cyclodextrin, glycogen; amino acids—L-arginine, L-asparagine, L-phenylalanine, L-serine, L-threonine, glycyl-L-glutamic acid. Each well of a single plate was filled with fresh lake water (150  $\mu$ l). The plates were incubated in darkness at a temperature of 22 °C for 72 hours. The absorbance was measured every 4 hours at 590 nm wavelength using a Biotek Synergy H1 plate reader (Biotek Corporation, USA). The observed absorbance of tetrazolium chloride reduced due to respiration increased during the incubation time. For analysis, we used the maximal colour development rates ( $V_{max}$ ). For the calculation of  $V_{max}$  values, Gen 5 software (Biotek Corporation, USA) was used. We identified the slope for every four consecutive reads.  $V_{max}$  was calculated using a linear regression model by determining the maximum slope during the 72-hour incubation time. The maximal rate of colour development,  $V_{max}$  (mOD/min), was treated as proportional to the microbial respiration of a single carbon source. To determine the respiration rate of the microbial community without the addition of any carbon substrate, control wells were used.

To determine the influence of each additional carbon source on the respiration rate, the difference in  $V_{max}$  between carbon-containing wells and control wells was calculated. When the difference ( $\Delta V_{max}$ ) was higher than zero, we used these results for further analysis. When  $\Delta V_{max}$  was lower than zero or equal to zero, we assumed that the source was not utilized effectively by the microorganism community, and we set this value of  $V_{max}$  to zero. For the analyses of community-level physiological fingerprinting, we used the relative values of the respiration rate for each carbon source, calculated as the percentage share of each carbon source in the sum of  $V_{max}$  for the whole plate.

**Statistical analysis.** To group lakes according to their physicochemical parameters, a PCC analysis based on a correlation matrix was used (Statistica, StatSoft, TIBCO Software Inc.).

To group the lakes according to the microorganism metabolic and phylogenetic differences, Bray-Curtis-based NMDS (non-parametric multidimensional scaling) was performed. PAST3 software<sup>59</sup> was used for the Bray-Curtis similarity matrix calculation. Statistica software (StatSoft, TIBCO Software Inc.) was used for the NMDS analysis and data visualization. To determine whether the differences between eutrophic and meso-eutrophic lakes were statistically significant, the ANOSIM test was used<sup>60</sup>. To assess which taxa were primarily responsible for observed difference between groups, the SIMPER test was used<sup>60</sup>.

For diversity indices, we used richness (number of taxa or used carbon sources) and Simpson's diversity index (1- $D$ ) to measure the 'evenness' (PAST3 software).

To find the relationship between functional and phylogenetic diversity, the Mantel permutation test was used (PAST3 software). Mantel tests were performed to reveal the correlation between two Bray-Curtis similarity-based matrices of relative metabolic and relative phylogenetic data.

Nonparametric Spearman correlations were calculated using PAST3 software. For table and bar graph preparation, Excel (Microsoft) and Origin (OriginLab) were used.

## Data Availability

All sequencing reads have been deposited at the NCBI Archive - BioProject: PRJNA509870 (<https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA509870>). The remaining datasets generated and/or analyzed during the current study are available in the manuscript or from the corresponding author on reasonable request.

## References

- Smith, V. H. & Schindler, D. W. Eutrophication science: where do we go from here? *Trends Ecol. Evol.* **24**, 201–207 (2009).
- Zwisler, W., Selje, N. & Simon, M. Seasonal patterns of the bacterioplankton community composition in a large mesotrophic lake. *Aquat. Microb. Ecol.* **31**, 211–225 (2003).
- Warnecke, F., Sommaruga, R., Sekar, R., Hofer, J. S. & Pernthaler, J. Abundances, identity, and growth state of actinobacteria in mountain lakes of different UV transparency. *Appl. Environ. Microbiol.* **71**, 5551–5559 (2005).
- Wan, Y., Ruan, X., Zhang, Y. & Li, R. Illumina sequencing-based analysis of sediment bacteria community in different trophic status freshwater lakes. *Microbiologyopen* **6**, 1–15 (2017).
- Mackenzie, R., Barros, J. A. & Martínez, M. A. Characterization of aerobic heterotrophic bacteria in cold and nutrient-poor freshwater ecosystems. *World J. Microbiol. Biotechnol.* **27**, 2499–2504 (2011).
- Schiaffino, M. R. *et al.* Bacterial community structure in a latitudinal gradient of lakes: the roles of spatial versus environmental factors. *Freshw. Biol.* **56**, 1973–1991 (2011).
- Phillips, K., Zaidan, F., Elizondo, O. R. & Lowe, K. L. Phenotypic characterization and 16S rDNA identification of culturable non-obligate halophilic bacterial communities from a hypersaline lake, La Sal del Rey, in extreme South Texas (USA). *Aquat. Biosyst.* **8**, 5 (2012).
- Bollmann, A., Bullerjahn, G. S. & McKay, R. M. Abundance and Diversity of Ammonia-Oxidizing Archaea and Bacteria in Sediments of Trophic End Members of the Laurentian Great Lakes, Erie and Superior. *PLoS One* **9**, e97068 (2014).
- Zwirgmaier, K., Keiz, K., Engel, M., Geist, J. & Raeder, U. Seasonal and spatial patterns of microbial diversity along a trophic gradient in the interconnected lakes of the Oesterseen Lake District, Bavaria. *Front. Microbiol.* **6** (2015).
- Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D. & Bertilsson, S. *A guide to the natural history of freshwater lake bacteria. Microbiology and molecular biology reviews: MMBR* **75** (2011).
- Kennedy, K., Hall, M. W., Lynch, M. D. J., Moreno-Hagelsieb, G. & Neufeld, J. D. Evaluating bias of Illumina-based bacterial 16S rRNA gene profiles. *Appl. Environ. Microbiol.* **80**, 5717–5722 (2014).
- Feigl, V., Ujaczki, E., Vaszita, E. & Molnár, M. Influence of red mud on soil microbial communities: Application and comprehensive evaluation of the Biolog EcoPlate approach as a tool in soil microbiological studies. *Sci. Total Environ.* **595**, 903–911 (2017).
- Rutgers, M. *et al.* Monitoring soil bacteria with community-level physiological profiles using Biolog<sup>TM</sup> ECO-plates in the Netherlands and Europe. *Appl. Soil Ecol.* **97**, 23–35 (2016).
- Leflaive, J. *et al.* Nutrient effects on the genetic and functional diversity of aquatic bacterial communities. *FEMS Microbiol. Ecol.* **66**, 379–390 (2008).
- Lewis, D. E. *et al.* Soil Functional Diversity Analysis of a Bauxite-Mined Restoration Chronosequence. *Microb. Ecol.* **59**, 710–723 (2010).



16. Leon, C., Campos, V., Urrutia, R. & Mondaca, M.-A. Metabolic and molecular characterization of bacterial community associated to Patagonian Chilean oligotrophic-lakes of quaternary glacial origin. *World J. Microbiol. Biotechnol.* **28**, 1511–1521 (2012).
17. Jin, X. *et al.* Environmental Factors Influencing the Spatial Distribution of Sediment Bacterial Community Structure and Function in Poyang Lake. *Res. Environ. Sci.*, <https://doi.org/10.13198/j.issn.1001-6929.2017.01.51> (2017).
18. Comte, J. & del Giorgio, P. A. Linking the patterns of change in composition and function in bacterioplankton successions along environmental gradients. *Ecology* **91**, 1466–1476 (2010).
19. Dickerson, T. L. & Williams, H. N. Functional Diversity of Bacterioplankton in Three North Florida Freshwater Lakes over an Annual Cycle. *Microb. Ecol.* **67**, 34–44 (2014).
20. Wacnik, A. From foraging to farming in the Great Mazurian Lake District: palynological studies on Lake Miłkowskie sediments, northeast Poland. *Veg. Hist. Archaeobot.* **18**, 187–203 (2009).
21. Kauppinen, E. Trophic state of the Great Masurian Lakes system in the past, present and future - causes, mechanisms and effects of changes. PhD thesis, <https://depotuw.ceon.pl/bitstream/handle/item/467/Elsi%20Kauppinen%20thesis.pdf?sequence=1> (2013).
22. Siuda, W., Grabowska, K., Kaliński, T., Kiersztyn, B. & Chróst R. J. Trophic state, eutrophication and the threats for water quality of the Great Mazurian Lakes System. In: Korzeniewska, E. & Harnisz, M. (eds), Polish River Basins and Lakes, Part I. Hydrology and Hydrochemistry. In print, ISBN 978-3-030-12123-5 (SpringerVerlag, 2019).
23. Carlson, R. E. A trophic state index for lakes I. *Limnol. Oceanogr.* **22**, 361–369 (1977).
24. Chróst, R. J. & Siuda, W. Microbial production, utilization, and enzymatic degradation of organic matter in the upper trophogenic layer in the pelagial zone of lakes along a eutrophication gradient. *Limnol. Oceanogr.* **51**, 749–762 (2006).
25. Kiersztyn, B., Kauppinen, E. S., Kaliński, T., Chróst, R. & Siuda, W. Quantitative description of respiration processes in meso-eutrophic and eutrophic freshwater environments. *J. Microbiol. Methods* **149**, 1–8 (2018).
26. Krstulović, N., Šolić, M. & Marasović, I. Relationship between bacteria, phytoplankton and heterotrophic nanoflagellates along the trophic gradient. *Helgoländer Meeresuntersuchungen* **51**, 433–443 (1998).
27. Paerl, H. W. & Huisman, J. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. *Environ. Microbiol. Rep.* **1**, 27–37 (2009).
28. Reynolds, C. & Petersen, A. C. The distribution of planktonic Cyanobacteria in Irish Lakes in relation to their trophic states. *Hydrobiologia* **424**, 91–99 (2000).
29. Wood, S. A. *et al.* Trophic state and geographic gradients influence planktonic cyanobacterial diversity and distribution in New Zealand lakes. *FEMS Microbiol. Ecol.* **93**, fiv234 (2017).
30. Liu, L., Yang, J., Yu, Z. & Wilkinson, D. M. The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. *ISME J.* **9**, 2068–2077 (2015).
31. Goodfellow, M. & Williams, S. T. Ecology of Actinomycetes. *Annu. Rev. Microbiol.* **37**, 189–216 (1983).
32. Zwart, G., Crump, B. C., Kamst-van Agterveld, M. P., Hagen, F. & Han, S. K. Typical freshwater bacteria: An analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat. Microb. Ecol.* **28**, 141–155 (2002).
33. Glockner, F. O. *et al.* Comparative 16S rRNA Analysis of Lake Bacterioplankton Reveals Globally Distributed Phylogenetic Clusters Including an Abundant Group of Actinobacteria. *Appl. Environ. Microbiol.* **66**, 5053–5065 (2000).
34. Parveen, B., Mary, I., Vellet, A., Ravet, V. & Debroas, D. Temporal dynamics and phylogenetic diversity of free-living and particle-associated Verrucomicrobia communities in relation to environmental variables in a mesotrophic lake. *FEMS Microbiol. Ecol.* **83**, 189–201 (2013).
35. He, S. *et al.* Ecophysiology of Freshwater Verrucomicrobia Inferred from Metagenome-Assembled Genomes. *mSphere* **2** (2017).
36. Kolmonen, E., Sivonen, K., Rapala, J. & Haukka, K. Diversity of cyanobacteria and heterotrophic bacteria in cyanobacterial blooms in Lake Joutikas, Finland. *Aquat. Microb. Ecol.* **36**, 201–211 (2004).
37. Haukka, K. *et al.* Effect of nutrient loading on bacterioplankton community composition in lake mesocosms. *Microb. Ecol.*, <https://doi.org/10.1007/s00248-005-0049-7> (2006).
38. Paver, S. F. & Kent, A. D. Temporal Patterns in Glycolate-Utilizing Bacterial Community Composition Correlate with Phytoplankton Population Dynamics in Humic Lakes. *Microb. Ecol.* **60**, 406–418 (2010).
39. Šimek, K. *et al.* Maximum growth rates and possible life strategies of different bacterioplankton groups in relation to phosphorus availability in a freshwater reservoir. *Environ. Microbiol.* **8**, 1613–1624 (2006).
40. Dai, Y. *et al.* Spatiotemporal variation of planktonic and sediment bacterial assemblages in two plateau freshwater lakes at different trophic status. *Appl. Microbiol. Biotechnol.* **100**, 4161–4175 (2016).
41. Jankowski, K., Schindler, D. E. & Horner-Devine, M. C. Resource Availability and Spatial Heterogeneity Control Bacterial Community Response to Nutrient Enrichment in Lakes. *PLoS One* **9**, e86991 (2014).
42. Bell, T., Newman, J. A., Silverman, B. W., Turner, S. L. & Lilley, A. K. The contribution of species richness and composition to bacterial services. *Nature* **436**, 1157–1160 (2005).
43. Raes, E. J., Bodrossy, L., van de Kamp, J., Bissett, A. & Waite, A. M. Marine bacterial richness increases towards higher latitudes in the eastern Indian Ocean. *Limnol. Oceanogr. Lett.* **3**, 10–19 (2018).
44. Crossetti, L. O., Bicudo, D. C., Bicudo, C. E. M. & Bini, L. M. Phytoplankton biodiversity changes in a shallow tropical reservoir during the hypertrophication process. *Braz. J. Biol.* **68**, 1061–1067 (2008).
45. Delgado-Baquerizo, M. *et al.* Lack of functional redundancy in the relationship between microbial diversity and ecosystem functioning. *J. Ecol.* **104**, 936–946 (2016).
46. Fuentes, N., Güde, H., Wessels, M. & Straille, D. Allochthonous contribution to seasonal and spatial variability of organic matter sedimentation in a deep oligotrophic lake (Lake Constance). *Limnologica* **43**, 122–130 (2013).
47. Ng, H.-T., da Motta Marques, D., Jeppesen, E. & Søndergaard, M. Bacterioplankton in the littoral and pelagic zones of subtropical shallow lakes. *Hydrobiologia* **646**, 311–326 (2010).
48. Chróst, R. J. & Overbeck, J. Kinetics of alkaline phosphatase activity and phosphorus availability for phytoplankton and bacterioplankton in lake plu/see (North German Eutrophic Lake). *Microb. Ecol.* **13**, 229–248 (1987).
49. Siuda, W. & Güde, H. The role of phosphorus and organic carbon compounds in regulation of alkaline phosphatase activity and orthophosphate regeneration processes in lakes. *Pol Arch Hydrobiol.* **41**, 171–187 (1994).
50. Kuang, J. *et al.* Predicting taxonomic and functional structure of microbial communities in acid mine drainage. *ISME J.* **10**, 1527–1539 (2016).
51. Amin, S. A. *et al.* Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature* **522**, 98–101 (2015).
52. Porter, K. G. & Feig, Y. S. The use of DAPI for identifying and counting aquatic microflora. *Limnol. Oceanogr.* **25**, 943–948 (1980).
53. Kalinowska, K., Guśpiel, A., Kiersztyn, B. & Chróst, R. J. Factors controlling bacteria and protists in selected Mazurian eutrophic lakes (North-Eastern Poland) during spring. *Aquat. Biosyst.* **9**, 9 (2013).
54. Koroleff, F. Determination of phosphorus. In Grasshoff, K., Ehrhardt, M. & Kremling, K. (eds) *Methods of Seawater Analysis*, 125–139 (Verlag Chemie, 1983a).
55. Arar, E. J. & Collins, G. B. *In vitro determination of chlorophyll a and pheophytin a in marine and freshwater algae by fluorescence.* US Environmental Protection Agency Method 445.0 Revision 1.2, <https://doi.org/10.1007/BF02229570> (1997).
56. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
57. Beckers, B. *et al.* Performance of 16s rDNA Primer Pairs in the Study of Rhizosphere and Endosphere Bacterial Microbiomes in Metabarcoding Studies. *Front. Microbiol.* **7** (2016).



58. Newton, R. J. & McLellan, S. L. A unique assemblage of cosmopolitan freshwater bacteria and higher community diversity differentiate an urbanized estuary from oligotrophic Lake Michigan. *Front. Microbiol.* **6** (2015).
59. Hammer, Ø., Harper, D. A. & Ryan, D. D. Past: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontol. Electron. T. Harper. Geol. Museum*, <https://doi.org/10.1016/j.bcp.2008.05.025> (2001).
60. Clarke, K. R. Non-parametric multivariate analyses of changes in community structure. *Austral Ecol.* **18**, 117–143 (1993).

### Acknowledgements

We thank Ewelina Rogulska for her laboratory work. Part of field studies were performed in the Hydrobiological Station in Mikołajki, Nencki Institute of Experimental Biology of Polish Academy of Science, 3 Pasteur Str. 02-093 Warsaw, Poland. This work was supported by National Science Centre, Poland [Grant Number 2015/17/B/NZ9/01552].

### Author Contributions

B.K. conceived and planned the experiments, carried out the experiment, processed the experimental data, wrote the main manuscript text, prepared figures, R.C. supervised the project, T.K. performed the measurements, W.S. performed the measurements, A.B. analysed the data, G.K. verified the analytical methods, K.G. processed the experimental data, performed the analysis, drafted the manuscript. All authors reviewed the manuscript.

### Additional Information

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019