Reciprocal transplant experiment in lakes with disparate trophic states show resolution dependent partial functional redundancy

Kshitij Tandon[1,2,3], Min-Tao Wan[4], Chia-Chin Yang[1], Shan-Hua Yang[5], Bayanmunkh Baatar[1], Chih-Yu Chiu[1], Jeng-Wei Tsai[6], Wen-Cheng Liu[7], Chen Siang Ng[3] and Sen-Lin Tang[1,2]*

[1]Biodiversity Research Center, Academia Sinica, Taipei 11529, Taiwan

[2]Bioinformatics Program, Institute of Information Science, Taiwan International Graduate Program, Academia Sinica, Taipei 11529, Taiwan

[3]Institute of Molecular and Cellular Biology, National Tsing Hua University, Hsinchu 300, Taiwan

[4]EcoHealth Microbiology Laboratory, WanYu Co., Ltd., Chiayi 600, Taiwan

[5]Institute of Fisheries Science, National Taiwan University, Taipei 10617, Taiwan

[6]Department of Biological Science and Technology, China Medical University, Taichung, Taiwan

[7]Department of Civil and Disaster Prevention Engineering, National United University, Miao-Li, Taiwan

* Corresponding author

Corresponding author Email: sltang@gate.sinica.edu.tw

Abstract:

Microbial communities have long been assumed to have functional redundancy, but very few studies have tested this empirically. Here, we performed a time series *in situ* reciprocal transplant experiment in two lakes with disparate trophic states to empirically test for functional redundancy. We determined that the two lakes had different community compositions, and these did not change even after incubating in a different environment. However, functional attributes were different between read- and predicted gene-level analyses of the communities, highlighting the importance of resolution when delineating community functional relationships. Furthermore, there was a linear relationship between community composition and functional attributes, with a broad range of similarities in the former and a narrower range and overall more similarity in the latter, providing support for partial functional redundancy. Ecoplate analysis suggests that the total metabolic activity of a community is influenced by its local environment, even though broad functions like carbon metabolism are widespread. Together, our results support the presence of resolution-dependent partial functional redundancy in aquatic environments.

Introduction

Understanding the diversity and role of microbial communities is important because microbes are ubiquitous[1] and drive ecosystem processes by governing biogeochemical cycles[2,3,4]. In recent years, researchers have studied community compositions and functions and attempted to understand the community function relationship in different ecosystems ranging from oceans[5,6,7] to freshwater ecosystems[8,9,10], soil[11,12], plants[13,14], and the human gut[15,16]. These studies have put forward two plausible scenarios delineating community function relationships: 1) community and function are coupled so that changes in the community composition will change microbial-mediated processes, and 2) where community and function appear decoupled, microbe-mediated processes are somewhat independent of changes in community composition[17]. The latter is also known as functional redundancy.

Lately, functional redundancy has been categorized into strict and partial functional redundancy. Strict functional redundancy implies that different taxa or phylotypes can conduct the same set of metabolic processes and can readily replace each other[17,18,19]. Partial redundancy implies that certain organisms can coexist while sharing specific functions, but differing in other functions or ecological requirements[20]. The paradigm of strict functional redundancy does not hold true in all ecosystems, as shown in soil[21,22] and, recently, the oceans[7]. Furthermore, studies applying different molecular approaches to understand the community function relationship have yielded different conclusions. Studies using taxonomic profiling with the 16S rRNA gene or housekeeping genes[23,24,25] have reported a disconnect between community composition and function, whereas studies focused on ecology and microbial molecular evolution have reported a strong relationship between community composition and functional traits[17,26,27,28].

Terrestrial plants and animal ecosystem studies have provided the most insights into the community function relationship[9] in recent years. Unlike terrestrial ecosystems, aquatic ecosystems have different and unique features[29]. Studies of aquatic ecosystems, especially rivers and lakes, in both the field[30,31,32] and laboratory[33,34,35] have been central for deducing community function relationships; some of these studies support functional redundancy[10,36,37,38] and others challenge it[39,40,41].

Langenheder S *et al* (2005) used water samples from four lakes as an incubating medium and inocula of combinations of each inside controlled laboratory bioreactors to test the community function relationship in the aquatic environment. They reported that communities with different compositions can have the same functions, and hence there is strong functional redundancy in the aquatic environment. Whereas Delgado-Baquerizo M *et al* (2016), in their independent microcosm experiment, reported that a decline in microbial diversity has significant and direct consequences for both broad and specific ecosystem functions, concluding that there is either little or no functional redundancy in lake ecosystems. It is important to note that these studies and many others are confined to specific functions (biomass, respiration rate, growth rate, etc.)[8,9,10], and results of direct 16S rRNA gene sequencing rarely align with those of whole metagenomic approaches[42]. Although these studies have broadened our understanding of community function relationships in aquatic ecosystems, none have mimicked the *in vivo* dynamics of natural ecosystems.

Efforts have been made using *in vitro* reciprocal transplant experiments to test the hypothesis of functional redundancy in aquatic environmental studies, but most of these studies were performed for short durations of time and in artificially controlled laboratory environments[8,10,43,44]. *In situ* transplant experiments have been used to study shifts in bacterial

community composition as a response to nutrient availability[8] and link functional redundancy with metabolic plasticity in aquatic environments[45]. These studies are limited to measuring broad functions such as cumulative microbial respiration, dissolved organic carbon, biomass duplication time, and specific growth rate as proxies for the metabolic potential of a community when it is transplanted into a different environment. Hence, empirical research is needed to test the hypothesis of functional redundancy in aquatic ecosystems using high throughput approaches encompassing both community composition and function.

In this study, we performed an *in situ* experiment in which water (surface and bottom) from one lake was reciprocally transplanted into another lake. Inoculated water was maintained inside dialysis tubes resembling "microbial cages" that are exposed to ambient nutrient concentrations from the incubating lake[46]. We performed a multi-approach analysis using the 16S rRNA gene and gene transcripts for total and active community compositions, respectively. We used whole metagenome sequencing and 16S rRNA-based pseudo-metagenome for functional profiles and Biolog Ecoplates for metabolic activity of the microbial communities in the two lakes. Community function comparisons were performed to test the hypothesis of functional redundancy and also to study how changing the local environment affects the metabolic activity of microbial communities.

Materials and methods

Sampling sites

Two close proximity (21 km), sub-alpine lakes in northern Taiwan were chosen as experimental sites: Tsuei-Feng Lake (TFL) (24° 30' 52.5'', 121° 36' 24.8'' E) and Yuan-Yang Lake (YYL) (24° 34' 33.6'' N, 121° 24' 7.2'' E), at an altitude of 1,840 and 1,670 m,

respectively. These lakes, despite being at similar altitudes and geographical locations, have different trophic states, with TFL being oligotrophic[47] and YYL mesotrophic[48]. Both lakes are within a protected area, and are therefore introduced to minimal or no anthropogenic activity year-round. These lakes have been at the focus of limnological and meteorological research in Taiwan over the past years.

Experimental setup, sample collection, and preparation

Water samples from the surface (0.5 m deep) and bottom (1 m above the sediment) of both lakes were collected every two weeks from January-February 2015 (four times total) (Figure 1D). These samples were collected at the center of the lakes (also the deepest point) using a boat. Two L of lake water (BG) (Figure 1A) was filtered using a piece of gauze (to remove large debris, e.g., leaves), then sand and dust particles were removed using an 11-µm filter. Finally, 200 ml of water was filtered using a 0.22-µm filter (diameter: 47 mm, Advantec); samples were kept at -20° C before DNA extractions in the laboratory. For each sample, three replicates (n = 3) were collected.

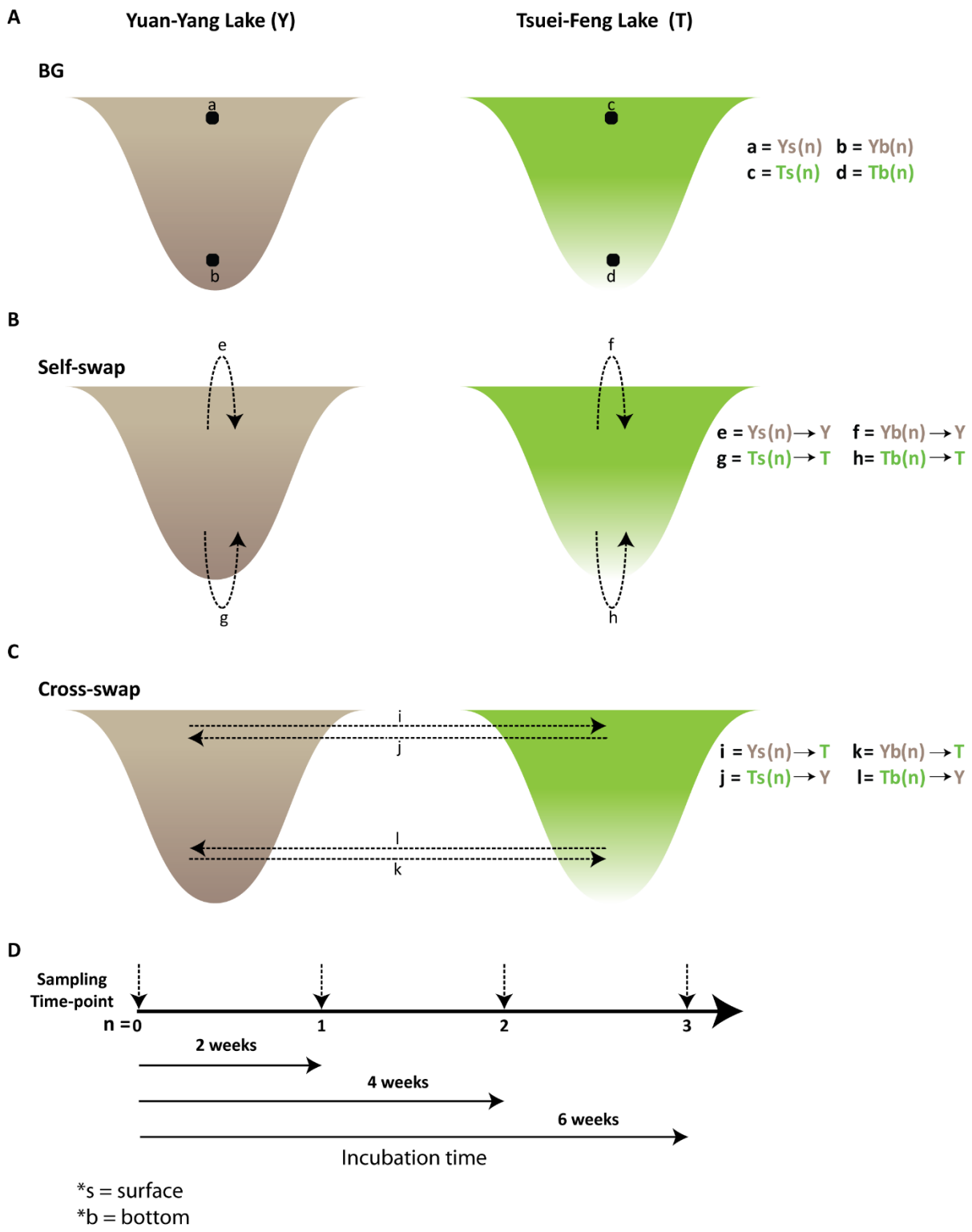

Figure 1: Schematic representation of the experiment. 1A) Sampling of background (BG) samples from the surface and bottom of Yuan-Yang (Y) and Tsuei-Feng (T) Lakes (a, b, c, d).

1B) Self-swap scheme with sample representation (e, f, g, h). **1C)** Cross-Swap scheme with sample representation (i, j, k, l). **1D)** Timeline for sample collection and the overall duration of the experiment.

To create an artificial disturbance (for both the surface and bottom), 6 L from each places in the lakes (surface and bottom) was swapped and incubated in the same and different lakes (e.g. TFL water was swapped and incubated in TFL (Self-swapped) and YYL (Cross-swapped)) (Figure 1B, 1C). The dialysis tube (Sigma-Aldrich, St. Louis, MO, USA) allows only ions (Mol. wt. <14 KDa) to pass through. These dialysis tubes were placed in specially designed metal cages.

Seventy-two dialysis tubes (36 surface and 36 bottom), each filled with water samples (~500 ml each), were incubated for six weeks. Every two weeks, samples from 12 tubes from each lake (6 self-swapped and 6 cross-swapped) were collected and the rest of the dialysis tube-containing water samples were replaced by new tubes to avoid algal growth on their outer surface. Collected water samples (~500 ml each tube) were filtered as described above and stored at -20°C before DNA extraction.

Nucleic acid extraction, PCR amplification, and sequencing

Total genomic DNA from filtered water samples was extracted using the UltraClean Soil DNA Kit (MoBio, Solana Beach, CA, USA) following the manufacturer's protocol. RNA from water samples was extracted using UltraClean Soil RNA Kit (MoBio, Solana Beach, CA, USA), and residual DNA was removed using a TURBO DNA-Free Kit (Invitrogen, Carlsbad, CA, USA). The nucleic acid yield was determined by NanoDrop ND-1000 UV-Vis Spectrophotometer (Nano-drop Technologies, Inc., Wilmington, DE, USA).

Complementary DNA (cDNA) was synthesized from purified RNA using the SuperScript IV First-Strand Synthesis System for RT-qPCR (Invitrogen, Carlsbad, CA, USA) following the manufacturer's protocol. To prepare 16S rRNA gene transcripts, PCR amplification was conducted to target the V6-V8 hypervariable regions of the bacterial 16S rRNA gene using the bacterial universal primers -968F (5'AACGCGAAGAACCTTAC-3')[49] and 1391R (5'-ACGGGCGGTGWGTRC-3')[50]. The reaction mixture contained 1 μl of 5 U TaKaRa *Ex Taq* HS (Takara Bio, Otsu, Japan), 5 μl of 10× *Ex Taq* buffer, 4 μl of 2.5 mM deoxynucleotide triphosphate mixture, 1 μl of each primer (10 μM), and 1-5 μl (10–20 ng) of template DNA in a volume of 50 μl. PCR was performed under the following conditions: 94°C for 5 min, 30 cycles of 94°C for 30 s, 52°C for 20 s, and 72°C for 45 s, with a final extension of 72°C for 10 min.

DNA-tagging PCR was used to tag each PCR product of the bacterial 16S rRNA gene's V6-V8 region[49]. The tag primer was designed with four overhanging nucleotides; this arrangement ensured 256 distinct tags—at the 5′ end of the 968F and 1391R primers—for bacterial DNA. The tagging PCR conditions comprised an initial step of 94°C for 3 min, 5 cycles at 94°C for 20 s, 60°C for 15 s, 72°C for 20 s, and a final step of 72°C for 2 min.

Illumina sequencing was performed on pooled 40-ng lots of uniquely marked samples by Yourgene Biosciences, Taiwan using the Miseq system. Six TruSeq DNA-PCR Free libraries (3: 16S rRNA gene and 3: 16S rRNA gene transcripts) were prepared for 2×300 bp paired-end reads by Yourgene Biosciences. Raw reads were sorted and primers removed before further analyses.

For whole-metagenome sequencing, the total genomic DNA was extracted using the UltraClean Soil DNA Kit (MoBio, Solana Beach, CA, USA) following the manufacturer's protocol. Extracted DNAs from the 41 samples were sent to Yourgene Biosciences (Taipei,

Taiwan) for library preparation without amplification (PCR Free library) and sequencing by the Illumina MiSeq system (USA).

Community and functional attribute analysis

Illumina sequenced 16S rRNA gene reads were sorted into samples using an in-house script. Sequences in each sample were screened by MOTHUR v1.3.81[51] to keep those 1) of lengths >390 bp and <450 bp, 2) with 0 ambiguous bases, and 3) with homopolymers <8 bp. UCHIME with USEARCH v8[52] (parameters: reference mode, rdp gold database, and mindiv of 3) was used to remove chimeric reads. After dereplication and removal of singletons, the reads were clustered at a 97% identity threshold to get cross-sample representative sequences as operational taxonomic units (OTUs) using the UPARSE[53] pipeline. The taxonomic classification of OTUs was performed by MOTHUR using SILVA 16S rRNA database v128[54,55]. 16S rRNA gene transcript read datasets were processed using the same approach outlined above.

Metagenomic reads obtained from the Illumina MiSeq system were preprocessed and the adapters removed using Cutadapt v1.4.2[56], quality trimming of reads was performed with Seqtk v. 1.2.-r94 (phred score <20, short read cut off <35 bp)[57], singlet reads were discarded. Metagenomes data were processed using two approaches (Figure S1). 1) Predicted ORF level: reads were *de-novo* assembled individually for each sample using CLC Genomics workbench version 1.10.1 (Qiagen) with default parameters, scaffolding was not performed, ORF prediction was performed using Prodigal[58] wrapped in Prokka[59] on contigs >200 bp long with metagenome mode enabled. 2) Reads level: quality filtered paired reads were merged using FLASH version 1.2.11[60]. These merged reads were used for the functional analysis.

Predicted ORFs and merged reads from each metagenome were searched for similarity using DIAMOND[61], implementing blastx[62] against the conserved domain database (CDD) of NCBI (last downloaded: 2018)[63] with e-value <1e-5 and aligned length of 50 amino acids. Annotated ORFs and reads were categorized at COG (Clusters of Orthologous Groups) Class and Family levels from the blastx results using in-house scripts. At the reads level, COG-Family represented in 90% of samples with at least 10 reads each, were used for the downstream analysis. Furthermore, previously-obtained 16S rRNA gene transcript abundance profiles were used to obtain imputed-metagenome functional profiles (at the COG Family level only) with PICRUSt2 (https://github.com/picrust/picrust2).

Ecoplates as the proxy for metabolic activity

The carbon assimilation potential of the microbial communities in the two lakes was assessed with Biolog EcoPlates™ (Biolog Inc., Hayward, CA, USA), hereafter written as ecoplates. Undiluted water (150 μl) from each sample was added to ecoplate wells, each containing 31 different carbon sources (and one control well with no carbon source). Plates were incubated at 25°C and the absorbance of each well was measured at 590 nm using an automatic optical density microplate reader (Spectramax M2, Molecular Devices, San Jose, CA, USA) every 24 hours for 7 days. On day 7, absorbance values from ecoplates were used to calculate average well color development (AWCD) in the plate, as expressed by the equation below.

$$AWCD = (\Sigma\,(R - C)/93)$$

where C and R represent the absorbance value in the blank well and substrate, respectively. Negative values were converted to 0. The absorbance values were then treated and normalized using AWCD.

Statistical analysis

All statistical analyses were performed in R[64]. Abundance profiles from the 16S rRNA gene and 16S rRNA gene transcript amplicon datasets were calculated using an in-house R script and processed with R packages phyloseq[65], vegan[66], and ggplot[67] for downstream analysis and visualization, including ordination, alpha and beta diversity calculations. Alpha diversities (Richness, Shannon, and Inv. Simpson) were measured using the estimate_richness function from phyloseq in R, with subsampling libraries with replacement 100 times and averaging the diversity estimates from each trial. Rarefaction curves were drawn using the ggrare function from the ranacapa[68] package on unrarefied samples. All the abundance profiles were log (x+1)-transformed to achieve normality. Beta diversity was performed based on an ordination analysis using the Bray-Curtis distance method unless otherwise stated. Permutational multivariate analyses of variance (PERMANOVA) were performed using the Adonis function in the vegan package. Ecoplate normalized absorbance values were also log (x+1)-transformed for PCA analysis and functional clusters were determined using gap statistics from the cluster[69] package in R with the K-means clustering method. Effects of the transplant experiment on carbon assimilation was tested using unpaired Student's t-test. The Mantel's Pearson correlation test was performed with Bray-Curtis distance for the community, functional data, and with Euclidean distance for ecoplate substrate utilization profiles. Microbial community and function differences

were quantified using *L*-divergence[70]. The linear relation between community and function was statistically tested with linear regression analysis and ANOVA.

Data Availability

All the sequencing data generated in the study has been deposited in the NCBI under the BioProject PRJNA596577. Other miscellaneous information's are available from the corresponding author upon request.

Results

Diversity analysis and its change

We sampled water from the surface and bottom of TFL and YYL every two weeks from January-February 2015. BG samples acted as the control for community composition and functional attributes of undisturbed lake water in their natural setting. Unrarefied samples were applied to generate rarefaction curves and determine how well the sequencing effort represented for the microbial diversity. Most of the samples from the 16S rRNA gene and 16S rRNA amplicons appeared to plateau, indicating that the sequencing adequately captured most of the microbial diversity in the two lakes (Figure S2, S3).

Alpha diversity (Shannon and Inv. Simpson) was measured after rarifying 16S rRNA gene samples at 27,689 (surface) and 26,238 (bottom) reads, and 16S rRNA samples at 25,513 (surface) and 23,131 (bottom) reads. YYL BG samples (Ys and Yb) had higher Shannon and Inv. Simpson diversities than TFL BG samples (Ts and Tb) (Table S1). We observed a mixed pattern in which diversity (Shannon and Inv. Simpson) in self-swap and cross-swap samples decreased from two to four weeks, and later increased from four to six weeks (Table S1).

Community function relationship in BG samples: hints of resolution dependence

The total community composition of BG samples from the two lakes was significantly different at the class and genus taxonomic levels. At the genus level, PCoA axis1 and axis2 explained >75% of the total variance in the surface and bottom levels, with inoculum (lake water) being the significant factor (Surface: $ADONIS_{inoculum}$: $F_{1, 6} = 16.01$, $p < 0.05$; Bottom: $ADONIS_{inoculum}$: $F_{1, 6} = 15.71$, $p < 0.05$) (Figure S4A).

There was no significant difference between the COG-Class based functional profile (ORF level) of BG samples from the two lakes (Table S2). Since COG-Class is a general classification for functional attributes, we performed a functional profile comparison at the finest resolution possible (i.e., COG-Family). At COG-Family (ORF-level), we observed that inoculum had a weak but significant effect driving the functional difference between BG samples from the lakes (Surface: $ADONIS_{inoculum}$: $F_{1, 6} = 2.85$, $p < 0.05$, Bottom: $ADONIS_{inoculum}$ $F_{1, 6}$: 2.69, $p < 0.05$) (Table S2).

We also performed a functional profile analysis on merged reads at the COG-Class and -Family levels. With reads at the COG-Family level, inoculum (lake water) had a strong and significant effect on the functional profile (COG-Family) of the BG samples (Surface: $ADONIS_{inoculum}$ $F_{1, 6} = 6.24$, $p < 0.05$; Bottom: $ADONIS_{incoculum}$: $F_{1, 6} = 4.152$, $p < 0.05$) (Figure S4B, Table S2). No significant difference was observed in reads at the COG-Class level (Table S2).

The different functional profiling outcomes for BG samples from the two lakes, based on the level of analysis (ORF vs Reads; COG-Class vs COG-Family), suggest a possible resolution dependency when delineating community function relationships. Hence, both approaches were

used (Metagenome level: ORF and Read; COG: Class and Family level) to compare the functional profiles and test the hypothesis that these lake microbial communities have functional redundancy.

The community functional relationship is resolution-dependent

The *in situ* reciprocal transplant experiment was conducted to empirically test the community function relationship in a temporal manner. The total community composition of the cross-swap samples was similar to their self-swap counterparts for both the surface and bottom samples. At the genus level, PCoA axis1 and axis2 explained >75% of the total variance (Figure 2A). Only the inoculum (not the incubating lake) had a significant effect (Surface: $ADONIS_{inoculum*incubating-lake}$: $F_{1, 8 (inoculum)} = 18.90$, $p < 0.05$; Bottom: $ADONIS_{inoculum*incubating-lake}$: $F_{1, 8 (inoculum)} = 20.44$, $p < 0.05$) (Table S2). This clustering of cross-swap samples with their self-swap counterparts spoke to the resistance and/or resilience of the bacterial community to external perturbations over time.

There was no significant difference in the functional profile of swapped samples (cross- and self-swap) between COG-Class and -Family when analyzed with ORF-level (Table S2). Nevertheless, when analyzed with reads-level at the COG-Family level, the inoculum had a significant effect on the clustering of the samples (Surface: $ADONIS_{inoculum*incubating-lake}$: $F_{1, 8 (inoculum)} = 4.02$, $p < 0.05$; Bottom: $ADONIS_{inoculum*incubating-lake}$: $F_{1, 8 (inoculum)} = 5.41$, $p < 0.05$) (Figure 2B). No significant difference was observed at the COG-Class level for the surface samples (Table S2). Furthermore, we observed that cross-swap samples tended to shift away from their self-swap counterparts with an increase in incubation time (Figure 2B marked by dashed arrows).

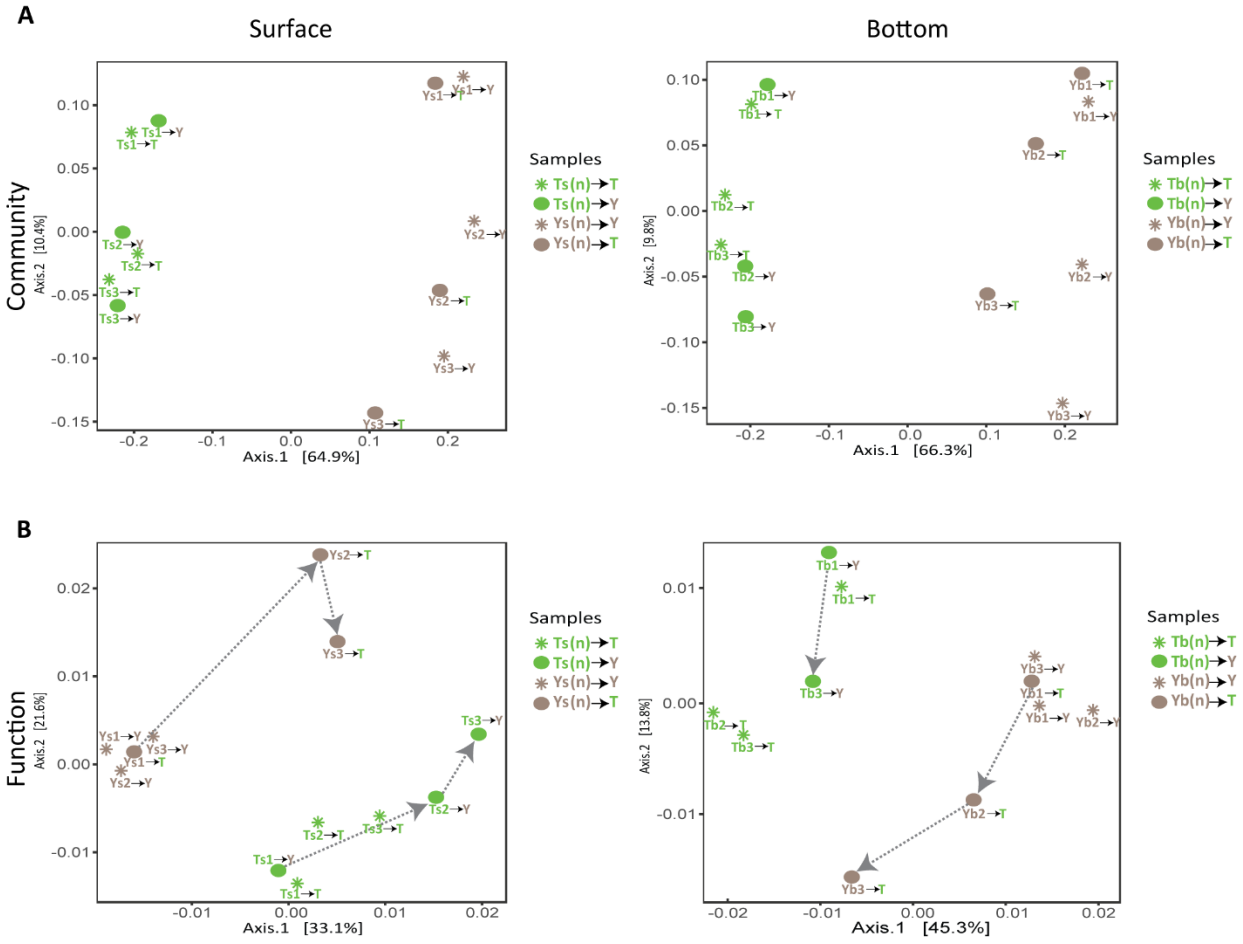

Figure 2: Principal Coordinate Analysis (PCoA) plots. 2A) 16S rRNA gene-based bacterial community composition was significantly different between the two lakes, and cross-swap samples clustered tightly with their self-swap counterparts**. 2B)** Functional profile (read-based COG-Family) derived from whole-metagenome sequencing was also significantly different between the swapped samples from the two lakes. Dotted grey arrows represent the shift in cross-swap samples away from their self-swap counterparts, indicating that the local environment had an effect on the functional profile.

Differences in community compositions and the results of the functional profile for swapped samples confirmed that community function relationships are resolution dependent. Furthermore, shifts in the functional profile of cross-swap samples away from their self-swap counterparts after two and four weeks of incubation in different lakes suggest that their community functions are dynamic. These shifts also suggest that the local environment plays an important role in shaping the community's functional profile.

Active community, its functional attributes, and partial functional redundancy

16S rRNA gene-based community composition and whole metagenome-based functional profiles represents the entire bacterial community and overall functional profile of the ecosystem, respectively. As not all bacteria are metabolically active at the same time, it is important to determine the community function relationship based on active community profiles, as metabolically active bacteria are one of the major driving forces in the relationship. Hence, we analyzed community profiles with 16S rRNA gene transcript amplicons and derived imputed pseudo-metagenome functional profiles (COG-Family level) from the active community composition using PICRUSt2.

The active community composition profile was positively correlated with the 16S rRNA gene-based total community profile ($R_{surface}$: 0.9318, $p < 0.01$; R_{bottom}: 0.9246, $p < 0.01$) (Table S3). Two lakes had significantly different community compositions at the genus level (Figure S5). PICRUSt2 derived a pseudo-metagenome-based functional profiles at the COG-Family level, and was also positively correlated with the whole metagenome-based functional profile (COG-Family: Reads level) ($R_{surface}$: 0.3430, $p < 0.01$; R_{bottom}: 0.2895, $p < 0.01$) (Table S3). PERMANOVA analysis showed that only the inoculum had a significant effect on sample

clustering (Surface: $ADONIS_{inoculum*incubating\text{-}lake}$: $F_{1,\ 8\ (inoculum)} = 16.65$, $p < 0.05$; Bottom: $ADONIS_{inoculum*incubating\text{-}lake}$: $F_{1,\ 8\ (inoculum)} = 15.05$, $p < 0.05$) (Table S2).

We performed a linear regression analysis based on the Bray-Curtis similarity (1-Bray-Curtis distance) on the active community and the PICRUSt2-based functional profile to test whether there is a linear relationship between bacterial community composition and functional profile at any similarity level. We found that similarities in community composition profiles were strongly correlated with functional attributes ($R^2_{surface} = 0.65$, $R^2_{bottom} = 0.63$). This linear relationship between community and function was based on a broad range of similarities in the community (>20% to <70%) and a relatively narrow but higher range of similarity in functional attributes (>90% to <97%). This result supports the idea that functional redundancy and functional dissimilarity co-exist in this ecosystem (i.e., partial functional redundancy) (Figure 3A).

Quantifying functional redundancy

L-divergence, which scales from 0 (identical) to 2 (highly divergent), was used to quantify functional redundancy. We identified a highly divergent community ($L\text{-divergence}_{surface}$: 1.02±0.48; $L\text{-divergence}_{bottom}$: 0.91±0.44) based on the 16S rRNA gene and a relatively similar functional profile ($L\text{-divergence}_{surface}$: 0.04±0.1; $L\text{-divergence}_{bottom}$: 0.039±0.41) based on whole metagenomes in the two lakes (Figure 3B). Furthermore, we also calculated *L*-divergence for community and functional profiles using different datasets from this study and obtained similar results (Table S4).

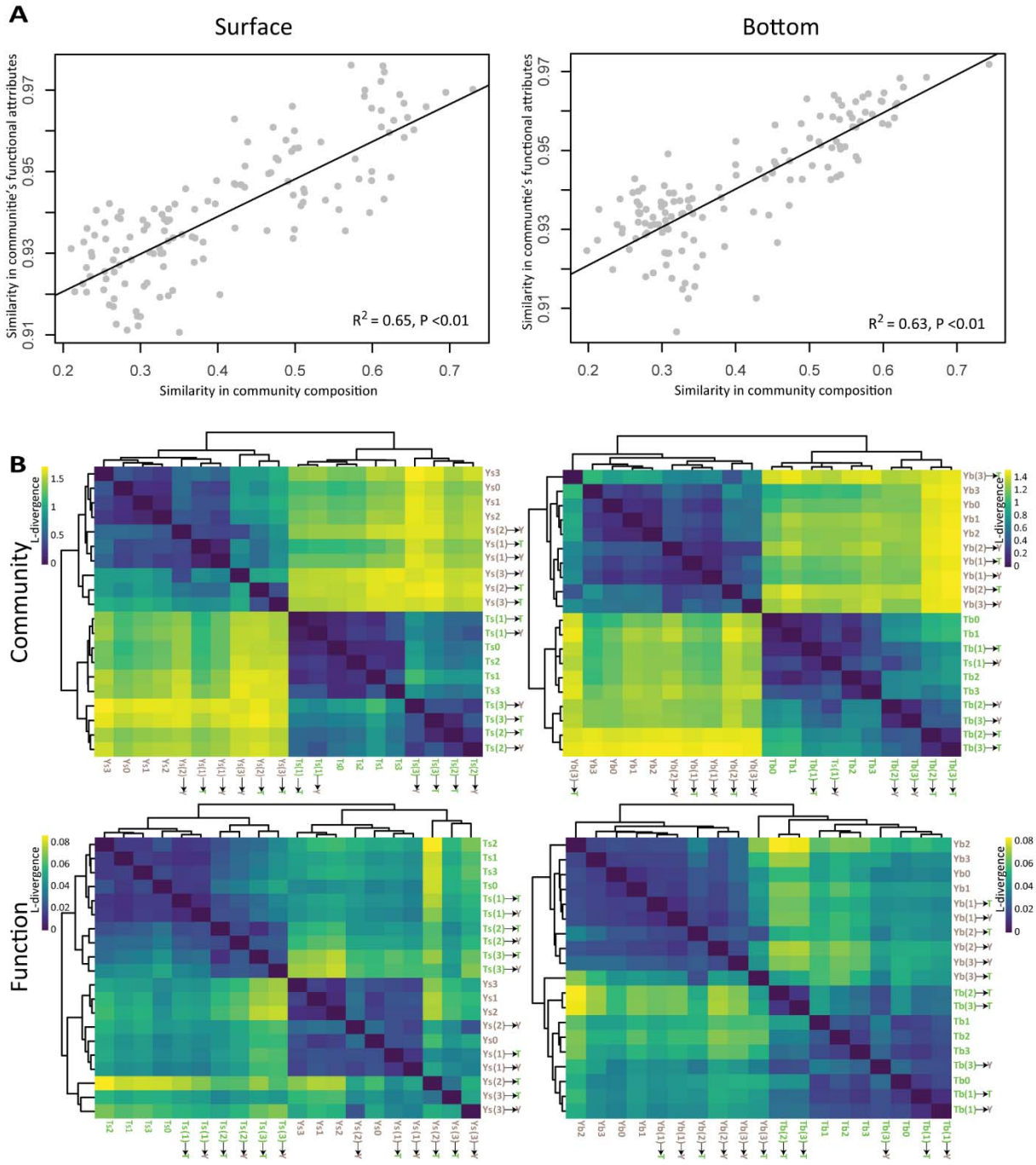

Figure 3: Community function relationship and *L*-divergence as a measure to quantify the functional redundancy. 3A) The relationship between the similarities in communities' (16S rRNA gene transcript-based) overall functional attributes (PICRUSt2-derived imputed

metagenome) and community composition of surface (R^2 = 0.65, p < 0.01, F test for the overall significance of the linear regression) and bottom samples (R^2 = 0.63, p < 0.01, F test for the overall significance of the linear regression). **3B)** Heat maps based on *L*-divergence measurements for community and overall functional attributes of surface and bottom samples.

Metabolic activity is influenced by the local environment

We used ecoplates as a proxy to both measure the metabolic activity (in terms of carbon metabolism) of the microbial communities in the two lakes and test temporal-physiological changes in metabolic activity during the reciprocal transplant experiment. Mean AWCD was used as an indicator of metabolic activity, and we observed that YYL lake BG samples had a higher mean AWCD than did TFL lake samples. Different values of AWCD were observed at different sampling time-points (Figure S6).

Swapped samples were used to test whether incubating the samples from the low nutrient environment (TFL) in the high nutrient environment (YYL) and vice versa affected metabolic activity. Two categories were used to compare samples. 1) Same inoculum in different incubating lakes (surface: Ts→Y vs Ts→T, Ys→T vs Ys→Y; bottom: Tb→Y vs Tb→T, Yb→T vs Yb→Y) 2) Different inoculum in the same incubating lake (surface: Ts→Y vs Ys→Y, Ys→T vs Ts→T; bottom: Tb→Y vs Yb→Y, Yb→T vs Tb→T).

While comparing samples with the same inoculum in different incubating lakes, we observed that the metabolic activity of Ts→Y samples was significantly higher than in Ts→T samples (Figure 4A) across all time points. However, Ys→T samples had significantly lower metabolic activity (except Ys3→T vs Ys3→Y) than Ys→Y samples (Figure S7A). In the bottom samples, the metabolic activity of Tb1→Y samples was significantly higher than Tb1→T

samples (Figure 4B), and no significant difference was observed at other sampling time points. Yb1→T samples had a lower metabolic activity than Yb1→Y samples, and no significant difference was observed between Yb2→T and Yb2→Y. However, Yb3→T had significantly higher metabolic activity than Yb3→Y (Figure S7B).

When the inoculum was different in the same incubating lake, we observed that Ys→T samples had higher metabolic activity than Ts→T samples across all sampling time points (Figure 4**C**), and Ys→Y samples had higher metabolic activity than Ts→Y samples (except Ts3→Y vs Ys3→Y) (Figure S7C). Of the bottom samples, Yb→T metabolic activity was higher than Tb→T (Figure 4D), and Yb→Y was higher than Tb→Y (except Tb3→Y) (Figure S7D).

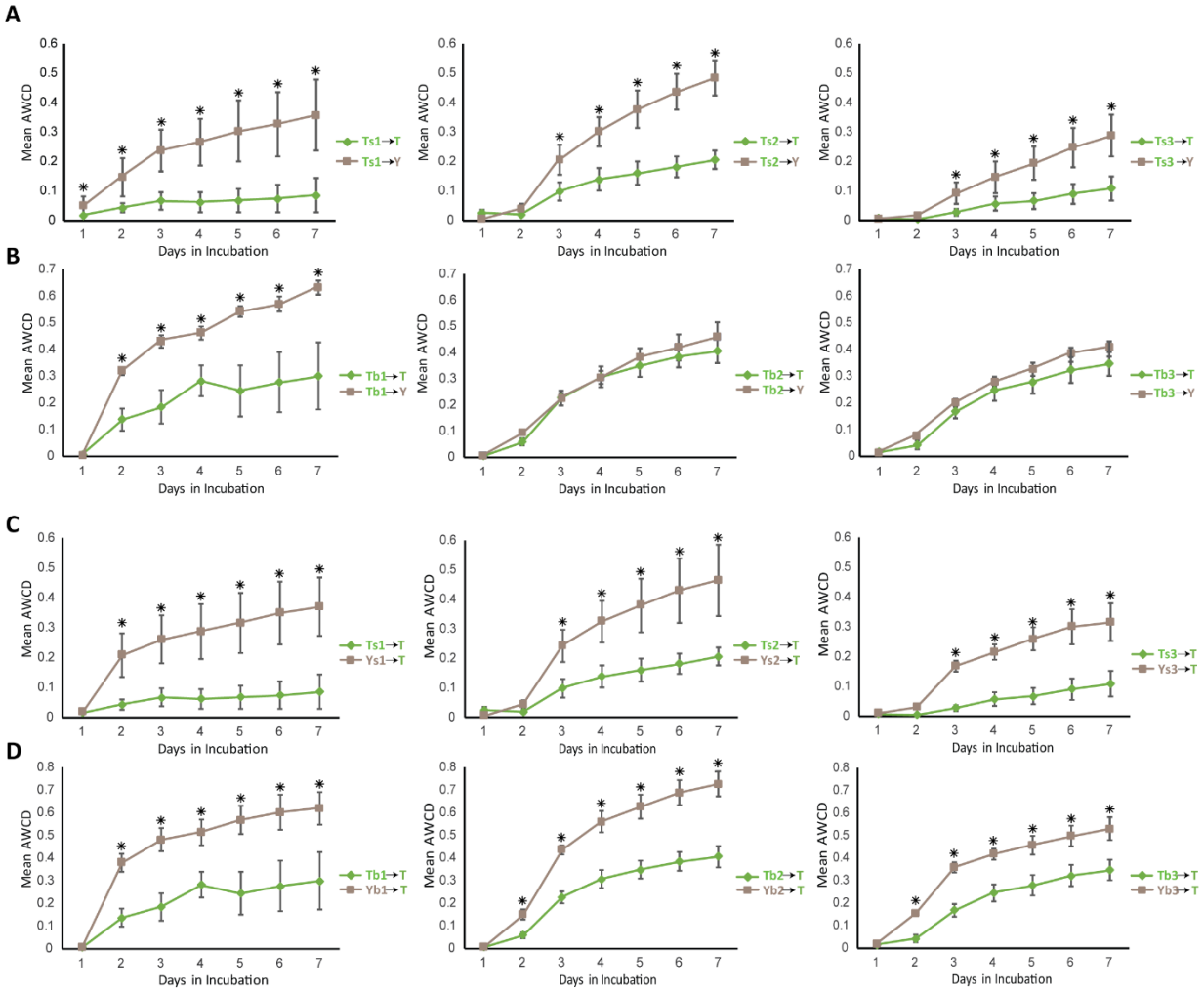

Figure 4: Effect of average well color development (AWCD) as a measure of metabolic activity of the community with a change in the local environment**. A)** and **B)** TFL self-swap samples (Ts(n)→T; Tb(n)→T) compared with cross-swap samples (Ts(n)→Y; Tb(n)→Y). **C)** and **D)** TFL self-swap samples (Ts(n)→T; Tb(n)→T) compared with cross-swap samples (Ys(n)→T; Yb(n)→T). Significant difference tested with Student's t-test $p < 0.05$. Error bars indicate standard deviation.

Discussion

In recent years, the focus of microbial ecology studies has shifted from understanding the community diversity to understanding what microbes are doing and how they respond to changes in the environment, and delineating community function relationships in diverse ecosystems[5,6,7,14,71,72]. With advancements in sequencing technology and relative improvements in metagenomic analyses, efforts have been made to establish community function relationships in pristine and disturbed ecosystems. Recently, laboratory and field studies from different environments testing community function relationships have provided insightful results[5,6,7,10,37,41,44,45,73,74,75,76]. Yet, our understanding of community function relationships and their magnitudes in natural environments remains poor because of the lack of *in situ* empirical time series studies applying high-throughput sequencing technology alongside functional approaches. In this setting, the main motivation of our study was to empirically test the hypothesis of functional redundancy in the aquatic ecosystem with an *in situ* reciprocal transplant experiment. We used high throughput amplicon, whole metagenome sequencing, and ecoplate analysis to determine the community function relationship in the aquatic ecosystem. To the best of our knowledge, this is the first study that examines the function redundancy hypothesis in surface and bottom lake waters with time series reciprocal transplant experiments *in situ.*

The 16S rRNA gene-based bacterial community composition of BG samples was significantly different between lakes (Figure S4A), which was expected, as the two lakes have different trophic states[47,48,77]. However, results from the functional profile analysis have unexpected implications for resolution dependency and the community function relationship. ORF-based functional profiles (BG samples) for COG-Class and -Family were not significantly different between the two lakes, suggesting that functional redundancy is present (Table S2).

Functional profiles supported functional redundancy in previous studies as well when analyzed at the ORF level[6,78]. Read-based functional profiles, however, were significantly different for the two lakes (Figure S4B), challenging our conclusions about functional redundancy and suggesting that resolution dependency helps determine the community function relationship. Recently, *Galand PE et al*, (2018) using sequenced reads from a three-year survey, determined a functional profile and identified a lack of strict functional redundancy in the oceans. They argued that the functional profile, when examined at the predicted gene level, can significantly undermine the overall functional profile of ecosystems, as predicted genes are extrapolated from the similarity between cultured and taxonomic annotations of environmental samples. Nonetheless, we propose that clustering reads directly to assign functions can also give spurious results that have little significance without taxonomy. Thus, we used reads directly to get functional profiles of the lake water community using *a priori* knowledge of operational databases.

Studies in recent years have explored community responses and functional profiles in different aquatic environments using reciprocal transplant experiments and other approaches[9,10,44,45,79]. We identified a significant difference in microbial communities in two lakes, even after reciprocal transplantation and incubation for six weeks. Cross-swap samples clustered with their self-swap counterparts for both surface and bottom samples (Figure 2A). Furthermore, inoculum had a significant effect on the clustering of the samples (Table S2). Our results are similar to *Langenheder S et al*, (2006); thus, we suggest that the bacterial community was more or less resistant to change, which is consistent with previous studies[17,74]. We cannot, however, rule out the possibility that the bacterial communities are resilient and can recover from a disturbance to an original or alternate stable state[80] because our sampling interval was two weeks—which is longer than most sample time point experiments—and also because of the

shifts observed within the cluster of cross-swap samples (Figure 2A, Figure S5). We believe that the community might have varied immediately or very shortly after transplantation, but then recovered to an alternate stable state in less than two weeks. Bacterial communities in a lake ecosystem was demonstrated to recover quickly after natural disturbance (lake mixing)[76]. Earlier studies have shown bacterial communities to be both resilient and resistant in various ecosystems[17,74].

We hypothesized that transplanting an inoculated sample from one lake into another lake with an entirely different trophic state would influence the sample's functional profile. A similar hypothesis was tested and confirmed in an earlier study[3]. Interestingly, we observed shifts in the functional profile of cross-swap samples away from their self-swap counterparts (Figure 2B, marked as dashed arrows). This phenomenon was observed for Ys→T and Yb→T samples, but was not as clearly for Ts→Y and Tb→Y samples (Figure 2B). During our experiment, the Tb2→Y sample (after four weeks of incubation) was lost, and hence the shift in Tb→Y samples cannot be determined accurately. Furthermore, transplant samples behaved differently in the two lakes, suggesting that local environment influences the functional profile of a community, maybe due to the nutrient availability in that environment and how it differs from that of the original environment[47,48]. Although these shifts are not quantified, they support our hypothesis that local environment influences the functional profile of the swapped bacterial community, as was previously reported in plants[14] and animals[81]. A long-time series analysis would help determine whether communities in a new environment eventually settle into an alternate stable state.

Partial functional redundancy has been reported in different ecosystems[7,10,41]. Therefore, we speculated that there is a linear relationship between community composition and ecosystem function. Regression analysis using active community composition and the imputed metagenome

functional profile identified a linear relationship between the community and its function in the ecosystem of both lakes (Figure 3A). Our result contrasts with a study on the marine ecosystem, where the range of similarities between community composition and function profile was of a similar magnitude[7]; this might be due to the aquatic environment studied and the duration and frequency of sampling. Furthermore, a relatively highly divergent community (L-divergence$_{surface}$: 1.02±0.48; L-divergence$_{bottom}$: 0.91±0.44) and relatively low divergent functional profile (L-divergence$_{surface}$: 0.04±0.1; L-divergence$_{bottom}$: 0.039±0.41) (Figure 3B) between the two lakes also confirms a partial functional redundancy, whose magnitude varies based on the resolution of the analysis (Table S4). Significant high positive correlations were obtained between the 16S rRNA gene and 16S rRNA gene transcript-based community composition and also with whole metagenome and PICRUSt2 derived imputed metagenome functional profiles (Table S3). This implies that either of the approaches can be used to determine the community function relationship in our study, and their results will be not differing significantly.

Natural communities can use diverse carbon and energy sources for growth and survival[82]. Microbes, especially bacteria, often overcome changes in the environment by expressing a range of metabolic capabilities or mixotrophy[83,84]. Therefore, ecoplates containing carbon sources of different complexities were used to test whether a change in the local environment has any consequence on the total metabolic activity of the bacterial community; they identified a similar potential substrate utilization pattern to other studies[40,43]. Over seven days of incubation and observation, and using carbon metabolism as a proxy for overall metabolism, we determined that the total metabolic activity of Ts→Y samples were higher than Ts→T ones (Figure 4A), and that of Ys→T samples was lower (except Ys3→T) than Ys→Y

ones (Figure S7A). Similar results were obtained for bottom samples (Figure 4B and Figure S7B). Furthermore, Ys→T and Yb→T samples had higher metabolic activity than Ts→T and Tb→T, respectively (Figure 4C, 4D), and Ts→Y and Tb→Y had lower metabolic activity than Ys→Y and Yb→Y, respectively (Figure S7C, S7D). These changes in total metabolic activity highlight the effect that local environment can have on the total metabolic activity of a bacterial community. These results also provide important evidence that metabolic activity differs distinctly between the surface and bottom waters of the lakes. We also demonstrate that inoculum has either a very weak or no effect on functional profiles in terms of carbon substrate utilization (Figure S8, Table S5), which is similar to findings from other studies[8,10,39,85]. Lack of correlation between bacterial community functional attributes and composition has been interpreted as evidence of functional redundancy[17,86]. However, the significantly high correlation suggests that functional redundancy may be constrained. We observed no or very small positive correlation between total or active bacterial community and substrate utilization profiles based on ecoplates (Table S3). A similar result was observed in boreal freshwater bacterioplankton communities and coastal waters[40,87]. This may again suggest that some substrates cause functional redundancy.

In the present study, we used a reciprocal transplant experiment in which the bacterial community from one lake was grown in another, very different lake. Dialysis tubing bags, which acted as microbial cages, provided bacteria access to the nutrients in their transplant environment. Although dialysis bags are not perfect cages, they are much more realistic than any experimental bottles or reactors. Our sampling strategy did not determine whether the bacterial communities in the two lakes were resilient to change; future experiments aiming to answer this question should use shorter intervals between sampling. Functional redundancy is a difficult

concept to quantify, and we still do not know its magnitude and variability among different ecosystems, so care should be taken when applying results from this study to other environments.

Conclusion

Establishing the community function relationship has been central to understanding the complexity of microbial diversity and their functions in different ecosystems. Laboratory and field studies have made important findings on this subject in recent years. Yet, few *in situ* long term experiments have been conducted empirically to test the hypothesis of functional redundancy. Furthermore, most previous studies were limited to measuring specific functions, and hence they could not make conclusions about overall ecosystem dynamics. This study is the first to test the community function relationship through reciprocal transplant experiments in lake waters (surface and bottom) using a time series, utilizing amplicon sequencing complemented with whole metagenome sequencing and metabolic analysis using ecoplates. The insights offered by this work deciphering the community function relationship of the freshwater aquatic ecosystem will add much-needed knowledge in this field.

We found that bacterial community composition is resistant and/or resilient to change in the local environment. Whereas, the functional profile is influenced by the local environment. Further, the community function relationship is dependent on the resolution of the analysis. At the predicted ORF level, we observed no significant difference in the community functional profile of the two lakes. However, at the reads level, a significant difference in the community functional profile of the two lakes was observed. Moreover, a significant correlation between community composition (broad similarity range) and community functional attributes (small but higher similarity range) supports the presence of partial functional redundancy in aquatic environments. Metabolic activity analysis using ecoplates also provides strong support for the

presence of functional redundancy in the aquatic environments and also provides evidence for the influence of the local environment on the metabolic activity. Further, the two lakes and their surface and bottom waters show differing effects on the community and functional profile behavior when disturbed artificially.

Acknowledgements

The authors would like to thank Dr. Ching-Hung Tseng, Dr. Jiun-Yan Ding, Mr. Cheng-Yu Yang, Dr. Sonny T. M. Lee, and Dr. Carol Enumi Lee for their assistance with sampling. K.T would like to acknowledge the Taiwan International Graduate Program (TIGP) for its fellowship. The authors also express their gratitude to the anonymous reviewers for their insightful comments and suggestions to improve the quality of the manuscript, and to Mr. Noah Last of Third Draft Editing for his English language editing. This study was supported by the Thematic Project at Academia Sinica (AS-103-TP-B15-3).

Contributions

S.L.T conceived of the idea for this study. K.T performed all the analyses and wrote and revised the manuscript. M.W and C.Y performed sampling and experiments. S.Y played an active role in the discussion, interpretation of results, and writing of the manuscript. B.B, C.C, and J.T helped with sampling and provided logistical support for the experimental setup. C.S.N provided critical comments on the manuscript. All authors read and approved the current version of the manuscript. All authors also declare that they have no conflict of interest.

Competing Interests

All authors declare no competing interests.

References

1. Hug, L. A. *et al.* A new view of the tree of life. *Nat. Microbiol.* **1**, 1–6 (2016).

2. Ren, Z. *et al.* Taxonomic and functional differences between microbial communities in Qinghai Lake and its input streams. *Front. Microbiol.* **8**, 1–14 (2017).

3. Graham, E. B. *et al.* Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes? *Front. Microbiol.* **7**, 1–10 (2016).

4. Falkowski, P. G., Fenchel, T. & Delong, E. F. The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science.* **320**, 1034-1039 (2008).

5. Tringe, S. G. *et al.* Comparative metagenomics of microbial communities. *Science.* **308**, 554–557 (2005).

6. Louca, S., Parfrey, L. W. & Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome. *Science.* **353**, 1272-1277 (2016).

7. Galand, P. E., Pereira, O., Hochart, C., Auguet, J. C. & Debroas, D. A strong link between marine microbial community composition and function challenges the idea of functional redundancy. *ISME J.* **12**, 2470–2478 (2018).

8. Gasol, J. M. *et al.* A transplant experiment to identify the factors controlling bacterial abundance, activity, production, and community composition in a eutrophic canyon-shaped reservoir. *Limnol. Oceanogr.* **47**, 62–77 (2002).

9. Peter, H. *et al.* Function-specific response to depletion of microbial diversity. *ISME J.* **5**, 351–361 (2011).

10. Langenheder, S., Lindström, E. S. & Tranvik, L. J. Weak coupling between community

composition and functioning of aquatic bacteria. *Limnol. Oceanogr*. **50**, 957–967 (2005).

11. Griffiths, B. S. *et al.* The relationship between microbial community structure and functional stability, tested experimentally in an upland pasture soil. *Microb. Ecol.* **47**, 104–113 (2004).

12. Bahram, M. *et al.* Structure and function of the global topsoil microbiome. *Nature* **560**, 233–237 (2018).

13. Reich, P. B. *et al.* Impacts of biodiversity loss escalate through time as redundancy fades. *Science.* **336**, 589–592 (2012).

14. Louca, S. *et al*. Functional structure of the bromeliad tank microbiome is strongly shaped by local geochemical conditions. *Environ. Microbiol.* **19**, 3132–3151 (2017).

15. Turnbaugh, PJ. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 222–227 (2009).

16. Moya, A. & Ferrer, M. Functional Redundancy-Induced Stability of Gut Microbiota Subjected to Disturbance. *Trends Microbiol*. **24**, 402–413 (2016).

17. Allison, S. D. & Martiny, J. B. H. Resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci.* USA. **105**, 11512–11519 (2008).

18. Yin, B., Crowley, D., Sparovek, G., De Melo, W. J. & Borneman, J. Bacterial functional redundancy along a soil reclamation gradient. *Appl. Environ. Microbiol*. **66**, 4361–4365 (2000).

19. Fuhrman, J. A., Cram, J. A. & Needham, D. M. Marine microbial community dynamics and their ecological interpretation. *Nat. Rev. Microbiol.* **13**, 133–146 (2015).

20 Jurburg, S.D., & Salles, J.F. Functional Redundancy and Ecosystem Function- The Soil Microbiota as a Case Study. *Ecosystems-Linking Strucuture and Function.* p29-42. InTech Oen Science. Rijeka (2015)

21. Strickland, M. S., Lauber, C., Fierer, N. & Bradford, M. A. Testing the functional significance of microbial community composition. *Ecology* **90**, 441–451 (2009).

22. Fierer, N. *et al.* Prairie Soils in the United States. *Science.* **342**, 621–625 (2013).

23. Jaspers, E and Overmann, J. Ecological Significance of Microdiversity: Identical 16S rRNA Gene Sequences Can Be Found in Bacteria with Highly Divergent Genomes and Ecophysiologies. *Appl. Environ. Microbiol.* **70**, 4831-4839 (2004).

24. Doolittle, WF and Zhaxybayeva, O. On the origion of prokaryotic species. *Genome Res* **19**(5), 744-756 (2009).

25. Maharjan, R., Seeto, S., Notley-McRobb, L. & Ferenci, T. Clonal adaptive radiation in a constant environment. *Science.* **313**, 514–517 (2006).

26. Gravel, D. *et al.* Experimental niche evolution alters the strength of the diversity-productivity relationship. *Nature* **469**, 89–94 (2011).

27. Martiny, J. B. H. *et al.* Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* **4**, 102–112 (2006).

28. Martiny, A. C., Treseder, K. & Pusch, G. Phylogenetic conservatism of functional traits in microorganisms. *ISME J.* **7**, 830–838 (2013).

29. Giller, P. S. *et al.* Biodiversity effects on ecosystem functioning: Emerging issues and their experimental test in aquatic environments. *Oikos* **104**, 423–436 (2004).

30. Shade, A. *et al.* Interannual dynamics and phenology of bacterial communities in a eutrophic lake. *Limnol. Oceanogr.* **52**, 487–494 (2007).

31. Jones, S. E., Newton, R. J. & McMahon, K. D. Evidence for structuring of bacterial community composition by organic carbon source in temperate lakes. *Environ. Microbiol.* **11**, 2463–2472 (2009).

32. Nelson, C. E. Phenology of high-elevation pelagic bacteria: The roles of meteorologic variability, catchment inputs and thermal stratification in structuring communities. *ISME J.* **3**, 13–30 (2009).

33. Judd, K. E., Crump, B. C. & Kling, G. W. Variation in dissolved organic matter controls bacterial production and community composition. *Ecology* **87**, 2068–2079 (2006).

34. Kritzberg, E. S., Langenheder, S. & Lindström, E. S. Influence of dissolved organic matter source on lake bacterioplankton structure and function - Implications for seasonal dynamics of community composition. *FEMS Microbiol. Ecol.* **56**, 406–417 (2006).

35. Newton, R. J. & Mcmahon, K. D. Seasonal differences in bacterial community composition following nutrient additions in a eutrophic lake. *Environ. Microbiol.* **13**, 887–899 (2011).

36. Fernandez, A. S. *et al.* Flexible community structure correlates with stable community function in methanogenic bioreactor communities perturbed by glucose. *Appl. Environ. Microbiol.* **66**, 4058–4067 (2000).

37. Cheaib, B., Boulch, M. Le, Mercier, P. L. & Derome, N. Taxon-function decoupling as an adaptive signature of Lake Microbial metacommunities under a chronic polymetallic

pollution gradient. *Front. Microbiol.* **9**, 1–17 (2018).

38. Tseng, C. H. *et al.* Microbial and viral metagenomes of a subtropical freshwater reservoir subject to climatic disturbances. *ISME J.* **7**, 2374–2386 (2013).

39. Findlay, S. E. G., Sinsabaugh, R. L., Sobczak, W. V. & Hoostal, M. Metabolic and structural response of hyporheic microbial communities to variations in supply of dissolved organic matter. *Limnol. Oceanogr.* **48**, 1608–1617 (2003).

40. Traving, S. J. *et al.* Coupling bacterioplankton populations and environment to community function in coastal temperate waters. *Front. Microbiol.* **7**, 1–13 (2016).

41. Delgado-Baquerizo, M. *et al.* Lack of functional redundancy in the relationship between microbial diversity and ecosystem functioning. *J. Ecol.* **104**, 936–946 (2016).

42. Lima-Mendez, G. *et al.* (2015) Determinants of community structure in the grobal plankton interactome.pdf. *Ocean Plankt.* **348**, 1–10 (2015).

43. Kirchman, D. L., Dittel, A. I., Findlay, S. E. G. & Fischer, D. Changes in bacterial activity and community structure in response to dissolved organic matter in the Hudson River, New York. *Aquat. Microb. Ecol.* **35**, 243–257 (2004).

44. Shade, A., Chiu, C.-Y. & McMahon, K. D. Differential bacterial dynamics promote emergent community robustness to lake mixing: an epilimnion to hypolimnion transplant experiment. *Environ. Microbiol.* **12**, 455–66 (2010).

45. Comte, J., Fauteux, L. & Del Giorgio, P. A. Links between metabolic plasticity and functional redundancy in freshwater bacterioplankton communities. *Front. Microbiol.* **4**, 1–11 (2013).

46. Herndl, G.J., Kaltenbok, E., Muller-Niklas, G. Dialysis bag incubation as a non radiolabeling technique to estimate bacterioplankton production in situ. *Handbook of Methods in Aquatic Microbial Ecology*. 495-503. Lewis Publisher, Florida (1993).

47. Wang, L. C. *et al.* Increased precipitation during the Little Ice Age in northern Taiwan inferred from diatoms and geochemistry in a sediment core from a subalpine lake. *J. Paleolimnol.* **49**, 619–631 (2013).

48. Tsai, J. W. *et al.* Seasonal dynamics, typhoons and the regulation of lake metabolism in a subtropical humic lake. *Freshw. Biol.* **53**, 1929–1941 (2008).

49. Chen, C. P., Tseng, C. H., Chen, C. A. & Tang, S.-L. The dynamics of microbial partnerships in the coral Isopora palifera. *ISME J.* **5**, 728–740 (2011).

50. Jorgenson, S.L. *et .al.* Correlating microbial community profiles with geochemical data in highly stratified sediments from the Arctic Mid-Ocean Ridge. *Proc. Natl. Acad. Sci*. **109**(42), 2846-2855. (2012).

51. Schloss, P. D. *et al.* Introducing mothur: Open-source, platform-independent, community supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).

52. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).

53. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–998 (2013).

54. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data

processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).

55. Yilmaz, P. *et al.* The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* **42**, D643-D648 (2014).

56. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12 (2011).

57. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* **11**, 1–10 (2016).

58. Hyatt, D., Chen, G.L., Locascia, P.F., Land, M.L.,Larimer, F.W., & Hauser, L.J. Prodigal: prokaryotic gene recognition and translation initiation site identificatin. *BMC Bioinformatics* **11**, 119. (2010).

59. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).

60. Magoč, T. & Salzberg, S. L. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).

61. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2014).

62. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

63. Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**, D222–D226 (2015).

64. R Developmetn Core Team. R: A Language and Environment for Staistical Compting. R

Foundation for Statistical Computing. Vienna, Austria. (2016).

65. McMurdie, P. J. & Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**, (2013).

66. Okansen, J. *et al.* vegan: Community Ecology Package: Ordination, Diversity and Dissimilarities. R package 2.5-5 (2011).

67. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. R package version 3.2.0. Spinger-Verlag New York. (2016)

68. Kandlikar, G. S. *et al.* Ranacapa: An R package and shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. *F1000Research* **7**, 1–19 (2018).

69. Maechler, M. Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. cluster: Cluster Analysis Basics and Extensions. R package version 2.1.0. (2019).

70. Lin, J. Divergence Measures Based on the Shannon Entropy. *IEEE Trans. Inf. Theory* **37**, 145–151 (1991).

71. Mooshammer, M. *et al.* Decoupling of microbial carbon, nitrogen, and phosphorus cycling in response to extreme temperature events. *Sci. Adv.* **3**, 1–14 (2017).

72. Orland, C. *et al.* Microbiome functioning depends on individual and interactive effects of the environment and community structure. *ISME J.* **13**, 1–11 (2019).

73. Debroas, D. *et al.* Metagenomic approach studying the taxonomic and functional diversity of the bacterial community in a mesotrophic lake (Lac du Bourget - France). *Environ. Microbiol.* **11**, 2412–2424 (2009).

74. Shade, A. *et al*. Resistance, resilience and recovery: Aquatic bacterial dynamics after water column disturbance. *Environ. Microbiol.* **13**, 2752–2767 (2011).

75. Baho, D. L., Peter, H. & Tranvik, L. J. Resistance and resilience of microbial communities - Temporal and spatial insurance against perturbations. *Environ. Microbiol.* **14**, 2283–2292 (2012).

76. Shade, A. *et al*. Lake microbial communities are resilient after a whole-ecosystem disturbance. *ISME J.* **6**, 2153–2167 (2012).

77. Tandon, K. *et al*. Bacterial community in water and air of two sub-alpine lakes in taiwan. *Microbes Environ.* **33**, 120–126 (2018).

78. Sunagawa, S. *et al*. Structure and function of the global ocean microbiome. *Science,* **348**, 1261359 (2015).

79. Langenheder, S., Lindström, E. S. & Tranvik, L. J. Structure and function of bacterial communities emerging from different sources under identical conditions. *Appl. Environ. Microbiol.* **72**, 212–220 (2006).

80. Pimm, S. L. The complexity and stability of ecosystems. *Nature* **307**, 321–326 (1984).

81. Hou, D. *et al*. Environmental factors shape water microbial community structure and function in shrimp cultural enclosure ecosystems. *Front. Microbiol.* **8**, 1–12 (2017).

82. Eiler, A. Evidence for the ubiquity of mixotrophic bacteria in the upper ocean: Implications and consequences. *Appl. Environ. Microbiol.* **72**, 7431–7437 (2006).

83. Meyer, A. F., Lipson, D. A., Martín, A. P., Schadt, C. W. & Schmidt, S. K. Molecular and Metabolic Characterization of Cold-Tolerant Alpine Soil Pseudomonas Sensu Stricto.

Appl. Environ. Microbiol. **70**, 483–489 (2004).

84. Swingley, W. D. *et al.* The complete genome sequence of Roseobacter denitrificans reveals a mixotrophic rather than photosynthetic metabolism. *J. Bacteriol.* **189**, 683–690 (2007).

85. Tranvik, L.J. & Hofle, M.G. bacterial Growth in Mixed Cultures on Dissoloved Organic Carbon from Humic and Clear Waters. *Appl. Environ. Micrbiol.* **53**, 482-488 (1987).

86. Rosenfeld, J. S. & Mall, M. Functional redundancy in ecology and conservation. *Fish. Res.* **1**, (2002).

87. Ruiz-Gonzalez, C., Nino-Garcia, J. P., Lapierre, J. F. & del Giorgio, P. A. The quality of organic matter shapes the functional biogeography of bacterioplankton across boreal freshwater ecosystems. *Glob. Ecol. Biogeogr.* **24**, 1487–1498 (2015).