Abstract The microbial community composition in meromictic Lake Kivu, with one of the largest CH$_4$ reservoirs, was studied using 16S rDNA and ribosomal RNA (rRNA) pyrosequencing during the dry and rainy seasons. Highly abundant taxa were shared in a high percentage between bulk (DNA-based) and active (RNA-based) bacterial communities, whereas a high proportion of rare species was detected only in either an active or bulk community, indicating the existence of a potentially active rare biosphere and the possible underestimation of diversity detected when using only one nucleic acid pool. Most taxa identified as generalists were abundant, and those identified as specialists were more likely to be rare in the bulk community. The overall number of environmental parameters that could explain the variation was higher for abundant taxa in comparison to rare taxa. Clustering analysis based on operational taxonomic units (OTUs at 0.03 cutoff) level revealed significant and systematic microbial community composition shifts with depth. In the oxic zone, Actinobacteria were found highly dominant in the bulk community but not in the metabolically active community. In the oxic–anoxic transition zone, highly abundant potentially active Nitrospira and Methylococcales were observed. The co-occurrence of potentially active sulfur-oxidizing and sulfate-reducing bacteria in the anoxic zone may suggest the presence of an active yet cryptic sulfur cycle.

Keywords Bacteria · Archaea · Pyrosequencing · Active · Bulk · qPCR · Network · Abundant and rare · Meromictic lake

Introduction

Meromictic lakes are good model systems for microbial ecology research due to the high vertical stability of the water masses and physicochemical gradients (particularly oxygen) that lead to relatively constant stratification of microbial populations [1–3]. Since microbes are key players in biogeochemical cycles, investigations on microbial diversity and community composition are important to understand the ecological functioning of lakes. Environmental heterogeneity was shown to produce community differences in lakes [4, 5]. According to Lennon and Jones [6], core community shifts, including active, dormant, and dead cells, may vary widely along vertical gradients, contributing to the community dynamics as well as to the maintenance of ecosystem biodiversity and ecosystem stability. Currently, high-throughput sequencing technology provides an opportunity to detect greater microbial diversity than the one detected by previous techniques [7–9]. It has revealed that the core community is composed of abundant and rare taxa that are in continuous exchange based on the
environmental gradients [10]. However, most recent studies were based on bulk community composition (i.e., DNA-based 16S ribosomal RNA (rRNA) genes) and did not distinguish between functionally active (i.e., RNA-based 16S rRNA genes) or dormant populations [11]. Even though potential biases can be introduced during DNA and RNA extraction, complementary DNA (cDNA) synthesis, and PCR amplification [12], parallel comparisons of DNA- and RNA-based pyrosequencing approaches can help differentiate whether abundant taxa are active or whether rare but active taxa exist. In addition, stronger evidence on the links between biogeochemical processes and microbial communities might be obtained if active communities are monitored rather than bulk communities [13–16].

Lake Kivu in East Africa is a meromictic lake with a permanent density stratification separating the oxic mixolimnion from a deep anoxic monolimnion rich in dissolved salts, carbon dioxide (CO2), and methane (CH4) [17]. In spite of the presence of high amounts of CH4 at the bottom of Lake Kivu, the CH4 concentration in the oxic zone is surprisingly low compared to other lakes globally, due to intense microbial methane oxidation [18, 19]. The vertical stratification of microbial taxa and their potential role in biogeochemical cycles in Lake Kivu were recently studied for both archaeal and bacterial counterparts using DNA-based analysis, and important key players of carbon, nitrogen, and sulfur cycles were identified [20, 21]. Furthermore, evaluation of the effect of temporal variations on abundant versus rare taxa in different layers revealed that abundant taxa were more stable between two sampling times than rare ones, indicating a potentially higher contribution of rare taxa to biogeochemical processes [20]. Therefore, in the present study, Lake Kivu was investigated as a model ecosystem by comparative DNA- and RNA-based pyrosequencing of bacterial and archaeal communities in two different seasons (i) to check if bulk (DNA) and potentially active (RNA) microbial communities have the same composition through the water column and (ii) to explore if rare bulk and potentially active taxa respond to environmental changes as abundant taxa do in Lake Kivu. Furthermore, the abundance of key genes related to the carbon (particulate methane monooxygenase gene (pmoA) and methyl coenzyme-M reductase gene (mcrA)) and nitrogen (ammonia monooxygenase gene (amoA), nitrite reductase gene (nirK) and nitrous oxide reductase (nosZ)) biogeochemical cycles was measured to evaluate the functional potential and their response to environmental changes.

Materials and Methods

Study Site and Sampling Meromictic and oligotrophic Lake Kivu is located between Rwanda and the Democratic Republic of the Congo at 1463 m above sea level. It has a surface area of 2370 km², a total volume of 580 km³, and a maximum depth of 485 m. Surface waters are considered to be oligotrophic with moderate primary production compared to other African lakes [22, 23]. Further details on the hydrology, physicochemistry, and biology of the lake are published elsewhere [22–25]. In order to study the microbial communities in the water column of Lake Kivu, water samples were collected from the upper 100 m of the water column in the North basin (off Gisenyi; 29.24° E, −1.72° N). Water samples were collected during two sampling campaigns during both rainy (RS, February 2012) and dry seasons (DS, September 2012). Up to 20 discrete depths were sampled along a vertical profile between 1 and 100 m to cover the whole gradient of oxygen concentrations (from oxic to anoxic waters). Water samples for chemical and microbiological analyses were collected using a 7.5-L Niskin bottle and stored in 4-L plastic containers for chemical analyses (except for CH4 and HS−) and 2-L Nalgene plastic bottles for biological analyses. Water samples for DNA and RNA extractions were immediately passed through 5.0-μm pore size filters (ISOPORE, Millipore, MA) to remove particulate debris as well as large protozoa. Eluents were then passed through 0.22-μm pore size filters (ISOPORE, Millipore, MA) to retain free-living prokaryotes. Filters for DNA extraction were preserved in Lysig Buffer as previously described [21], whereas filters for RNA extraction were preserved in 300 μL of RNAlater (Ambion) and all stored frozen until further analyses.

Chemical Analyses Temperature, conductivity, pH, and dissolved oxygen (DO) vertical depth profiles were measured in situ with a YSI 6600 V2 (Yellow Spring Instruments, USA) multiparametric probe. According to Wright et al. [26], the upper 100 m of the water column of Lake Kivu was split into three distinct vertical layers: an oxic surface layer (DO>90 μM), a transition zone (DO between 1 and 90 μM), and a deep anoxic zone (DO<1 μM; Fig. 1). The concentration of methane (CH4) was measured using the headspace technique with a gas chromatograph with a flame ionization detector as previously described [18]. Samples for NOx and SO42− were filtered directly through 0.2-μm pore size cellulose acetate syringe filters. NH4+ concentrations were determined using the dichloroisocyanurate-salicylate-nitroprussiate colorimetric method [27]. NO3− concentrations were determined by the sulfanilamide coloration method [28]. NO3− concentrations were determined after vanadium reduction to nitrite and quantified in this form following the nitrite determination procedure [28, 29]. SO42− concentrations were measured using ion chromatography. Samples for HS− determination were not filtered but preserved instead with zinc acetate and stored frozen. HS− concentrations were measured spectrophotometrically [30]. The detection limits for these methods were 0.5 nM for CH4 and 0.3, 0.03, 0.1, 2, and 0.5 μM for NH4+, NO2−, NO3−, SO42−, and HS−, respectively. Samples for particulate organic carbon concentration (POC) were filtered on precombusted
(overnight at 450 °C) 25-mm glass fiber filters (Advantec GF-75; 0.3 μm) and dried. These filters were later decarbonated with HCl fumes for 4 h and then dried and packed in silver cups. POC was determined by elemental analysis-isotope ratio mass spectrometry (EA-IRMS, ThermoFlashHT with Thermo Delta V Advantage). POC was calibrated with International Atomic Energy Agency (IAEA-C6) and acetanilide.

Chlorophyll a (Chla) concentrations were determined using high-performance liquid chromatography (HPLC). At each sampling depth, 3 L of water were filtered on a Macherey–Nägel GF-5 filter (nominal porosity, 0.4 μm). Pigment extraction was carried out in 10 mL of 90 % HPLC grade acetone (Fisher Scientific). After two 15-min sonication steps separated by an overnight period at 4 °C, the extracts were stored in 2-mL amber vials at −25 °C. HPLC analysis was performed following the method described in Sarmento et al. [23]. Commercial external standards (DHI Lab Products) were used for calibration.

Nucleic Acid Extraction DNA was extracted using a combination of enzymatic cell lysis and the cetyltrimethyl ammonium bromide (CTAB) extraction protocol as previously described [20, 31]. Dry DNA pellets were finally rehydrated in 50 μL of 10 mM Tris–HCl buffer (pH 7.4).

Total RNA was extracted using a mirVana RNA isolation kit (Ambion, Austin, TX, USA; [32]). RNA samples were treated with the Turbo DNA-free kit (Ambion, Austin, TX, USA), and cDNA synthesis was performed with the iScript reverse transcription supermix (Bio-Rad) in a MasterCycler 5331 Gradient (Eppendorf, Hamburg, Germany). Possible DNA contamination of RNA templates was routinely monitored by PCR amplification of RNA aliquots that were not reverse transcribed. No contaminating DNA was detected in any of these reactions.

Real-Time Quantitative PCR qPCR was used to quantify the presence of distinct functional genes. The following genes were used as proxies for quantification of DNA: the bacterial alpha subunit of the particulate methane monooxygenase enzyme (pmoA) for methanotrophs, the archaeal alpha subunit of methyl-coenzyme M reductase enzyme gene (mcrA) for methanogens, the alpha subunit of the archaeal ammonia monooxygenase enzyme gene (amoA) for archaeal ammonia oxidation, the bacterial copper-containing nitrite reductase gene (nirK), and the cytochrome cd1-containing nitrite reductase gene (nosZ) for bacterial denitrification. All qPCR assays were performed in a StepOne Real-Time PCR system (Applied Biosystems, Foster City, CA, USA) using the primers and thermal conditions described in Table 1. All reactions were performed in triplicate for both serial dilutions of titrated standards and unknown templates using a 20 μL reaction mixture consisting of the 1× PowerSYBR green PCR master mix (Applied Biosystems), each primer (Table 1), and DNA templates of known concentrations of standards or 20 ng DNA extracted from water samples. The specificity of the amplification products was further confirmed by melting curve analyses, and the expected sizes of the amplified fragments were checked on a 1 % agarose gel. Overall, average efficiencies for
<table>
<thead>
<tr>
<th>Primer pair&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sequence (5′–3′)</th>
<th>Cycles&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Denaturation</th>
<th>Annealing</th>
<th>Elongation</th>
<th>Signal detect.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pmoA [200 nM]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[33]</td>
</tr>
<tr>
<td>A189f</td>
<td>GGN GAC TGGGAC TTC TGG</td>
<td>40</td>
<td>95</td>
<td>30</td>
<td>56</td>
<td>30</td>
<td>72</td>
</tr>
<tr>
<td>mb661r</td>
<td>CCGGMGCAA CGT CYT TAC C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mcrA [200 nM]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[34]</td>
</tr>
<tr>
<td>LuF</td>
<td>GGT GGTGTM GGA TTC ACA CAR TAY GCW ACA GC TTC ATT GCR TAG TTW GGR TAG TT</td>
<td>40</td>
<td>95</td>
<td>30</td>
<td>55</td>
<td>30</td>
<td>72</td>
</tr>
<tr>
<td>LuR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>amoA [200 nM]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[21]</td>
</tr>
<tr>
<td>AOA-amoAf</td>
<td>CTG AYT GGG CYT GGA CAT C</td>
<td>60</td>
<td>94</td>
<td>39</td>
<td>59</td>
<td>40</td>
<td>72</td>
</tr>
<tr>
<td>AOA-amoAr</td>
<td>TTC TTC TTT GTT GCC CAG TA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nirK [500 nM]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[35]</td>
</tr>
<tr>
<td>nirK876</td>
<td>ATY GGC GGV CAY GGC GAnirK1040</td>
<td>6/40</td>
<td>94</td>
<td>15</td>
<td>63–58/60</td>
<td>30</td>
<td>72</td>
</tr>
<tr>
<td>nirK1040</td>
<td>GCC TCG ATC AGR TTR TGG TT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nosZ [750 nM]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[35]</td>
</tr>
<tr>
<td>nosZ2F</td>
<td>CGC RAC GGC AAS AAGGTS MSS GT</td>
<td>6/40</td>
<td>94</td>
<td>15</td>
<td>65–60/60</td>
<td>30/20</td>
<td>72</td>
</tr>
<tr>
<td>nosZ2R</td>
<td>CAK RTG CAK SGC RTGGCA GAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Concentration for each primer pair is given in brackets

<sup>b</sup> All quantitative PCR reactions have an initial denaturation step at 95°C during 3 min
all quantification reactions ranged from 0.88 to 0.97 with \( R^2 \) values >0.99. Standard curves were generated from serial dilutions of previously titrated suspensions of the desired genes from isolates or environmental clones, purified (QIAquick; Qiagen), and quantified. Statistical analyses (ANOVA and Tukey’s HSD) were performed to compare differences between water zones (oxic, transition, anoxic) [36].

**16S Tag-Encoded FLX-Titanium Amplicon Pyrosequencing**

Bacterial and archaeal tag-encoded FLX amplicon pyrosequencing (bTEFAP and aTEFAP, respectively), analyses by means of a Roche 454 FLX instrument with titanium reagents, were performed at the Research and Testing Laboratory (Lubbock, TX, USA) as described previously [37, 38]. The PCR primers for FLX amplicon pyrosequencing were chosen to span the variable V1–V3 regions in the 16S rRNA gene: 27 F (5′-GAGTTTTGATCNGTGGCTCAG-3′) and 519R (5′-GWNTACNGCGGCKGCTG-3′) for bacteria and V3–V4 regions ARCH 349 F (5′-GYGCASCAGKCGMGAAW-3′) and ARCH 806R (5′-GGACTACVSGGGGTATCTAAT-3′) for archaea. These primers cover about 78 and 70 % of publicly available 16S rRNA for bacteria and archaea, respectively (check using TestPrime tool available at SILVA webpage (http://www.arb-silva.de/search/testprime/)).

**Pyrosequencing Data Analyses**

All sequences generated in this study can be downloaded from the National Center for Biotechnology Information (NCBI) Short Read Archive, accession number: SRP021176. Pyrosequencing data were processed using Mothur [40]. To minimize the effects of random sequencing errors, a denoising algorithm (shhh.flows) included in the pipeline is used and low-quality sequences were removed by eliminating those without an exact match to the forward primer, without a recognizable reverse primer, with a length shorter than 200 nucleotides, and those containing any ambiguous base calls. We trimmed the barcodes and primers from the resulting sequences. Chimeric sequences were removed using the Uchime software [39] implemented in Mothur [40]. The latest Greengenes and SILVA databases were used for classification of bacterial and archaeal 16S rRNA gene sequences, respectively, at a 80 % confidence threshold using Mothur. After taxonomic assignment of the sequences to the phylum, class, and genus level, relative abundance of a given phylogenetic group was set as the number of sequences affiliated with that group divided by the total number of sequences per sample. The sequences retrieved were grouped based on oxygen stratification as oxic, transition, and anoxic zones.

Sequences were clustered into operational taxonomic units (OTUs) by setting a 0.03 distance limit (equivalent to 97 % similarity). A data matrix was created based on the relative abundance of genus and OTU. The relative abundances were square root transformed. Dendrograms were generated based on the Bray–Curtis similarity index using complete linkage clustering. Furthermore, the scatterplot depicted the relationship between the depths and observed species-based Bray–Curtis similarity. The discrimination of microbial assemblages based on time and stratification was tested with one-way analysis of similarities (ANOSIM). Unless otherwise stated, all the analyses were performed using PRIMER 6 [41].

To measure habitat specialization, the niche breadth approach was used, as described previously [42, 43]. The niche breadth \( (B_j) \) was calculated using the following equation:

\[
B_j = \frac{1}{\sum_{i=1}^{K} P_{ij}^2}
\]

in which \( P_{ij} \) is the proportion of individuals belonging to species \( j \) present in a given habitat \( i \). OTUs with a high \( B_j \) value indicate a wide range of habitats, which can be considered as habitat generalists. Similarly, OTUs with a low value can be assigned as habitat specialists.

Furthermore, the data sets were split into two groups: Taxa represented by more than 15 sequences per sample were defined as abundant taxa, and taxa that had fewer than 15 sequences were considered rare. Venn diagrams of OTUs were constructed to illustrate the number of shared OTUs (both subsampled and not subsampled) for the two sampling campaigns based on DNA and RNA. A relative abundance matrix for OTUs was created for sampling times and points pooled, and those OTUs that appeared less than five times were removed from the matrix. For network inference, Pearson’s rank correlations were calculated between OTUs with CoNet [44], when coefficient \( (r) \) was both above 0.8 and statistically significant (adjusted \( p \) value <0.01 with Benjamini–Hochberg). The data matrix was translated into an association network using Cytoscape 2.6.3 [44, 45]. Cytoscape depicts data sets as nodes (tribes) connected by lines that denote the positive correlation. In the network, the size of each node is proportional to the number of connections (degree) and the node border color indicates if the OTU is either abundant or rare.

We further quantified the influences of environmental variables (\( \text{pH} + \text{DO} + \text{NO}_x + \text{POC} + \text{CH}_4 + \text{Chl-a} \)), depth, and season on bacterial community variation using variation partitioning analyses [46], with the varpart function of the vegan package [47, 36]. If any colinearity was present, it would be revealed by the analysis.

**Results**

**Environmental Parameters**

Lake Kivu is characterized by vertical gradients in temperature, \( \text{pH} \), conductivity, \( \text{DO} \), \( \text{CH}_4 \), and \( \text{HS}^- \) (Fig. 1). Surface waters were characterized by high \( \text{DO} \) content and low \( \text{NO}_x \).
and CH₄ concentrations, whereas the anoxic deep waters were characterized by high concentrations of NH₄⁺ and dissolved CO₂ and CH₄ [18]. NOₓ⁻ showed maximum values in the oxic–anoxic transition zone, whereas increasing concentrations with depth of HS⁻ (>50 μM) and CH₄ (>100 μM) were only observed in the anoxic zone (Fig. 1). The mixed oxic zone was deeper during the DS than during the RS with a more gradual decrease in DO with depth and a steady increase in NOₓ values toward the chemocline (Fig. 1). The pH decreased with depth over the first 100-m depth (data not shown). POC concentration was higher in the oxic water than in the anoxic zones during both sampling campaigns and was slightly higher in the oxic water compartments in DS than in RS (data not shown).

qPCR In the current study, functional genes involved in key processes of carbon and nitrogen cycles were measured by genomic DNA qPCR. In relation to the carbon cycle, the pmoA gene was used to quantify the bacterial methanotrophs and the mcrA gene to quantify archaeal methanogens and methanotrophs. Concerning the N cycle, the amoA gene was used to quantify archaeal ammonia oxidizers (AOA) while the nirK and nosZ genes were used to quantify bacterial denitrifiers.

In RS, pmoA gene copies ranged between 1.13×10³ and 2.92×10⁴ gene copies mL⁻¹, whereas in DS, the values varied between 4.19×10² and 9.63×10⁴ gene copies mL⁻¹ through the water column. In DS, more pmoA gene copies were measured in the transition zone (p<0.05, Fig. 2) than in RS. Concerning archaeal methanogens and methanotrophs, mcrA gene copies were only detected in the anoxic zone with abundance ranging between 8.00×10² and 3.53×10³ gene copies mL⁻¹ in RS (Fig. 2), whereas lower values were recovered in DS (4.15×10² to 2.21×10³ gene copies mL⁻¹). The difference between sampling dates was not statistically significant.

The numbers of archaeal amoA gene copies were used to investigate the vertical distribution and abundance of AOA in relation to the NOₓ profile available for both sampling campaigns. The quantitative distribution varied with depth, and

Fig. 2 Depth profiles of gene copy numbers for pmoA, mcrA, amoA, and nirK/nosZ genes determined by quantitative PCR for two sampling campaigns (February (RS) and September (DS)). Error bars represent the standard deviation from triplicates. The transition zones were indicated with a gray bar.
the highest copy numbers were mainly observed in the oxic and transition zones (from 30.0 to 42.5 m in RS and from 40.0 to 57.5 m in DS). The maximum values were 4.2×10^5 and 2.5×10^4 gene copies mL^{-1} in RS and DS, respectively (Fig. 2). Furthermore, functional genes involved in bacterial denitrification (i.e., nirK and nosZ genes) showed similar depth patterns during both seasons with rather low values (never exceeding 1×10^3 gene copies mL^{-1}; Fig. 2). Oxic waters had significantly higher densities of nirK than in transition zone (p<0.05) and higher densities of nosZ than anoxic water in September (p<0.05). During the RS, no clear pattern with depth was observed, but higher values were observed at 27.5 and 45 m, whereas during the DS, maximum values were observed at 10 m and decreased regularly with depth, with a slight increase at the oxic–oxic transition zone.

**Pyrosequencing Overview and Community β-Diversity**

After removing the noise and poor-quality reads with Mothur, about 34,097 bacterial DNA reads and 27,095 bacterial RNA reads were used for subsequent analyses. Rarefaction analysis indicated that most of the bacterial libraries may require deeper sequencing to avoid underestimation of microbial diversity in the samples (Table S1). Samples were randomly subsampled to the smallest sampling size using Mothur in order to reduce bias in species richness due to the differences in the number of sequences. The number of bacterial OTUs increased with depth for both DNA and RNA-based bacterial communities (Table S1). Richness estimates of the RNA-based bacterial community in the deepest anoxic waters were slightly higher than the DNA-based community in RS, whereas richness estimates of the DNA-based bacterial community in anoxic waters were higher than the estimates of the potentially active bacterial community in DS. The number of observed bacterial species per depth ranged from 127 to 240 in RS and 237 to 603 in DS for bulk and active bacterial communities, respectively (Table S1). The Cha01 estimator predicted richness values in the range of 324–988 in RS and 492–2234 in DS. Moreover, rank-abundance curves of DNA- and RNA-based bacterial communities showed a power-law distribution, which had a few abundant OTUs and a long tail of low-abundant species (Fig. S1). Venn diagrams also suggested that highly abundant taxa were common in a high percentage between DNA- and RNA-based communities. In contrast, a high proportion of rare species was determined to be specific to either DNA- and RNA-based community (Fig. 3).

In turn, 16,700 archaean DNA reads and 2573 archaean RNA reads were recovered for subsequent analyses. No high-quality sequences were retrieved from 15, 25, 35, and 50 m RNA samples in RS and from 10, 25, 40, 50, and 53 m DNA and RNA samples in DS. Rarefaction analysis based on OTUs indicated that most of the archaean libraries reached the plateau level with the exception of the 40 m RNA sample in RS and the 80 m RNA sample in DS. In general terms, richness was higher in the anoxic zone than in the oxic zone during both sampling times. During both seasons, archaean richness estimates (i.e., Cha01 estimator) predicted lower richness values in the oxic and oxic–oxic transition zone than in the anoxic waters (Table S2). No clear richness differences were observed between DNA- and RNA-based archaean communities in oxic and oxic–oxic transition zones, whereas richness estimates were higher for DNA samples than RNA samples in the anoxic zone (Table S1).

**Bacterial Community Composition Analysis**

Between 7 and 42 % of the bacterial sequences analyzed were unclassified, with an increasing proportion toward the anoxic zone. Altogether, 34 distinct bacterial phyla were recovered from DNA samples and 29 phyla from RNA samples in RS, while 43 phyla were recovered from DNA samples and 29 phyla from RNA samples in DS. *Actinobacteria*, *Proteobacteria* (*Alphaproteobacteria*, *Betaproteobacteria*, *Epsilonproteobacteria*, and *Gammaproteobacteria*), *Cyanobacteria*, *Planctomycetes*, *Bacteroidetes*, *Chlorobi*, *Chloroflexi*, and *Nitrospirae* were detected both in the DNA- and RNA-based bacterial communities (Fig. 4, S2A–I). In DNA- and RNA-based communities, candidate bacterial phyla such as *GN02*, *OP3*, *OP8*, *NC10*, *NKB19*, *TM6*, and *WS1* were mostly detected in the anoxic zone but never exceeding relative abundance of 1 %. The comparison of the DNA- and RNA-based bacterial communities revealed similar relative abundance of the numerically dominant taxa with the exception of a lower relative abundance of *Actinobacteria* (Fig. 4, S2B) and a higher abundance of *Planctomycetes* (Fig. 4, S2C) and *Bacteroidetes* (Fig. 4) in the potentially active bacterial community composition (BCC) with respect to DNA-based bacterial community in the oxic zone. The BCC is further described in the co-occurrence patterns of the bacterial taxa section.

Bray–Curtis similarity trees were constructed based on total OTU composition of the 47 bacterial RNA and DNA samples to determine the associations among the communities. Clustering and ANOSIM analyses revealed the stratification (three zones based on DO concentrations) as the most important factor in structuring both DNA- and RNA-based BCC (Fig. 5). Although the clustering showed the grouping of the samples by DO zones, no similarity higher than 60 % was observed between samples. When the DNA- and RNA-based bacterial communities were compared based on Bray–Curtis distances against depth, significant negative correlations were observed (Fig. 6), with higher slopes (m) and linear correlation coefficients (R) for DNA- than RNA-based BCC (m=−6.48, R=0.77 for the bulk community and m=−3.58, R=0.47 for the potentially active community), suggesting more rapid changes in the DNA-based BCC with depth than in the RNA-based community. Moreover, the BCC differed significantly and systematically depending on whether DNA
or RNA was analyzed (Fig. 4). Based on clustering at the OTU level and ANOSIM analysis of the bulk community, a significant temporal effect (DS vs. RS) was also observed (Table 2).

Archaeal Community Composition Analysis At an 80 % confidence threshold, all of the reads could be assigned to the Archaea using the SILVA database classifier. Most of all, the nonchimeric and good-quality archaeal amplicons retrieved were affiliated with archaeal phyla, and only two samples from the deepest depths had 7 % unclassified Archaea. All archaeal amplicons retrieved spanned over archaeal lineages covering Thaumarchaeota, Crenarchaeota, and Euryarchaeota within both cultured (e.g., marine Crenarchaeota group, Methanosaeta, or Candidatus Methanoregula) and uncultured (e.g., Miscellaneous Crenarchaeotic Group (MCG) and GOM-Arc-I) groups.

At both sampling times, one OTU dominated the DNA- and RNA-based communities of the oxic zone, which belonged to the marine group I, tentative ammonia-oxidizing archaea (AOA). AOA were also dominant in the bulk communities in the transition zone, whereas the potentially active archaeal community composition (ACC) was mainly composed of AOA, Methanomicrobia, and Methanobacteria in the transition zone (Fig. S2J). The DNA- and RNA-based communities thriving in the anoxic zone were dominated by the two methanogens mentioned above and by members of the MCG group.

A Bray–Curtis similarity tree was constructed based on total OTU composition of the 40 archaeal RNA and DNA samples to determine the associations among the communities. Clustering and ANOSIM analyses revealed sampling date as the main grouping factor in structuring the ACC (Fig. S3, Table 2) and nucleic acid type as the second grouping factor. Oxygen had the lowest global structuring effect on ACC, butoxic, and transition water samples clearly segregated from anoxic samples (Fig. S3).

Bacterial Abundance Classification and Overlap Estimates Bulk and potentially active bacterial taxa from oxic, transition, and anoxic zones were grouped per season (Fig. 3). Venn diagrams of OTUs were constructed to illustrate the number of shared OTUs for the two sampling campaigns based on DNA and RNA. The DNA and RNA samples over the two seasons shared less than 1 % of OTUs in the rare
Forty-six, 19, and 35% of the abundant OTUs were shared in the oxic, transition, and anoxic zone, respectively. Rare species were also found to be specific to the sampling season and nucleic acid-type pool for both subsampled and not-subsampled data (i.e., DNA or RNA).

**Co-occurrence Patterns of Bacteria**  
Co-occurrence patterns of the microbes present in Lake Kivu were only analyzed with the bacterial data set due to the lack of archaeal information at some depths. The co-occurrence patterns for the potentially active core microbiome detected by Venn diagrams were

![Diagram](image)

**Fig. 4** Taxonomic classification of DNA and RNA bacterial reads at the phylum level in Lake Kivu for the two sampling campaigns (February (left) and September (right)). *CFB Bacteroidetes*

Co-occurrence patterns of the microbes present in Lake Kivu were only analyzed with the bacterial data set due to the lack of archaeal information at some depths. The co-occurrence patterns for the potentially active core microbiome detected by Venn diagrams were

![Diagram](image)

**Fig. 5** Hierarchical clustering of Bray–Curtis similarities of the bacterial community composition at OTU level retrieved by pyrosequencing of both DNA (black symbols) and RNA (gray symbols) nucleic acid pools in (a) February and (b) September. Samples were grouped according to oxygen content as oxic (triangles), transition (dots), and anoxic (squares).
further assessed using network inference based on strong and significant Pearson correlations ($r>0.8$, $p<0.01$). The most complex interaction was found between potentially active bacteria in the anoxic zone. The microbial network for both the oxic and transition zones was composed of 66 nodes (OTUs) with moderate interconnection (2.3 edges per node on average), whereas in the network of the anoxic zone, there were ca. three times more nodes with a higher degree of connection (5.6 edges per node on average; Fig. 3). Aside from many unclassified bacteria, interactions between different known taxa were observed. For instance, in the oxic zone, there was a fairly complex subnetwork including Actinobacteria, Bacteroidetes, Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Cyanobacteria, and OPB56. In the transition zone, a subnetwork was observed between OTUs affiliated with N-fixing methane-oxidizing Methylocaldum, hydrogen-oxidizing Hydrogenophaga, and nitrifying bacteria Nitrospira with members of the unclassified proteobacterial class (Fig. 3a). In the anoxic zone, a co-occurrence pattern was observed between sulfate reducers (Desulfobacca, Desulfocapsa, and Desulfobacterium), sulfur oxidizers (Sulfurimonas), and hydrogen oxidizers (Dehalococcoidetes).

Co-occurrence Patterns of Potentially Active Bacteria and Archaea in the Anoxic Zone Diverse and complex interactions between potentially active bacteria were found in the anoxic zone; therefore, their correlations with archaea were further investigated. The co-occurrence patterns of the potentially active microbiome with a relative abundance higher than 0.5% were defined using network inference based on strong and significant Pearson correlations ($r>0.8$, $p<0.001$). The microbial network in the anoxic zone consisted of 31 moderately interconnected nodes (2.9 edges per node on average; Fig. S4). A clear interaction between bacterial sulfate reducers, sulfur oxidizers, nitrogen-fixing methanotrophs, archaeal methanogens, and anaerobic methane oxidizers was observed. In addition, another interaction between members of Methanomicrobia, Chlorobicaea, and Syntrophaceae was also observed.

Niche Breadth Comparison Levins’ niche breadth was calculated as a proxy to understand how particular taxa relate to their environment. The bulk bacterial community had a strong linear relationship between niche breadth and their relative abundance, while generalist and specialist OTUs in the potentially active bacterial community were more equally distributed within rare and abundant taxa. In addition, niche breadth in the DNA-based bacterial community was higher than in the RNA-based community. Potentially active bacterial OTUs able to develop in either the oxic or anoxic zone were mainly assigned at different taxonomic levels to Burkholderiales, Pseudanabaenaceae, Acidobacteria, and Methylocaldum. In

---

**Table 2** ANOSIM test on bacterial and archaeal community composition at the OTU level between the three analyzed water compartments

<table>
<thead>
<tr>
<th>Test for differences between DNA/RNA across all water compartments</th>
<th>Bacteria</th>
<th>Archaea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global test</td>
<td>$R$</td>
<td>$p$</td>
</tr>
<tr>
<td>Oxic-transition</td>
<td>0.370</td>
<td>0.03</td>
</tr>
<tr>
<td>Oxic–anoxic</td>
<td>0.735</td>
<td>0.001</td>
</tr>
<tr>
<td>Transition–anoxic</td>
<td>0.454</td>
<td>0.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test for differences between sampling dates across all DNA/RNA</th>
<th>Bacteria</th>
<th>Archaea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global test</td>
<td>$R$</td>
<td>$p$</td>
</tr>
<tr>
<td>Oxic-transition</td>
<td>0.370</td>
<td>0.03</td>
</tr>
<tr>
<td>Oxic–anoxic</td>
<td>0.735</td>
<td>0.001</td>
</tr>
<tr>
<td>Transition–anoxic</td>
<td>0.454</td>
<td>0.03</td>
</tr>
<tr>
<td>Global test</td>
<td>0.145</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Significant values were indicated in bold
contrast, no significant relationship was observed between niche breadth and DNA- and/or RNA-based archaeal relative abundance. Archaeal OTUs detected in both the oxic and anoxic zones were assigned to MGI and Miscellaneous Crenarchaeotic Group. The niche breadth observed for the RNA-based bacterial and archaeal community was narrower than for the DNA-based bacterial and archaeal community (Fig. S5).

**Linking Bacterial Communities to Environmental Parameters** Variance partitioning analysis (VPA) was performed to quantify the relative contribution of season, depth, and environmental parameters to the taxonomic structure of the bacterial and archaeal communities. The analysis of environmental variables was restricted to a subset of all variables acquired during the field cruises including: DO, pH, NO$_3$, POC, CH$_4$, and Chl-$\alpha$. Clear differences were observed in variance partitioning patterns, if all bacterial OTUs or both abundance categories (abundant/rare) were considered. In all cases, the impact of the environmental parameters was found to be the highest, followed by either season or depth (Fig. 7). The parameters analyzed (depth, season, and environmental variables) in total explained 50 % of the variation observed for the bulk bacterial community, and they explained only 30 % of the variation observed for the potentially active bacterial community. Furthermore, the total relative contributions of variables increased when VPA was performed for only abundant bacterial species, whereas it clearly decreased for rare species. The sum of the parameters analyzed (depth, season, environmental variables) explained 73 % of the variation observed for the bulk archaeal community and only 18 % for the potentially active archaeal community (data not shown).

**Discussion**

In the present study, DNA- and RNA-based amplicon pyrosequencing was performed to further investigate the temporal and vertical effects on the bulk and potentially active microbial biosphere of Lake Kivu and to gain a more complete view of the microbial dynamics in Lake Kivu than the one afforded by analyzing either nucleic acid pool alone.

**Community Composition at a High Taxonomic Level** Due to the strong vertical chemical gradients and zonation of redox reactions with depth, a clear vertical stratification of microbial communities is usually observed for meromictic lakes [48–51]. In spite of the different extraction protocols used, the relative abundances of potentially active and bulk bacterial and archaeal communities were mostly similar at the phylum level, with some notable exceptions. These differences were mostly observed in the oxic zone. For instance, *Actinobacteria* were highly dominant in the bulk bacterial community of surface waters (around 30 %); however, this dominance was lower in the potentially active community, as previously shown in a eutrophic lake [52]. *Actinobacteria* might experience lower

![Fig. 7 Variation partitioning of bacterial community composition. The figure depicts the fractions (%) of variation in bacterial community composition that are explained by environmental parameters (pH+DO+NO$_3$+POC+CH$_4$+Chl), depth, and season or that remain unexplained for bulk and active bacterial communities.](image-url)
mortality rates than other bacterial lineages due to the resistance to protozoan grazing [53] and small cell size [54]. Therefore, *Actinobacteria* might not require rapid growth to maintain high cell densities in the water column [55]. This might be also due to the difference in lysis methods used for DNA and RNA extraction. *Bacteroidetes* (CFB) and *Planctomycetes* were substantial components of potentially active BCC in theoxic waters of Lake Kivu in the rainy and dry season, respectively (Fig. 4). Many studies have reported high relative abundance of *Bacteroidetes* or *Planctomycetes* following algal blooms [56–60]. Both phyla are known to be present and attached to aggregates, and they have various extracellular enzymes that can degrade complex organic molecules and give a competitive advantage with regards to other microbial groups [61–63]. Presence of a high percentage of these two phyla in different seasons might be due to the competition between these two phyla for organic matter or dominance of different phytoplankton species in two seasons, as mainly a higher dominance of diatoms was observed in the DS and *Cyanobacteria* were more dominant in the RS (Fig. 4; [23]). Besides, *Cyanobacteria* represent a big fraction of the active microbial community in oxic and illuminated waters (Fig. 4 and Fig. S2) due to their oxic photoautotrophic metabolism. A recent study done in Lake Kivu suggested that a variety of ecological niches of heterotrophic bacteria can be supported by the diversity of molecules excreted by the phytoplankton community [64]. Detection of a low abundance of potentially active *Planctomycetes*-affiliated sequences in the anoxic zone supported the idea that members of this group have anaerobic metabolic abilities, as suggested by previous studies showing that they rely on carbohydrate fermentation and sulfur reduction for growth and survival under anaerobic conditions [65]. Furthermore, sequences affiliated to *Chlorobi* phylum were recovered from both nucleic acid pools of either oxic or anoxic waters and sampling campaigns. Their recovery from oxic waters is an interesting feature and deserves further investigations, but recent evidences pointed toward new *Chlorobi* representatives able to develop under aerobic conditions performing phototrophic metabolisms [66, 67].

Furthermore, higher abundance of AOA was shown in the bulk community in comparison to the potentially active community. Such discrepancy between DNA- and RNA-based studies of surface waters has also been observed in freshwater ecosystems [4, 68]. The denitrification genes *nirK* and *nosZ* were slightly higher in the oxic zone than in the anoxic zone (Fig. 2). Previously, growth of facultative anaerobic denitrifiers in aerobic environments was demonstrated, which suggests a wider distribution of potential denitrifiers in the aerobic growth conditions of this lake [69, 70]. The good correlation found between *nirK* and *nosZ* genes might be due to the fact that these genes are usually located in the same operon [71, 72].

Detection of highly abundant potentially active *Nitrospira* in the transition zone confirmed a previous study of the nitrite-oxidizing community in the transition zone of Lake Kivu (Figs. 4 and S2D; [20]). The CH$_4$ oxidation peak in the transition zone was also strongly correlated with the relative abundance of active methanotrophs and abundance of *pmoA* gene (p<0.05; data not shown). Detection of potentially active *Methylococcales* and *Methylomassiliis* indicates aerobic CH$_4$ oxidation as the main process preventing CH$_4$ from escaping to the atmosphere (Fig. S2E). The co-occurrence of CH$_4$ oxidizing *Methylocalldum* and hydrogen-oxidizing *Hydrogenophaga* supports recent findings of CH$_4$ consumption in O$_2$-limiting conditions by *Methylocalldum*, leading to the release of several organic molecules (e.g., acetate, lactate, succinate, and hydrogen) that can be used by *Hydrogenophaga* [73].

Several archaeal groups able to produce CH$_4$ have also been detected in both DNA and RNA-based communities mainly in the anoxic waters. Indeed, their presence even in the transition zone is consistent with recent studies that showed the presence of potentially active methanogens in the oxygenated water column of an oligotrophic lake attached to photoautotrophic microbes that might be enabling aerobic growth and supply of methanogenic substrates [74, 75]. However, the presence of acetoclastic methanogens (*Methanosarcinales*) in the oxic–anoxic transition zone might also indicate the presence of syntrophy between methanogenic and fermentative microbes, through the occurrence of a coupled mutualistic interaction between hydrogen-/formate-producing and hydrogen-/formate-using microorganisms [76]. A consistent relationship between uncultured OTUs of *Methanomicrobia*, *Chlorobicaeae*, and *Syntrophaceae* was also observed in the anoxic zone (Fig. S3), which could indicate the presence of syntrophic relationships between these microbes in hydrocarbon or complex high molecular weight compound degradation, as previously shown in the sediments of the River Tyne [77].

Network analysis revealed a complex interaction between active *Dehalococcoidetes*, sulfate-reducing bacteria (SRB; *Desulfuocapsa, Desulfohabacca*), and sulfur-oxidizing bacteria (SOB; *Sulfurimonas, Sulfuricurvum*) in the anoxic zone (Fig. 3c). The co-occurrence of active SOB and SRB in the anoxic zone may further suggest the presence of an active yet cryptic sulfur cycle [20]. It has been evidenced that *Dehalococcoidetes* may have a nutritional dependence on other organisms [78, 79]: the co-occurrence of this active but rare bacterial class with SRB might thus be linked to its role in hydrocarbon degradation (Fig. 3c). Recent investigations have described complex interactions between *Dehalococcoides*, SRB, and methanogens depending on their metabolic activities and nutrient (e.g., hydrogen, acetate) or vitamin (e.g., B12) concentrations [80], suggesting their multiple metabolic capabilities.

Bulk Versus Potentially Active Community Structure

Although the samples analyzed had mostly similar taxonomic characterization at the phylum level, clustering analysis at the OTU level highlighted the importance of which nucleic acid
pool (i.e., DNA or RNA) was analyzed. The presence of sequences exclusively found in one of the two nucleic acid pools suggests that using both nucleic acids provides a more complete overview of the microbial diversity present in a given environment since on average, only 20% of the OTUs (18–22%) were shared per DO zone and per season. As previously suggested [81], sequences found only in the DNA pool might be related to metabolically less active, yet abundant populations, whereas sequences detected only in RNA are associated with active members of the community but may be present in relatively low numbers. The type of populations associated with shared and unshared sequences could be the result of differences in their respective overall abundance and/or a reflection on the overall metabolic status of a population within a given OTU. The detection of taxa in only one pool and not in the other can also be due to insufficient pyrosequencing in the effort to reveal complete diversity, and this had obviously more effect on the rare community.

Inactive cells are seen as a potential seed bank, which enables a rapid community adaptation to changing environments, so that the overall community always prevails [82, 83]. DNA and RNA profiles were found to be more similar in the anoxic zone than in the oxic or transition zones. The results agree with the fact that anaerobic processes are less energetic than aerobic respiration, thus suggesting slow growth rates in the anoxic water compartment of Lake Kivu. Besides, the anoxic water compartment of Lake Kivu was also more diverse in types as pointed out by the richness estimations ($p<0.02$). It was previously suggested that anoxic environments maintain a higher diversity of energetic pathways and that this complexity permits the retention of higher metabolic and thus ecological diversity [3, 84].

**Abundant and Rare Taxa** Both bulk and potentially active microbial communities present in the water column of Lake Kivu comprised few very abundant OTUs and a huge number of low-abundant and rare OTUs (data not shown). The fact of even recovering a long tail of rare low-abundant OTUs within the potentially active members of the microbial community evidenced the existence of an active rare biosphere in Lake Kivu or a pool of dormant microbes with enough RNA pool to recover from this situation when necessary [6]. Venn diagrams revealed that bulk and potentially active bacterial communities shared a high fraction of those highly abundant taxa. In contrast, rare OTUs were detected in a high proportion exclusively in either bulk or potentially active bacterial communities. As suggested for estuarine and coastal ocean waters, the metabolically active rare biosphere of Lake Kivu might be potentially important for the functioning and the temporal dynamics of microbial communities as they shift between seasons [13–15].

Levins’ niche width was used to classify the taxa retrieved in Lake Kivu as generalists or specialists based on the environmental conditions in which they developed but not directly on their intrinsic biological properties [85]. Most bacterial taxa identified as generalists were abundant, whereas those identified as specialists were more likely to be rare (mainly in the bulk community), and this pattern was less evident for the potentially active community. Such a significant relationship could not be observed for archaeal taxa, which might be due to missing depths in our analysis. The detection of some OTUs in either bulk or potentially active communities in both oxic and anoxic zones might reflect some metabolic flexibility and capacity to cope with the different redox conditions present in the lake’s water column.

Moreover, variation partitioning analyses were performed to assess the contributions of depth, season, and environmental parameters to microbial community structure. The pool of environmental parameters analyzed here could explain the highest part of the variation in bacterial communities. The results of this study showed that DO had the most significant effect in shaping the BCC, as shown previously [86]. In addition to DO, bacterial community variance was also significantly related to pH in the lake. Any significant deviation in environmental pH should impose stress on single-cell organisms, since intracellular pH of most microorganisms is usually within 1 pH unit of neutral [87, 88]. The significant effect of pH on the BCC was also shown in a range of aquatic environments [88]. The overall number of environmental variations that could explain the presence of taxa was also higher for abundant taxa in comparison to rare taxa. This might indicate that abundant taxa are sufficient to describe the mechanisms underlying the beta-diversity [6]. However, it is possible that rare species were more affected by environmental parameters such as concentrations of micronutrients, hydrocarbons, as well as stochastic dispersal, immigration, and predation, which were not measured in this study and thus not taken into account in the statistical analysis [6]. Furthermore, a higher degree of the variation was explained for the bulk BCC than for the potentially active BCC. It might indicate that the seed bank likely represents the species pool of the complex water habitat and enables the shift in rapidly changing environments and also potentially stabilize ecosystem.

**Conclusions**

Combination of complementary methods (16S rRNA, rDNA pyrosequencing, and qPCR) provided insights into the active community and functional genes present in Lake Kivu, where biogeochemical cycling appeared to be functioning synergistically. Detection of key players in biogeochemical cycles in the various water layers were also supported by the vertical chemical gradients. The experimental evaluation suggests that the quantification of functional genes by qPCR was mostly in agreement with the pyrosequencing of complex microbial communities.
communities. This study showed that rare taxa (mainly specialists), especially in the anoxic zone, cannot solely be characterized as a seed bank or dormant cells in Lake Kivu but that a significant portion of the rare community is potentially active, which supports the results of various studies conducted in freshwater and marine environments [13, 14]. However, the impact of environmental parameters measured was stronger for the generalists (mainly abundant taxa), which indicates the presence of different ecological rules for different groups.

Acknowledgments In addition to the authors of this paper, the Lake Kivu consortium includes the following individuals: S. Bouillon (Katholieke Universiteit Leuven), M.-V. Cammarie, F. A. E. Roland (Université de Liège), B. Leporcq, K. de Saedeleer (Université de Namur), A. Anzil, S. Vanderschueren, C. Michiels (Université Libre de Bruxelles), and G. Alunga (DR Congo team). The consortium gratefully acknowledges the Rwanda Energy Company and Michel Halbwachs for free access to their industrial platform off Gisenyi. R. Trias (Institut de Physique du Globe de Paris, France) is acknowledged for supplying some of the qPCR-positive controls. This work was funded by the Fonds National de la Recherche Scientifique (FNRS) under the MICKI (Microbial diversity and processes in Lake Kivu) project and the Belgian Federal Science Policy Office EAGLES (East African Great Lake Ecosystem Sensitivity to changes, SD/AR/02A) project, and contributes the European Research Council starting grant project AFRIVAL (African river basins: Catchment-scale carbon fluxes and transformations, 240002).

Conflict of Interest The authors declare no conflict of interest.

References


67. aaaaaaaaaaaaa