RESEARCH ARTICLE

Investigating variability in microbial community composition in replicate environmental DNA samples down lake sediment cores

John K. Pearman[1]*, Georgia Thomson-Laing[1], Jamie D. Howarth[2], Marcus J. Vandergoes[3], Lucy Thompson[1], Andrew Rees[2], Susanna A. Wood[1]

1 Coastal and Freshwater Group, Cawthron Institute, Nelson, New Zealand, 2 Victoria University of Wellington, Wellington, New Zealand, 3 GNS Science, Lower Hutt, New Zealand

* john.pearman@cawthron.org.nz

🔓 OPEN ACCESS

Abstract

Lake sediments are natural archives that accumulate information on biological communities and their surrounding catchments. Paleolimnology has traditionally focussed on identifying fossilized organisms to reconstruct past environments. In the last decade, the application of molecular methodologies has increased in paleolimnological studies, but further research investigating factors such as sample heterogeneity and DNA degradation are required. In the present study we investigated bacterial community heterogeneity (16S rRNA metabarcoding) within depth slices (1-cm width). Sediment cores were collected from three lakes with differing sediment compositions. Samples were collected from a variety of depths which represent a period of time of approximately 1,200 years. Triplicate samples were collected from each depth slice and bacterial 16S rRNA metabarcoding was undertaken on each sample. Accumulation curves demonstrated that except for the deepest (oldest) slices, the combination of three replicate samples were insufficient to characterise the entire bacterial diversity. However, shared Amplicon Sequence Variants (ASVs) accounted for the majority of the reads in each depth slice (max. shared proportional read abundance 96%, 86%, 65% in the three lakes). Replicates within a depth slice generally clustered together in the Non-metric multidimensional scaling analysis. There was high community dissimilarity in older sediment in one of the cores, which was likely due to the laminae in the sediment core not being horizontal. Given that most paleolimnology studies explore broad scale shifts in community structure rather than seeking to identify rare species, this study demonstrates that a single sample is adequate to characterise shifts in dominant bacterial ASVs.

Introduction

Understanding the lasting impacts of environmental decision making is fundamental to protecting and remediating Earth's biosphere. Long term insights are crucial to identifying and

understanding ecological processes and inherent variability; however, biological monitoring rarely extends beyond a few decades [1]. Lake sediments accumulate information about autochthonous biological communities [2, 3] and their surrounding catchments [4, 5], enabling temporal insights into ecological processes [1]. Paleolimnology has traditionally focussed on identifying and enumerating organisms' remains, but in the last decade, advances in molecular methodologies have permitted the identification of soft-bodied organisms that typically do not leave "footprints" in the sediment record [6 and references within].

The genetic material present in environmental samples, such as sediment, can be defined as environmental DNA (eDNA). This DNA originates from a complex mixture of living cells, inactive cells, fragments of dead cells and extracellular DNA from a variety of sources including faeces, urine and saliva [7]. Techniques such as metabarcoding, which enable the identification of taxa from samples containing a mix of species using high throughput sequencing and DNA reference databases, have recently been applied to characterise bacteria, Archaea, microbial eukaryotes, plants and vertebrates in lake sediments [8–12]. The application of these techniques has allowed new insights into temporal trends over decades, centuries and even millennium. Their use has the potential to allow testing of ecological theories which require long term data sets [e.g. evolutionary dynamics, anthropogenic impacts 6, 13].

Before molecular techniques can be robustly applied to the analysis of biological communities in sediment cores, there are a number of methodological considerations that need to be addressed. For example, numerous studies using metabarcoding approaches have shown that DNA extraction efficiency and taxonomic coverage of microbial species in sediments is substantially different depending on the DNA extraction method applied [14–17]. A comprehensive review by Lear et al., [18] suggested that the most appropriate method for small sediment volumes was the DNeasy PowerSoil kit (Qiagen, Germany), which was used in the current study.

Beyond technical considerations, lake sediments have different compositions, which can affect the absorption of DNA; for example, humic acids and clay materials have a stronger binding capacity for DNA than other sediments [19]. Mineralogy, porewater pH, and temperature can also impact DNA retention and degradation in sediments [20, 21]. Reworking of sediments, and subsequent DNA transport, is known to occur through the movement of water, active growth of organisms and bioturbation [22]. However, initial research indicates that DNA leaching is minimal in saturated aquatic sediments where pore water movement is minimal [4, 23].

An area of methodological consideration that requires further investigation in lake sediment cores is heterogeneity in eDNA across the small spatial scales of depth slices of the cores. Spatial heterogeneity within surface sediment samples has been previously shown [24, 25] but while some work has indicated clusters of replicates down sediment cores [26], there has been limited investigation of how variability amongst replicates changes with depth. The investigation of variation within depth slices in sediment cores is especially vital as a large proportion of studies base their analysis on one or two DNA extractions from each depth slice [27–31].

The aim of the present study was to investigate bacterial community heterogeneity (16S rRNA metabarcoding) within depth slices. Sediment cores were collected from three lakes with differing sediment compositions. The deepest sediment in the cores was estimated to be up to 1,200 years old. Triplicate DNA extractions followed by bacterial 16S rRNA metabarcoding were undertaken on samples collected from core depth slices (1 cm width). We hypothesised that: (1) Replication would provide a better representation of the ASV richness within core depth slices; (2) a more complete inventory of microbial richness (accumulation curves plateau) would be obtained with increasing depth; (3) community structure and composition would be more similar within depth slices than between depth slices; and (4) that there would

be greater variability in the triplicates with increasing age of sediment due to differences in DNA degradation over time.

Methods

Sampling sites

The three lakes were chosen to cover a wide range of sediment types (S1 Fig). The map was produced in ggmap [32] using Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.

Lake Nganoke is a small (~3.1 ha), shallow (max. depth 2.3 m) lake located at 17 m above sea level (a.s.l) at the southern end of the Wairarapa Valley in the lower North Island of New Zealand (41˚21′20″ S, 175˚11′10″ E). The catchment of the lake (~174 ha) is incised into a terrace that has a maximum elevation of 100 m and is vegetated with high producing exotic grassland used for sheep and beef farming. The catchment lithology is characterised by early Pleistocene river gravels and late Pleistocene beach deposits [33]. Sediments accumulating on the lake floor are highly organic silts formed largely from autochthonous sedimentation.

Lake Paringa is a medium sized lake (~475 ha) situated at 16 m a.s.l with a maximum depth of 58 m that has formed in an over-deepened glacial trough (~17,000 years ago). It is located on the West Coast of the South Island of New Zealand (43˚43'10"S, 169˚24'8"E), within a large catchment (~6,000 ha) that drains the ~1420 m a.s.l high frontal range of the Southern Alps. The catchment is primarily vegetated in undisturbed temperate podocarp rainforest [34]. The lithology of the lake catchment is predominantly cataclasites, mylonites and schists of the Rakaia Terrane east of the Alpine Fault and Greenland group metasediments to the west of the fault [35]. Lake Paringa's sedimentary fill is characterized by a repeating sequence of three deposits comprising (i) co-seismic megaturbidites formed by shaking-induced subaqueous mass wasting; (ii) post-seismic hyperpycnite stacks formed during periods of elevated fluvial sediment flux from earthquake-induced landsliding; and (iii) inter-seismic layered silts formed between earthquakes when the catchment is relatively geomorphically quiescent [34].

Lake Pounui is a small (~46 ha), shallow (max. depth 9.6 m), lowland coastal lake situated at 10 m a.s.l about 30 km northeast of Wellington, New Zealand (41˚20'34"S, 175˚6'48"E). Lake Pounui is thought to have formed ~3,000 years ago by the damming of a stream valley by beach sediment and alluvial gravels [36]. Lake Pounui's catchment (627 ha) extends to an elevation of 470 m in the foot hills of the Rimutaka Ranges. The majority of hillslopes are vegetated by unmodified indigenous beech-podocarp forest (96%), with the remainder in pastoral land cover [37]. The catchment lithology is composed on late Pleistocene marine benches underlying beach deposits, loess and alluvial gravels south east of the Wharekauhau Thrust fault and Esk Head sand and mudstones north west of the fault [33, 36]; material eroded from these sources enters Lake Pounui [36, 38, 39]. The sediments accumulating on the lake floor are a mixture of allochthonous silts and very fine sands and autochthonous organic silts [36].

Sample collection

In lakes Pounui and Nganoke sediment cores were retrieved using a Uwitech (Mondsee, Austria) gravity corer with 2-m polyvinyl chloride (PVC; 65-mm dia.) barrels. In Lake Paringa a 6-m sediment core was collected using a Mackereth corer (50-mm dia.; [40]). Barrels were cleaned with 2% bleach prior to coring. The collected sediment cores were stored at 4˚C in darkness for up to 4 weeks until sub-sampling for lakes Nganoke and Paringa. The samples from Lake Pounui were collected, subsampled and stored frozen (-80˚C) within 6 hours. A single core per lake was used for molecular analysis.

In the laboratory, the cores were sliced along the longitudinal plane, photographed, and described in detail including colour, texture and structure [41]. To prevent contamination, the top 2–3 mm of each half of the core was carefully removed with a sterile spatula. Slices of the cores were sampled at various depths from the top of the core (S1 Table and S2 Fig). Triplicate sub-samples (~ 0.5 g) were taken near the centre of the half-core using a sterile spatula and placed in separate tubes. Samples were kept frozen (-20°C) and in the dark until DNA extraction. Data on the age of sediment were sourced from other studies [34, 42, 43; S1 Table].

DNA extraction, PCR and high throughput sequencing

Each step of the molecular analyses (i.e. DNA extraction, Polymerase Chain Reaction [PCR] set-up, template addition, PCR analysis) was conducted in dedicated separated sterile laboratories, with sequential workflow to ensure no cross-contamination. Rooms dedicated to DNA extraction, amplification set-up and template addition were equipped with ultra-violet (UV) sterilisation. UV sterilisation of the room and equipment was undertaken for 15 min before and after each use. The PCR set-up and template addition were always undertaken in laminar flow cabinets with HEPA filtration. Aerosol barrier tips were used throughout.

DNA was extracted from approximately 0.25 g of sediment using the DNeasy PowerSoil Kit (Qiagen, Germany) following the manufacturer's instructions on a QIAcube sample preparation robot (Qiagen). A negative extraction control, where no sediment was added to the extraction tube, was included for every 23 samples.

The V3-V4 region of the bacterial 16S rRNA gene was amplified by PCR using the bacterial specific primers 341F: 5'-CCT ACG GGN GGC WGC AG-3' and 805R: 5'-GAC TAC HVG GGT ATC TAA TCC-3' [44]. The primers included Illumina™ overhang adapters to allow dual indexing as described in Kozich et al., [45]. Triplicate PCR reactions were undertaken on each sample in 20-μL volumes. The reaction mixture consisted of 10 μL of 2 × PCR MyFi™ Mix (Bioline), 1 μL of each primer (10 μM), 1.5 μL of bovine serum albumin (20 mg mL^{-1}) and 1.5 μL of template DNA. Cycling conditions were an initial 1 min 30 sec at 94°C before 30 cycles of 94°C for 30 s, 52°C for 30 s, 72°C for 45 s, with a final extension step at 72°C for 5 min. Negative PCR controls were run alongside the samples. The triplicate PCR replicates were pooled and 20 μL was cleaned and normalized using SequalPrep Normalization plates (ThermoFisher Scientific, USA), resulting in a concentration of ~1 ng mL^{-1}. The cleaned samples were sent to Auckland Genomics Facility for paired-end (2 x 250 base pairs (bp)) sequencing on an Illumina Miseq™ platform. Sequence libraries were prepared as detailed in the Illumina 16S metagenomics library prep manual (https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf). One deviation from the protocol was that, after the indexing PCR, 5 μL of each sample was pooled, and clean-up was undertaken on the pooled library instead of samples being individually cleaned. The concentration and quality of the library was quantified using a bioanalyzer. The library was diluted to 4 nM and denatured, and a 15% PhiX spike-in was added. The library was further diluted to a final loading concentration of 7 ρM; raw sequence reads were deposited in the NCBI short read archive under the accession number: PRJNA655562.

Primers were removed from the raw reads with cutadapt [46] allowing one mismatch. Sequences without primer sequences were discarded. Remaining sequences were processed with DADA2 [47] within the R framework [48]. Forward reads were truncated to 230 bp while reverse reads were truncated to 228 bp. The number of allowed maximum "expected errors" (maxEE) was two and four, to account for differences in sequence quality, for the forward and reverse sequences respectively. The first 10^8 bp were used to construct a parametric error

matrix. Sequences were dereplicated and amplicon sequence variants (ASVs) were inferred based on this error matrix. After, inference singletons were discarded and remaining paired-end reads were merged with a maximum of 1 bp mismatch and a required minimum overlap of 10 bp. Chimeric sequences were removed from the analysis using the removeBimeraDenovo function within DADA2.

The resulting ASVs were taxonomically classified in DADA2 using the rdp classifier [49] against the SILVA 132 database [50] with a bootstrap of 70. The results were combined into a phyloseq object [51], and any ASVs classified as eukaryotes, chloroplasts or mitochondria were removed. Negative controls were assessed and read numbers for ASVs found in the negative controls were removed from the samples via subtraction [52]. After this quality check step on average greater than 99% of reads were retained in the lakes Nganoke and Pounui samples, and greater than 95% for Lake Paringa samples (S2 Table).

Accumulation curves were plotted showing the number of ASVs against sequencing depth for each replicate and for the replicates merged together using the R package ranacapa [53]. To compare samples within lakes, subsampling to an even sequencing depth was undertaken for each sample per lake (Lake Nganoke: 14,900 reads, Lake Paringa:18,800 reads, Lake Pounui: 9,800 reads). The number of observed ASVs was calculated per sample, using the phyloseq package [51] and plotted against depth to see changes in richness down the core. The number of shared ASVs and reads amongst the replicates was calculated to assess the similarity of the replicates within a slice. The community composition of each depth slice was assessed at the phyla level of taxonomic classification. Multivariate analysis, to see the clustering of replicates within slices, was undertaken on the samples using Bray Curtis dissimilarity matrices and visualized with non-metric multidimensional scaling (nMDS) plots. Each depth slice was assigned a number, increasing with depth, to assist with visualization of data. Smaller numbers represent the top of the core and larger numbers deeper slices in the core. Exact depths/ages are detailed in S1 Table. Comparisons of the mean similarity (1- Bray Curtis distance and 1- Jaccard distance) within and between depth slices were calculated using a Kruskal-Wallis test in R. The mean similarity (1- Bray Curtis distance) of the triplicates for each depth slice was calculated and plotted with a linear regression calculated to assess differences in similarity with depth. Figures were produced in R using the package ggplot2 [54].

Results and discussion

The use of molecular techniques to characterise biological communities in environmental samples has been increasing over the past decade. There is an ongoing need to ensure an in-depth knowledge of the robustness and caveats of these techniques for specific applications. A total of 5,317,829, 1,880,055 and 1,827,617 sequences were obtained for sediment core samples from lakes Nganoke, Paringa and Pounui, respectively. After bioinformatically subsampling to an even depth, this resulted in 45,139 ASVs for Lake Nganoke, 19,540 ASVs for Lake Paringa, and 51,202 ASVs for Lake Pounui.

Metabarcoding samples often contain a high number of low abundance ASVs [55, 56], and large sequencing efforts are required to reveal the full level of diversity. This was observed in the present study and is highlighted by the accumulation curves which, in general, did not reach a plateau, especially in samples near the top of the core and for the merged replicates (Fig 1 and S3–S5 Figs). Similar results have been observed in other lake sediment studies that have used molecular methodologies [57, 58]. These pipelines have a finer scale of resolution than previous operational taxonomic unit methods [47]. It is however feasible that a portion of the diversity observed is due to PCR/sequencing errors that were not detected in the bioinformatics pipeline. The accumulation curves in the present study suggest that taking three
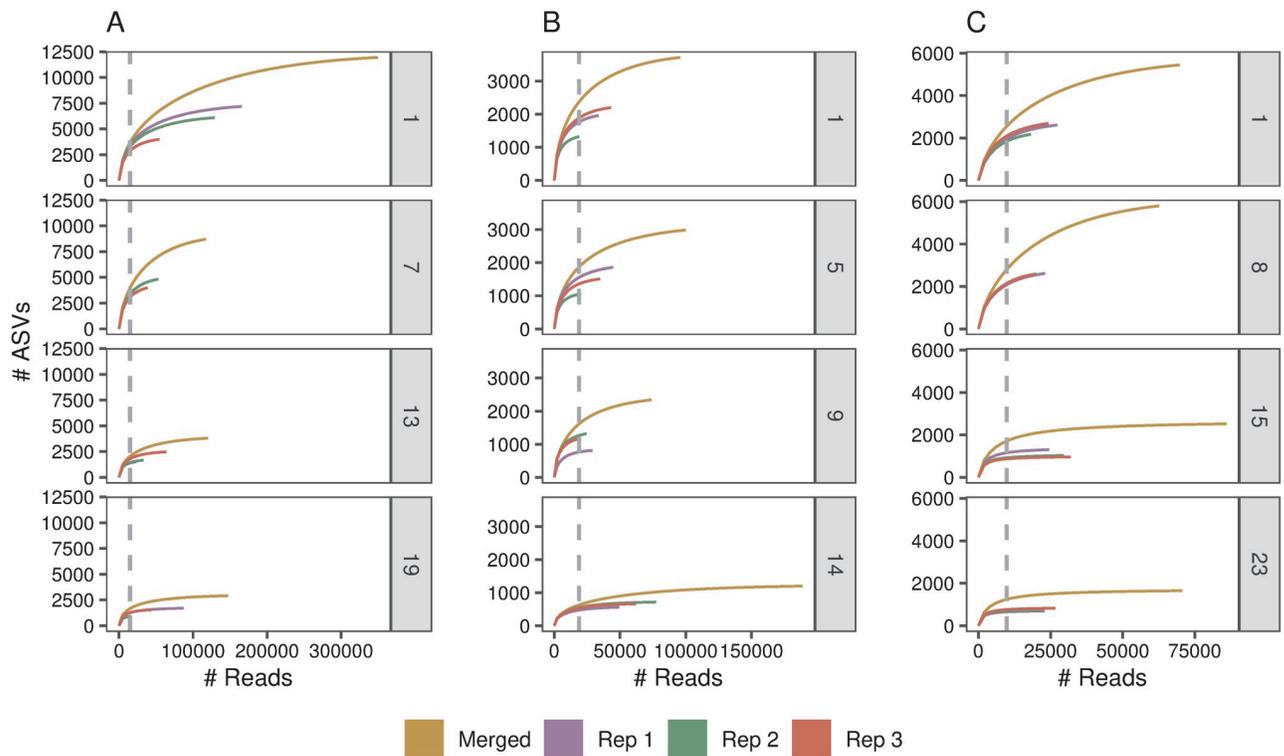
Fig 1. Accumulation curves displaying the observed number of Amplicon Sequence Variants (ASVs) with increasing number of reads for four distinct depth slices (top of core, two slices at intermediate core depths and bottom of core (see S1 Table for depths of slices)) in sediment core samples taken from; A) Lake Nganoke, B) Lake Paringa, and C) Lake Pounui. Accumulation curves are depicted for individual replicates and as a combination of replicates (merged). The level of sub sampling to an even depth used in the further analysis was indicated by the dashed vertical line. Plots for all slices are in S3–S5 Figs. Note: some lines overlap and therefore not all three are visible.

https://doi.org/10.1371/journal.pone.0250783.g001

replicate samples is insufficient to capture the full extent of bacterial diversity in most sediment core slices, although a fuller inventory of the ASVs in deeper samples was achieved (Fig 1). Further investigation including more replication as well as the inclusion of technical replicates is required to improve the estimation of diversity within sediment core samples, especially in shallower samples.

The more complete inventory of ASVs in the deeper sections of the core could be due to the lower ASV richness observed with depth in the core (S6 Fig). This trend could be the result of a couple of factors. Firstly, at deeper depths there may be increased damage to the DNA present in the sediment through strand breakage, abasic sites, miscoding lesions and crosslinks [59, 60]. This can result in fewer amplifiable templates and, as the majority of the ASVs present in metabarcoding samples are rare and of low abundance, it is likely that a proportion of these would no longer be detected [6, 58]. Rates of DNA degradation are affected by a combination of abiotic (e.g. mineralogical composition, organic load, temperature and stable stratification) and biotic factors (e.g. degradation via DNase activity) and thus the effect of DNA degradation is likely to vary between lakes [6, 61]. Secondly, the samples taken in this study cover a time scale ranging from ~800 years (Pounui) to ~1200 years (Paringa) before present. Thus, it cannot be ruled out that alpha diversity is being affected by temporal changes, driven by pressures such as climate change and anthropogenic impacts, in the ecology of the lakes. Indeed, it has previously been shown that climate change has led to an increase in the diversity of photosynthetic microbial taxa [62, 63]. An in-depth analysis of the composition of the changes in

microbial communities is the subject of other research, but a high-level analysis shows temporal changes. For example, the phylum Bacteroidetes increases in relative abundance at the top of the core in all three lakes (S7 Fig). Other phyla, such as Cyanobacteria (Pounui) and Nitrospirae (Nganoke) also increase in relative abundance in recent times, indicating shifts in lake dynamics possibly related to increases in temperature [63] and nutrients [64], respectively. A combination of methods could be applied to obtain a better indication of sediment core richness, including increasing sequencing depth within a sample, undertaking multiple DNA extractions at a depth, or increasing the number of PCR replicates undertaken [30, 65].

The number of ASVs shared amongst the three replicates was lowest in Lake Pounui (6.5–12.2%), with Lake Paringa (11.7–25.3%) and Lake Nganoke (13.7–23.4%) having higher proportions of shared ASVs. There was no significant trend in the number of shared ASVs with depth (p > 0.05; S8A–S8C Fig). Although the number of ASVs shared amongst the three replicates within a depth slice was relatively low, the majority of reads were accounted for by shared ASVs. This indicates that within a depth slice the abundant community is being detected in all replicates, however there is a large number of rare and low abundance ASVs present as has been previously observed in environmental metabarcoding studies [55, 56]. In Lake Paringa, shared ASVs amongst replicates within a depth slice accounted for a maximum of 96% (range 76–96%) of the reads, compared to 69–86% in Lake Nganoke and 45–65% in Lake Pounui. Bioturbation, namely organism-induced mixing of sediment in both vertical and horizontal directions, might affect the numbers of shared reads within each lake. Lakes Nganoke and Pounui both have kākahi (freshwater mussel; *Echyridella menziesi*) which would likely result in higher bioturbation rates than in Lake Paringa. With lower levels of bioturbation, the stratigraphy would be less disturbed leading to more homogeneous replicates at any depth and consequently a higher shared number of reads. While lakes Nganoke and Pounui have higher levels of bioturbation, likely reducing the number of shared reads amongst triplicates, there is lower sedimentary layering in Lake Nganoke. Bioturbation is therefore likely to be mixing similar layers and distinct changes in the community would be reduced.

Similarity comparisons were undertaken to further explore differences within and between depth slices. The mean similarity of pairwise comparisons within depth slices (based on relative abundance data) was significantly higher than the mean of pair wise similarity comparisons between slices (1 –Bray Curtis distance; Nganoke: p < 0.001; Paringa: p < 0.001; Pounui: p < 0.001; Fig 2A–2C). A similar trend was noted when the presence/absence of ASVs was assessed (1 –Jaccard distance; Nganoke: p < 0.001; Paringa: p < 0.001; Pounui: p < 0.001; Fig 2D–2F). This indicates that the variability in bacteria communities within a depth slice is lower than amongst depths. Thus, this suggests that temporal trends can be discerned from microbial communities across small spatial scales in agreement with previous studies.

The non-metric multidimensional (nMDS) plots showed that depth slice replicates generally clustered together (Fig 3) although there was a temporal trend observed in all lakes. We anticipated that there would be marked differences in bacterial community composition and structure across depths in all lakes. Over the last ~1,000 years these lakes have been subjected to multiple natural and anthropogenic disturbances which would have impacted their microbial communities ([34, 43], Andrew Rees, Victoria University unpub. data). Studies on these taxonomic shifts and their potential drivers are the focus of ongoing studies. There were several notable exceptions to the clustering, for example, triplicates were relatively widely dispersed in Lake Paringa slices 9–11. Reasons for the lower homogeneity in these samples is unknown, and technical origins of this variation cannot be ruled out, but there could be greater sediment heterogeneity due to rapid changes in the landscape around the lakes (caused by impacts of human settlement or natural disturbances, i.e., earthquakes). In Lake Pounui slices 11, 19 and 20 (Fig 3) had a wider spread. Slices 19 and 20 correspond to a region of the
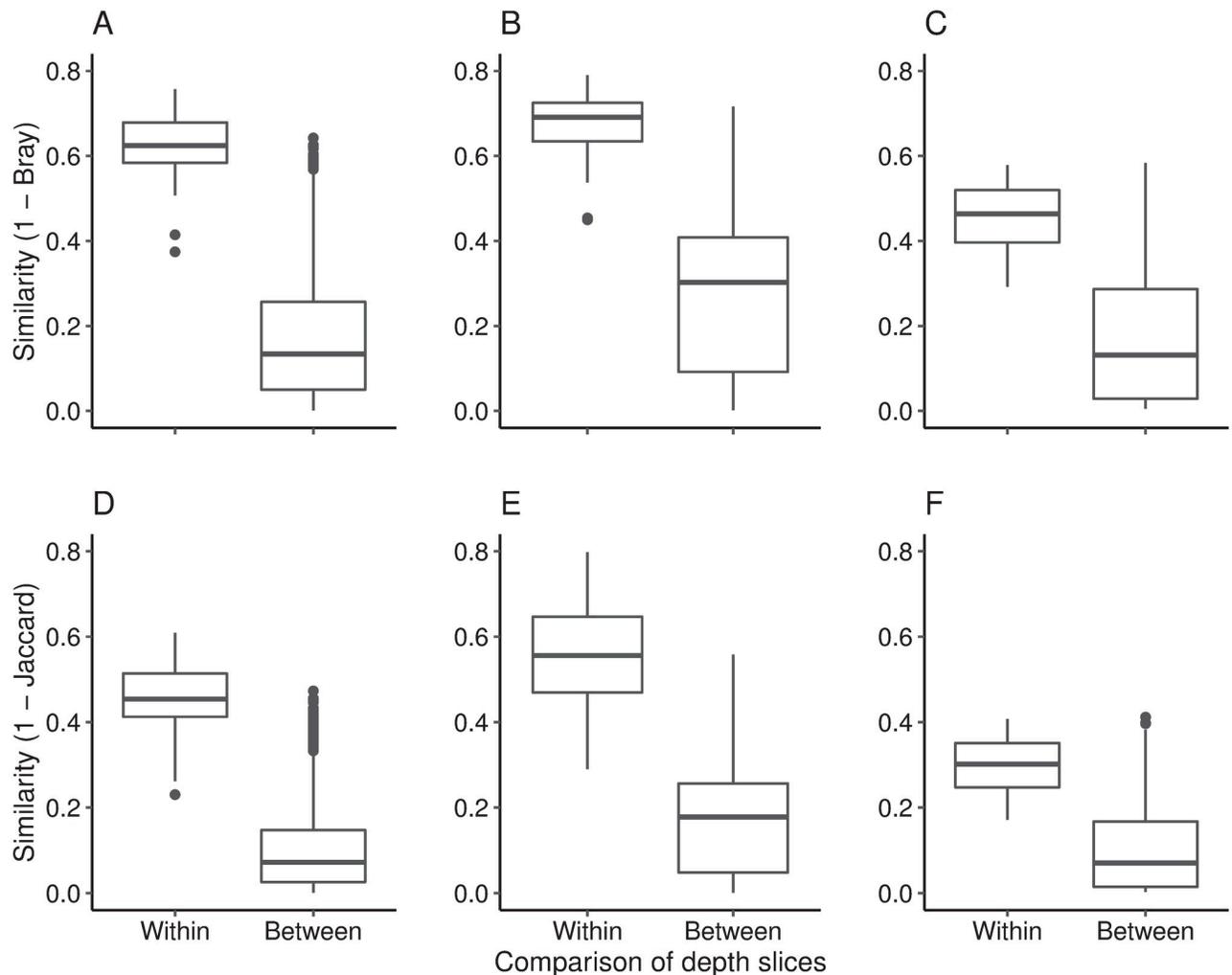
Fig 2. Similarity comparisons (1—Bray Curtis distance) within and between depth slices for Lake Nganoke (A), Lake Paringa (B), and Lake Pounui (C). Similarity comparisons (1—Jaccard distance) within and between depth slices for Lake Nganoke (D), Lake Paringa (E), and Lake Pounui (F). Boxes display the first and third quartile spread of the data, with the line in the box indicating the median, the whiskers denoting the minimum and maximum values and the dots as outliers of the data.

core where the laminae are not horizontal and have a dip of 20–30 degrees (S2 Fig). Replicates taken in this part of the core may be sampling across the stratigraphy which would consequently lower the similarity and may explain the lack of clustering. In Lake Nganoke, one of the replicates from slice 6 was very similar in composition to the slice above. Reworking of sediment by living organisms has been documented and could account for this observation although there was 5 cm between these layers [22].

Sediments from a range of depths/ages was analysed to determine whether there was greater variability (lower similarity between triplicates) in older sediments. We hypothesised that greater variability would occur in older sediments and that this might be caused by stochastic differences in DNA degradation. However, this hypothesis was not supported in lakes Nganoke and Paringa, although in Lake Paringa there was a significant positive relationship between similarity and slice depth (p = 0.0241, r^2 = 0.357; Fig 4) and a significant increase in the proportion of shared reads with depth was observed (p < 0.001, r^2 = 0.658; S8E Fig). This could indicate that DNA degradation in the core is removing the rarer low abundance ASVs,
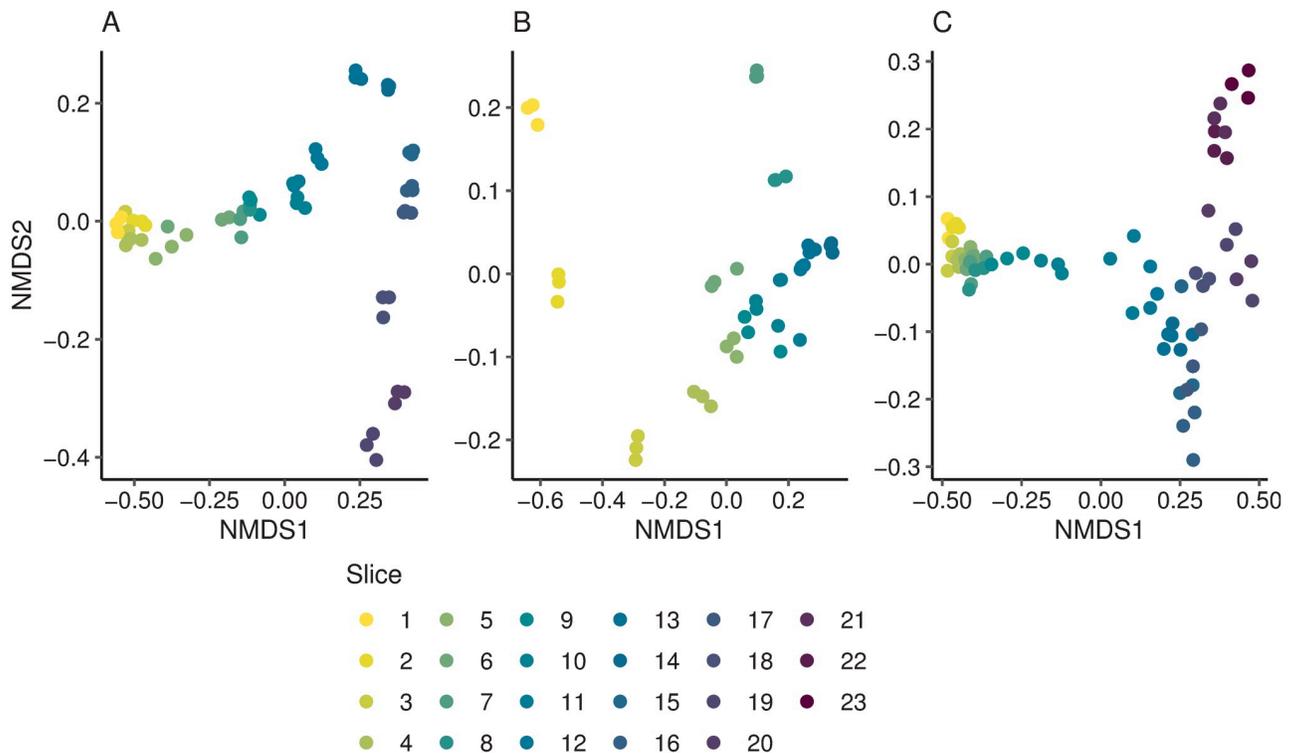
Fig 3. Non-metric multidimensional scaling plot based on Bray Curtis distance matrices for samples taken from Lake Nganoke (A), Lake Paringa (B), and Lake Pounui (C). Points are coloured by depth slice. Approximate ages of slices vary between lakes and are given in S1 Table. Stress values for the non-metric multidimensional scaling (nMDS) plots were Nganoke: 0.048; Paringa: 0.066 and Pounui: 0.055.

and that these deeper slices are thus dominated by abundant ASVs that are shared amongst all replicates. It is unknown why this significant trend is only observed in Lake Paringa and further investigation are required including assessing the effect of how different sediment matrices affect DNA degradation [20, 21].
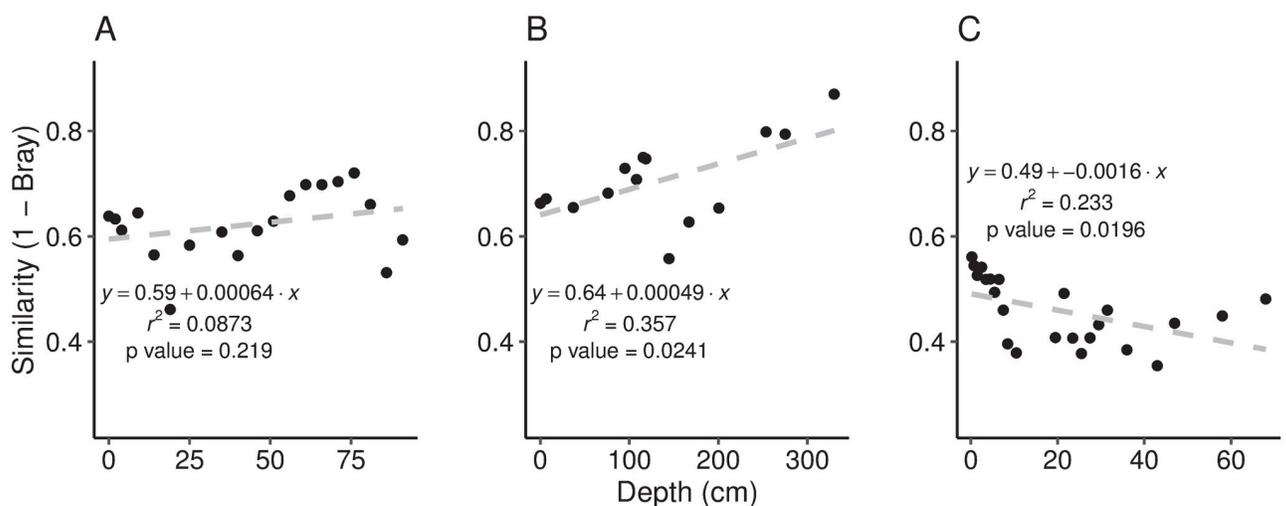
Fig 4. Linear regression of similarity (1-Bray Curtis) and depth slice in Lake Nganoke (A), Lake Paringa (B), and Lake Pounui (C). Depth slices are organised by depth (1 being the most recent). Approximate ages of depth slices vary between lakes and are given in S1 Table.

Only Lake Pounui showed a significant decline in similarity within a slice with depth (p = 0.0196, r^2 = 0.233; Fig 4). The decrease in similarity at these depths is most likely due to the non-horizontal laminae deeper in the core (S2 Fig), which would likely result in the replicates representing different time periods. If laminae are not horizontal then increased variability in replication may be observed and results would need to be carefully considered as replicates may represent different temporal periods.

Conclusion

In summary, replication improved the detection of rare species in the sediment samples, however the full diversity was not captured even when triplicate samples were used in the surface depth slices. However, most paleolimnology studies do not aim to identify rare species, but rather explore broad scale shifts in community structure [6, 28]. There may be exceptions to this when investigating rare or invasive species [66], and in these instances it might be more appropriate to use targeted methods such as droplet digital PCR or refrain from subsampling if using metabarcoding [67].

The community structure and composition were more similar within depth slices than between them. However, there were exceptions where one replicate was markedly different. Reasons for this are unknown but could be related to bioturbation or sediment heterogeneity. This highlights the need for caution when interpreting these types of data and further replication and increased sampling of depths immediately above and below may be required. Variability among replicates remained relatively constant among depth and did not decrease as anticipated. The exception was Lake Pounui were there was a decrease in similarity with depth which was most likely related to laminae not being horizontal at these depths.

Supporting information

S1 Fig. Map of the position of each lakes. The map was produced in ggmap [32] using Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.
(TIF)

S2 Fig. Photographs of the sediment cores sampled with the sampled depth slices depicted.
(TIF)

S3 Fig. Accumulation curves displaying the observed number of Amplicon Sequence Variants (ASVs) with increasing number of reads for each distinct depth slice (see S1 Table for depths of slices) in sediment core samples taken from Lake Nganoke.
(TIF)

S4 Fig. Accumulation curves displaying the observed number of Amplicon Sequence Variants (ASVs) with increasing number of reads for each distinct depth slice (see S1 Table for depths of slices) in sediment core samples taken from Lake Paringa.
(TIF)

S5 Fig. Accumulation curves displaying the observed number of Amplicon Sequence Variants (ASVs) with increasing number of reads for each distinct depth slice (see S1 Table for depths of slices) in sediment core samples taken from Lake Pounui.
(TIF)

S6 Fig. The total richness per slice (triplicates combined) with depth down the core. A—Lake Nganoke; B—Lake Paringa; C- Lake Pounui.
(TIF)

S7 Fig. Relative proportional composition of the community at phyla level with depth down the core. A—Lake Nganoke; B—Lake Paringa; C- Lake Pounui. Phyla that did not account for on average > 1% of the community were not show so bars do not add up to 100%.
(TIF)

S8 Fig. The proportion of shared Amplicon Sequence Variants (ASVs; A—Lake Nganoke; B—Lake Paringa; C—Lake Pounui) and reads (D—Lake Nganoke; E—Lake Paringa; F—Lake Pounui) amongst replicates within a depth slice plotted against depth. Note the different scales.
(TIF)

S1 Table. Sub-bottom depth and age of each depth slice for lakes Nganoke, Paringa and Pounui.
(XLSX)

S2 Table. Number of reads before and after removing contaminating sequences from controls by subtraction.
(XLSX)

Author Contributions

Conceptualization: Jamie D. Howarth, Marcus J. Vandergoes, Susanna A. Wood.

Data curation: John K. Pearman.

Formal analysis: John K. Pearman.

Funding acquisition: Marcus J. Vandergoes, Susanna A. Wood.

Investigation: Jamie D. Howarth, Susanna A. Wood.

Methodology: Georgia Thomson-Laing, Lucy Thompson, Andrew Rees.

Project administration: Marcus J. Vandergoes, Susanna A. Wood.

Visualization: John K. Pearman.

Writing – original draft: John K. Pearman, Susanna A. Wood.

Writing – review & editing: John K. Pearman, Georgia Thomson-Laing, Jamie D. Howarth, Marcus J. Vandergoes, Lucy Thompson, Andrew Rees, Susanna A. Wood.

References

1. Seddon AWR, Mackay AW, Baker AG, Birks HJB, Breman E, Buck CE, et al. Looking forward through the past: identification of 50 priority research questions in palaeoecology. J Ecol. 2014; 102: 256–267.

2. Capo E, Debroas D, Arnaud F, Domaizon I. Is planktonic diversity well recorded in sedimentary DNA? Toward the reconstruction of past protistan diversity. Microb Ecol. 2015; 70: 865–875. https://doi.org/10.1007/s00248-015-0627-2 PMID: 26022714

3. Millet L, Arnaud F, Heiri O, Magny M, Verneaux V, Desmet M. Late-Holocene summer temperature reconstruction from chironomid assemblages of Lake Anterne, northern French Alps: The Holocene. 2009. https://doi.org/10.1177/0959683608100576

4. Giguet-Covex C, Pansu J, Arnaud F, Rey P-J, Griggo C, Gielly L, et al. Long livestock farming history and human landscape shaping revealed by lake sediment DNA. Nat Commun. 2014; 5: 1–7. https://doi.org/10.1038/ncomms4211 PMID: 24487920

5. Pansu J, Giguet-Covex C, Ficetola GF, Gielly L, Boyer F, Zinger L, et al. Reconstructing long-term human impacts on plant communities: an ecological approach based on lake sediment DNA. Mol Ecol. 2015; 24: 1485–1498. https://doi.org/10.1111/mec.13136 PMID: 25735209

6. Domaizon I, Winegardner A, Capo E, Gauthier J, Gregory-Evans I. DNA-based methods in paleolimnology: new opportunities for investigating long-term dynamics of lacustrine biodiversity. J Paleolimnol. 2017; 58: 1–21.

7. Taberlet P, Bonin A, Zinger L, Coissac E. Environmental DNA: For Biodiversity Research and Monitoring. Oxford University Press; 2018.

8. Alsos IG, Sjögren P, Edwards ME, Landvik JY, Gielly L, Forwick M, et al. Sedimentary ancient DNA from Lake Skartjørna, Svalbard: Assessing the resilience of arctic flora to Holocene climate change. 2016. https://doi.org/10.1177/0959683615612563

9. Kisand V, Talas L, Kisand A, Stivrins N, Reitalu T, Alliksaar T, et al. From microbial eukaryotes to metazoan vertebrates: Wide spectrum paleo-diversity in sedimentary ancient DNA over the last 14,500 years. Geobiology. 2018; 16: 628–639. https://doi.org/10.1111/gbi.12307 PMID: 30035382

10. Parducci L, Bennett KD, Ficetola GF, Alsos IG, Suyama Y, Wood JR, et al. Ancient plant DNA in lake sediments. New Phytol. 2017; 214: 924–942. https://doi.org/10.1111/nph.14470 PMID: 28370025

11. Poulain AJ, Aris-Brosou S, Blais JM, Brazeau M, (Bill) Keller W, Paterson AM. Microbial DNA records historical delivery of anthropogenic mercury. ISME J. 2015; 9: 2541–2550. https://doi.org/10.1038/ismej.2015.86 PMID: 26057844

12. Vuillemin A, Ariztegui D, Leavitt PR, Bunting L. Recording of climate and diagenesis through sedimentary DNA and fossil pigments at Laguna Potrok Aike, Argentina. Biogeosciences. 2016; 13: 2475–2492. https://doi.org/10.5194/bg-13-2475-2016

13. Bálint M, Pfenninger M, Grossart H-P, Taberlet P, Vellend M, Leibold MA, et al. Environmental DNA time series in ecology. Trends Ecol Evol. 2018; 33: 945–957. https://doi.org/10.1016/j.tree.2018.09.003 PMID: 30314916

14. Holmsgaard PN, Norman A, Hede SChr, Poulsen PHB, Al-Soud WA, Hansen LH, et al. Bias in bacterial diversity as a result of Nycodenz extraction from bulk soil. Soil Biol Biochem. 2011; 43: 2152–2159. https://doi.org/10.1016/j.soilbio.2011.06.019

15. Knauth S, Schmidt H, Tippkötter R. Comparison of commercial kits for the extraction of DNA from paddy soils. Lett Appl Microbiol. 2013; 56: 222–228. https://doi.org/10.1111/lam.12038 PMID: 23252687

16. Young JM, Rawlence NJ, Weyrich LS, Cooper A. Limitations and recommendations for successful DNA extraction from forensic soil samples: A review. Sci Justice. 2014; 54: 238–244. https://doi.org/10.1016/j.scijus.2014.02.006 PMID: 24796953

17. Pearman JK, Keeley NB, Wood SA, Laroche O, Zaiko A, Thomson-Laing G, et al. Comparing sediment DNA extraction methods for assessing organic enrichment associated with marine aquaculture. PeerJ. 2020; 8: e10231. https://doi.org/10.7717/peerj.10231 PMID: 33194417

18. Lear G, Dickie I, Banks J, Boyer S, Buckley H, Buckley T, et al. Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. N Z J Ecol. 2018; 42: 10. https://doi.org/10.20417/nzjecol.42.9

19. Dell'Anno A, Stefano B, Danovaro R. Quantification, base composition, and fate of extracellular DNA in marine sediments. Limnol Oceanogr. 2002; 47: 899–905. https://doi.org/10.4319/lo.2002.47.3.0899

20. Levy-Booth DJ, Campbell RG, Gulden RH, Hart MM, Powell JR, Klironomos JN, et al. Cycling of extracellular DNA in the soil environment. Soil Biol Biochem. 2007; 39: 2977–2991. https://doi.org/10.1016/j.soilbio.2007.06.020

21. Torti A, Lever MA, Jørgensen BB. Origin, dynamics, and implications of extracellular DNA pools in marine sediments. Mar Genomics. 2015; 24: 185–196. https://doi.org/10.1016/j.margen.2015.08.007 PMID: 26452301

22. Rawlence NJ, Lowe DJ, Wood JR, Young JM, Churchman GJ, Huang Y-T, et al. Using palaeoenvironmental DNA to reconstruct past environments: progress and prospects. J Quat Sci. 2014; 29: 610–626. https://doi.org/10.1002/jqs.2740

23. Anderson-Carpenter LL, McLachlan JS, Jackson ST, Kuch M, Lumibao CY, Poinar HN. Ancient DNA from lake sediments: Bridging the gap between paleoecology and genetics. BMC Evol Biol. 2011; 11: 30. https://doi.org/10.1186/1471-2148-11-30 PMID: 21272315

24. Brannock PM, Halanych KM. Meiofaunal community analysis by high-throughput sequencing: Comparison of extraction, quality filtering, and clustering methods. Mar Genomics. 2015; 23: 67–75. https://doi.org/10.1016/j.margen.2015.05.007 PMID: 26001512

25. Lejzerowicz F, Esling P, Pawlowski J. Patchiness of deep-sea benthic Foraminifera across the Southern Ocean: Insights from high-throughput DNA sequencing. Deep Sea Res Part II Top Stud Oceanogr. 2014; 108: 17–26. https://doi.org/10.1016/j.dsr2.2014.07.018

26. Ibrahim A, Capo E, Wessels M, Martin I, Meyer A, Schleheck D, et al. Anthropogenic impact on the historical phytoplankton community of Lake Constance reconstructed by multimarker analysis of sediment-core environmental DNA. Mol Ecol. 2020;n/a. https://doi.org/10.1111/mec.15696 PMID: 33070403

27. Monchamp M-E, Spaak P, Domaizon I, Dubois N, Bouffard D, Pomati F. Homogenization of lake cyanobacterial communities over a century of climate change and eutrophication. Nat Ecol Evol. 2018; 2: 317–324. https://doi.org/10.1038/s41559-017-0407-0 PMID: 29230026

28. Capo E, Debroas D, Arnaud F, Perga M-E, Chardon C, Domaizon I. Tracking a century of changes in microbial eukaryotic diversity in lakes driven by nutrient enrichment and climate warming. Environ Microbiol. 2017; 19: 2873–2892. https://doi.org/10.1111/1462-2920.13815 PMID: 28585365

29. Monchamp M-E, Walser J-C, Pomati F, Spaak P. Sedimentary DNA reveals cyanobacterial community diversity over 200 years in two perialpine lakes. Appl Environ Microbiol. 2016; 82: 6472–6482. https://doi.org/10.1128/AEM.02174-16 PMID: 27565621

30. Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, De Barba M, et al. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Mol Ecol Resour. 2015; 15: 543–556. https://doi.org/10.1111/1755-0998.12338 PMID: 25327646

31. Parducci L, Väliranta M, Salonen JS, Ronkainen T, Matetovici I, Fontana SL, et al. Proxy comparison in ancient peat sediments: pollen, macrofossil and plant DNA. Philos Trans R Soc Lond B Biol Sci. 2015; 370: 20130382. https://doi.org/10.1098/rstb.2013.0382 PMID: 25487333

32. Kahle D, Wickham H. ggmap: Spatial Visualization with ggplot2. R J. 5: 144–161.

33. Begg JG, Johnston MR. Geology of the Wellington area 1:250 000 geological map 10. Institute of Geological and Nuclear Sciences; 2000. https://natlib.govt.nz/records/22189402

34. Howarth JD, Fitzsimons SJ, Norris RJ, Jacobsen GE. Lake sediments record cycles of sediment flux driven by large earthquakes on the Alpine fault, New Zealand. Geology. 2012; 40: 1091–1094. https://doi.org/10.1130/G33486.1

35. Rattenbury M, Jongens R, Cox S. Geology of the Haast Area 1:250 000. Geological Map 14. Institute of Geological & Nuclear Sciences; 2010.

36. Turner GM, Lillis DA. A palaeomagnetic secular variation record for New Zealand during the past 2500 years. Phys Earth Planet Inter. 1994; 83: 265–282. https://doi.org/10.1016/0031-9201(94)90093-0

37. Perrie A, Milne JR. Lake water quality and ecology in the Wellington region: State and trends. Wellington, NZ: Greater Wellington Regional Council; 2012. Report No.: GW/EMI-T-12/139.

38. Mcfadgen BG. Archaeology of the Wellington Conservancy: Wairarapa. A study in tectonic archaeology. Wellington, New Zealand: Department of Conservation; 2003. https://www.semanticscholar.org/paper/Archaeology-of-the-Wellington-Conservancy%3A-A-study-Mcfadgen/73b4da315cbcb3c6f95add49b34fd9d9aeadca8c

39. Palmer AS, Vucetich CG, McGlone M, Harper MA. Last Glacial loess and early Last Glacial vegetation history of Wairarapa Valley, New Zealand. N Z J Geol Geophys. 1989; 32: 499–513. https://doi.org/10.1080/00288306.1989.10427557

40. Mackereth FJH. A portable core sampler for lake deposits. Limnol Oceanogr. 1958; 3: 181–191. https://doi.org/10.4319/lo.1958.3.2.0181

41. Schnurrenberger D, Russell J, Kelts K. Classification of lacustrine sediments based on sedimentary components. J Paleolimnol. 2003; 29: 141.

42. Cochrane L. Reconstructing ecological change, catchment disturbance, and anthropogenic impact over the last 3000 years at Lake Pounui, Wairarapa, New Zealand. Victoria University, Wellington. 2017. http://researcharchive.vuw.ac.nz/handle/10063/6665

43. Parish J. Reconstructing the effects of riparian planting in mitigating agriculturally intensified run off from paleo lake records: A case study of Lake Nganoke. MSc thesis unpublished, Victoria University, Wellington. 2020.

44. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 2013; 41. https://doi.org/10.1093/nar/gks808 PMID: 22933715

45. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol. 2013; 79: 5112–5120. https://doi.org/10.1128/AEM.01043-13 PMID: 23793624

46. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011; 17: 10–12. https://doi.org/10.14806/ej.17.1.200

47. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016; 13: 581. https://doi.org/10.1038/nmeth.3869 PMID: 27214047

48. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. 2020.

49. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007; 73: 5261–5267. https://doi.org/10.1128/AEM.00062-07 PMID: 17586664

50. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. 2007; 35: 7188–7196. https://doi.org/10.1093/nar/gkm864 PMID: 17947321

51. McMurdie PJ, Holmes S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PloS One. 2013; 8: e61217. https://doi.org/10.1371/journal.pone.0061217 PMID: 23630581

52. Bell KL, Burgess KS, Botsch JC, Dobbs EK, Read TD, Brosi BJ. Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. Mol Ecol. 2019; 28: 431–455. https://doi.org/10.1111/mec.14840 PMID: 30118180

53. Kandlikar GS, Gold ZJ, Cowen MC, Meyer RS, Freise AC, Kraft NJB, et al. ranacapa: An R package and Shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. F1000Research. 2018; 7. https://doi.org/10.12688/f1000research.16680.1 PMID: 30613396

54. Wickham H. ggplot2: elegant graphics for data analysis. Springer; 2016.

55. Lynch MD, Neufeld JD. Ecology and exploration of the rare biosphere. Nat Rev Microbiol. 2015; 13: 217. https://doi.org/10.1038/nrmicro3400 PMID: 25730701

56. Pedros-Alio C. The rare bacterial biosphere. Annu Rev Mar Sci. 2012; 4: 449–66. https://doi.org/10.1146/annurev-marine-120710-100948 PMID: 22457983

57. Pearman JK, Biessy L, Thomson-Laing G, Waters S, Vandergoes MJ, Howarth JD, et al. Local factors drive bacterial and microeukaryotic community composition in lake surface sediment collected across an altitudinal gradient. FEMS Microbiol Ecol. 2020; 96. https://doi.org/10.1093/femsec/fiaa070 PMID: 32310266

58. Wurzbacher C, Fuchs A, Attermeyer K, Frindte K, Grossart H-P, Hupfer M, et al. Shifts among Eukaryota, Bacteria, and Archaea define the vertical organization of a lake sediment. Microbiome. 2017; 5: 41. https://doi.org/10.1186/s40168-017-0255-9 PMID: 28388930

59. Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S. Ancient DNA. Nat Rev Genet. 2001; 2: 353–359. https://doi.org/10.1038/35072071 PMID: 11331901

60. Lindahl T. Instability and decay of the primary structure of DNA. Nature. 1993; 362: 709–715. https://doi.org/10.1038/362709a0 PMID: 8469282

61. Corinaldesi C, Beolchini F, Dell'Anno A. Damage and degradation rates of extracellular DNA in marine sediments: implications for the preservation of gene sequences. Mol Ecol. 2008; 17: 3939–3951. https://doi.org/10.1111/j.1365-294X.2008.03880.x PMID: 18643876

62. Keck F, Millet L, Debroas D, Etienne D, Galop D, Rius D, et al. Assessing the response of micro-eukaryotic diversity to the Great Acceleration using lake sedimentary DNA. Nat Commun. 2020; 11: 1–8.

63. Paerl HW, Huisman J. Blooms like it hot. Science. 2008; 320: 57. https://doi.org/10.1126/science.1155398 PMID: 18388279

64. Li F, Zhang X, Xie Y, Wang J. Sedimentary DNA reveals over 150 years of ecosystem change by human activities in Lake Chao, China. Environ Int. 2019; 133: 105214. https://doi.org/10.1016/j.envint.2019.105214 PMID: 31665682

65. Lanzén A, Lekang K, Jonassen I, Thompson EM, Troedsson C. DNA extraction replicates improve diversity and compositional dissimilarity in metabarcoding of eukaryotes in marine sediments. PLoS ONE. 2017; 12. https://doi.org/10.1371/journal.pone.0179443 PMID: 28622351

66. Savichtcheva O, Debroas D, Kurmayer R, Villar C, Jenny JP, Arnaud F, et al. Quantitative PCR Enumeration of Total/Toxic Planktothrix rubescens and Total Cyanobacteria in Preserved DNA Isolated

from Lake Sediments. Appl Environ Microbiol. 2011; 77: 8744–8753. https://doi.org/10.1128/AEM.06106-11 PMID: 21984244

67. von Ammon U, Wood SA, Laroche O, Zaiko A, Tait L, Lavery S, et al. Combining morpho-taxonomy and metabarcoding enhances the detection of non-indigenous marine pests in biofouling communities. Sci Rep. 2018; 8: 1–11.